Oversvømmelsesvarsling i Vejle

Prognosticering af vandstande med neurale netværk

Benjamin Refsgaard og Thomas Aagaard Jensen

OVERSVØMMELSESVARSLING I VEJLE

– PROGNOSTICERING AF VANDSTANDE MED NEURALE NETVÆRK

Benjamin Refsgaard og Thomas Aagaard Jensen

KANDIDATSPECIALE

AALBORG UNIVERSITET
STUDENTERRAPPORT

INSTITUT FOR BYGGERI OG ANLÆG
AALBORG UNIVERSITET

Juni 2018

AALBORG UNIVERSITET
STUDENTERRAPPORT

9.-10. semester ved Det Teknisk-Naturvidenskabelig Fakultet
Institut for Byggeri og Anlæg
Thomas Manns Vej 23
9220 Aalborg Øst
http://www.byggeri.aau.dk/

Titel:

Oversvømmelsesvarsling i Vejle

Undertitel:

Prognosticering af vandstande med neurale netværk

Projektperiode:

September 2017 - Juni 2018

Forfattere:

Benjamin Refsgaard
Thomas Aagaard Jensen

Vejledere:

Michael Robdrup Rasmussen
Jesper Ellerbæk Nielsen
Søren Liedtke Thorndahl

Denne rapport er produktet af det arbejde, som er udført på 9. og 10. semester af kanditatuddannelsen i Vand og Miljø under Institut for Byggeri og Anlæg. Med Vejle som projektlokalitet, omhandler rapporten udviklingen af et prognoseværktøj baseret på datadrevet modellering, der kan forudsige vandstande i Grejs Å-systemet, så der kan foretages de nødvendige foranstaltninger mod oversvømmelser.

Prognoseværktøjet optimeres ved forskellige forudsigelsestider, da det viser sig, at fordusigelsestiden er styrende for metodevalget. Når forudsigelsestiden er under en time, fungerer en model baseret på den aktuelle vandstand og den foregående tendens bedst. Herefter er det nødvendigt at kende de styrende processer for vandstanden i danske vandløb, så der kan opsættes neurale netværk, der kan beskrive relationen mellem nedbør og vandstand på en simpel og effektiv måde.

Generelt har neurale netværk vist sig velegnede til at modellere komplekse sammenhænge, hvilket især er anvendeligt til at beskrive vandstandsændringer i Omløbsåen, hvor oversvømmelser historisk set har forvoldt store problemer.

Antal sider i hovedrapport: 105
Appendices: 9
Færdiggjort: 8. juni 2018

Predicting the future isn't magic, it's artificial intelligence.

– Dave Waters

Abstract

This report is the product of the work done by two graduates in the 3rd and 4th semesters of the master in Water and Environmental Engineering at Aalborg University under the Department of Civil Engineering. The report surrounds the danish city of Vejle, where flooding has been a major issue for many years. Because of the flooding issues, the municipality of Vejle has built a sluice and a distribution control structure, which can control the water level in the stream of *Omløbsåen*, which flows through the city centre.

The main objective of this Master's Thesis is to develop data-driven forecast-methods to predict the water levels in Omløbsåen, and in the stream of *"Grejs Å"*, which covers the runoff from most of the catchment area. The forecast-methods are developed at different lead times, where the use of Neural Networks has proven to be very effective for establishing complex relations - for example between measurements of rainfall and water levels, which historically have been described by deterministic modelling of the freshwater system through the use of governing equations.

Predictions of water levels at different lead-times are valuable for a number of reasons. Short-term water level predictions in Omløbsåen have an influence on decision making and real-time control of the sluice and the distribution structure, while long-term water level predictions can be used to warn and prepare the citizens of Vejle against increased runoff from the catchment of Grejs Å.

Analyses show that the best data-driven models for short-term predictions are based on the actual water level and the tendency in the measurements. The municipality desires to keep the sluice open as much as possible, because closure of the sluice will prevent fish from travelling upstream. The short-term predictions can act as a tool to not only determine the closure of the sluice in critical situations, but also to indicate whether it is safe to open the sluice again.

The best models for long-term predictions are also based on the actual water level, but they include the rainfall, as this is an indication to whether the water level is rising or falling. It is concluded that models for long-term predictions, which includes rainfall, performs best, when the catchment consists of mostly a rural environment, where the runoff is linked to natural hydrological processes.

Forord

Denne rapport er udarbejdet af Benjamin Refsgaard og Thomas Aagaard Jensen i perioden september 2017 til juni 2018 som et led i afslutningen af ingeniøruddannelsen for Vand og Miljø, og det er derfor kulminationen af fem års studiearbejde på Institut for Byggeri og Anlæg.

Baggrunden for projektet er et område i Vejle Midtby, der har været plaget af oversvømmelser som følge af forhøjet vandstand i Omløbsåen. Oversvømmelserne har historisk set været forårsaget af stormflod, langvarige/koblede nedbørshændelser og snesmeltning i vandløbsoplande eller kortvarig intens nedbør over byen. Vandstanden i Omløbsåen kan reguleres af et fordelerbygværk i den nordlige del af Vejle og af et nybygget sluseanlæg ved udløbet til Vejle Å, som det søges at styre på bedst mulig vis. Heraf fremkommer ønsket om at forudsige forhøjede vandstande i Omløbsåen på korte og lange tidshorisonter, som kan anvendes til at optimere styringen af fordelerbygværk og sluseanlæg, til at klargøre beredskabet og til at varsle byens borgere.

I projektet er der fokus på anvendelsen af datadrevne modeller til prognoser af vandstande i Vejle Midtby og opland, hvor især brugen af neurale netværk viser sig at være egnet til at beskrive komplekse sammenhænge mellem de forskellige datasæt. Her har modellernes evne til at efterligne målte vandstande været en kilde til gentaget begejstring.

Der rettes stor tak til professor Michael Robdrup Rasmussen, lektor Jesper Ellerbæk Nielsen og lektor Søren Liedtke Thorndahl for altid åbne døre, og for inspirerende og konstruktiv vejledning.

Læsevejledning

Der vil i løbet af rapporten optræde kildehenvisninger efter Harvard-metoden. Kilder vil fremgå som (Efternavn År) i tekst.

Bagerst i rapporten findes litteraturlisten, hvor kilderne er listet i alfabetisk orden efter efternavnet på førstnævnte forfatter.

Relevante appendices fremgår bagerst i rapporten, hvor eksempelvis relevante Matlab-koder fremgår. I de elektroniske bilag ligger de anvendte data.

Figurer og tabeller nummereres efter kapitel og figurnummer i kapitlet, så første figur i kapitel 7 er navngivet figur 7.1, mens den næste vil hedde figur 7.2 og så fremdeles. Alle figurer og tabeller er tilført forklarende tekst, således de kan forstås uafhængigt af brødteksten.

Af benyttede programmer nævnes ArcGIS, QGis, Matlab, Microsoft Excel, MIKE 11 og diverse CAS-værktøjer. Al kortmateriale stammer fra Kortforsyningen. (Kortforsyningen 2016)

Rapporten er udarbejdet af:

Benjamin Refsgaard	Thomas Aagaard Jensen

Indhold

I Del I - Introduktion

1 Indledning 3

1.1 Vejle: *Vadestedet ved vandet* 4

1.2 Regulering af vandstanden i Omløbsåen 7

1.3 Machine Learning - det nye sort 12

1.4 Formål med projektet 13

2 Datapræsentation 15

2.1 Anvendt data 15

2.2 Måleperioder, synkronisering og fejlfinding af data 22

3 Forundersøgelse af data 25

3.1 Styrende processer for vandstanden i de danske vandløb 25

3.2 Forundersøgelse af forhold i Omløbsåen 28

3.3 Forundersøgelse af forhold gældende i Grejs Å ved Ny Grejsdalsvej 32

3.4 Forundersøgelse af forhold i Grejs Å ved Grejsdalens Planteskole 39

II Del II - Modellering med neurale netværk

4 Præsentation af neurale netværk 45
4.1 Anvendelse af neurale netværk som prognoseværktøj 45
4.2 Modellering af sinus-afhængige data med neurale netværk 46
4.3 Sammenfatning af modelleringsprocedure med neurale netværk 53

5 Opsætning af neurale netværk 55
5.1 Anvendelse af neurale netværk til modellering af vandstande 55
5.2 Optimering af netværkskonfigurationer til modellering af vandstande 60
5.3 Sammenfatning på optimering af netværkskonfigurationer 64

III Del III - Vandstandsprognoser i Vejle Midtby

6 Korte forudsigelser 67
6.1 Formålet med korte forudsigelser 67
6.2 Korte forudsigelser af vandstanden i Grejs Å ved Grejsdalens Planteskole 68
6.3 Korte forudsigelser af vandstanden i Grejs Å ved Ny Grejsdalsvej 72
6.4 Korte forudsigelser af vandstanden i Omløbsåen ved Aagade 75
6.5 Anbefaling til styring 84

7 Lange forudsigelser 87
7.1 Formål med lange forudsigelser 87
7.2 Nedbør som input til neurale netværk 88
7.3 Forkastning af Manningtal og potentiel fordampning som input til neurale netværk 91
7.4 Anvendelighed af radardata som input til neurale netværk 92
7.5 Modelkvalitet ved lange forudsigelser 93

IV Del IV - Afrunding

8 Konklusion 97

9 Perspektivering 99

V Appendices

A Udtræk af radardata 109

B Skift mellem sommer- og vintertid 115

C Forsinkelse mellem vandspejlskoter 117

D Eksempel på gengivelse af sinuskurve 123

E Opsætning af neurale netværk i MATLAB 131

F Forudsigelse på baggrund af kurveforløb 149

G Modelresultater 153

H NAR - et alternativ til tendens-modellen 167

I Radardata over Vejle Midtby 169

J Tolkning af data fra fordelerbygværket 173

Del I - Introduktion

1 Indledning 3

1.1 Vejle: *Vadestedet ved vandet*
1.2 Regulering af vandstanden i Omløbsåen
1.3 Machine Learning - det nye sort
1.4 Formål med projektet

2 Datapræsentation 15

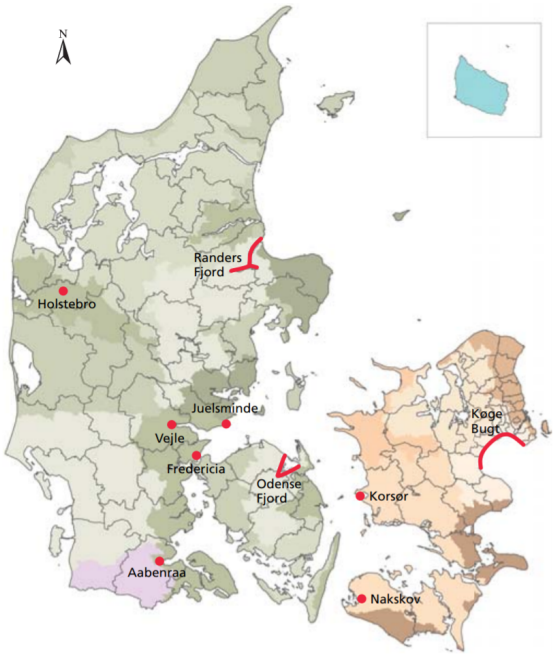
2.1 Anvendt data
2.2 Måleperioder, synkronisering og fejlfinding af data

3 Forundersøgelse af data 25

3.1 Styrende processer for vandstanden i de danske vandløb
3.2 Forundersøgelse af forhold i Omløbsåen
3.3 Forundersøgelse af forhold gældende i Grejs Å ved Ny Grejsdalsvej
3.4 Forundersøgelse af forhold i Grejs Å ved Grejsdalens Planteskole

1. Indledning

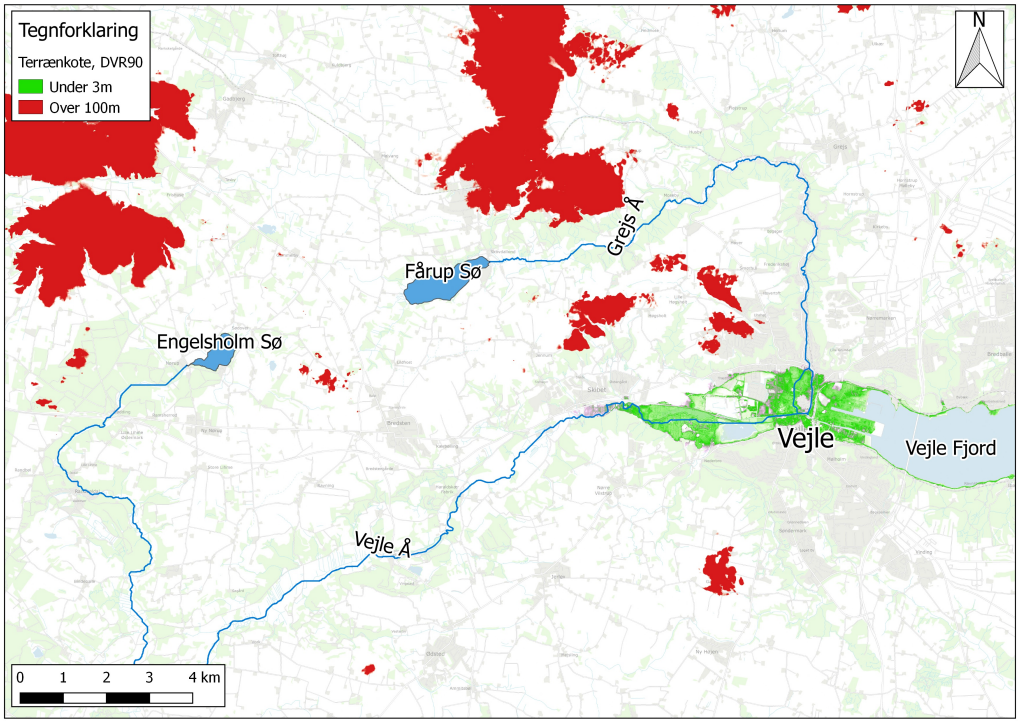
I årene 1998-2002 oplevede Centraleuropa flere ekstreme oversvømmelsessituationer. Denne periode danner grundlag for EU's oversvømmelsesdirektiv, som har pålagt alle medlemslande at planlægge for ekstreme oversvømmelser, som kan medføre væsentlige negative følger for sundhed, miljø, kulturarv eller økonomi. I Danmark har Naturstyrelsen og Kystdirektoratet fået ansvaret for planlægningen og har på baggrund af en screening af potentielle konsekvenser udpeget byen Vejle som et af 10 afgrænsede risikoområder for ekstreme oversvømmelser, som vist på figur 1.1. (Naturstyrelsen og Kystdirektoratet 2011)



Figur 1.1: Vejle er udvalgt som ét af ti afgrænsede risikoområder for ekstreme oversvømmelser i Danmark. (Naturstyrelsen og Kystdirektoratet 2011)

1.1 Vejle: *Vadestedet ved vandet*

Vejle er en by i den sydøstlige del af Jylland. Denne del af Jylland er et af de mest kuperede områder af Danmark, og byen er placeret i et lavtliggende knudepunkt mellem Grejs Å mod nord, Vejle Å mod vest og Vejle Fjord mod øst, hvilket ses på figur 1.2. Navnet *Vejle* kommer af det olddanske ord *wæthel*, som betyder vadested, hvilket i vikingetiden blev betragtet som en passende karakteristik af området. (Trap 1904)



Figur 1.2: Vejle ligger som knudepunkt mellem viste vandløb og Vejle Fjord i et område, hvor terrænet skråner kraftigt mod byen.

Som et lavtliggende knudepunkt i terrænet har Vejle historisk set været usædvanligt hårdt ramt af oversvømmelser, fordi byen både er oversvømmelsestruet fra fjorden og fra vandløbene, der løber gennem midtbyen. De ældste eksempler fra år 1924 og 1941 er nærmere beskrevet i et uddrag fra Vandløbsbogen, se figur 1.3. Af teksten fremgår det, at hændelserne forekom ved høje vandføringer i vandløbene forårsaget af en kombination af snesmeltning i det bakkede opland og kraftig nedbør.

I April 1924 havde de ommeldte Vandløb foraarsaget en alvorlig Oversvømmelse, der endda med Afbrydelser varede i 3 Dage, ligesom der i 1926 havde været ret betydelige Oversvømmelser.

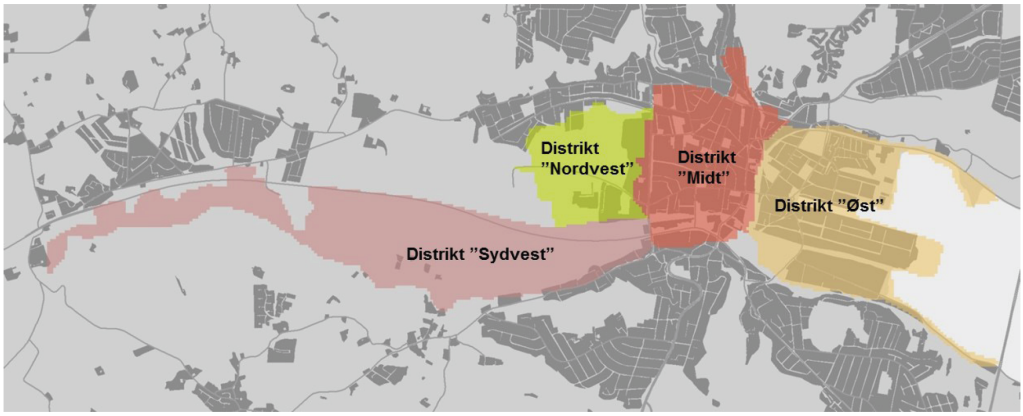
Oversvømmelserne er hidtil fremkommet derved, at der paa Bakkerne i de nævnte Vanløbs Opland, et Areal paa ca. 8150 ha. samles betydelige Mængder Is og Sne, der ved pludselig Stigning af Temperaturen - navnlig af Dagtemperaturen - hurtigt smelter, og i Forbindelse med stærk Nedbør sætter betydelige Vandmasser i Gang. Den maximale Vandføring skal saaledes i 1924 have være 21 m^3/Sek. og i 1941 32 m^3/Sek.

Figur 1.3: Uddrag fra Vandløbsbogen som beskriver en alvorlig oversvømmelse i 1924 og 1941. (Vejle Byråd 2014)

Et andet eksempel fra 1973 viser, at det ikke altid har været snesmeltning og kraftig nedbør, der har forårsaget oversvømmelser, men at også en høj vandstand i Vejle Fjord alene kan være årsag til oversvømmelse i Vejle. På figur 1.4 ses en bus, der pløjer gennem vandmasserne langs kysten.

Figur 1.4: I 1973 var der oversvømmelse i den østlige del af byen forårsaget af høj vandstand i Vejle Fjord. (Vejle Stadsarkiv 2018)

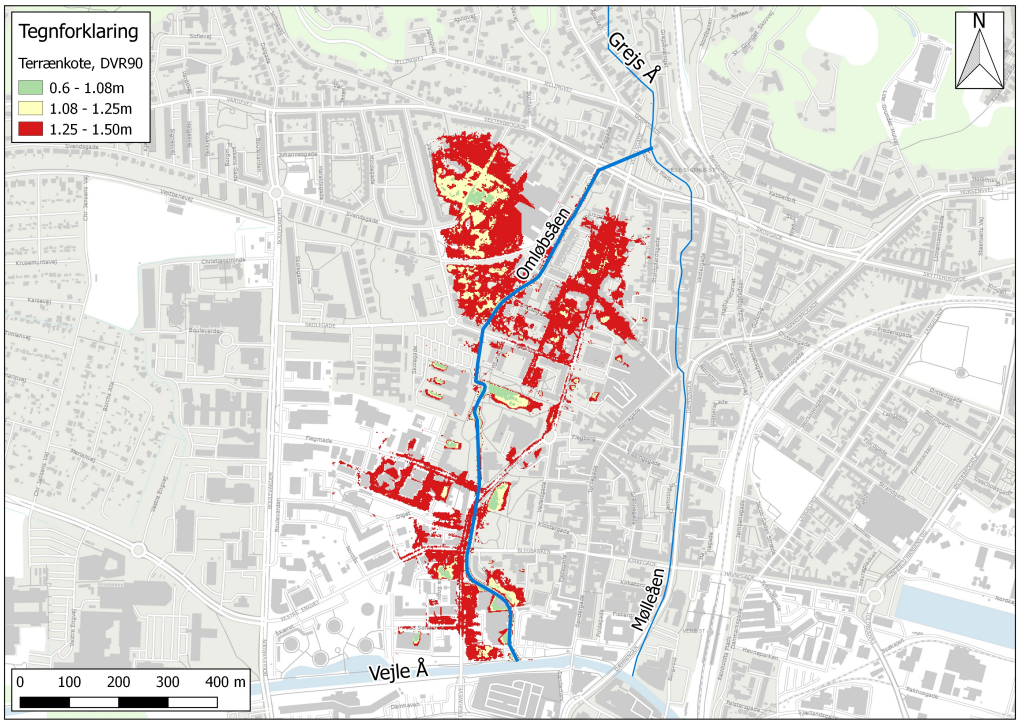
Historien afslører, at kilden til oversvømmelserne ikke altid er den samme. Det gør det særligt problematisk at kortlægge de potentielle konsekvenser, fordi omfanget og udbredelsesmønstret er afhængigt af kilden. I særlige tilfælde er beboerne så uheldige, at byen både oversvømmes fra fjord og vandløb på samme tid. Derfor har Vejle Kommune udarbejdet en risikostyringsplan i 2014, som analyserer konsekvenserne på baggrund af statistiske scenarier for vandføringer i vandløbene og vandstande i fjorden. Deraf er der blevet udpeget fire separate distrikter, som ses på figur 1.5. (Vejle Byråd 2014)



Figur 1.5: Der er udpeget fire distrikter i risikostyringsplanen på baggrund af statistiske scenarier. (Vejle Byråd 2014)

Af risikostyringsplanen fremgår det, at *Distrikt "Midt"* er det mest udsatte, da området oplever oversvømmelser ved samtlige scenarier. I den nordlige del af distriktet deles Grejs Å i de mindre vandløb Omløbsåen og Mølleåen, hvorefter de løber hver sin vej rundt om bymidten og slutteligt ud i Vejle Å. Særligt området langs Omløbsåen fremhæves i risikostyringsplanen som problematisk. (Vejle Byråd 2014)

Det ses på figur 1.6, at terrænnet langs Omløbsåen nogle steder er i niveau med den daglige vandstand i åen, hvilket er markeret med grønt. Det hænder også, at vandspejlskoten i Omløbsåen overstiger koterne markeret med gul og rød, hvor rød symboliserer højere koter end brinkkanten midt på strækningen af vandløbet. I disse tilfælde er der risiko for påvirkning af større dele af midtbyen.



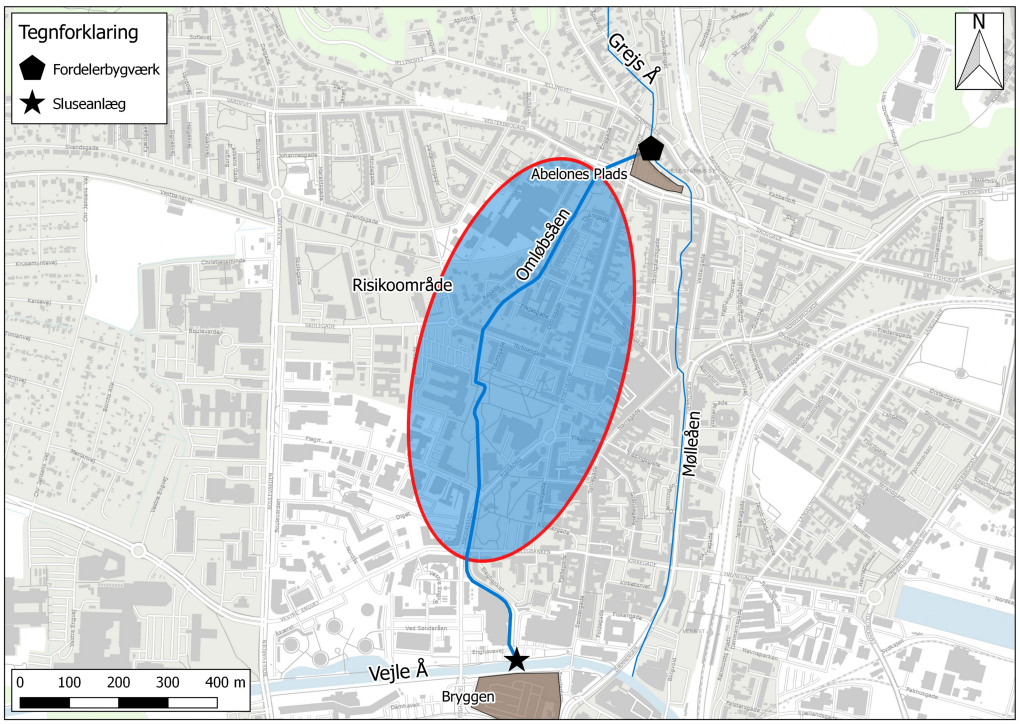
Figur 1.6: Når vandspejlskoten i Omløbsåen når en kote, som er højere end brinkkanten, er store dele af Vejle Midtby oversvømmelsestruet.

Omløbsåen strømmer gennem et område med lav og tæt bebyggelse, og det fremgår af risikostyringsplanen, at de økonomiske konsekvenser på denne strækning er den største på tværs af alle de fire distrikter (Vejle Byråd 2014). Det ses af figur 1.7, at flere tilstødende ejendomme har kælder. Enkelte steder på strækningen langs Omløbsåen er der også anlagt rekreative grønne arealer, hvor oversvømmelser ikke har samme konsekvenser.

Figur 1.7: Flere af husene langs Omløbsåen er bygget med kælder (til venstre). Andre steder er der anlagt grønne arealer, som eksempelvis i den nordlige del af vandløbet (til højre).

Risikoområdet langs Omløbsåen var ligeledes ramt under de føromtalte hændelser i 1924, 1941 og 1973, og der forefindes også nyere eksempler, som bekræfter, at det ikke for alvor er lykkedes at løse oversvømmelsesproblemerne.

Med afsæt i risikostyringsplanen har Vejle Kommune udarbejdet en klimatilpasningsplan, hvor der er lavet forslag til indsatsområder forskellige steder i kommunen. Her figurerer også en prioritering af indsatser, hvor risikoområdet på figur 1.8 er angivet som 1. prioritet. Indsatserne er inddelt i fysiske, undersøgende og strategiske tiltag, hvor konkrete fysiske tiltag allerede er udført i form af et fordelerbygværk ved Abelones Plads og et sluseanlæg ved Bryggen. (Vejle Kommune 2014)



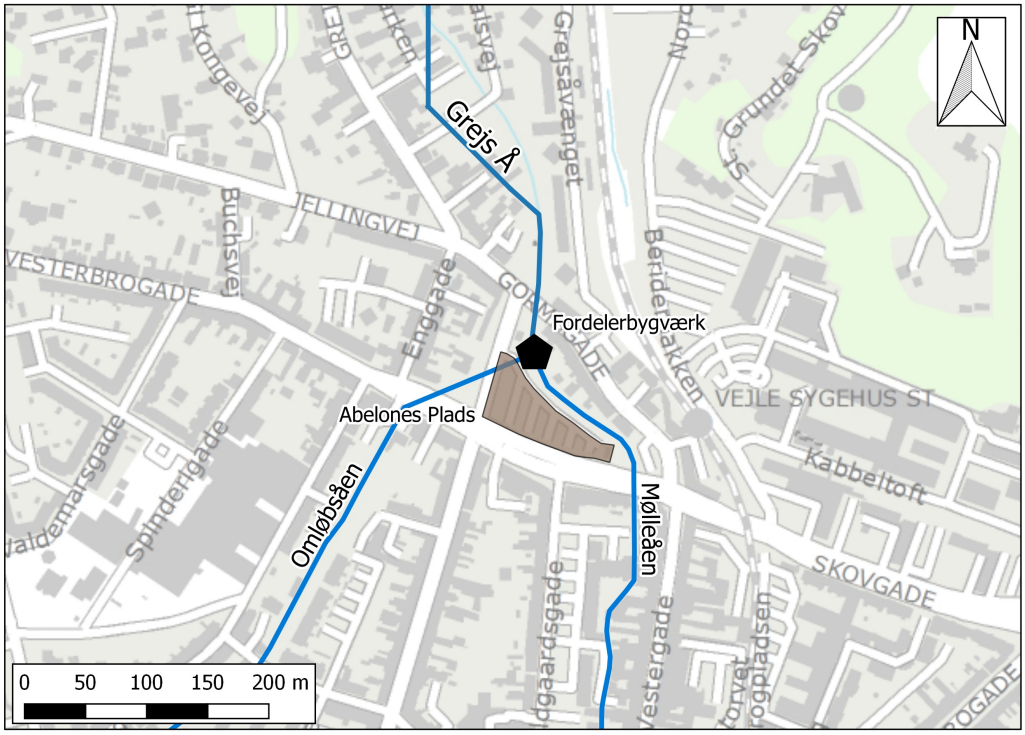
Figur 1.8: De nye tiltag for reducering af oversvømmelsesrisiko i risikoområdet rummer et fordelerbygværk ved Abelones Plads og et sluseanlæg ved Bryggen. (Vejle Kommune 2014)

Fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen har til opgave at regulere vandstanden i Omløbsåen, når der er risiko for oversvømmelse.

1.2 Regulering af vandstanden i Omløbsåen

1.2.1 Fordelerbygværket ved Abelones Plads

I den nordlige del af Vejle Midtby ved Abelones Plads deles Grejs Å i Mølleåen og Omløbsåen via et fordelerbygværk, som vist på figur 1.9. Fordelerbygværkets funktion er at kunne styre vandtilstrømningen fra Grejs Å, hvilket især er brugbart ved kraftig nedbør i oplandet.



Figur 1.9: Grejs Å deles i Mølleåen og Omløbsåen ved Abelones Plads i den nordlige del af risikoområdet.

Fordelerbygværket udgøres af tre justerbare spjæld, der kan regulere vandtilstrømningen til Omløbsåen. Når de lukkes i, stuver vandet op og mere vand vil løbe over i Mølleåen. Før spjældene er der monteret riste, som fanger flydegods.

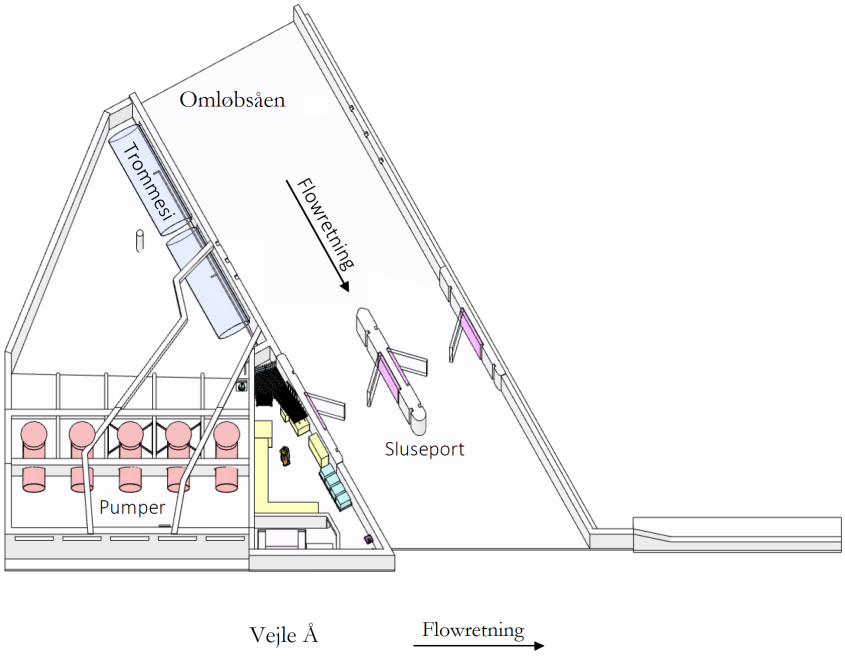
Figur 1.10: Vandet fordeles mellem Mølleåen (til venstre) og Omløbsåen (til højre) af fordelerbygværket, som består af tre spjæld bag ristene til højre.

1.2.2 Sluseanlægget ved Bryggen

Som en konsekvens af de mange oversvømmelser valgte Vejle Kommune i samarbejde med Vejle Spildevand A/S at etablere et sluseanlæg (figur 1.11), der danner spydspids i projektet *"Fjordbyen - det gode liv og møde med vandet"* (Vejle Spildevand A/S 2017) (Faarup 2016). Sluseanlægget er placeret ved Omløbsåens udmunding i Vejle Å, hvilket illustreres på figur 1.8.

Figur 1.11: Sluseanlægget fra modsatte side af Vejle Å ved Bryggen.

Ved kraftig vandtilstrømning fra Grejs Å kan sluseanlægget kunstigt sænke vandstanden i Omløbsåen ved at pumpe vandet fra vandløbet og ud i Vejle Å. Modsat fordelerbygværket, som fordeler vandføringen mellem Omløbsåen og Mølleåen, kan sluseanlægget øge kapaciteten af Omløbsåen. Sluseportene kan desuden lukke i og afskærme Omløbsåen for stormflod fra Vejle Fjord.



Figur 1.12: Redigeret systemtegning af sluseanlægget. (Vejle Kommune 2015b)

Når sluseanlægget er i kraft, reduceres sandsynligheden for oversvømmelser langs Omløbsåen, men samtidig blokeres naturlige strømningsforhold, hvilket forhindrer fisk i at vandre op i systemet.

1.2.3 Fiskene i fokus

Nord for Abelones Plads består Grejs Å af omtrent 80 km opstrøms vandløbsstrækninger, hvor vandkvaliteten er god, og der er gode gydeforhold for havørreder og andre laksefisk. Siden den sydlige del af Mølleåen blev rørlagt i 1931 (figur 1.13), har Omløbsåen fungeret som den bedste fiskepassage til Grejs Å. (Vejle Kommune 2015a)

Figur 1.13: Den sydlige del af Mølleåen rørlægges i 1931. (Vejle Stadsarkiv 2018)

Der er fra Vejle Kommunes side gjort tiltag for, at disse forhold skulle bevares og i nogle tilfælde forbedres. Således blev der i sommeren 2015 fjernet et 9 m bredt stemmeværk ved Abelones Plads, der var blevet udpeget af staten i en tidligere vandplan som en spærring mellem Omløbsåen og Grejs Å. Stemmeværket havde et fald på op til én meter, hvilket blev udlignet ved en restaurering af vandløbsbunden, hvor der blev anlagt et stryg over en strækning på 90 m. Stryget er anlagt med et jævnt fald for at sikre gunstige strømningsforhold, hvilket ses på figur 1.14 til venstre.

Figur 1.14: Stemmeværket blev i 2015 erstattet af et stryg over en strækning på 90 m nord for Abelones Plads for at tilgodese fiskevandring fra eksempelvis havørreder.

Desuden blev der i år 2004 genslynget en strækning i den øvre del af Grejs Å i et område, der nu fungerer som et hyppigt besøgt område for lystfiskere (Naturstyrelsen 2006) (Vejle Sportsfiskerforening 2011).

Disse tiltag viser tydeligt, at Vejle Kommune har interesse i at bevare gode forhold for fisk i Grejs Å og samtidig reducere risikoen for oversvømmelse i midtbyen, hvilket kræver en samlet styringsstrategi af fordelerbygværk og sluseanlæg.

1.2.4 Samlet styringsstrategi og udvikling af prognoseværktøjer

Fordelerbygværket og sluseanlægget udgør de vigtigste forudsætninger for at undgå oversvømmelser langs Omløbsåen. Hvis ambitionen er helt at undgå oversvømmelser i risikoområder, kan styring hypotetisk foregå på følgende ekstreme måder:

- Fordelerbygværkets tre spjæld lukkes i, så alt vand ledes i Mølleåen ⇒ Mølleåen oversvømmer, og problemet er dermed skubbet til et andet område.
- Sluseportene lukkes, og pumperne afvander Omløbsåen unaturligt ⇒ Fisk kan ikke søge op i systemet, og alle tidligere projekter er betydningsløse. Det er desuden en fordel økonomisk at have pumperne til at køre så sjældent som muligt.

Ingen af disse løsninger er at foretrække, og derfor søger Vejle Kommune at lave en samlet styringsstrategi mellem fordelerbygværket og sluseanlægget, som:

- Holder risikoområdet fri af oversvømmelser.
- Sørger for at udnytte magasineringskapaciteten i Omløbsåen og Mølleåen optimalt.
- Holder sluseportene åbne så vidt muligt, så fisk frit kan passere.

Siden indvielsen af sluseanlægget i september 2016 har ønsket været at udvikle et prognoseværktøj, der skal kunne forudsige høje vandstande, så styringen af fordelerbygværket og sluseanlægget kan optimeres, hvilket har udmøntet sig i projektet *SMART Vand Vejle*, som omfatter udviklingen af et realtids varslings- og styringssystem af vandløbene i Vejle (Miljøstyrelsen 2018). Varslingssystemet skal give Højtvandsgruppen, som består af repræsentanter fra Vejle Kommunes beredskab, vandløbsmyndighederne, vejchefen og Vejle Spildevand A/S, et bedre overblik og beslutningsgrundlag i kritiske situationer, samt varsle og forberede den enkelte borger. (Vejle Spildevand A/S 2017)

Parterne i projektet *SMART Vand Vejle* er Vejle Spildevand A/S, Dansk Hydraulisk Institut (DHI og Aalborg Universitet. Projektet omfatter udvikling af en vidtgående deterministisk model til modellering og prognosticering af vandstand i vandløbene, som strømmer gennem Vejle Midtby. Denne model omfatter en kobling af følgende submodeller:

- Grundvandsmodel over vandløbsoplandene (MIKE SHE).
- Afstrømningen i vandløbene (MIKE 11).
- Interaktion mellem vandløb, afløbssystemet i Vejle Midtby og Vejle Fjord (MIKE URBAN).

De tre modeller fungerer som randbetingelser for hinanden og baseres på målinger og prognoser af nedbør, temperatur, potentiel evapotranspiration og vandstanden i Vejle Havn. For at den samlede model kan fungere i realtid er det nødvendigt at have en model, som hurtigt kan afvikles, når nye målinger forefindes. Dette har i henhold til projektets slutrapport medført, at det har været nødvendigt med forsimplinger af især MIKE URBAN-modellen. (Miljøstyrelsen 2018)

Denne sammensætning af modeller er den naturlige og logiske indgangsvinkel for de fleste hydraulikere, men kompleksiteten og skalaen medfører mange ukendte variabler, hvilket bidrager til en model, som formentlig både er vanskelig og tidskrævende at få til at stemme overens med virkeligheden. Udfordringerne bliver ikke mindre af, at det hele skal foregå i realtid. Derfor forsøges det i dette afgangsprojekt at gå nye veje og takle udfordringerne på en alternativ måde.

1.3 Machine Learning - det nye sort

Deterministiske modeller har i mange år været grundpillen i styring og dimensionering af afløbssystemer eller klimaløsninger. En alternativ vej, der er forholdsvist ny på vandområdet, men efterhånden har været kendt i en del år, er *Machine Learning*. (Samuel 1959)

Machine Learning er et felt indenfor computervidenskab, der bygger på ideen om, at en *maskine* er i stand til at opbygge en kunstig intelligens gennem *læring*. Læringen sker ved, at maskinen præsenteres for store mængder data, og maskinen forsøger at finde sammenhæng i det præsenterede data. Maskinen vil herefter forhåbentligt på baggrund af sammenhængen i det præsenterede data være i stand til at se sammenhænge i nye datasæt, som den i realiteten aldrig har set før.

I ordet *læring* ligger også, at maskinen bliver klogere og klogere efterhånden, som datamængden vokser, da det nye data løbende kan adapteres. Det vil sige, at maskinen kan tilpasse sig ændringer i det hydrologiske system i form af eksempelvis kloakseparering, etablering af forsinkelsesbassiner eller klimaforandringer, så længe maskinen præsenteres for ny data. Dette står i kontrast til de traditionelle deterministiske modeller, der kræver manuelle opsætninger, klimafremskrivninger på nedbør og løbende genkalibreringer.

I dette projekt er fokus rettet mod en gren indenfor Machine Learning kaldet *Artificial Neural Networks*, herefter neurale netværk eller NN. Som ved alle typer af modeller er der ved neurale netværk både fordele og ulemper. Neurale netværk:

+ Kan beskrive mønstre mellem data meget nøjagtigt.
+ Bliver klogere og klogere efterhånden som datagrundlaget øges.
+ Kan tilpasse sig ændringer automatisk gennem præsentation af ny data.
+ Kan opsættes uden det store forarbejde, hvis passende mængder data er tilgængelig.
+ Er ikke afhængige af ikke-varierende fysiske parametre.
+ Er beregningseffektive og egner sig derfor godt til realtidsmodellering.
+ Kræver ikke licens til kommercielt software.

÷ Kræver en passende mængde data til oplæring af netværket.
÷ Mister hurtigt korrekthed, hvis der skal beskrives situationer, som ligger udenfor spændet af præsenteret data.
÷ Kan potentielt beskrive mønstre i præsenteret data for nøjagtigt, så den generelle sammenhæng mistes.
÷ Vil forsøge at finde sammenhænge i data, selvom de ikke findes.
÷ Indlærer fejl i data.
÷ Oversætter et signal til et andet og har dermed ingen hukommelse.
÷ Har en ugennemsigtig beregningsprocedure, og det kan derfor være svært at fejlfinde.

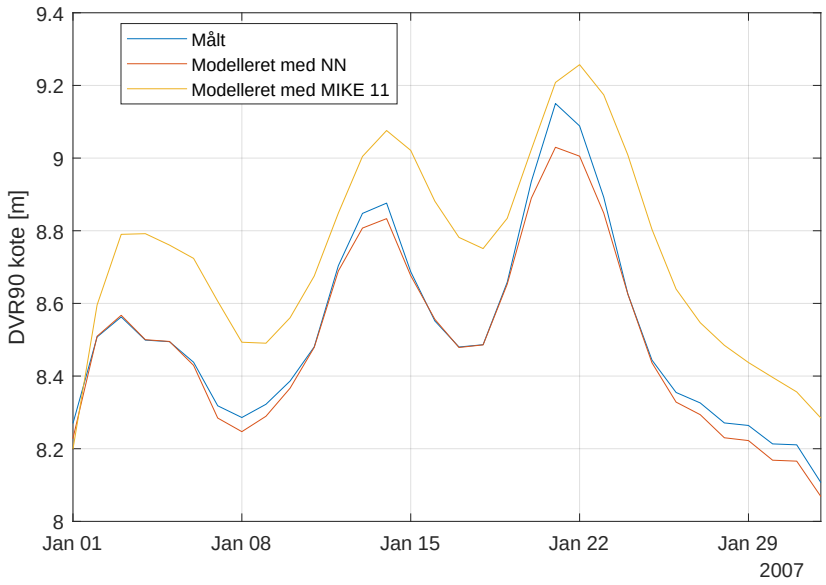
I følgende afsnit vises et simpelt eksempel, der skal illustrere nogle af fordelene ved neurale netværk, og samtidig antyder anvendeligheden indenfor modellering af vandstande.

1.3.1 Eksempel: Modellering af vandstande i Storå

På 2. semester af masteruddannelsen i Vand og Miljø blev der udfærdiget et semesterprojekt omhandlende klimasikring af midtbyen i Holstebro. I programmet MIKE 11 blev der lavet en numerisk 1D-strømningsmodel af Storå, som løber gennem Holstebro. Formålet med denne model var at simulere døgnmiddelvandstande i midtbyen. Modellen havde input i form af vandføringen ved en måler placeret opstrøms i Storå og en distribueret oplandsafstrømning på strækningen mellem opstrøms måler og Holstebro. (Achton-Boel m.fl. 2017)

Der blev lagt et stort arbejde i at kalibrere modellen op imod vandstandsmålinger fra en måler placeret i Holstebro Midtby. Kalibreringen skete ved at skrue på Manningtallet både stedligt og tidsligt, så modelresultaterne kom tættere på måledata.

At omdanne et bestemt input til et bestemt output (i dette tilfælde vandføringer til vandstand) er netop styrken ved et neuralt netværk, hvilket er illustreret på figur 1.15. Der er her vist resultatet af et neuralt netværk, der har fået samme input som den anvendte MIKE 11-model, og som uden yderligere arbejde kan give et resultat, der ligger markant tættere på måledata.



Figur 1.15: Det neurale netværk formår uden stort forarbejde at modellere døgnmiddelvandstande, der stemmer bedre overens med de målte værdier end MIKE 11-modellen.

I det tidligere semesterprojekt skulle den numeriske model anvendes til at modellere scenarier i fremtiden, hvor vandføringen i Storå forventes at stige grundet klimaforandringer. Enkelte af disse hændelser ville i så fald have en størrelse på vandføringen, der hidtil ikke forekommer målinger på, og her vil resultatet af brugen af det neurale netværk være meget usikkert. Der skulle også modelleres scenarier, hvor tværsnittet af åen blev udvidet. Dette ville på tilsvarende vis ikke kunne modelleres med det opstillede neurale netværk, da netværket ikke ville have data på eksempelvis tidligere tværsnitsændringer. På trods af disse begrænsninger, er et neuralt netværk effektivt og præcist til at gengive måledata med stor nøjagtighed.

1.4 Formål med projektet

Det er vist i eksemplet, at der foreligger et potentiale i at modellere vandstande med neurale netværk. Formålet med dette projekt er at eftervise potentialet ved at udvikle datadrevne modeller, der kan prognosticere vandstande i Vejle. Det ønskes at gøre dette flere steder i Grejs Å og Omløbsåen på forskellige tidsskalaer, således den fælles styring af fordelerbygværket og sluseanlægget kan optimeres, og beredskab og borgere kan alarmeres. Af dette formål udspringer problemformuleringen:

"I hvilket omfang egner datadrevne modeller sig til at prognosticere vandstande i Vejle, og hvorledes kan dette anvendes til at forbedre beslutningsgrundlaget for styringen af fordelerbygværk og sluseanlæg?"

Bag problemformuleringen gemmer sig en række underspørgsmål, som det ønskes at besvare løbende gennem rapporten for at bevare fokus og beskrive hvert enkelt analyses bidrag til den samlede besvarelse. Underspørgsmålene lyder som følgende:

- Hvilke processer er overordnet set styrende for vandstanden i danske vandløb?

- Hvilke forhold har indflydelse på vandstanden i Grejs Å og Omløbsåen?
- Hvordan optimeres netværkskonfigurationerne af de neurale netværk til målrettet at kunne modellere vandstande i Grejs Å og Omløbsåen?
- Hvordan kan datadrevne modeller på bedst mulig vis anvendes til at forudsige vandstande i Grejs Å og Omløbsåen på korte og lange tidshorisonter?
- Hvorledes kan modellering af vandstande i Omløbsåen anvendes til at forbedre beslutningsgrundlaget for styringen af fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen?

1.4.1 Strategi for projektarbejdet

Fundamentet for vandstandsprognoserne er datadrevne modeller, og det erfares, at der uden det store forarbejde kan opsættes gode modeller. De fremragende modeller kræver derimod et godt indblik i de målte tidsserier, og derfor afsættes første del af rapporten til præsentation af det anvendte data og forundersøgelser af data fra forskellige målestationer i Vejle Midtby og opland. Forundersøgelserne skal medvirke til en bedre forståelse af de styrende processer for vandstanden i Grejs Å og Omløbsåen og på den måde forbedre mulighederne for at opbygge neurale netværk, der har har en praktisk anvendelighed.

Anden del af rapporten omhandler en præsentation af fremgangsmåden for modellering med neurale netværk. Først præsenteres grundprincipperne bag modellering med neurale netværk for at give læseren en forståelse af modelleringsproceduren. Dernæst beskrives opsætningen af neurale netværk rettet specifikt mod at optimere netværkskonfigurationerne, så modellerne målrettes prognosticering af vandstande i Vejle.

Tredje del af rapporten fokuserer på prognoser af vandstande, som skal anvendes til en optimeret styringsstrategi af fordelerbygværk og sluseanlæg og varsling mod forhøjet afstrømning fra opland. Prognoserne udføres med afsæt i forundersøgelsen af data og de optimale netværkskonfigurationer. Prognoser af vandstande er i denne del af rapporten delt i korte og lange forudsigelsestider, da det viser sig, at de optimale netværkskonfigurationer varierer alt efter hvilken forudsigelsestid, der arbejdes med.

Opsummeret i punktform kan strategien for projektarbejdet beskrives således:

- Første del: Forundersøgelser og præsentation af det anvendte data, som skal bidrage til bedre forståelse af de styrende processer for vandstanden i Grejs Å-systemet.
- Anden del: Præsentation af modelleringsproceduren med neurale netværk og en beskrivelse af, hvordan netværkskonfigurationer kan målrettes modellering af vandstande.
- Tredje del: Anvendelse af datadrevne modeller til at lave prognoser af vandstande i Grejs Å og Omløbsåen ved korte og lange forudsigelsestider.

Til slut samles de tre dele i en afrunding, hvor projektarbejdes hovedkonklusioner fremgår.

2. Datapræsentation

En af de vigtigste forudsætninger for at lave modellering med neurale netværk er, at der er et passende datagrundlag. I Vejle forefindes historiske data om nedbør, vandstande og andre tekniske forhold såsom åbningsstatus på sluseanlægget. I dette kapitel præsenteres det anvendte data. I kapitlet beskrives placeringerne af målerne, hvilken tidsperiode der haves data i, og hvilke referencer der anvendes, når der gennem rapporten henvises til en måler.

2.1 Anvendt data

I Vejle Midtby og opland forefindes data, som eksempelvis beskriver geologi, topografi, og nedbør. Der er som en naturlig del af processen udvalgt data, som er fundet anvendeligt i projektet.

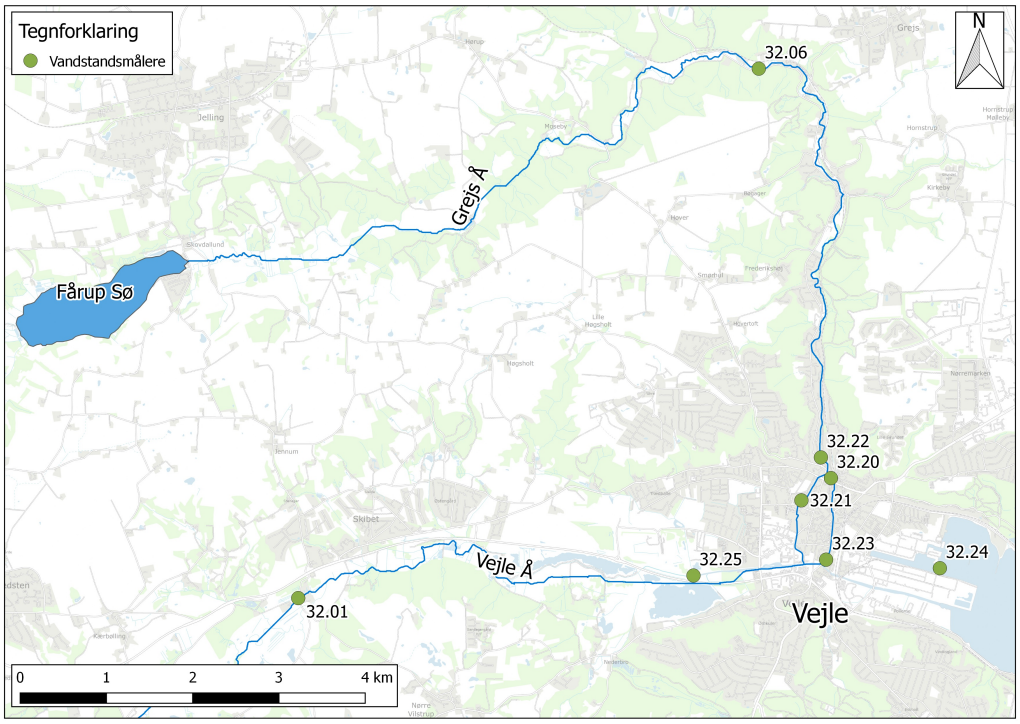
- Vandstandsmålinger:
  - I Vejle Midtby og opland.
  - I Vejle Fjord.
  - Ved fordelerbygværk og sluseanlæg.
- Målinger af åbningsstatus på spjældene ved fordelerbygværket og sluseportene ved sluseanlægget.
- Målinger af nedbør i form af:
  - SVK-målere.
  - Radarsignaler

Formålet med dette kapitel er at kortlægge placeringen af målerne, danne et overblik over måleperioderne og identificere eventuelle fejl. Det er relevant, blandt andet fordi systematiske fejl indlæres i de neurale netværk.

2.1.1 Vandstandsmålere

Langs vandløbene i området omkring Vejle er der placeret en række vandstandsmålere. I de øvre dele af vandløbene er der placeret enkelte målere, mens tætheden er større i midtbyen, hvilket formentlig skyldes et ønske om præcist at kunne sige, hvad vandstanden er i midtbyen på forskellige steder på

forskellige tidspunkter - eventuelt til styring af afløbssystemet. Placeringen af vandstandsmålerne ses på figur 2.1.



Figur 2.1: Kortet viser Grejs Å, Omløbsåen, Mølleåen og et udpluk af Vejle Å samt de vandstandsmålere, der er placeret langs åløbene. Hver målestation følger navngivningen fra hjemmesiden www.hydrometri.dk, drevet af Orbicon.

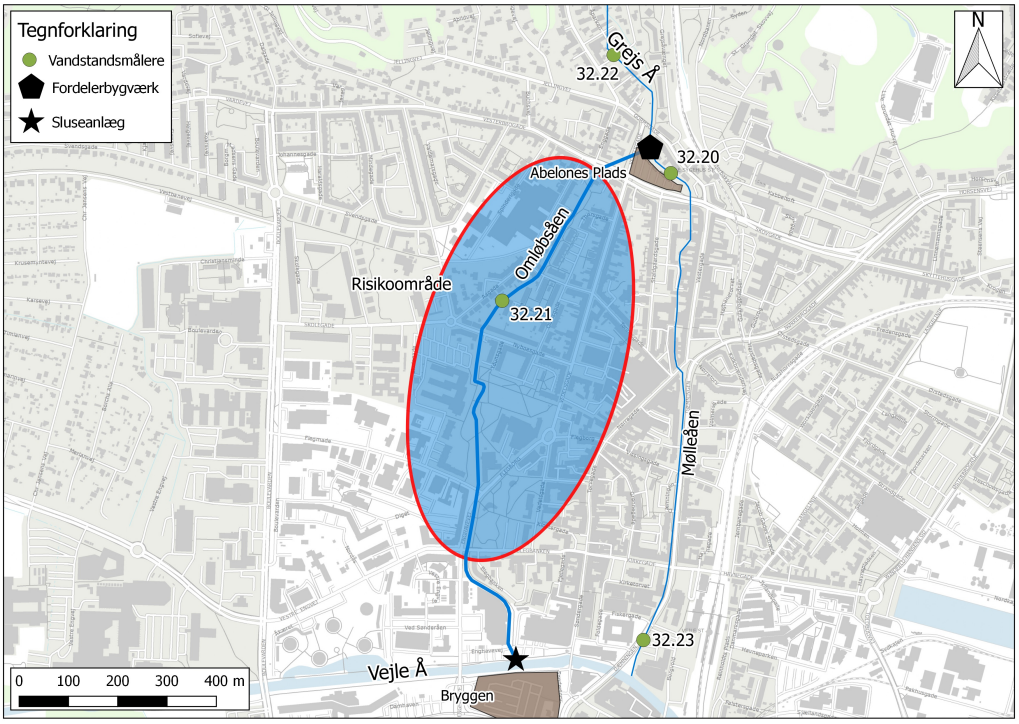
I tabel 2.4 er angivet informationer om vandstandsmålerne, og hvilke perioder der anvendes data fra. De angivne referencer vil løbende igennem rapporten fungere som betegnelse, når en målestation omtales.

Tabel 2.1: Tæt på Vejle forefindes data for vandstande fra i alt otte vandstandsmålere.

Reference	Periode	Lokation
ST. 32.01	01/12-2015 →	Haraldskær, Vejle Å
ST. 32.06	01/12-2015 →	Grejsdalens Planteskole, Grejs Å
ST. 32.20	01/12-2015 →	Abelones Plads, Mølleåen
ST. 32.21	01/12-2015 →	Aagade, Omløbsåen
ST. 32.22	01/12-2015 →	Ny Grejsdalsvej, Grejs Å
ST. 32.23	01/12-2015 →	Dæmningen, Mølleåen
ST. 32.24	06/27-2016 →	Vejle Havn, Vejle Fjord
ST. 32.25	01/11-2016 → 22/01-2017	Rosborg, Vejle Å

2.1.2 Andre målere

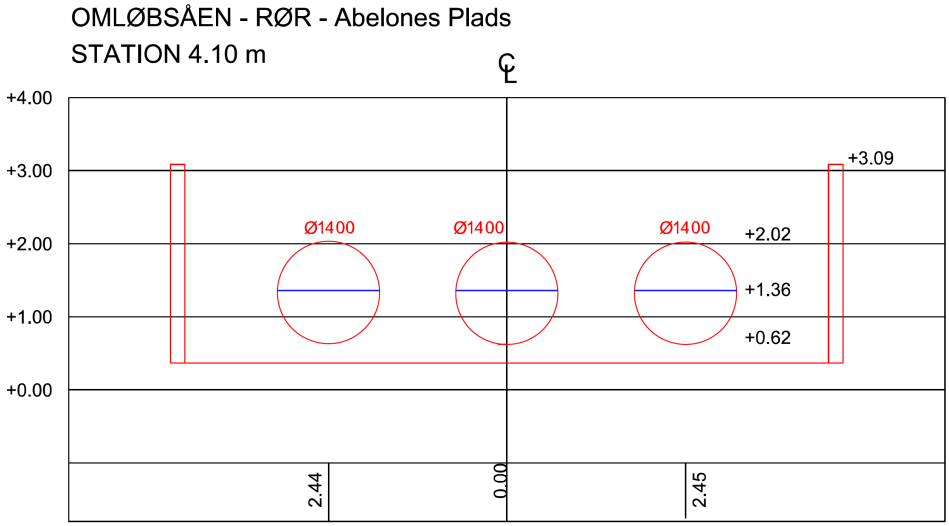
Udover vandstandsmålerne i vandløbene i Vejle midtby er der placeret en række ekstra sensorer ved fordelerbygværket og sluseanlægget. Den nærmere placeringen af vandstandsmålerne i Vejle midtby og lokationen af fordelerbygværk og sluseanlæg ses på figur 2.2. Det ses, at ST. 32.21 ligger centralt i risikoområdet midt i Omløbsåen.



Figur 2.2: I Vejle midtby er tætheden af vandstandsmålere højere end i øvrige dele af systemet.

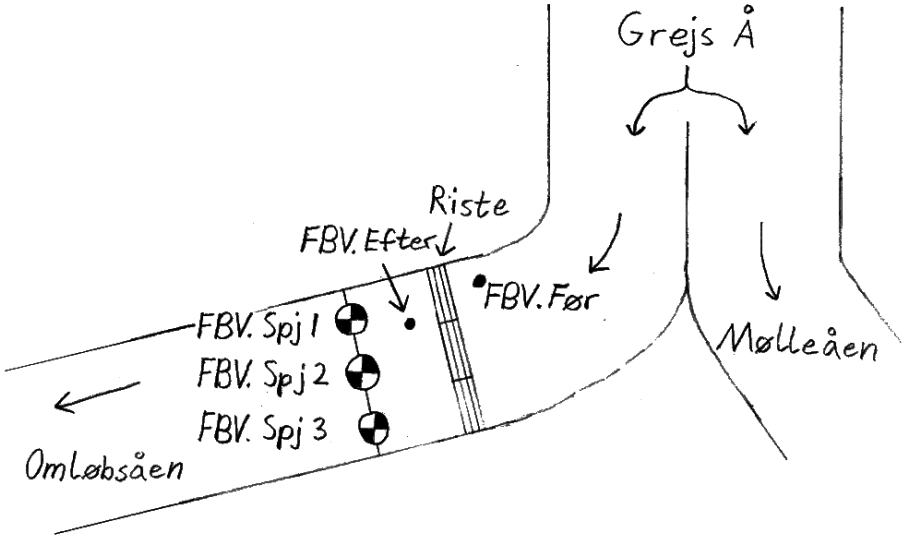
Målere ved fordelerbygværket ved Abelones Plads

Fordelerbygværket ved Abelones Plads har stor indvirkning på vandstanden i Omløbsåen såvel som i Mølleåen. Fordelerbygværket består af tre spjæld, som individuelt kan åbnes og lukkes til forskellige niveauer. På figur 2.3 ses et snit af fordelerbygværkets tre spjæld.



Figur 2.3: Fordelerbygværkets tre spjæld med koter og afstand fra centerlinje. I kote 3,09 løber vandet på terræn. (LIFA A/S Landinspektører 2014)

Ved fordelerbygværket måles vandstanden før og efter ristene, som begge måles før spjældene. Derudover er der data for åbningsprocenten for hvert af de tre spjæld. På figur 2.4 ses en skitsering af sensorernes placering ved fordelerbygværket.



Figur 2.4: Skitsering af sensorernes placering ved fordelerbygværket.

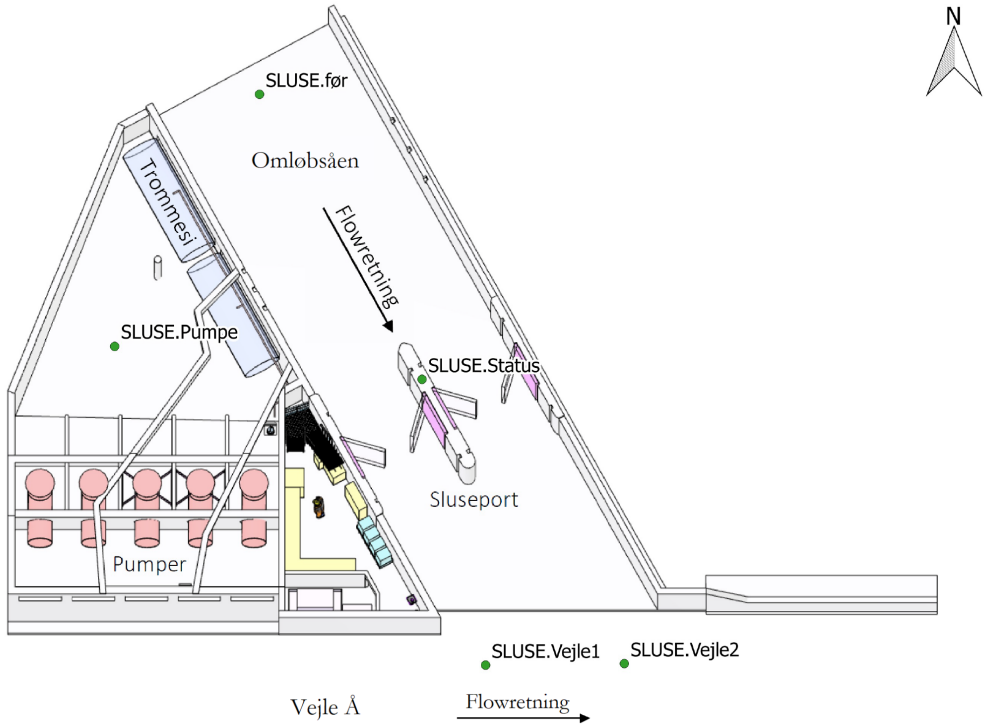
I tabel 2.2 er angivet informationer om sensorerne ved fordelerbygværket vist på figur 2.4. De angivne referencer vil løbende igennem rapporten fungere som betegnelse, når en sensor omtales.

Tabel 2.2: Ved fordelerbygværket måles åbningsprocenten af de tre spjæld, men også vandstande før og efter ristene måles for at tjekke tilklogning.

Fordelerbygværk v. Abelones Plads			
Reference	**Periode**	**Beskrivelse**	**Enhed**
FBV.Før	29/03-2017 →	Vandstand før rist	[m]
FBV.Efter	29/03-2017 →	Vandstand efter rist	[m]
FBV.Spj1	29/03-2017 →	Åbningsprocent af spjæld 1	[%]
FBV.Spj2	29/03-2017 →	Åbningsprocent af spjæld 2	[%]
FBV.Spj3	29/03-2017 →	Åbningsprocent af spjæld 3	[%]

Målere ved sluseanlægget ved Bryggen

Sluseanlægget ved Bryggen kan aktivt styres og har stor indvirkning på vandstanden i Omløbsåen. På grund af denne styring er det naturligt at have sensorer, som beskriver status tæt ved sluseanlægget. Sensorernes placering kan ses på figur 2.5.



Figur 2.5: Redigeret systemtegning af sluseanlægget, hvor sensorerenes placering er skitseret og markeret med grønne prikker. (Vejle Kommune 2015b)

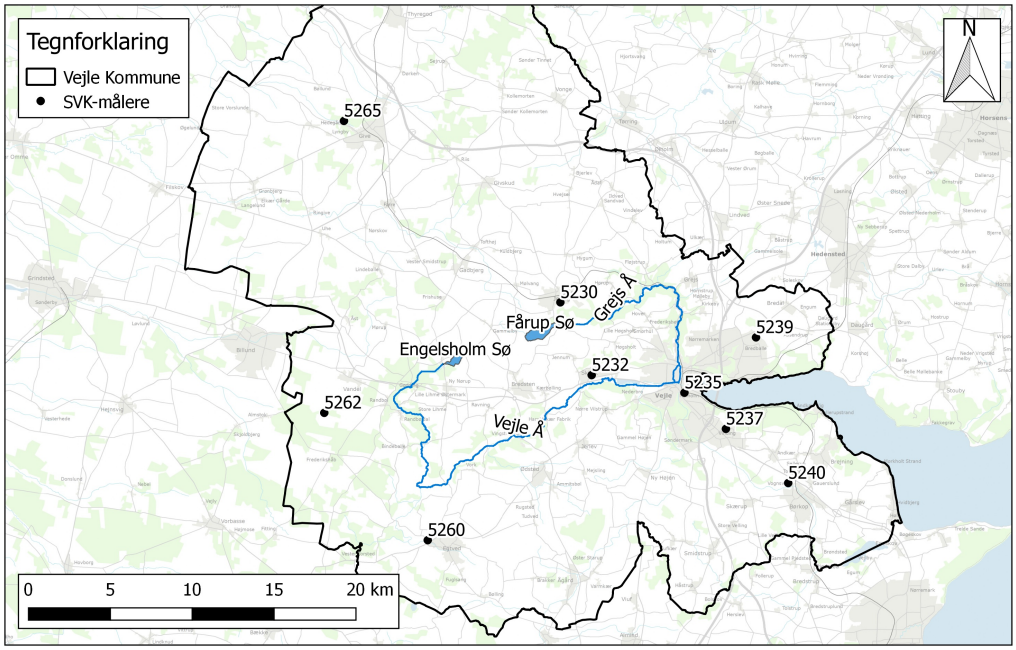
I tabel 2.3 er angivet informationer om sensorerne ved sluseanlægget vist på figur 2.5. De angivne referencer vil løbende igennem rapporten fungere som betegnelse, når en sensor omtales.

Tabel 2.3: Ved sluseanlægget forefindes vandstande og åbningsstatus i de angivne perioder.

Sluseanlæg v. Bryggen			
Reference	**Periode**	**Beskrivelse**	**Enhed**
SLUSE.Før	13/09-2016 →	Vandstand i Omløbsåen	[m]
SLUSE.Pumpe	13/09-2016 →	Vandstand i pumpesumpen	[m]
SLUSE.Status	14/02-2017 →	Åbningsstatus	[-]
SLUSE.Vejle1	13/09-2016 →	Vandstand i Vejle Å udenfor sluseportene 1	[m]
SLUSE.Vejle2	13/09-2016 →	Vandstand i Vejle Å udenfor sluseportene 2	[m]

2.1.3 Regnmålere

Nedbør er en vigtig faktor at have data for, når oversvømmelser i Vejle skal beskrives, fordi forhøjede vandstande kan forekomme som følge af kraftig nedbør. Der er placeret en række regnmålere i området omkring Vejle, hvilket ses på figur 2.6.



Figur 2.6: Kortet viser Vejle Å og Grejs Å samt de omkringliggende regnmålere. Hver nedbørsmåler er navngivet og nummereret af Spildevandskomitéen, og kortet vil fremover fungere som reference, når en regnmålermåler omtales.

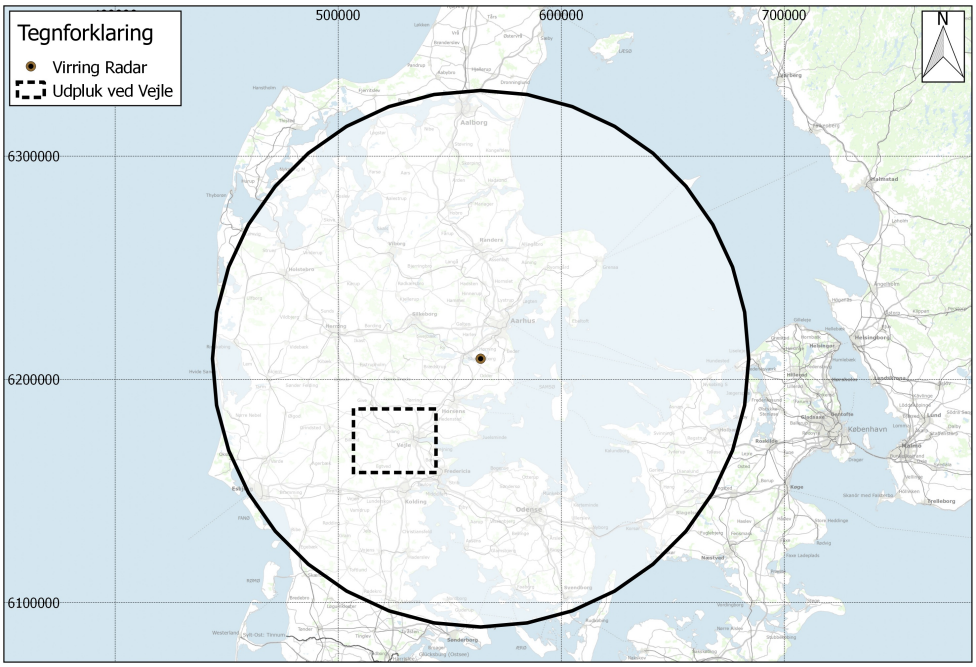
I tabel 2.4 er der angivet informationer om regnmålerne, og hvilke perioder de har målt i. De angivne referencer vil løbende i rapporten fungere som betegnelse, når en målestation omtales.

Tabel 2.4: I området omkring Vejle forefindes data for nedbør fra i alt ni regnmålere, hvoraf de otte stadig er aktive.

Reference	Periode	Lokation
SVK 5230	16/12-2009 →	Jelling
SVK 5232	06/10-2010 →	Skibet
SVK 5235	07/01-1979 →	Vejle Centralrenseanlæg
SVK 5237	20/12-2003 →	Vejle Pumpestation
SVK 5239	06/10-2010 →	Bredballe
SVK 5240	16/12-2009 →	Børkop Pumpestation
SVK 5260	16/12-2009 →	Egtved Renseanlæg
SVK 5262	22/02-1994 → 07/02-1999	Flyvestation Vandel
SVK 5265	22/12-2009 →	Give Renseanlæg

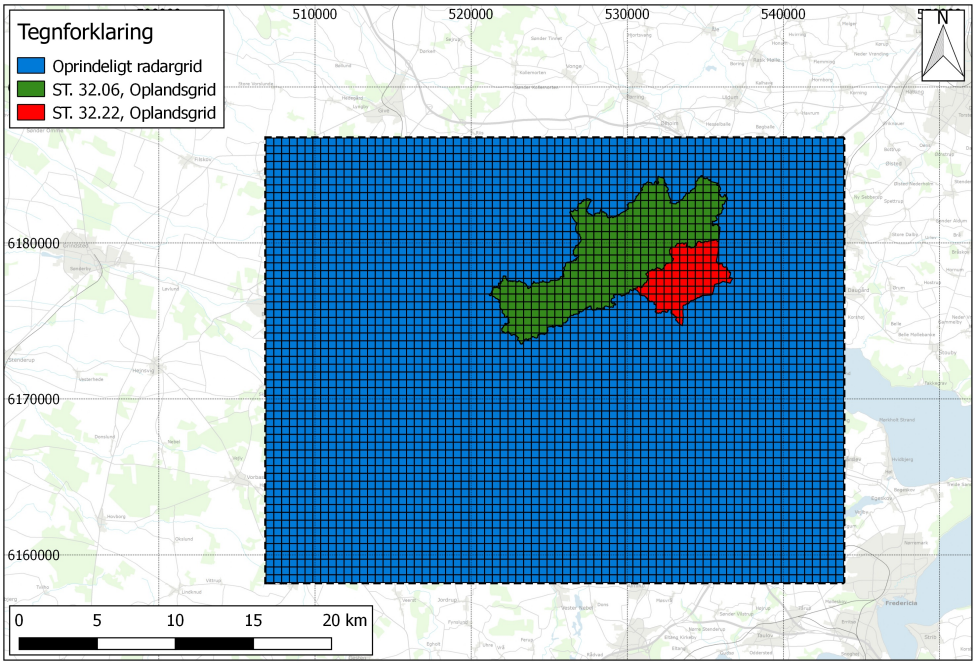
2.1.4 Nedbør i form af radar

Der anvendes signaler fra en radar i Virring i et område ved Vejle, som vist på figur A.1. Radaren er ejet af Dansk Meteorologisk Institut.



Figur 2.7: Radaren i Virring dækker en stor del af landet, men det er kun i området ved Vejle, hvor signalerne er relevante.

Der haves radardata i griddet, som er anvist som det oprindelige radargrid på figur 2.8. Da det kun ønskes at anvende radar i oplandet til Grejs Å-systemet, begrænses området hertil. Dette område er yderligere inddelt i to oplande til henholdsvis vandstandsmåleren ved Grejsdalens Planteskole (ST. 32.06) og vandstandsmåleren ved Ny Grejsdalsvej (ST. 32.22).



Figur 2.8: Det oprindelige radargrid dækker oplandet til alle vandløb, som løber til Vejle. Der er dog kun anvendt data indenfor de markerede oplandsgrid.

På figur 2.8 gennemskærer vandløbsoplandene mange af cellerne. Alle gennemskårne celler medtages som en del af oplandsgriddet.

I tabel 2.5 er der angivet informationer om radardata, og hvilke perioder, der anvendes data fra. De angivne referencer vil løbende i rapporten fungere som betegnelse, når en målestation omtales.

Tabel 2.5: I området omkring Vejle forefindes radardata i et større område, som afgrænses til oplandet til to vandstandsmålere i Grejs Å-systemet.

Reference	Periode	Lokation
Radar.Planteskole	01/02-2015 →	Oplandet til Grejsdalens Planteskole (ST. 32.06)
Radar.NyGrejsdalsvej	01/02-2015 →	Oplandet til Ny Grejsdalsvej (ST. 32.22)

Radardata er midlet over oplandet til de to vandstandsmålere, da det ønsket at anvende radar til direkte sammenligning med regnmålere. Derfor fortolkes Radar.Planteskole og Radar.NyGrejsdalsvej som to regnmålere, der i realiteten bare dækker et større område. En nærmere beskrivelse af metoden til at udtrække radar ses i appendiks A.

2.2 Måleperioder, synkronisering og fejlfinding af data

Når det ønskes at anvende datadrevne modeller, er det altafgørende, at datakvaliteten er god. Det skyldes eksempelvis, at en analyse baseret på dårlig rådata ofte vil udvande et modelresultat, uanset hvor godt det måtte være (Liaison Technologies 2017).

Datavalidering af målere er et emne, der ligger uden for rammerne af dette projekt, og derfor antages det som udgangspunkt, at alle målinger er korrekte, fordi det er en forudsætning for ordentlig brug af datadrevne modeller. Dog anskues data løbende med en kritisk indgangsvinkel, så åbenlyse fejl kan identificeres og om muligt frasorteres.

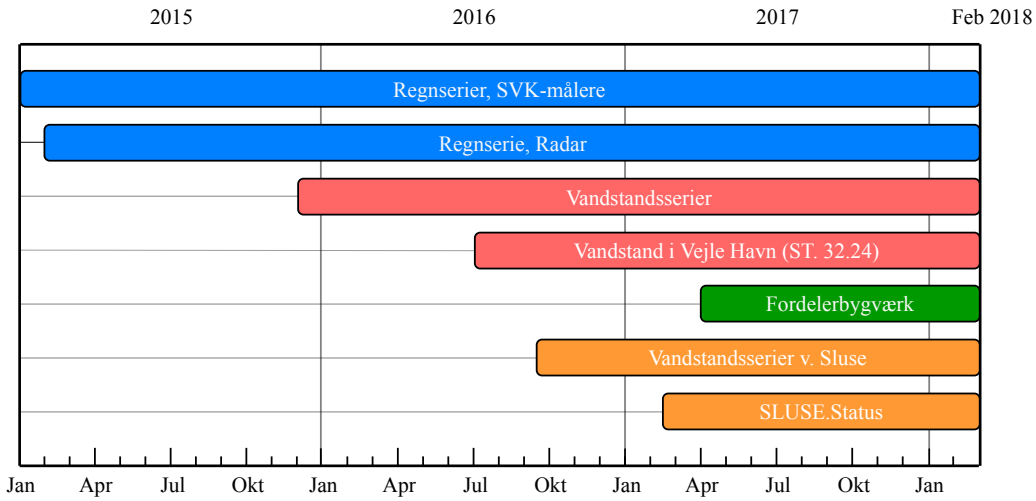
Når de forskellige typer af data skal anvendes i samspil med hinanden på en fornuftig måde, er det vurderet at være en fordel, at:

- Der er overblik over indbyrdes måleperioder.
- Afbrudte perioder og målefejl frasorteres.
- Data følger samme tidszone.
- Data har samme tidsopløsning.

Følgende delafsnit beskriver arbejdet, der er lagt i at efterkomme de ovenstående punkter.

2.2.1 Måleperioder

På figur 2.9 vises en oversigt over perioderne, som de førnævnte data strækker sig over. Det ses, at der er stor forskel i perioder, som dataene strækker sig over. Når der anvendes neurale netværk til modellering, er det den korteste af de anvendte måleserie, der sætter begrænsningen for datagrundlaget.



Figur 2.9: Perioderne som de forskellige data strækker sig over.

2.2.2 Afbrud i måleserierne

Enkelte tidsserier rummer huller, hvor måleperioden har været afbrudt. Disse afbrydelser udfyldes ved brug af lineær interpolation, hvor det vurderes at være en realistisk metode til at genskabe data. Således udfyldes afbrydelser på over to timer ikke, da det vurderes, at interpolation i disse tilfælde kan give fejlagtige bud på eksempelvis vandstandsvariationer. Når data plottes igennem rapporten vises perioder med afbrydelser på over to timer stadig som værenrede lineært interpoleret, men perioderne vil ikke blive anvendt til analyser og indlæring af de neurale netværk.

2.2.3 Tidszoner

En af ulemperne ved neurale netværk er, at systematiske fejl indlæres i modellen. Derfor er det væsentligt, at dataene fra forskellige kilder synkroniseres, så de følger samme tidszone. Tidsangivelsen på data fra SVK-målerne er i Coordinated Universal Time (UTC) (Cappelen 1993). For den resterende del af det tilgængelige data fremgår det ikke hvilken tidszone, der er anvendt. Derfor har det været nødvendigt at undersøge overgangene mellem vinter- og sommertid for at fastlægge den anvendte tidszone. Dette er beskrevet i appendiks B.

Undersøgelsen viser, at på nær data fra SVK-målere følger alt data Greenwich Mean Time (GMT) og inkluderer sommertid. Det vil sige, at data i vinterperioder angives i tidszonen GMT+1 og i sommerperioderne GMT+2. Ved at sammenholde måleserier med plots på www.hydrometri.dk er det dog fundet, at data fra Orbicons hydrometridatabase ikke skifter mellem vinter- og sommertid på det korrekte tidspunkt, men usystematisk indenfor et interval på $\pm$ ét døgn fra det korrekte tidspunkt. At dataene potentielt er forskudt en time i en kort periode omkring skiftene mellem sommertid og vintertid vil potentielt forringe kvaliteten af de neurale netværk, men det vurderes ikke at have en væsentlig betydning i det store billede, blandt andet fordi skiftene ikke forskyder nogle voldsomme afstrømningshændelser. Det er derfor valgt ikke at udbedre disse data.

For at have samme tidszone på alle dataene er det valgt at omskrive SVK-dataene til GMT+1 inklusiv sommertid.

2.2.4 Varierende tidsopløsning

Et neuralt netværk giver kun et output, hvis der findes et input. Da datagrundlaget omfatter tidsserier med forskellige, og somme tider varierende, tidsopløsninger, er det nødvendigt at få kontrol over

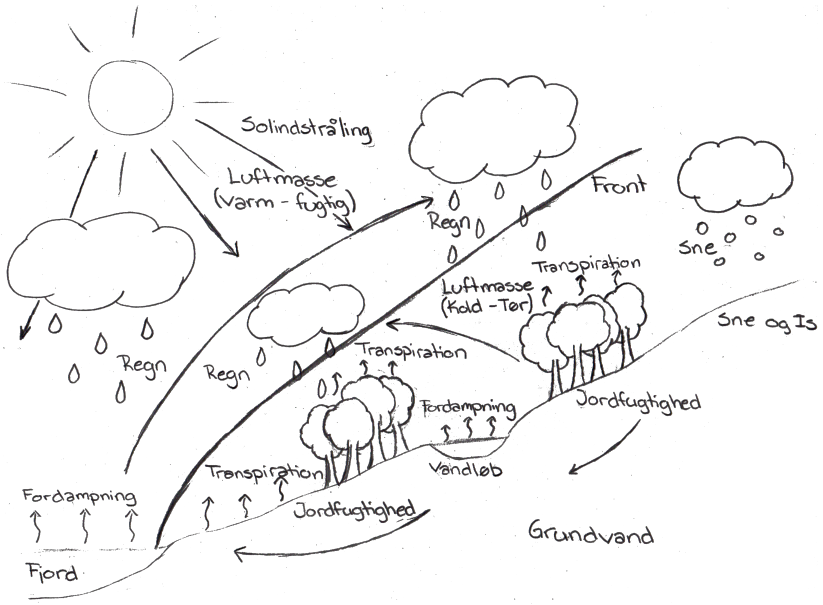
disse, så det neurale netværk kan foretage en én-til-én oversættelse. Det kommer eksempelvis til udtryk ved, at regnintensiteter optræder med en tidsopløsning på ét minut, mens vandstande optræder med en tidsopløsning på 5 eller 15 minutter. For at have det størst mulige datagrundlag, er der i disse tilfælde valgt at interpolere vandstandene lineært, så der haves en tidsserie med samme tidsopløsning som regnintensiteterne. Det er vurderet, at vandstandsvariationerne i langt de fleste tilfælde ikke er af betydende karakter indenfor en tidsskala på 15 minutter, og derfor vurderes det acceptabelt at interpolere mellem de målte værdier. Alt data ender dermed i en tidsopløsning på 1 minut.

3. Forundersøgelse af data

At kende de styrende processer for vandstande i vandløbene er afgørende for opsætningen af de neurale netværk. Formålet med forundersøgelsen er at danne et overblik over hvilke processer, der er styrende for vandstanden i risikoområdet, så der kan opsættes gode og brugbare modeller målrettet prognoser i Vejle Midtby og opland.

3.1 Styrende processer for vandstanden i de danske vandløb

I de følgende to delafsnit præsenteres de grundlæggende styrende processer for vandstanden i danske vandløb, hvilket inkluderer vandets vej til vandløbet og kapaciteten i vandløbet.



Figur 3.1: Skitsering af afstrømning til et vandløb inspireret af Vandforsyningsbogen. (Karlby m.fl. 2014)

3.1.1 Vandets vej til vandløbet

Vandet i danske vandløb stammer fra nedbør. Der er en række parametre, som er betydende for, hvor meget af nedbøren, der ender i vandløbet, og hvor hurtigt dette sker. De omtalte parametre er:

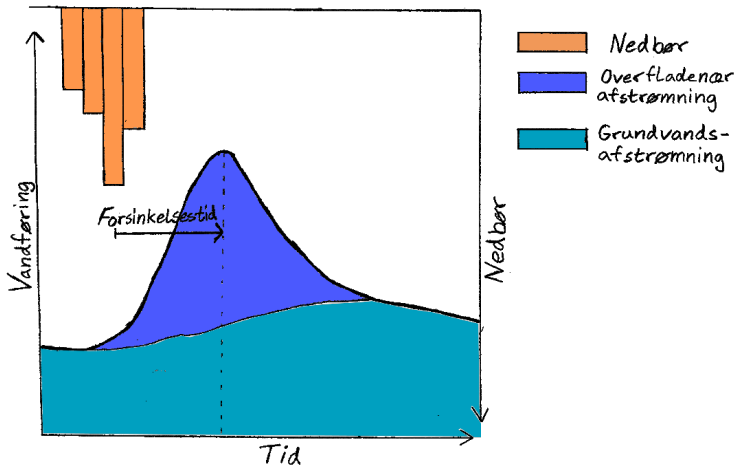
- Fordampning i form af evaporation fra åbne jordoverflader og transpiration fra vegetation.
- Nettoindstråling som er den mængde af solindstrålingen, der passerer atmosfæren.
- Temperaturen.
- Vindforhold.
- Infiltrationskapacitet.
- Jordbundsforhold.

Disse parametre yder i større eller mindre grad et bidrag til de styrende processer for vandstanden i et vandløb, og kredsløbet er skitseret på figur 3.1.

Nedbøren kan falde som både regn og sne. Når nedbøren rammer jordoverfladen i et ruralt område, vil den enten fordampe, infiltrere ned igennem jorden eller afstrømme på overfladen. I dette delafsnit beskrives disse processer kort, og hvad disse betyder for, hvordan vandføringen i vandløbet afhænger af nedbør. Beskrivelserne læner sig op ad kapitel 6 i rapporten *Ferkvandets kredsløb*. (Refsgaard m.fl. 2003).

En væsentlig del af den nedbør, som falder i Danmark, vil fordampe igen. Mængden af vand, der potentielt vil fordampe fra jordoverfladen afhænger blandt andet af nettoindstråling, temperatur, vind og vegetation. Disse parametre vil variere over året og af denne grund vil andelen af nedbøren, som potentielt vil kunne fordampe ligeledes variere over året. For at en fordampning skal kunne finde sted, skal der være vand tilstede, så den fordampning, der reelt finder sted, afhænger ligeledes af mængden af vand, der er tilgængelig ved jordoverfladen og dermed mængden af nedbør, der er faldet i perioden op til det aktuelle tidspunkt.

Som et udtryk for hvor meget vand, der kan infiltrere ned gennem jorden, anvendes udtrykket jordens infiltrationskapacitet. Infiltrationskapaciteten vil afhænge af jordbundsforholdene og vegetation. Infiltrationskapaciteten vil normalt være høj, hvis jorden ikke har været udsat for nedbør i længere tid, da jorden i så fald vil suge vandet til sig. Hvis jorden er frossen kan det omvendt resultere i, at infiltrationskapaciteten er meget lav. Infiltrationskapacitet afhænger dermed meget af vejret op til det aktuelle tidspunkt.



Figur 3.2: Noget af nedbøren vil afstrømmme hurtigt som overfladenær afstrømning, mens andet vil afstrømme langsomt som grundvandsafstrømning.

Vandet, der infiltrerer igennem jorden, vil enten igennem dræn blive transporteret til nærliggende vandløb eller infiltrere videre ned til grundvandsmagasinerne, hvor det ligeledes vil løbe til vandløb eller direkte i havet. Hvis intensiteten af nedbøren overskrider infiltrationskapaciteten kan vandet ikke infiltrere og vil derfor enten opmagasineres eller afstrømme på overfladen.

Afstrømningen på overfladen og transporten af vand gennem dræn kaldes den *overfladenære afstrømning* under en fælles betegnelse. Under danske forhold vil størstedelen af den overfladenære afstrømning ske ved dræning. Den overfladenære afstrømning kommer ofte til udtryk i vandløbet som en relativ kraftig og kortvarig stigning i vandføringen efter en nedbørshændelse. Den øgede grundvandsafstrømning vil komme til udtryk som en lille, men længerevarende stigning i vandføringen efter en nedbørshændelse. Dette er illusteret på figur 3.2.

Selv hvis det var muligt at beskrive den præcise afstrømning til et vandløb fra det dertilhørende opland, så vil der stadig ligge en udfordring i at vurdere, hvorvidt dette vil føre til oversvømmelse. Det afgørende for, hvorvidt et vandløb udgør en risiko for oversvømmelse i det omkringliggende område, er nemlig ikke vandføringen, men vandstanden. Sammenhængen mellem vandføring og vandstand afgøres af vandløbets kapacitet.

3.1.2 Vandløbets kapacitet

Sammenhængen mellem vandføring og vandstand, når det gælder rør, kan normalt findes ved anvendelse af Manningformlen, som er vist i ligning 3.1.

$$Q = A(h) \cdot M(h,t) \cdot R(h)^{2/3} \cdot I^{1/2} \tag{3.1}$$

Hvor

Q	Vandføringen	[m^3/s]
A	Det gennemstrømmede tværsnitsareal	[m^2]
M	Manningtallet	[$m^{1/3}/s$]
R	Den hydrauliske radius	[m]
I	Energilinjegradienten	[-]

Anvendes Manningformlen til beregning af vandføringen i rør, vil Manningtallet normalt antages at være uafhængigt og konstant. Anvendes Manningformlen til vandløb vil denne antagelse være behæftet med flere fejlkilder. Manningtallet vil afhænge af vanddybden h, da der må forventes varierende vegetation langs vandløbets kant. Manningtallet vil desuden også afhænge af tiden t, da vegetationen varierer over året.

Det gennemstrømmede tværsnitareal og den hydrauliske radius afhænger både i vandløb og rør af vanddybden h, men erosion og aflejringer kan resultere i, at korrelationen mellem vandstanden og henholdsvis det gennemstrømmede tværsnitsareal og den hydrauliske radius ikke er konstant.

Energilinjegradienten vil desuden i mindre grad kunne variere på grund af stuvning/sænkning, hvis der sker lokale ændringer på nogle af de andre parametre. Energilinjegradienten vil dermed i nogle tilfælde være afhængig af både vandstand og tid.

Grundet disse usikre parametre, som knytter sig til sammenhængen mellem vandstand og vandføring i vandløb, anvendes ofte en Q-h-relation i stedet for Manningformlen til at beskrive sammenhængen. Q-h-relationen vil normalt være bygget på målinger af vandstand og vandføring, hvor vandføringen er estimeret på baggrund af flowmålinger flere steder i vandløbets tværsnit. Estimeringen af vandføringen er dog behæftet med væsentlige usikkerheder:

- Turbulente strømninger i vandløbene gør flowmålingerne varierende.

- Flowmålingerne er punktmålinger, som ikke er fuldstændig repræsentative for hele tværsnittet.
- Grundet ændring i strømningsforhold og begrænset måledata ved de højeste vandføringer, er usikkerheden på estimeringen af vandføringen ekstra stor i disse tilfælde.
- Q-h-relationen vil gennem tiden ændre sig på grund af eksempelvis årstidsvariationer på Manningtallet og sedimenttransport.

Usikkerhederne ved Q-h-rationerne har tidligere givet anledning til diskussion af deres anvendelighed ved især høje vandføringer (Larsen 2016).

3.1.3 Fokus på vandstand

Der er valgt udelukkende at fokusere på målingerne for vandspejlskoten i analyserne i rapporten, da dette netop er den centrale parameter i forbindelse med oversvømmelsesvarsling. Ved ikke at tage de estimerede vandføringer med i analyserne kan det potentielt resultere i, at der mistes viden, der er inkorporeret i variationerne på Q-h-relationerne. Eksempelvis vil Q-h-relationen ændre sig, hvis der sker sedimenttransport eller en grødeskæring. For at fange sådanne ændringer skal der dog forekomme målinger af vandføringen relativt ofte. Det er en fordel i praksis, at vandstanden kan modelleres direkte, da det dermed ikke er nødvendigt med opdatering af Q-h-relationen ved løbende manuelle målinger. På grund af de nævnte usikkerheder, der er knyttet til estimeringen af vandføringen i vandløb, vil modelleringen af vandføring og en senere transformation til vandspejlskote ligeledes indføre usikkerheder. Ved at anvende rådata af vandspejlskoter indføres der med sikkerhed ikke ukorrekte transformerede data.

3.2 Forundersøgelse af forhold i Omløbsåen

ST. 32.21 er som beskrevet i kapitel 2 placeret ved Aagade centralt i Omløbsåen. Derfor er der fokus på netop at kunne prognosticere vandstanden i dette punkt.

Vejle Spildevand har defineret, hvornår vandstanden ved forskellige stationer er kritisk. Vandstanden inddeles i tre kategorier: Rød, gul, grøn. Vandstanden kategoriseres som grøn, når den ikke er problematisk. Vandstanden kategoriseres som gul, når den er kritisk og tæt på kanten, men ikke løber over. Vandstanden kategoriseres som rød, når den løber over kanten. I Omløbsåen ved Aagade (ST. 32.21) er koterne for kategorierne defineret således:

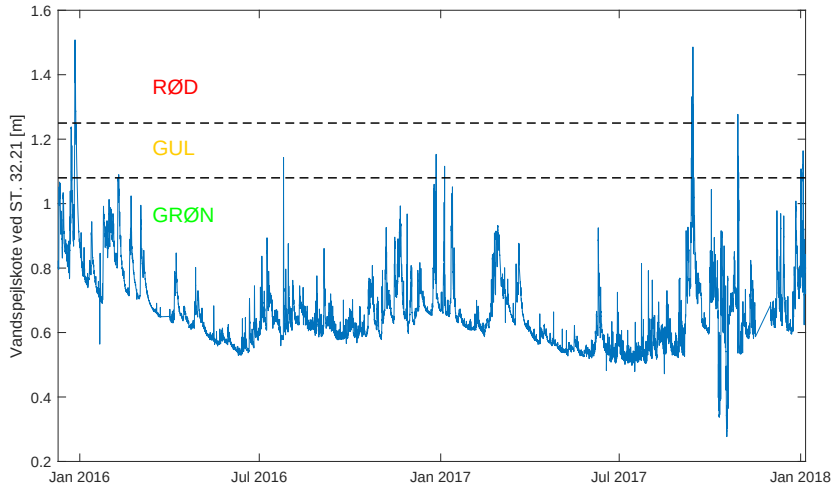
- Grøn: < 1,08
- Gul: 1,08 - 1,25
- Rød: > 1,25

På figur 3.3 er vandstanden i Omløbsåen vist sammen med de kritiske koter. Det ses, at fra december 2015 til januar 2018 har der været tre hændelser, hvor vandstanden har været målt i rød zone. De to hændelser med højest vandstand har ligeledes givet dokumenterede problemer med oversvømmelse i Vejle Midtby (www.112alarm.dk 2015; Vejle Amts Folkeblad 2017).

På baggrund af nærmere analyser af måleserien vist på figur 3.3, vurderes vandstanden i Omløbsåen primært at være styret af følgende forhold:

- Vandstanden i Vejle Havn.
- Afstrømningen i Vejle Å.
- Afstrømningen i Grejs Å.

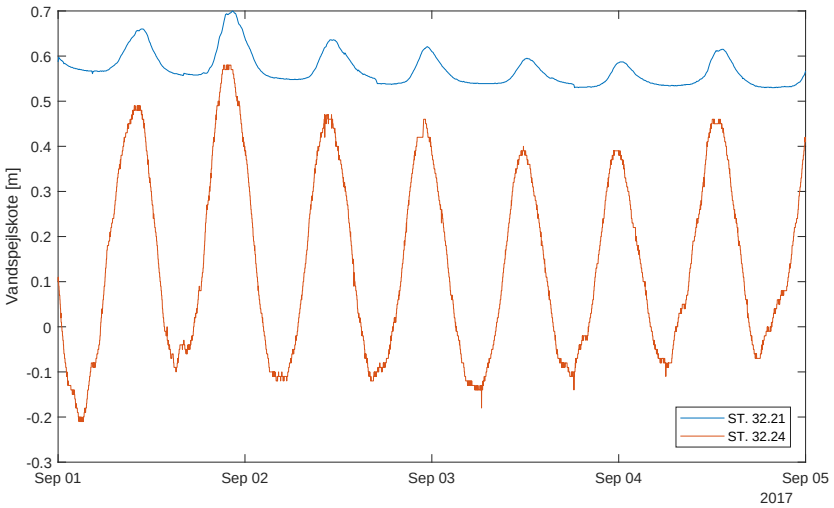
Udover disse naturlige forhold kan vandstanden i Omløbsåen styres af fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen, som nævnt tidligere i kapitel 1. Hvor stor en betydning, hvert enkelt element har for Omløbsåen, uddybes nærmere i det følgende.



Figur 3.3: Målt vandstandskote i Omløbsåen (ST. 32.21) fra 10/12-2015 til 6/1-2018 og grænserne for kritiske koter.

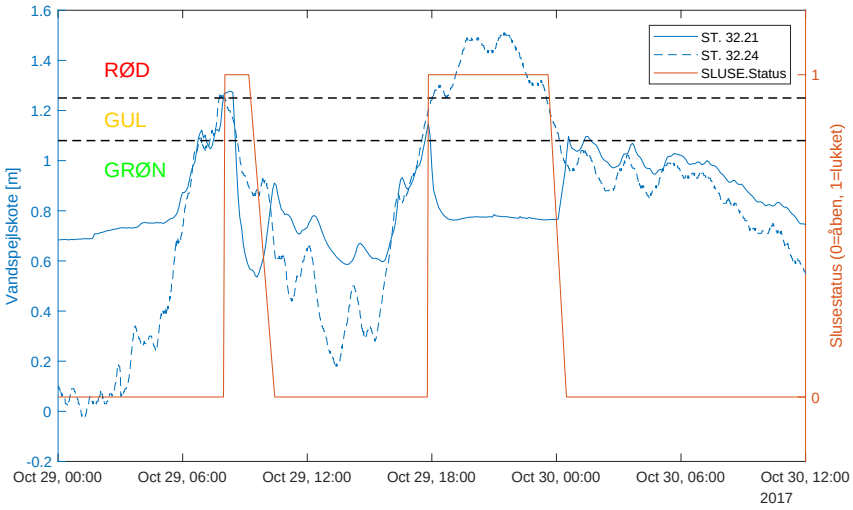
Vandstand i Vejle Havn

På figur 3.4 ses det tydeligt, at vandstanden i Vejle Havn (ST. 32.24) har påvirkning på vandstanden i Omløbsåen (ST. 32.21) ved daglig vandstand.



Figur 3.4: Vandstanden i Omløbsåen (ST. 32.21) påvirkes af vandstanden i Vejle Havn (ST. 32.24), hvilket i dette tilfælde illustreres ved tidevandet.

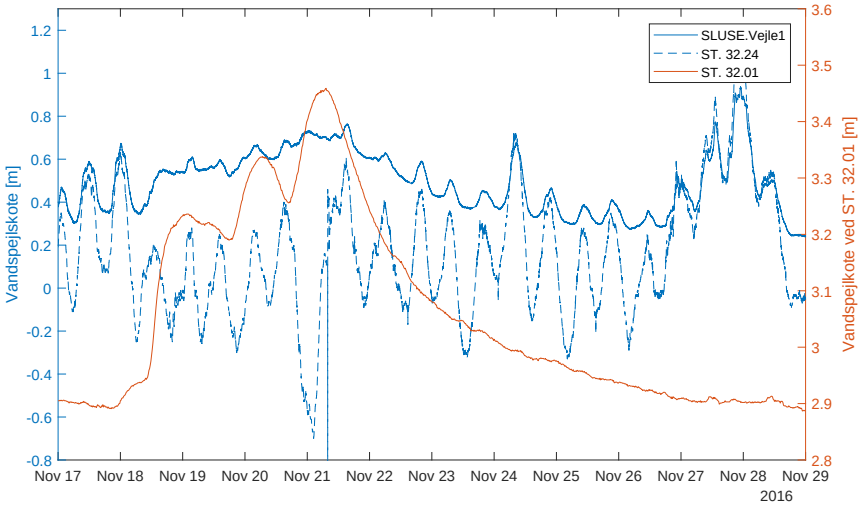
I visse tilfælde kan vandstanden i Vejle Havn alene være årsag til en kritisk vandstand i Omløbsåen, hvilket ses af eksemplet på figur 3.5. I dette tilfælde er sluseportene blevet lukket og vandstanden i Omløbsåen er dermed blevet holdt nede af pumperne. Hvis sluseanlægget ikke var blevet taget i brug, ses det af vandspejlskoten i Vejle Havn (ST. 32.24), at der med høj sandsynlighed ville være forekommet oversvømmelser langs Omløbåen.



Figur 3.5: Vandstanden i Vejle Havn (ST. 32.24) kan i visse tilfælde få vandstanden i Omløbsåen (ST. 32.21) til at nå kritiske koter.

Afstrømning i Vejle Å

Vandstanden, og dermed trykniveauet, ved udløbet fra Omløbsåen til Vejle Å har nødvendigvis betydning for kapaciteten af Omløbsåen. Det ses på figur 3.6, at der ved en forhøjet vandstand opstrøms i Vejle Å ved Haraldskær (ST. 32.01) også forekommer en forhøjet vandstand ved Omløbsåens udløb til Vejle Å. Vejle Havn (ST. 32.24) ses dog stadig som den mest betydningsfulde faktor for vandstanden ved sluseanlægget.

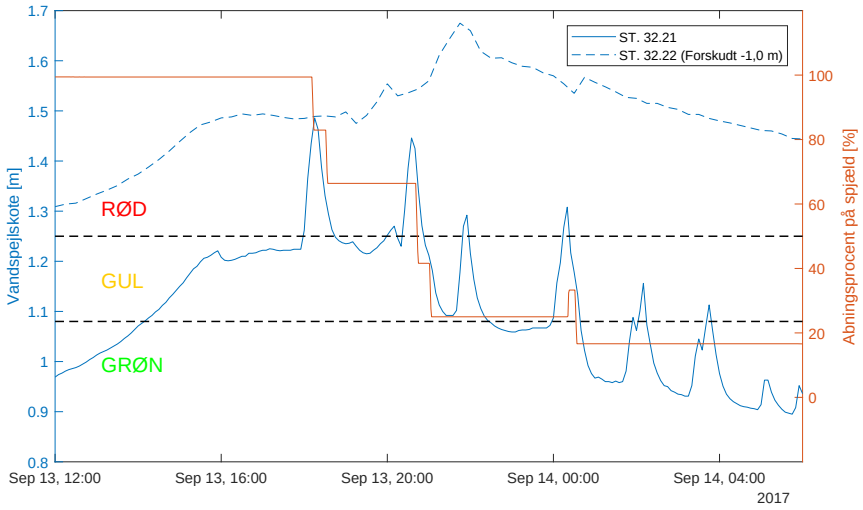


Figur 3.6: Vandstanden i Vejle Havn (ST. 32.24) er den mest betydningsfulde faktor for vandstanden ved Omløbsåens udløb til Vejle Å (SLUSE.Vejle1), men også afstrømningen fra oplandet til Vejle Å har en betydning (ST. 32.01).

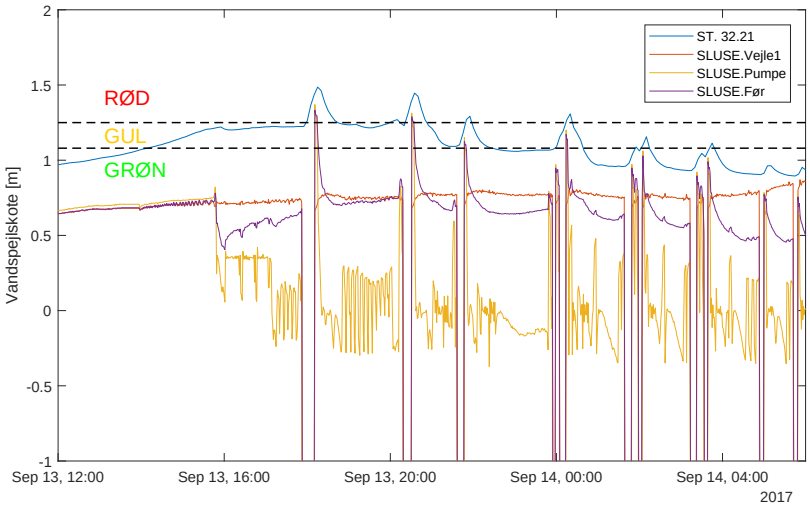
En stor afstrømning i Vejle Å vurderes ikke alene at kunne være årsag til kritiske vandstande i Omløbsåen. Dog formodes en stor afstrømning i Vejle Å sjældent at forekomme uden, at der også er en stor afstrømning i Grejs Å.

Afstrømning i Grejs Å

Den anden afgørende årsag til høje vandstande i Omløbsåen er høj afstrømning fra oplandet til Grejs Å. En forøget afstrømning fra oplandet kommer til udtryk ved en forhøjet vandstand opstrøms fordelerbygværket ved Ny Grejsdalsvej (ST. 32.22). I disse tilfælde kan vandstanden i Omløbsåen både styres af sluseanlægget og ved fordelerbygværket. På figur 3.7 ses et eksempel på en hændelse, hvor den gennemsnitlige åbningsprocent på spjældene blev reduceret for at reducere vandføringen i Omløbsåen og dermed forhindre en kritisk vandstand.



Figur 3.7: Åbningsprocenten på spjældene reduceres for at forhindre høj vandstand i Omløbsåen (ST. 32.21).



Figur 3.8: Strømsvigt og flydegods fanget i ristene ind til pumpebrønden medførte periodevis høje vandstande i Omløbsåen (ST. 32.21).

De tydelige peaks på vandstanden i Omløbsåen skyldes i tilfældet 13/09-2017, at der forekom midlertidige strømsvigt ved sluseanlægget, hvilket resulterede i, at pumperne var ude af drift, mens sluseportene var lukkede (Vejle Amts Folkeblad 2017). Dette medførte unaturlige vandstandsstigninger i Omløbsåen. Det ses på figur 3.8, at målerne placeret ved sluseanlægget i flere omgange ikke

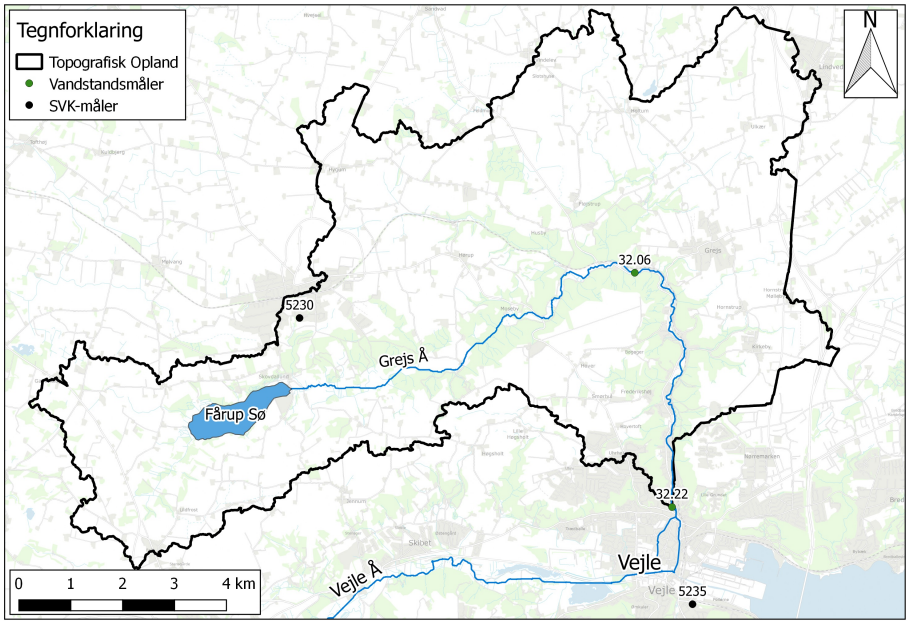
måler noget grundet strømsvigtene, og derfor forefindes markante negative værdier i disse perioder. Derudover var der fanget en del flydegods i ristene ind til pumpebrønden, hvilket kommer til udtryk ved den store forskel i vandspejlskoten mellem vandstanden før sluseportene og i pumpebrønden.

Sammenfatning af forundersøgelse af forhold i Omløbsåen

Forundersøgelsen af forholdene gældende i Omløbsåen viser, at en høj vandstand i Omløbsåen kan forårsages af enten høj vandstand i Vejle Havn eller en høj afstrømning i Grejs Å. Sluseanlægget ved Bryggen kan begrænse problemerne forårsaget af høj vandstand i Vejle Havn såvel som en høj afstrømning i Grejs Å. Fordelerbygværket ved Abelones Plads kan primært begrænse problemerne forårsaget af høj afstrømning i Grejs Å. Andre faktorer som en høj afstrømning i Vejle Å og flydegods i ristene ind til pumpebrøndene ved sluseanlægget kan i mindre grad være med til at forværre situationen i Omløbsåen.

3.3 Forundersøgelse af forhold gældende i Grejs Å ved Ny Grejsdalsvej

ST. 32.22 er placeret ved Ny Grejsdalsvej umiddelbart opstrøms fordelerbygværket ved Abelones Plads. Oplandet ses på figur 3.9.

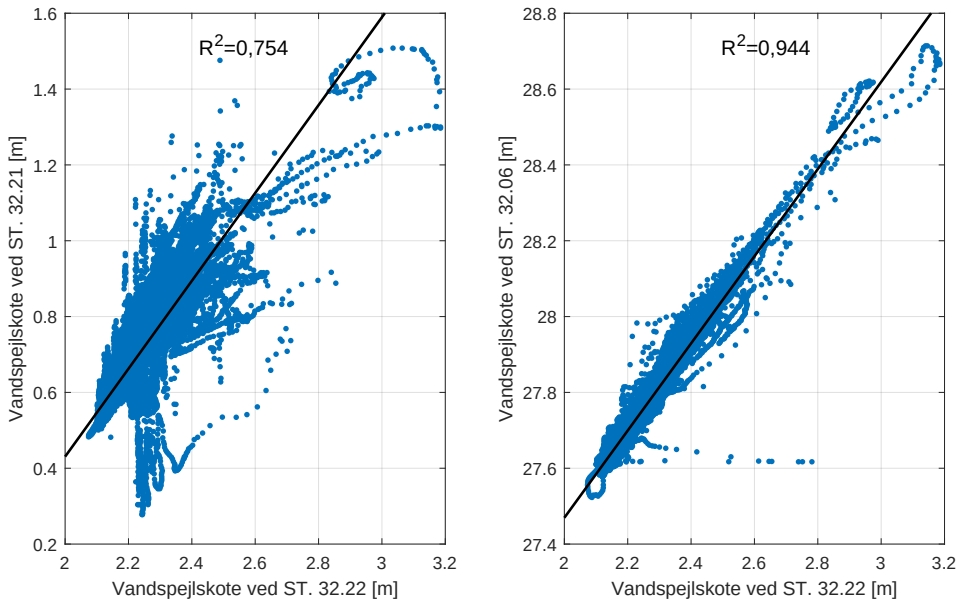


Figur 3.9: Oplandet til Grejs Å ned til målestationen ved Ny Grejsdalsvej (ST. 32.22) dækker et areal på ca. 80 km^2.

Hovedformålet med projektet er at kunne lave prognoser af vandstanden i Omløbsåen. I afsnit 3.2 er det gjort klart, at vandstanden i Omløbsåen er stærkt afhængig af afstrømningen fra Grejs Å. Strategien er derfor at lave prognoser af vandstanden ved Ny Grejsdalsvej, da det vurderes, at denne station fungerer som den bedste beskrivelse af afstrømningen i Grejs Å. Dette er vurderet, fordi stationen dækker hele oplandet til Grejs Å og ikke i samme grad som vandstanden ved Aagade i Omløbsåen er påvirket af fordelerbygværket, sluseanlægget og vandstandsændringer i Vejle Havn.

Opstrøms målestationen ved Ny Grejsdalsvej ligger en anden målestation ved Grejsdalens Planteskole (ST. 32.06), som dækker 80 % af oplandet til Grejs Å. Derfor forventes det, at vandstanden ved Grejsdalens Planteskole kan anvendes til at lave prognoser af vandstanden ved Ny Grejsdalsvej eller erstatte målingen, hvis den påvirkes af opstuvning fra fordelerbygværket.

For at underbygge ovenstående tankegang er det undersøgt, hvordan de tre målestationer er korreleret med hinanden. Korrelogrammerne er vist på figur 3.10. Der er plottet målepunkter for hvert 15. minut i perioden 10/12-2015 til 06/01-2018. Afbrudte måleperioder er frasorteret. Det er vigtigt at bemærke, at korrelationen mellem vandstandene ikke nødvendigvis behøver at være lineær, hvis eksempelvis udformningen af tværsnittene ved stationerne er meget forskellige.



Figur 3.10: Sammenhæng mellem vandspejlkoten målt ved Ny Grejsdalsvej (ST. 32.22) og Omløbsåen (ST. 32.21) (til venstre) er markant dårligere, end mellem Ny Grejsdalsvej og ved Grejsdalens Planteskole (ST. 32.06) (til højre).

Korrelationen mellem vandstanden i Omløbsåen og vandstanden ved Ny Grejsdalsvej er ikke perfekt, hvilket kan skyldes:

- Vandstanden i Omløbsåen er påvirket af vandstanden i Vejle Havn.
- Vandstanden i Omløbsåen er påvirket af indstillingen på spjældene ved fordelerbygværket og styring af sluseanlægget.
- Vandstanden i Omløbsåen er påvirket af afstrømning fra Vejle Midtby.

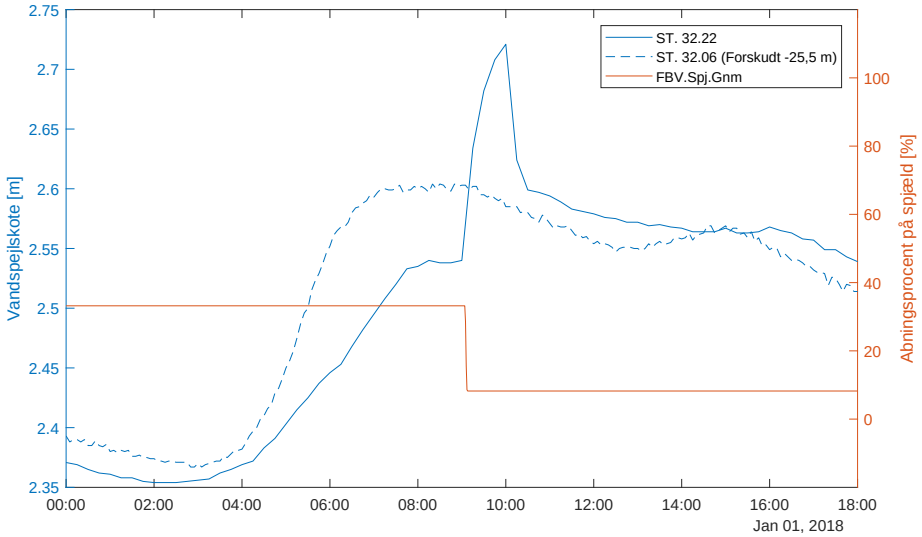
Den lineære korrelationen mellem vandstanden ved Ny Grejsdalsvej og ved Grejsdalens Planteskole er overraskende god. Afvigelser formodes at skyldes:

- Stedlig fordelt nedbør, som kan medføre, at vandføringen ved de to stationer ikke er lineært korreleret.
- Indstillingen på fordelerbygværket ved Abelones Plads kan have betydning for vandstanden ved Ny Grejsdalsvej.
- Varierende Manningtal eller tværsnit ved de to stationer kan medføre, at samme vandføring resulterer i forskellige vandstande.
- Der er en transporttid mellem de to stationer, så en bølge må derfor forventes at optræde senere ved Ny Grejsdalsvej end ved Grejsdalens Planteskole.

Det forsøges i følgende analyser at afdække hvilke af disse årsager, der er mest betydende for den manglende entydige sammenhæng mellem vandstanden målt ved Grejsdalens Planteskole og ved Ny Grejsdalsvej.

Fordelerbygværk

Åbningsprocenten på spjældene ved fordelerbygværket ved Abelones Plads viser sig i nogle tilfælde at påvirke vandstanden ved Ny Grejsdalsvej. På figur 3.11 er vist en hændelse 01/01-2018, hvor åbningsprocenten har en tydelig påvirkning. Det ses, at vandstanden ved Grejsdalens Planteskole (ST. 32.06) ikke er stigende efter klokken 07.00, og vandstandsstigningen ved Ny Grejsdalsvej (ST. 32.22) klokken 09.00 vurderes derfor at skyldes ændringen på spjældet og ikke øget afstrømning fra oplandet.

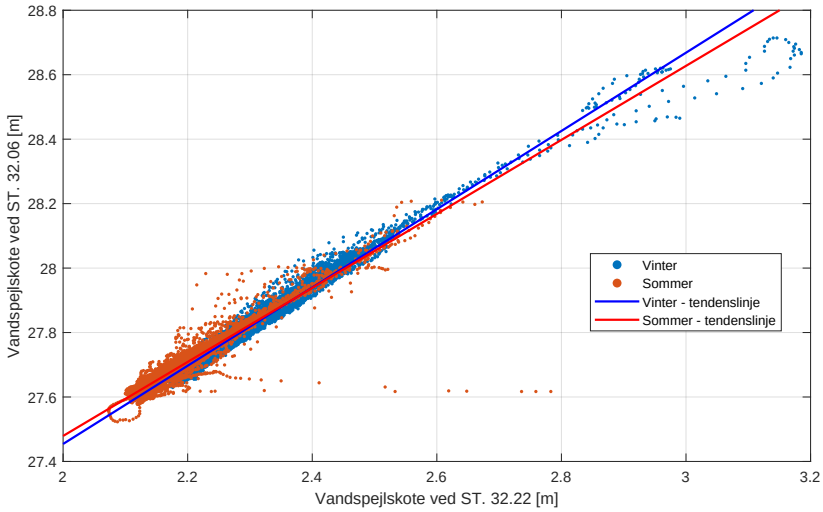


Figur 3.11: Åbningsprocenten på spjældene reduceres og medfører opstuvning til vandstandsmåleren ved Ny Grejsdalsvej (ST. 32.22).

En ændring af åbningsprocenten har dog ikke altid en påvirkning på vandstanden ved Ny Grejsdalsvej. Undersøgelsen af dataene har vist, at hvis der haves en relativt lav vandstand ved Ny Grejsdalsvej og spjældene justeres i intervallet af de "høje"åbningsprocenter, så ses der ingen påvirkning ved Ny Grejsdalsvej. Det giver god mening, fordi spjældene lukker oppe fra og ned.

Varierende Manningtal og tværsnit

Som beskrevet i afsnit 3.1.2 må manningtallet formodes at variere over året, men hvis dette er tilfældet ved begge stationer, kan det potentielt set betyde, at det ikke vil kunne ses på korrelationen mellem vandstanden ved de to stationer. På figur 3.12 ses korrelationen mellem vandstanden ved de to stationer henholdsvis om sommeren og om vinteren. Sommer- og vinterperioden er valgt ud fra, hvornår det henholdsvis laveste og højeste Manningtal forventes. De to tendenslinjer ligger især tæt i det interval, hvor størsteparten af målepunkterne forekommer. Dette tyder på, at årstidsvariationer på Manningtallet har relativt lidt at sige i forhold til korrelationen mellem vandstanden ved de to stationer. Afvigelsen mellem de to tendenslinjer stiger dog ved højere vandstande, men da datagrundlaget her tilsvarende er begrænset, og da det langt fra er sikkert, at korrelationen er lineær, vurderes den afvigende korrelation ved høje vandstande ikke at kunne begrundes af årstidsvariationer.

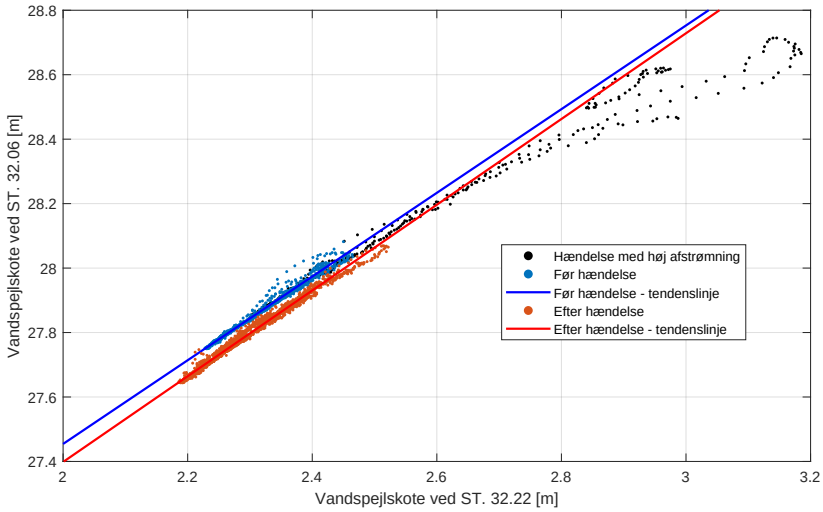


Figur 3.12: Sammenhæng mellem vandspejlskoten målt ved ST. 32.22 og ST. 32.06 henholdsvis om sommeren (juni 2016 - september 2016 og juni 2017 - september 2017) og om vinteren (december 2015 - marts 2016 og december 2016 - marts 2016) med indlagt lineær regression.

I analysen af korrelationen mellem vandstanden ved de to stationer er det fundet, at en hændelse med høj afstrømning kan medføre en tydelig forskydning af korrelationen. Dette er illustreret på figur 3.13. En sådan forskydning af korrelationen formodes at skyldes, at den høje afstrømning har ændret tværsnittet ved erosion eller fjernet grøde og dermed øget Manningtallet. Korrelationen efter hændelsen kan godt være et resultat af, at både vandstanden ved Grejsdalens Planteskole og Ny Grejsdalsvej er reduceret ved den samme vandføring og dermed fremgår den potentielle konsekvens af hændelsen ikke så tydeligt af korrelationsanalysen. Antaget at forskydningen er forårsaget af erosion eller et øget Manningtal, vil konsekvensen af hændelsen være tydeligere i en model, som baserer sig på sammenhængen mellem vandføring og vandstand.

Det kunne også være en mindre logisk ting, som har forårsaget forskydningen, som at vandet under hændelsen har nået et niveau og en hastighed, som potentielt kan have drejet vinklen på vandstandsmåleren en smule. Hvad der er grunden til forskydningen vides ikke med sikkerhed og er umuligt at fastlægge præcist på baggrund af det tilgængelige data. Dog er det værd at lægge mærke til, at sammenlignes figur 3.13 med korrelationen året efter, er det fundet, at korrelationen året efter har rykket sig tilbage til korrelationen før hændelsen. Dette kunne tyde på, at tværsnittet i Grejs Å ved sedimentation er vendt tilbage til åens oprindelige tilstand før hændelsen i slutningen af december 2015.

Drastiske og uforudsete ændringer i systemet som en ændring i Manningtal eller tværsnit på grund af en hændelse er udfordrende at inkorporere i de fleste modeller. I en traditionel model over vandløbet skal der eventuelt medtages erosion som en del af modellen eller ske en løbende kalibrering. Begge disse løsninger vurderes at være omkostningstunge. Den løbende kalibrering er netop en af styrkerne ved en model baseret på neurale netværk, som løbende kalibrerer sig selv på baggrund af målingerne. Hvis det vurderes, at der ofte forekommer tilfælde, hvor forholdene ændres drastisk, og disse ændringer ikke direkte kan kvantificeres som et input til modellen, er det også en mulighed at vægte de sidst målte data højere i træningen af modellen, så modellen hurtigt inkorporerer ændringerne i systemet.



Figur 3.13: Sammenhængen mellem vandspejlskoten målt ved Ny Grejsdalsvej (ST. 32.22) og Grejsdalens Planteskole (ST. 32.06) i december 2015 og januar 2016 før og efter en hændelse med høj afstrømning med indlagt lineær regression.

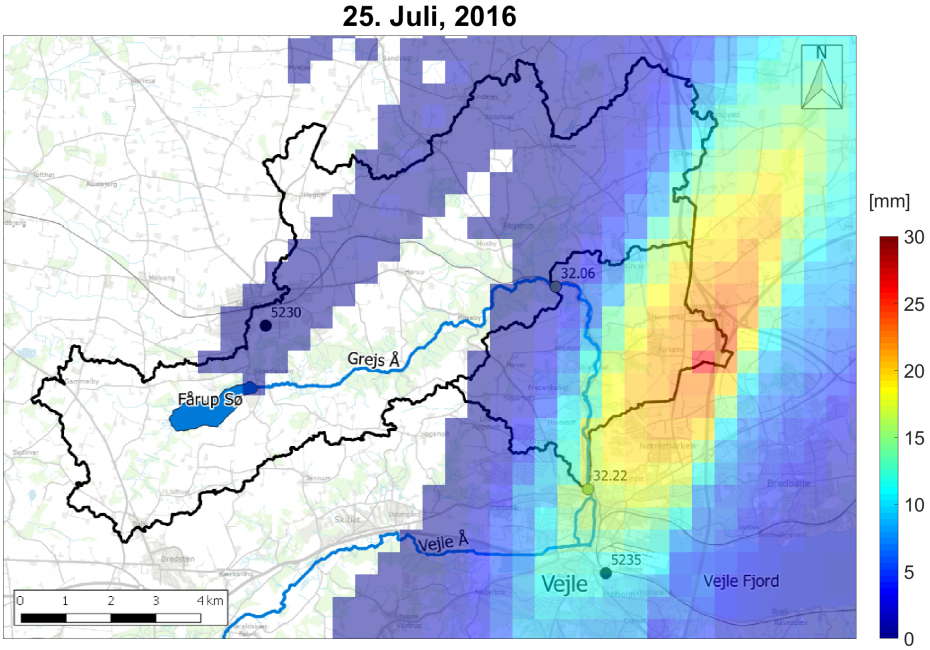
Forsinkelse mellem Grejsdalens Planteskole og Ny Grejsdalsvej

Da stationen ved Grejsdalens Planteskole ligger opstrøms stationen ved Ny Grejsdalsvej forventes det, at der er en forsinkelse på afstrømningen mellem de to stationer. Der er derfor lavet korrelationsanalyser ved forskellige forsinkelsestider. Analysen er vist i appendiks C og viser, at analyseres der på hele måleperioden, så forekommer den bedste lineære korrelation, hvis målingerne ved Grejsdalens Planteskole forsinkes med 26 minutter. Det viser sig dog, at R^2-værdien kun ændrer sig fra 0,944 til 0,945 ved at forsinke målingerne i forhold til at sammenligne målingerne direkte. Forsinkelsen i systemet vurderes derfor ikke at være den væsentligste faktor til begrundelse af afvigelsen på korrelationen mellem de to stationer vist til højre på figur 3.10.

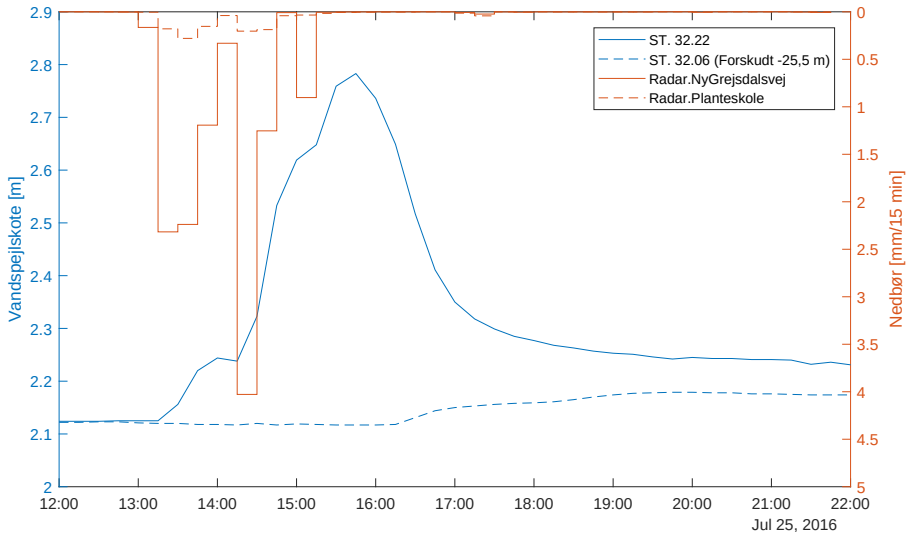
Selvom forsinkelsen ikke umiddelbart gør relationen mellem de to stationer markant bedre, kan det ikke benægtes, at en bølge opstrøms i systemet vil forplante sig nedstrøms. Dette vil formentlig være tydeligt, hvis vandstanden ændres drastisk, og denne information vil potentielt kunne varsle om høj vandstand i Vejle Midtby. Derfor er der lavet korrelationsanalyser på de 10 hændelser, hvor vandstanden ved Grejsdalens Planteskole har været højest, da der burde ses et tydeligere billede af forsinkelsen, hvis der kun analyseres på enkelte hændelser. Analyserne ses i appendiks C. Der viser sig at være en variation i, hvor lang tid målingerne ved Grejsdalens Planteskole skal forsinkes for at opnå den bedste lineære korrelation. Disse tider varierer mellem -45 minutter til 43 minutter. Det ses også, at vandspejlskoten ved Ny Grejsdalsvej er tydeligt påvirket af skift på spjældåbninger ved fordelerbygværket ved enkelte af de analyserede hændelser (eksempelvis figur C.10). Dette besværliggør en analyse af forsinkelsestiden og har uden tvivl en indvirkning på de fundne forsinkelsestider. Det kan desuden være en del af forklaringen på de fundne negative forsinkelser. Der er dog også et eksempel, hvor der findes en negativ forsinkelsestid på -23 minutter, hvor der ikke er tydelig påvirkning fra fordelerbygværket (figur C.3). Der forefindes ikke data om spjældåbinger fra denne hændelse, så det har ikke været muligt at validere, hvorvidt fordelerbygværket har haft en indvirkning. De varierende forsinkelsestider medfører, at det vurderes at være besværligt på en simpel måde at forudsige vandspejlskoten i Vejle Midtby på baggrund af målingerne ved Grejsdalens Planteskole. Det kan dog ikke benægtes på baggrund af denne analyse, at vandstanden ved Grejsdalens Planteskole ikke kan bidrage til modelleringen af Ny Gresjdalsvej, når der anvendes neurale netværk.

Stedlig fordeling af nedbør

Den tydelige afstikker til højre på figur 3.10, hvor vandspejlskoten ved Ny Grejsdalsvej er på 2,8 meter, mens vandspejlskoten ved Planteskolen er på 27,6 meter, skyldes nedbør den 25. juli 2016 i området mellem de to stationer. På figur 3.14 er vist den akkumulerede nedbør i oplandet den 25/07-2016. Hændelsen er ligeledes illustreret på figur 3.15. Der ses en tydelig forskel på mængden af nedbør, som faldt over oplandet.



Figur 3.14: Akkumuleret nedbør den 25/07-2016 i området omkring Vejle.

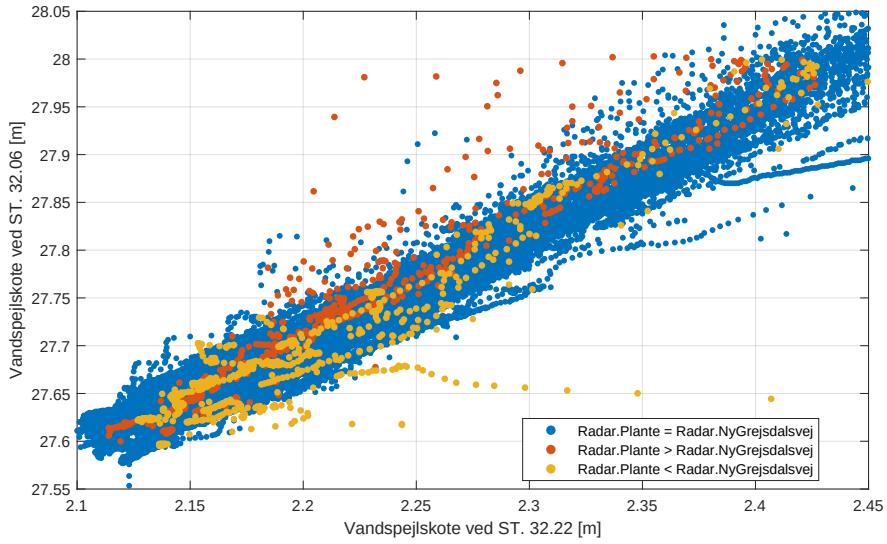


Figur 3.15: Der er tydelig forskel på den gennemsnitlige observerede nedbør i oplandet til ST. 32.06 (Radar.Planteskole) og oplandet mellem ST. 32.06 og ST. 32.22 (Radar.NyGrejsdalsvej).

På de to år, der anvendes data fra, er hændelsen 25/07-2016 en singularitet. Der er ikke andre eksempler på en tilnærmelsesvis så stor afvigelse fra korrelationen mellem vandspejlskoten ved Grejsdalens Planteskole og Ny Grejsdalsvej. Det er dog sandsynligt, at den stedlige fordeling af nedbør i andre tilfælde også i mindre grad har en indflydelse på korrelationen mellem vandspejlskoten ved de to stationer. For at give et billede af betydningen af den stedlige fordeling af nedbør er data omkring vandspejlskoterne på baggrund af radardata inddelt i tre kategorier:

- Der har været målt en nedbørsforskel på under 2 millimeter mellem oplandet til ST. 32.06 og oplandet mellem ST. 32.06 og 32.22 de sidste 10 timer. (Radar.Plante = Radar.NyGrejsdalsvej)
- Der har været målt to millimeter **mere** nedbør i oplandet til ST. 32.06 end i oplandet mellem ST. 32.06 og 32.22 de sidste 10 timer. (Radar.Plante > Radar.NyGrejsdalsvej)
- Der har været målt to millimeter **mindre** nedbør i oplandet til ST. 32.06 end i oplandet mellem ST. 32.06 og 32.22 de sidste 10 timer. (Radar.Plante < Radar.NyGrejsdalsvej)

På figur 3.16 ses det, at flere af punkterne, som falder markant udenfor den generelle korrelationen, hører ind under kategorien, hvor der har været forskel på nedbørsmængder målt i oplandene. Det tyder dog også på, at ikke alle hændelser med forskel i nedbørsmængder resulterer i tydelige afvigelser fra den generelle korrelation.



Figur 3.16: Flere af de tydelige kortvarige afvigelser fra korrelationen kan skyldes stedlig fordelt nedbør.

Sammenfatning af forundersøgelse af forhold gældende i Grejs Å ved Ny Grejsdalsvej

Forundersøgelsen af forholdene gældende ved Ny Grejsdalsvej viser, at der er en tæt korrelation mellem vandspejlskoten ved Grejsdalens Planteskole og vandspejlskoten ved Ny Grejsdalsvej. Dette indikerer, at vandstanden ved Ny Grejsdalsvej primært er styret af afstrømningen fra oplandet. Den tætte korrelation giver et fantastisk grundlag til at fastlægge, hvad der udover afstrømning påvirker vandspejlskoten ved Ny Grejsdalsvej.

Der ses eksempler på, at åbningsprocenten på spjældene ved fordelerbygværket ved Abelones Plads også har en påvirkning på vandspejlskoten ved Ny Grejsdalsvej. Denne påvirkning er kun tydelig, hvis der forekommer en kombination af høj vandstand og lav åbningsprocent.

Årstidsvariationer på Manningtallet ses ikke på korrelationen mellem vandspejlskoten ved Grejsdalens Planteskole og vandspejlskoten ved Ny Grejsdalsvej. Da der må formodes stadig at være en

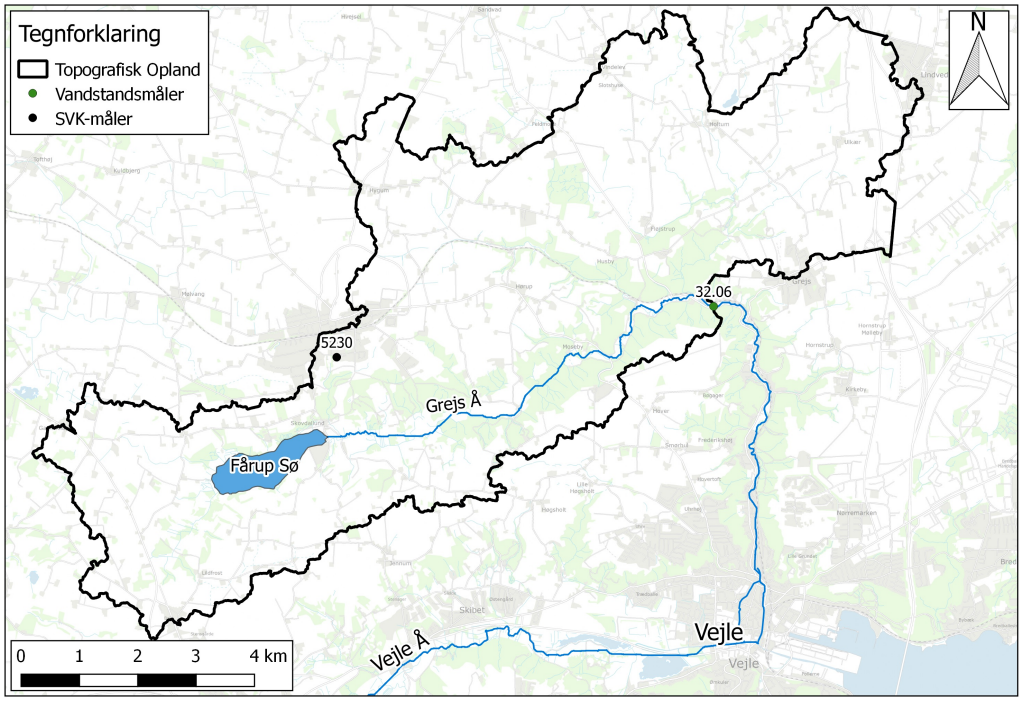
ændring af Manningtallet over året, må dette indikerer, at Manningtallet ikke ændrer sig uafhængigt ved de to målestationer på grund af årstidsvariationer. Det ses dog, at en hændelse med høj afstrømning kan ændre korrelationen mellem de to vandspejlskoter. Dette formodes at skyldes en ændring af Manningtal eller tværsnit grundet erosion.

Da Grejsdalens Planteskole ligger opstrøms Ny Grejsdalsvej, må der forventes at være en naturlig forsinkelse af afstrømningen mellem de to stationer. Den bedste lineære korrelation mellem de to vandspejlskoter er fundet, hvis målingerne fra Grejsdalens Planteskole forsinkes med 26 minutter. Den generelle korrelation forbedres dog kun minimalt, og analyseres der på enkelte hændelser, er der relativ stor variation i den optimale forsinkelsestid.

Den stedlige fordeling af nedbør viser sig at kunne være en væsentlig faktor, når tydelige kortvarige afvigelser fra korrelationen skal beskrives.

3.4 Forundersøgelse af forhold i Grejs Å ved Grejsdalens Planteskole

ST. 32.06 er placeret ved Grejsdalens Planteskole. Stationen ved Grejsdalens Planteskole er stationen, der ligger længst oppe i oplandet til Grejs Å. På figur 3.17 er oplandet til Grejsdalens Planteskole vist.



Figur 3.17: Oplandet til Grejs Å ned til målestationen ST. 32.06 dækker et areal på ca. 63 km^2

ST 32.06 er placeret udenfor større bebyggelsesområder og har ligeledes et opland, der ikke er væsentligt bebygget. Det forventes ikke, at stationen er påvirket af aktiv styring, og det må derfor forventes, at der forekommer en naturlig afstrømning fra oplandet.

Selvom stationen ved Grejsdalens Planteskole ligger langt fra det aktuelle risikoområde ved Omløbsåen, er det er valgt at analysere forholdende gældende ved Grejsdalens Planteskole på grund af to ting:

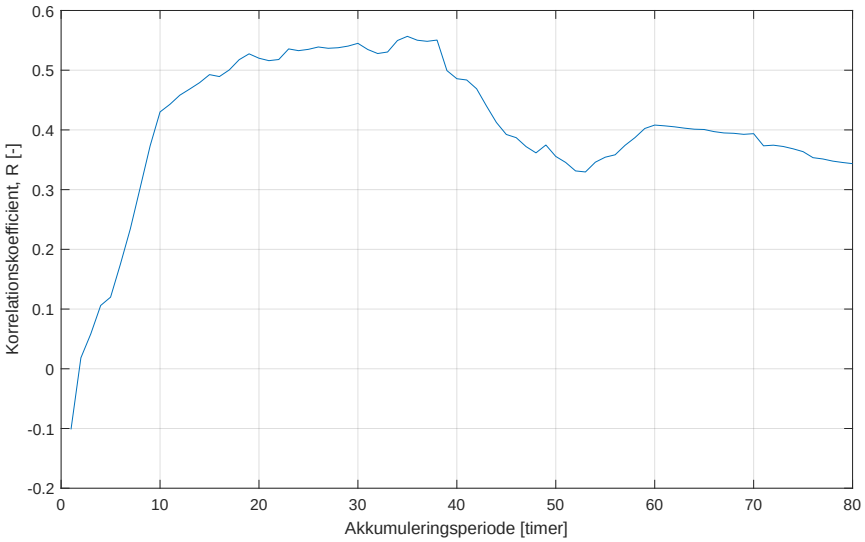
- Den naturlige afstrømning gør stationen fordelagtig til nogle af de indledende analyser af systemet og den grundlæggende opsætning af de neurale netværk.

- Da stationens opland udgør ca. 80% af oplandet til hele Grejs Å, vil resultater fra analysen af denne station kunne anvendes til beskrivelse af forhold nedstrøms.

Simpel sammenhæng mellem målt nedbør og peak på vandspejlskoter

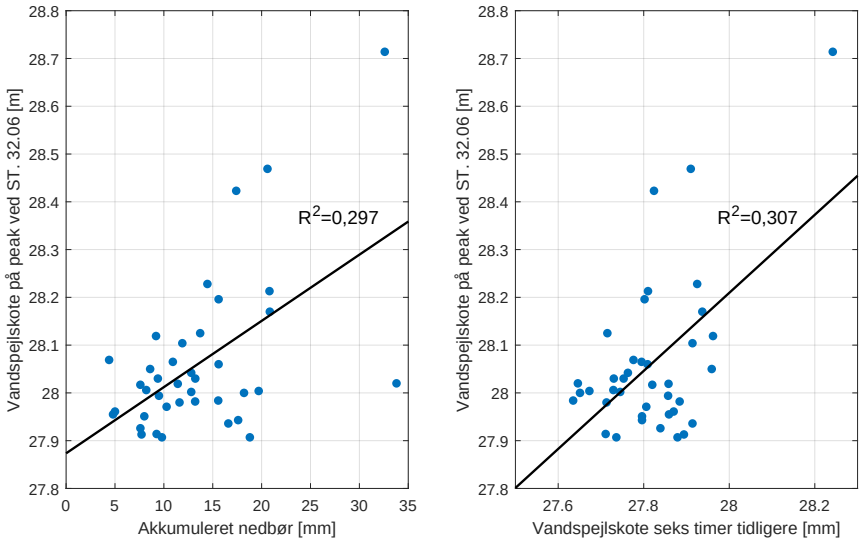
Da der forventes en naturlig afstrømning må udsvingene fra grundvandsafstrømningen ved denne station primært være forårsaget af nedbørsmængder og ikke af aktiv styring. Det er derfor undersøgt, hvorvidt der kunne findes en simpel sammenhæng mellem målt nedbør i Jelling (SVK 5230) og målte peaks på vandspejlskoten målt ved Grejsdalens Planteskole. En sådan sammenhæng ville være praktisk anvendelig i forbindelse med oversvømmelsesvarsling, da prognoser af nedbør er nemt tilgængelige. Som beskrevet i afsnit 3.1 er vandstanden dog et produkt af mange parametre og ikke bare nedbørsmængden. Derfor er det ikke forventet, at en simpel sammenhæng eksisterer, men analysen tydeliggør udfordringerne ved at finde sammenhæng mellem målt nedbør og målte peaks på vandspejlskoten.

Først er det undersøgt ved hvilken akkumuleringsperiode på nedbøren, der findes den bedste korrelation. Det er valgt at medtage alle peaks over kote 27,9 for at få en passende mængde datapunkter. På figur 3.18 ses, at inddragelse af hver enkelt time de første 10 timer øger kvaliteten på korrelationen markant. I intervallet mellem en akkumuleringstid på 10 og 40 timer er korrelationskoefficienten rimelig konstant. Akkumuleres der over 40 timer, aftager korrelationskoefficienten igen.



Figur 3.18: Korrelationskoefficient på den lineære korrelation mellem vandstanden målt ved peaks over kote 27,9 meter ved Grejsdalens Planteskole og akkumuleret nedbør tilbage i tiden fra den målte vandstand.

Plottes den akkumulerede nedbør over de sidste 30 timer op til peaket på vandstanden sammen med vandspejlskoten i peaket ses, at der er stor usikkerhed på den lineære korrelation. Dette er illustreret til venstre på figur 3.19. Til sammenligning findes en tilsvarende korrelationskoefficient ved en lineær korrelation mellem vandspejlskoten i peaket og vandspejlskoten målt seks timer tidligere. Dette bekræfter det forventede: At en simpel sammenhæng mellem nedbør og vandstand ikke findes.



Figur 3.19: Sammenhængen mellem målt kote ved peaks over 27,9 meter ved Grejsdalens Planteskole (ST. 32.06) og akkumuleret nedbør over 30 timer (til venstre). Sammenhængen mellem målt vandspejlskote ved peaks over 27,9 meter ved Grejsdalens Planteskole (ST. 32.06) og målt kote ved Grejsdalens Planteskole seks timer tidligere (til højre).

Sammenfatning af forundersøgelse af forhold i Grejs Å ved Grejsdalens Planteskole

På trods af naturlige afstrømningsforhold ved Grejsdalens Planteskole har det ikke været muligt at finde en simpel sammenhæng mellem nedbør og vandstand, hvilket synliggør behovet for en model, som kan modellere komplekse sammenhænge mellem måledata, som har åbenlyse forbindelser til hinanden. Det gøres traditionelt ved en deterministisk grundvandsmodel, men i dette projekt søges det at anvende et alternativ: Neurale netværk.

Del II - Modellering med neurale netværk

4 Præsentation af neurale netværk 45

4.1 Anvendelse af neurale netværk som prognoseværktøj

4.2 Modellering af sinus-afhængige data med neurale netværk

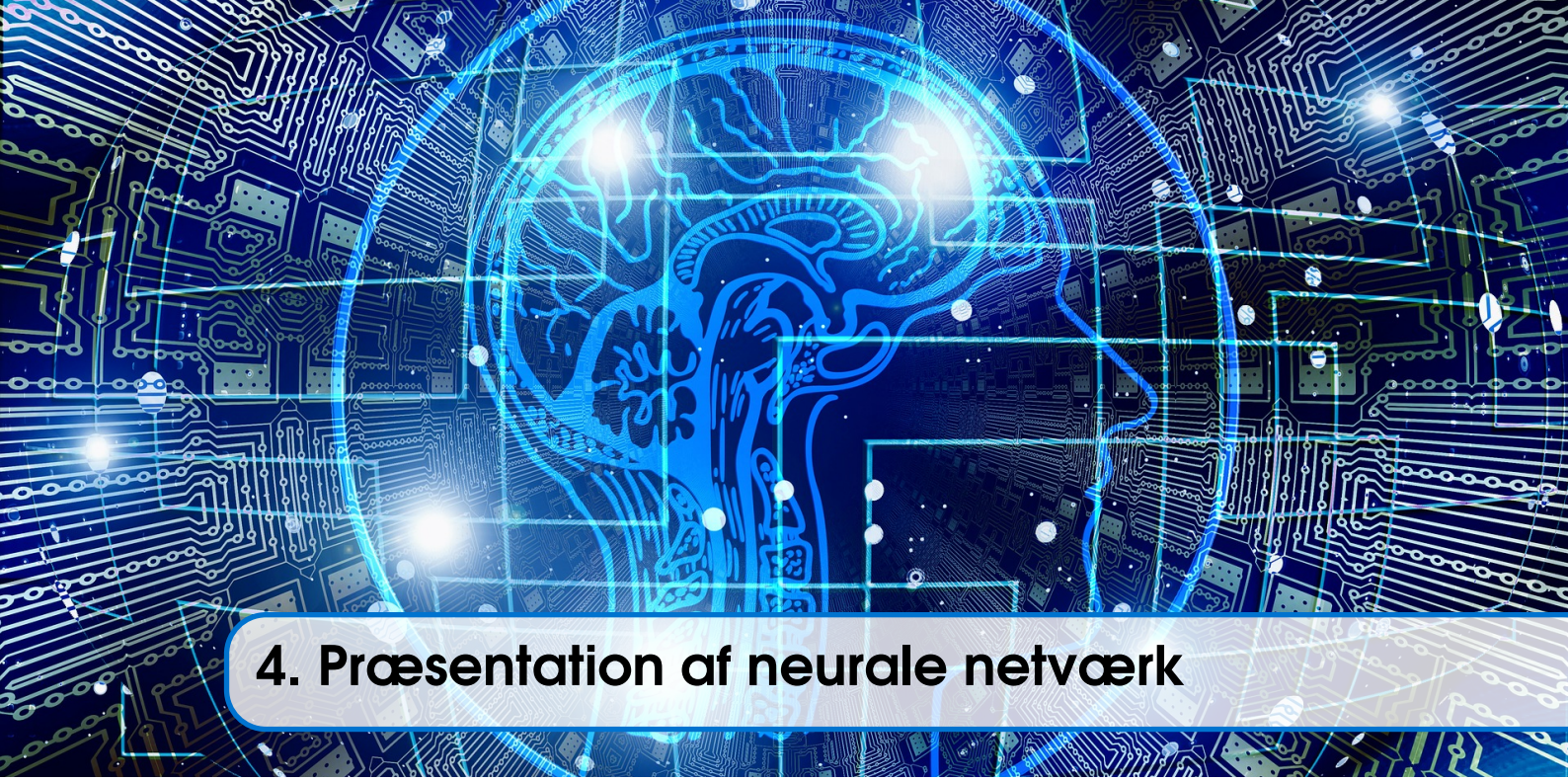
4.3 Sammenfatning af modelleringsprocedure med neurale netværk

5 Opsætning af neurale netværk 55

5.1 Anvendelse af neurale netværk til modellering af vandstande

5.2 Optimering af netværkskonfigurationer til modellering af vandstande

5.3 Sammenfatning på optimering af netværkskonfigurationer



4. Præsentation af neurale netværk

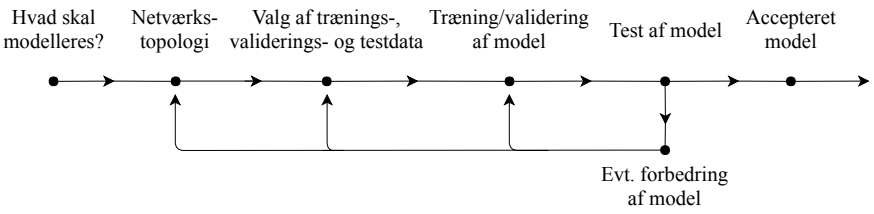
Brugen af neurale netværk til modellering af vandstande er utraditionel, og metoden er ny og alternativ. I dette kapitel beskrives fremgangsmåden gennem et eksempel med en sinus-kurve for at få grundbegreberne på plads, før de skal anvendes til modellering af vandstande i Vejle Midtby og opland.

4.1 Anvendelse af neurale netværk som prognoseværktøj

Anvendelse af neurale netværk til prognosticering af ikke-lineære tidsserier er igennem de senere år blevet mere og mere populær. Afhængigt af formålet, anvendes netværkene til prognoser på mange forskellige tidsskalaer af alt fra forudsigelse af elektricitetsbelastningen om få minutter til befolkningstilvæksten på årsbasis med nogle lovende resultater (Singh m.fl. 2018; Benzer 2015). Dog er det et relativt nyt felt indenfor vand og miljø-området, hvor metoden stadig er under udvikling. Der findes eksempler på modellering af vandføringer og vandstande i vandløb med neurale netværk, hvor prognosekvaliteten ved forskellige tidshorisonter udforskes, men hvor der ikke har været konkrete problemstillinger at løse (Kasiviswanathan m.fl. 2016; Chang m.fl. 2014). Det er ikke tilfældet i dette projekt, hvor det ønskes at prognosticere vandstanden i Vejle Midtby, så det er muligt at varsle oversvømmelser, så de enten undgås eller håndteres af beredskabet.

4.1.1 Fremgangsmåde for modellering med neurale netværk

Neurale netværk er noget, der oprindeligt blev udviklet til at lave en model af hjernen, som senere skulle vise sig at være anvendeligt til løsninger af ikke-lineære ligningssystemer og opbygning af kunstig intelligens (McCulloch m.fl. 1943). Derfor er den klassiske fremgangsmåde udviklet på baggrund af matematiske eksempler, som intet har at gøre med forudsigelse af vandstandsændringer i vandløb. Fremgangsmåden for modellering med neurale netværk er dog langt hen ad vejen ens på tværs af faggrupper. Fremgangsmåden for modellering med neurale netværk i dette projekt fremgår af figur 4.1.

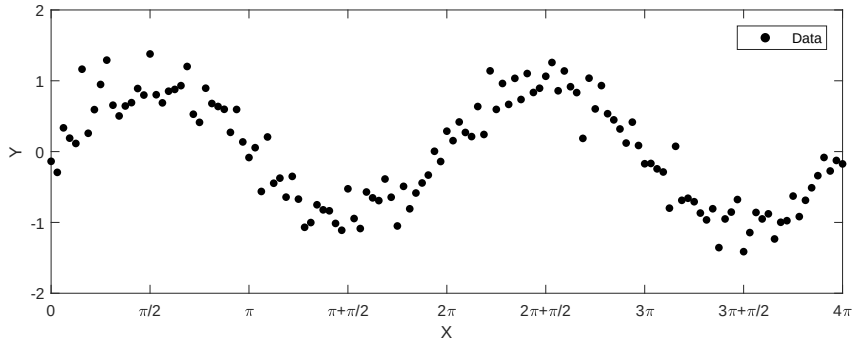


Figur 4.1: Fremgangsmåden for modellering med neurale netværk.

Fremgangsmåden gennemgås bedst gennem et eksempel.

4.2 Modellering af sinus-afhængige data med neurale netværk

I det følgende vil fremgangsmåden for modellering af neurale netværk blive gennemgået gennem et eksempel, hvilket i dette tilfælde er en sinuskurve med noget tilfældigt indlagt støj, som vist på figur 4.2. Eksemplet er ren matematik, men det er relevant for modellering af vandstande, da vandstandsmålinger basalt set også består af et signal og noget støj.



Figur 4.2: Tilfældigt genereret data, som det ønskes at modellere med et neuralt netværk.

Den grundlæggende modelleringsprocedure vil blive præsenteret med det formål at give en overordnet forståelse, og derfor vises bagvedliggende beregninger i appendiks D. Al modellering baseres på Matlab's *Neural Net Fitting Tool*, og der anvendes begreber, som er specifikke for neurale netværk (MathWorks 2017). For yderligere præcision af disse begreber henvises til bogen *"Neural Networks - A Systematic Introduction"*, som har dannet grundlag for følgende afsnit. (Rojas 1996)

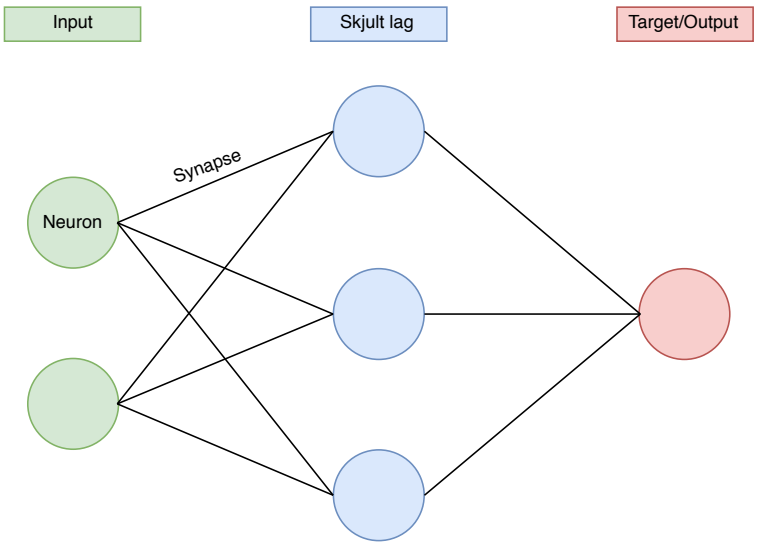
4.2.1 Hvad skal modelleres?

Første skridt i modelleringen er at fastlægge, hvad det ønskes at modellere, og på det punkt adskiller proceduren sig ikke fra andre modelleringsværktøjer. Den store forskel ligger i, at det kræver målinger af det, det ønskes at modellere. Disse målinger kaldes for ***target***.

4.2.2 Netværkstopologi

Når target er valgt, skal strukturen, også kaldet topologien, af den resterende del af det neurale netværk bygges op. Det kan gøres på flere forskellige måder, men i dette projekt fokuseres der på ***feedforward***-netværk, som i dette tilfælde består af to lag udover target. Dels et ***skjult lag***, som består af et prædefineret antal ***neuroner***, og dels af et ***inputlag***, som består af værdier, der på en eller anden måde er korrelerede med target. De tre lag bindes sammen af ***synapser***, som har til

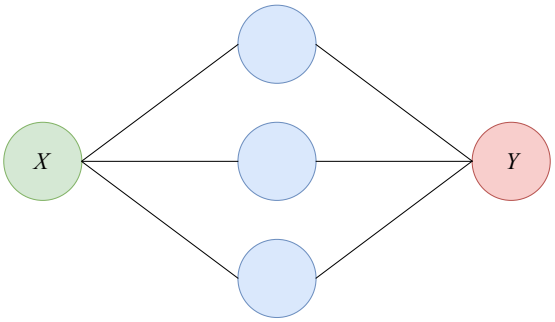
opgave at transmittere og vægte signalerne igennem netværket. En typisk netværkstopologi ses af figur 4.3.



Figur 4.3: Illustration af et feedforward-netværk. Forskellen mellem target og output er, at target anvendes til indlæring af netværket, mens outputtet er det egentlige resultat.

Af denne opbygning fremgår det også, at feedforward-netværket ikke rummer nogen hukommelse. Det vil sige, at outputtet altid baseres på input, og ikke på hvad tidligere output har vist. (M. R. Rasmussen 1992)

For at tage udgangspunkt i eksemplet med datapunkterne i figur 4.2, skal der vælges et input, der er korreleret med target. Det ses, at der er tydelige harmoniske udsving afhængigt af værdien på X-aksen. Derfor vælges X som input til netværket, som vist på figur 4.4.



Figur 4.4: Det vælges at anvende X som input.

Det vælges i første omgang at anvende tre neuroner i det skjulte lag, hvilket kan vise sig enten at være for mange eller for få. Antallet af neuroner styrer antallet af frihedsgrader i netværket, og derfor kan for mange neuroner medføre overfitting. For få neuroner har den omvendte effekt forstået på den måde, at der ikke er nok frihedsgrader til at beskrive mønsteret mellem target og input.

4.2.3 Valg af trænings-, valideringsdata og testdata

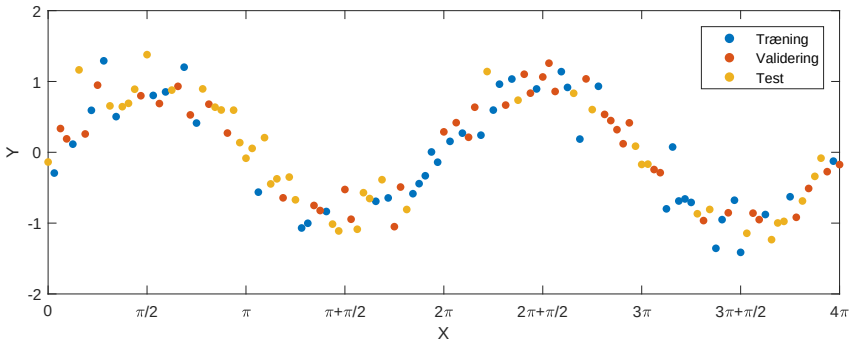
Indlæringen af et neuralt netværk bygger på en iterativ proces, hvor det forsøges at minimere fejlen mellem target og modeloutput. Til det formål udvælges tre dele af target: en til træning, en til validering og en til test af netværket.

Træningen er den del af target, som det neurale netværk forsøger at få til at stemme overens med modeloutputtet ved kalibrere frihedsgraderne.

Sideløbende med træningen afprøves netværket på valideringsdelen, som er en uafhængig del af træningen, ved hver iteration for at undersøge, om netværket er ved at ***overfitte***. Overfitting er en risiko, fordi det neurale netværk næsten altid kan finde frihedsgrader, der beskriver træningsdelen virkelig godt. Uden en løbende validering er den risiko endnu større, fordi valideringen sørger for, at træningen stopper, før netværket overfittes. Derfor giver det heller ikke mening at snakke om kalibrering af neurale netværk, fordi træningen og validering er to vidt forskellige ting, men som begge er elementer i kalibreringen.

Endelig haves testen, som, ligesom valideringen, er en uafhængig del af target. Forskellen fra valideringen er, at testen først foretages, når det neurale netværks frihedsgrader er justeret helt på plads. Derfor er testen det, som normalt kaldes for validering.

Traditionelt udvælges de tre dele tilfældigt, hvilket er udmærket i dette eksempel, men som er uhensigtmæssigt i forbindelse med modellering af vandstande. Det uddybes i kapitel 5, men eksempelvis er det fordi, at trænings- og valideringsvandstande kan minde så meget om hinanden, at valideringen ikke viser uregelmæssigheder, når der forekommer overfitting. Den tilfældige, men ligelige, inddeling af target ses af figur 4.5.



Figur 4.5: Af target udvælges tilfældige værdier til træning, validering og test, og der afsættes en tredjedel til hvert formål.

Herefter er næste skridt at træne/validere modellen.

4.2.4 Træning/validering af model

Træningen af netværket styres af en træningsalgoritme, som kan justere frihedsgraderne på plads. I dette projekt anvendes *Levenberg-Marquardt*-algoritmen, som er hyppigt anvendt til indlæring af feedforward-netværk baseret på tidsserier, blandt andet fordi det er en af de hurtigste (Levenberg 1944; Marquardt 1963; Okkan 2011). Algoritmen består af en fremadrettet beregning og en bagudrettet beregning. Den fremadrettede beregning er første skridt, hvor der gættes på et sæt frihedsgrader, hvorefter der haves et output, som den bagudrettede beregning kan tage udgangspunkt i. Her minimeres fejlen, i dette tilfælde *mean squared error* (MSE), mellem output og træningsdelen af target ved ***backpropagation*** og ***gradientmetoden***, og træningen stopper herefter, når den fejl er acceptabelt lav, eller at fejlen bliver større i valideringen (Straeter 1971).

Det er valgt at benytte MSE som evalueringsparameter, fordi den er følsom overfor fejlestimater. Således straffer MSE store fejl hårdt, hvilket fremgår af ligning 4.1.

$$MSE = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n} \tag{4.1}$$

Hvor

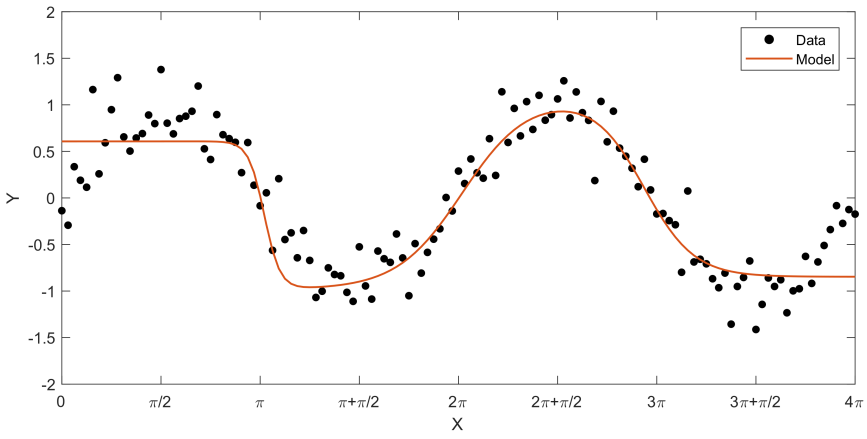
MSE	Mean squared error	[-]
n	Antal datapunkter	[-]
$\hat{y}$	Modelleret værdi	[-]
y	Observeret værdi	[-]

Ved evaluering af modeller til forudsigelser af vandstande er det brugbart, at store fejl straffes hårdere, da fejlene typisk er koncentreret omkring forhøjede vandstande, hvor det største fokuspunkt ligger. I det tilfælde omregnes MSE dog til den tilsvarende *root mean squared error* (RMSE), som er kvadratroden af MSE, for at omskrive evalueringsparameteren til enheden meter. I dette eksempel er den omskrivning ligemeget, fordi værdierne på y-aksen er enhedsløse, og at der er fokus på beskrivelsen af modelleringsproceduren, som ikke ændres på baggrund af evalueringsparameteren.

Levenberg-Marquardt-algoritmen bygger på nogle stopkriterier, der manuelt kan fastsættes før træningen påbegyndes. Eksempelvis kan der sættes en tidsbegrænsning eller et maksimalt antal ønskede iterationer eller et mål for hvor stor en MSE, der accepteres. Fordelen ved at anvende stopkriterier på denne måde er, at træningen forløber hurtigt. Til gengæld konvergerer træningen ikke altid mod et globalt minimum, eller det samme minimum, og det afhænger af, hvor godt frihedsgraderne i den fremadrettede del af beregningen gættes. (MATLAB 2006)

4.2.5 Test af model

Når træningen er afsluttet visualiseres modellen, så tydelige fejl kan identificeres. Det fremgår af figur 4.6, at selvom netværket er trænet på plads, og fejlen er minimeret, så betyder det ikke nødvendigvis, at resultatet er godt. Det kan diskuteres, hvordan kvaliteten af modellen skal evalueres, men i dette tilfælde ønskes en model, der kan gengive tendensen i data, uden den påvirkes for meget af det systematiske støjsignal.



Figur 4.6: Selvom netværket er trænet på plads, betyder det ikke nødvendigvis, at resultatet er godt.

Det ses, at modellen ikke kan gengive target i intervallet $[0, \pi]$, og derfor er der plads til forbedringer.

4.2.6 Eventuel forbedring af model

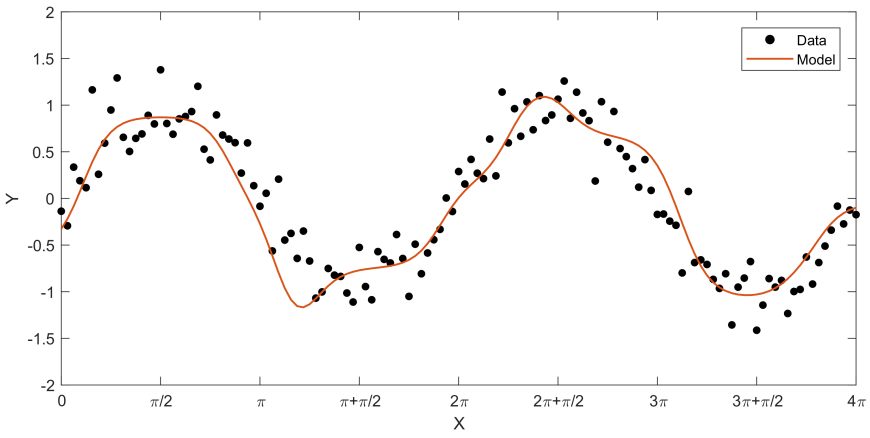
Af visualiseringen på figur 4.6 vurderes det, at netværket ikke gengiver mønstrene mellem input og target med den ønskede nøjagtighed, hvilket kan skyldes en eller flere af følgende ting:

- Dårlig netværkstopologi.
- Dårligt valg af input.
- Dårligt inddeling af target.
- Modellen er trænet ned i et lokalt minimum, hvor fejlen ikke er reduceret tilstrækkeligt.

Ændring af netværkstopologi

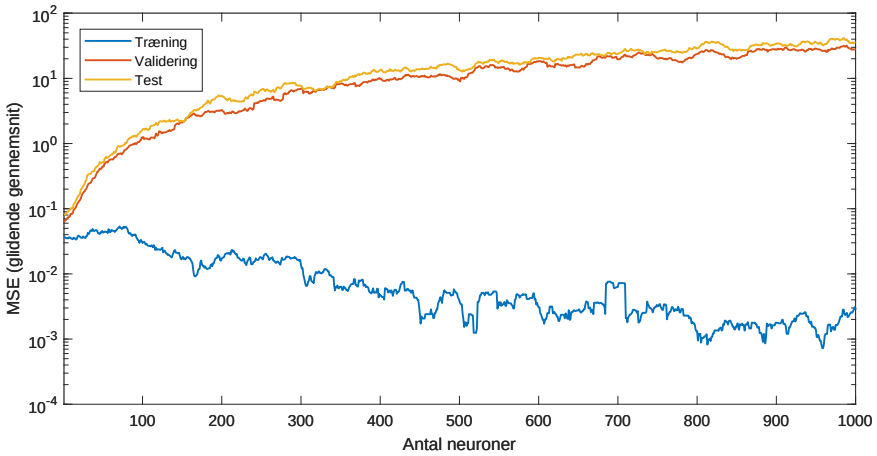
Med dårlig netværkstopologi menes, at der enten er for mange eller for få neuroner. I tilfældet på figur 4.6 vil det i givetfald være for få neuroner, og dermed for få frihedsgrader, fordi netværket ikke kan beskrive tendensen i data korrekt. Det forsøges derfor at lave en model med samme input og target, men hvor der nu er ti neuroner i det skjulte lag.

Med flere neuroner har netværket bedre mulighed for at vægte og transmittere inputsignalet videre i netværket. Det nye resultat ses på figur 4.7.



Figur 4.7: Resultatet af et netværk, hvor antallet af neuroner forøges fra tre til ti.

Ved visuel inspektion konstateres det, at et netværk med ti fremfor tre neuroner resulterer i en model, der bedre kan beskrive tendensen i data, men det ses også, at modellen forsøger at tilpasse sig noget af støjen i dette tilfælde. Det er dog ikke altid reglen, at flere neuroner giver en bedre model. For mange neuroner vil forårsage overfitting, fordi netværket har så mange frihedsgrader, at det uden problemer kan gengive træningsdataen præcist, mens valideringen og testen er dårlig. Udviklingen i MSE som funktion af antallet af neuroner ses på figur 4.8.



Figur 4.8: Forskellen mellem MSE i træningen, valideringen og testen bliver større, når antallet af neuroner øges. Ved 1000 neuroner er forskellen flere størrelsesordner.

Der vises et glidende gennemsnit, da det primært er tendensen i MSE, der er interessant. Udsvingene undervejs skyldes, at modellen lander i forskellige minima, mens den generelle tendens skyldes forøgelsen i antallet af neuroner i det skjulte lag. Det ses, at 1000 neuroner forbedrer MSE i træningen med to størrelsesordner fra 10^{-1} til 10^{-3}, mens valideringen og testen bliver tilsvarende dårligere. Det handler derfor om at finde en balance, hvor der er nok neuroner til at beskrive mønstrene i træningsdata korrekt, men hvor MSE i testen og valideringen holdes i nogenlunde samme størrelsesorden.

Ændring af input

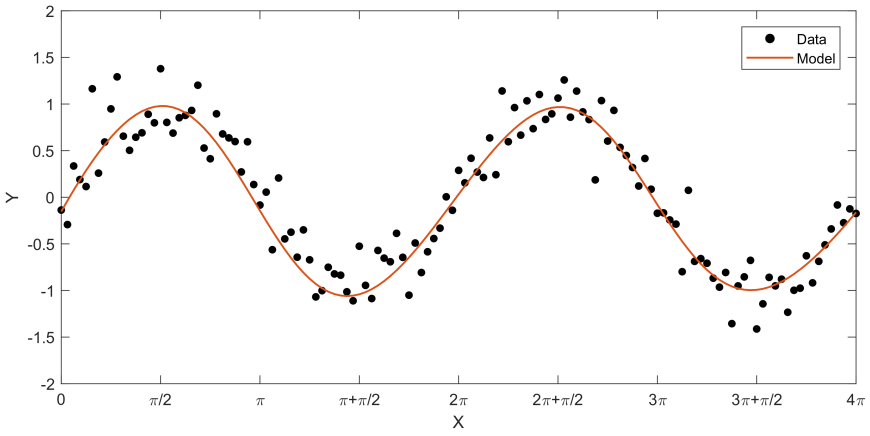
Vurderet ud fra figur 4.6 så er det ikke et forkert valg af input, der forårsager et dårligt netværk. Dermed er det ikke sikkert, at inputtet ikke er valgt forkert, men det viser sig først tydeligt, når netværket skal anvendes på sammenhængende testperioder, ligesom det er tilfældet ved modellering af vandstande.

Bedre inddeling af target

Inddelingen af target har en betydning forstået på den måde, at modellen ikke har mulighed for at gengive mønstre, som den ikke har set før. Betydningen kan ikke beskrives godt i dette eksempel, fordi dataene inddeles tilfældigt og usystematisk fra inddeling til inddeling, og derfor vil mere træningsdata ikke automatisk give en lavere MSE-værdi i træningsperioden, da den ligeså meget er bestemt af, hvordan træningsdata ligger, og hvilket minimum træningen er landet i. Derfor kan betydningen af inddelingen bedre beskrives ved modellering af vandstande, hvor der ses tydelige forbedringer ved at inddele target i perioder fremfor tilfældigt, hvorfor denne analyse rykkes til kapitel 5.

Lokalt/globalt minimum

Et problem der går igen ved løsninger af ikke-lineære ligningssystemer er risikoen for at lande i et lokalt minimum, hvor fejlen ikke er reduceret tilstrækkeligt. Det er ikke et unikt problem for neurale netværk, men alligevel noget der skal være fokus på under evalueringen. I de fleste tilfælde kan den problemstilling løses ved at gentræne netværket et par gange for at undgå et eventuelt dårligt netværk, hvilket giver netværket på figur 4.9.

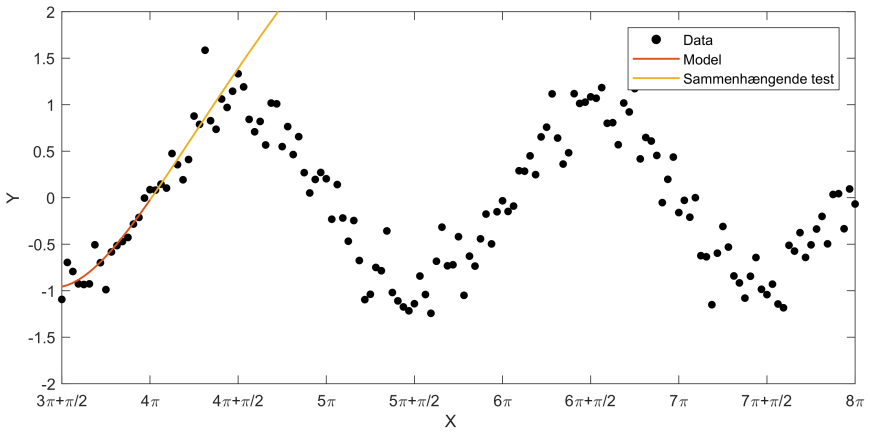


Figur 4.9: Resultatet af det samme netværk som vist på figur 4.6, men hvor netværket er gentrænet et par gange.

På den baggrund er konklusionen, at den oprindelige model var resultatet af en dårlig træning.

4.2.7 Accept af model

På baggrund af forsøgene på at forbedre netværket er det vurderet, at modellen vist på figur 4.9 er den bedste. Den gengiver tendensen i data og indlærer ikke støj. Støjen er dog noget der burde fjernes på forhånd, ligesom der i øvrigt burde laves datavalidering på vandstandsmålinger. Den bedste model afprøves på en sammenhængende test, der strækker sig ud over det oprindelige data. Resultatet ses på figur 4.10.

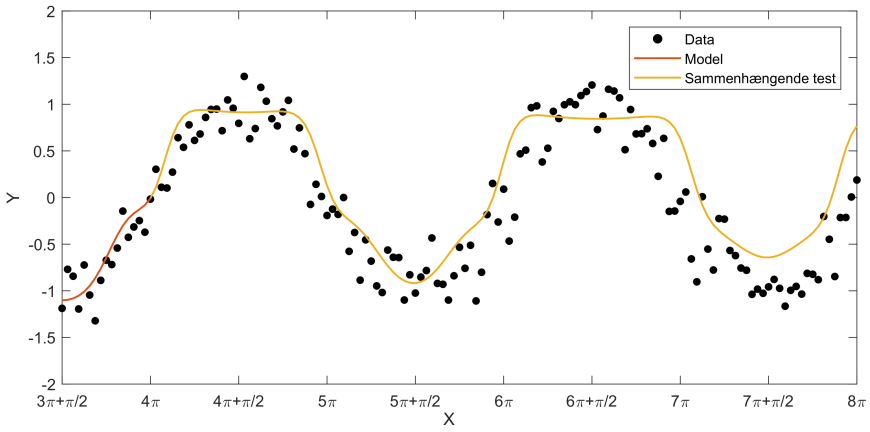


Figur 4.10: Den bedste model egner sig ikke godt til at beskrive en sammenhængende testperiode, fordi det er et mønster, netværket ikke har forudsætningerne for at genkende.

Det ses, at den bedste model ikke har forudsætningerne for at gengive data i en ny periode, hvilket er naturligt, da netværket ikke har været præsenteret for de mønstre før. Det giver god mening, for modellen kan sammenlignes med et netværk, hvor vandstandene beskrives på baggrund af datoen. Der vil det være naturligt, at modellen ikke kan operere på en sammenhængende efterfølgende test.

4.3 Sammenfatning af modelleringsprocedure med neurale netværk

Når der skal laves et godt neuralt netværk, er det en fordel at have valgt den rigtige netværkstopologi, undgå lokale minima og vælge trænings-, valideringss- og testdata fornuftigt. Det erfares, at den store gevinst ligger i at vælge de rigtige input, fordi netværket vil have bedre mulighed for at operere på data, som det ikke har været præsenteret for før. Figur 4.11 viser, hvordan modellen uden videre kan testes i en sammenhængende periode, når der tilføjes en inputneuron med sin(X). Det var ikke oplagt at tilføje et ekstra input, fordi det oprindelige accepterede netværk vist på figur 4.9 gengav target godt.



Figur 4.11: Et netværk, hvor sin(X) tilføjes som input, kan til dels beskrive en sammenhængende testperiode.

Betydningen af valget af input bliver tydeligere, når netværkene skal anvendes til at modellere vandstande. Det er en generel erfaring, at jo større korrelation der er mellem input og target, des bedre bliver netværket, og derfor ligger det primære fokus i at vælge nogle smarte input, når de optimale netværkskonfigurationer skal bestemmes.

Valg af smarte input kan være en udfordring, når der senere skal modelleres vandstande i stedet for et konstrueret eksempel. I afsnit 3.1 er de styrende processer for vandstanden i et vandløb beskrevet, og i de processer indgår nogle fysiske parametre, som ikke er målbare. Reelt måles kun vandstanden og nedbøren. Derfor vil det gennemgående fokuspunkt for den resterende del af rapporten primært være at få disse målinger omsat til brugbare input, selvom baggrundsviden siger, at relationen mellem nedbør og vandstand er langt mere kompleks.

5. Opsætning af neurale netværk

I dette kapitel bliver der redegjort for, hvordan neurale netværk kan anvendes til at modellere vandstande i Grejs Å. Der lægges vægt på at benytte erfaringerne fra kapitel 3 til at optimere modellerne - blandt andet ved at vælge input, da gode input har vist sig at have stor betydning for modelkvaliteten.

5.1 Anvendelse af neurale netværk til modellering af vandstande

Med udgangspunkt i fremgangsmåden beskrevet i kapitel 4 ønskes det at beskrive de metoder, som anvendes til at optimere netværkskonfigurationerne målrettet til brug af prognosticering af vandstande. Det ønskes at koble fremgangsmåden med den viden, som er indsamlet i kapitel 3, så metoden kan målrettes til forudsigelse af vandstande i Vejle Midtby og opland. På det grundlag undersøges det, hvordan:

- Vandstanden ved Grejsdalens Planteskole (ST. 32.06) kan anvendes til modellering af vandstanden ved Ny Grejsdalsvej (ST. 32.22).
- Nedbør kan anvendes til at forbedre prognosen af vandstanden ved Ny Grejsdalsvej (ST. 32.22).
- Netværkskonfigurationerne kan tilpasses, så indlæringen af de neurale netværk bliver bedst mulig.

Det direkte sammenligningsgrundlag i følgende analyser er RMSE-værdier med mindre det i enkelte tilfælde giver mening at vise eksempler. Det skyldes, at RMSE-værdierne erfares at være udmærket mål for, hvor godt modellerne kan ramme peak-vandstande, eksempelvis fordi RMSE straffer store fejl hårdt som beskrevet i afsnit 4.2.4.

Til undersøgelsen er der udviklet en standardkode, som anvendes til at udvikle alle modeller gennem rapporten. Den ses i appendiks E.

5.1.1 Brugen af vandstand som input til de neurale netværk

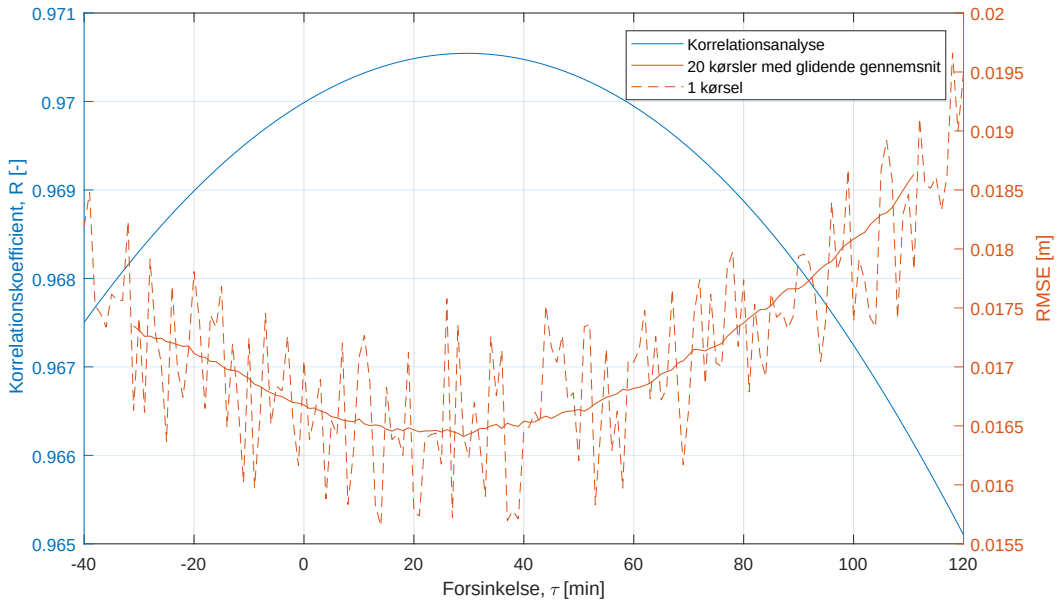
I følgende analyse gennemgås optimering af input i form af vandstande med udgangspunkt i modellering af vandstanden i Grejs Å ved Ny Grejsdalsvej (ST. 32.22) på baggrund af vandstanden ved Grejsdalens Planteskole (ST. 32.06).

I afsnit 3.3 er det beskrevet, at der set over hele serien forekommer den bedste korrelation mellem de to stationer ved en forsinkelse på 26 minutter af målingerne ved Grejsdalens Planteskole. Det testes derfor, om den information kan anvendes til at optimere inputtet fra Grejsdalens Planteskole. Netværkskonfigurationerne ses af tabel 5.1.

Tabel 5.1: Netværkskonfiguration til modellering af Ny Grejsdalsvej (ST. 32.22) på baggrund af Planteskolen (ST. 32.06).

Parameter	Konfiguration
Input	Planteskolen (ST. 32.06)
Target	Ny Grejsdalsvej (ST. 32.22)
Antal neuroner i det skjultelag	10
Træning/validering	10-Dec-2015 → 31-Dec-2016: 50% træning og 50% validering
Test	01-Jan-2017 → 06-Jan-2018: 100% test

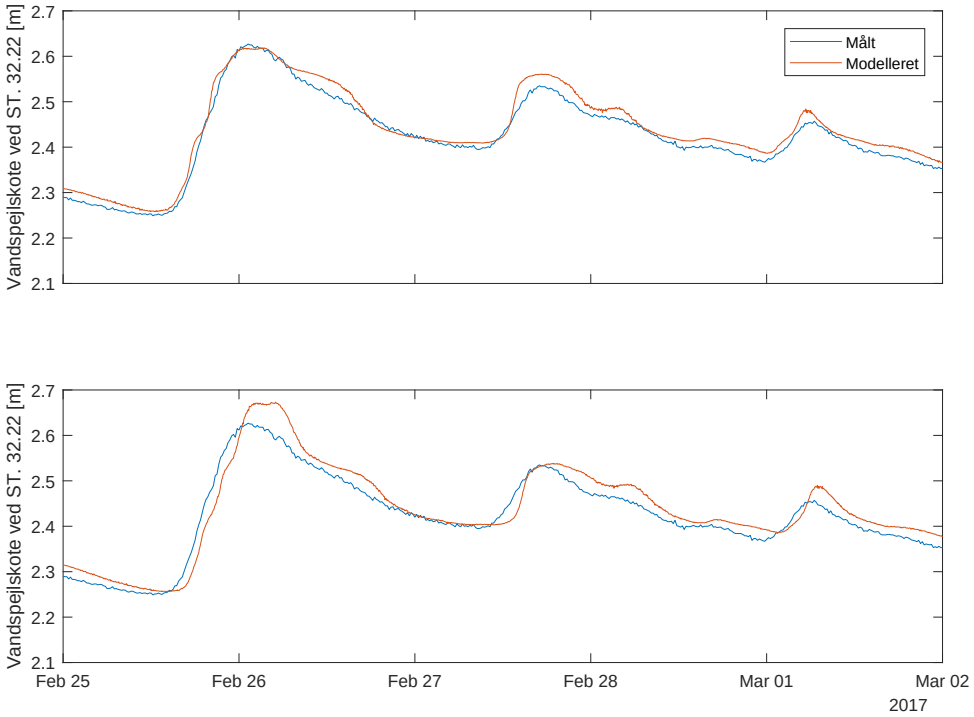
Figur 5.1 viser, at der er sammenfald mellem den forsinkelse, der giver størst korrelation mellem de to vandstandsmålere, og den forsinkelse, der overordnet set giver det bedste input.



Figur 5.1: Korrelation mellem vandstand målt i Grejs Å ved Grejsdalens Planteskole (ST. 32.06) og ved Ny Grejsdalsvej (ST. 32.22) i perioden 10/12-2015 til 06/01-2018.

Skalaen på y-aksen på figur 5.1 viser, at korrelationskoefficienten varierer mellem $R = 0,967$ og $R = 0,975$ i intervallet [-40, 120], og at der ligeledes er tale om variation i RMSE på tredje decimal i samme interval. Det kan virke voldsomt, at så små variationer skulle have en betydning for,

hvor god modellen er. Figur 5.2 viser målte og modellerede vandstande i perioden 25/02-2017 til 02/03-2017, som er en af de mest regnfulde perioder i hele testperioden, hvor vandstanden ved Grejsdalens Planteskole (ST. 32.06) forsinkes med det optimale $\tau = 26\,\text{min}$ øverst og derefter med den værste $\tau = 120 min$ i forhold til Ny Grejsdalsvej (ST. 32.22).



Figur 5.2: Forsinkelsen mellem Grejsdalens Planteskole (ST. 32.06) og Ny Grejsdalsvej (ST. 32.22) har en betydning for modellens timing. Øverst ses modellerede vandstande, når forsinkelsen er på $\tau = 26\,\text{min}$. Nederst ses modellerede vandstande, når forsinkelsen er $\tau = 120\,\text{min}$.

Det ses, at gevinsten ved at forsinke input i forhold til target på den måde, primært ligger ved en forbedring i modellens timing. Til gengæld fremgår det af appendiks C, at den optimale forsinkelse kan variere betydeligt i tilfælde af forhøjede vandstande, hvilket er problematisk, når det især er modellering af forhøjede vandstande, der er fokus på. Derfor er det heller ikke altid anvendeligt at benytte vandstanden ved Grejsdalens Planteskole som input til en model, der skal beskrive Ny Grejsdalsvej.

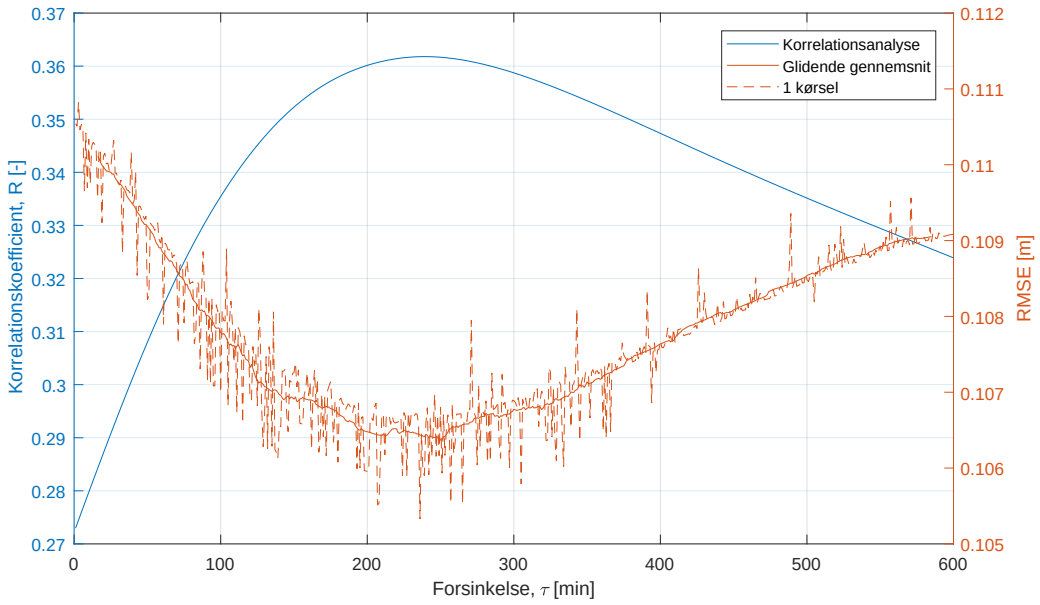
5.1.2 Brugen af nedbør som input til de neurale netværk

Med udgangspunkt i modellering af vandstanden ved Ny Grejsdalsvej (ST. 32.22) erfares det, at der er en forsinkelse mellem registreret nedbør i regnmåleren i Jelling (SVK 5230) til der måles en egentlig vandstandsændring. Det er vanskeligt at finde en simpel korrelation mellem målt nedbør og vandstand på et senere tidspunkt, da korrelationen er stærkt afhængig af afstrømningsforholdene. Det forsøges at lave et netværk, der anvender korrelationen mellem fire timers akkumuleret nedbør og vandstanden ved Ny Grejsdalsvej (ST. 32.22). Netværkskonfigurationerne ses i tabel 5.2.

Tabel 5.2: Netværkskonfiguration til modellering af Ny Grejsdalsvej (ST. 32.22) på baggrund af Jelling (SVK 5230).

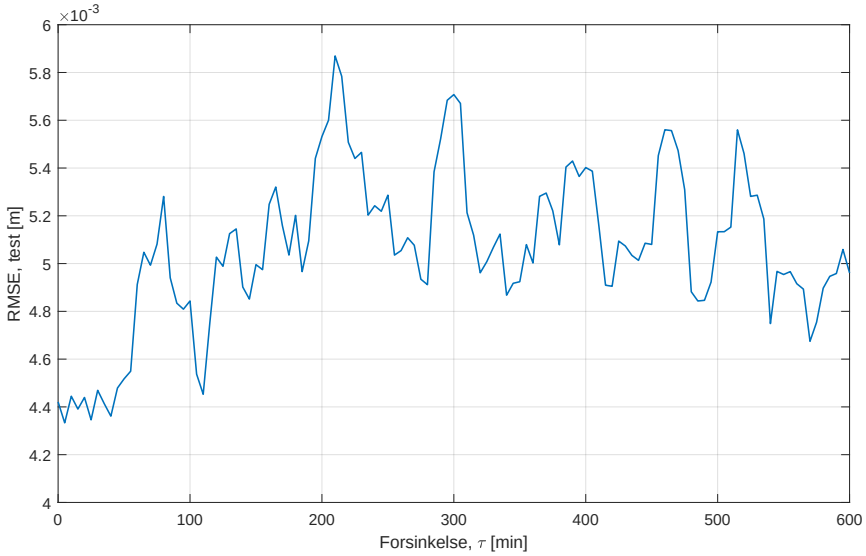
Parameter	Konfiguration
Input	Jelling (SVK 5230): 4t
Target	Ny Grejsdalsvej (ST. 32.22)
Antal neuroner i det skjultelag	10
Træning/validering	10-Dec-2015 → 31-Dec-2016: 50% træning og 50% validering
Test	01-Jan-2017 → 06-Jan-2018: 100% test

Figur 5.3 viser, at der er sammenfald mellem den forsinkelse, der giver størst korrelation mellem de to målere, og den forsinkelse, der overordnet set er det bedste input, ligesom det var tilfældet på figur 5.1. Ved sammenligning med figur 5.1 konstateres det, at korrelationen er markant ringere, og at RMSE-værdien er blevet en størrelsesorden ringere. Det hænger blandt andet sammen med, at der også måles vandstande i perioder, hvor det ikke regner. I disse perioder er korrelationen 0.



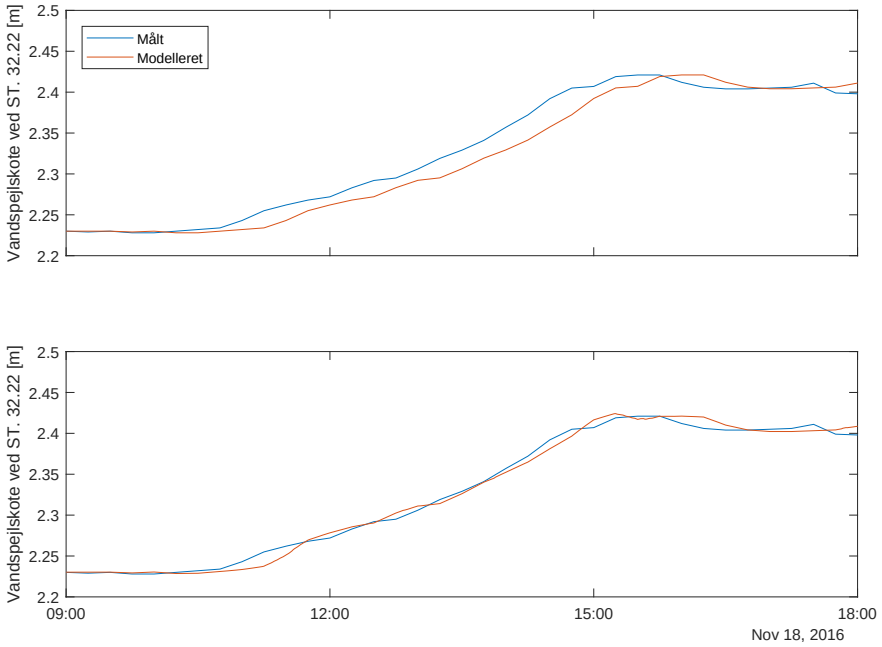
Figur 5.3: Korrelation mellem fire timers akkumuleret nedbør målt i SVK 5230 og vandstanden målt i Grejs Å ved Ny Grejsdalsvej (ST. 32.22) i perioden 10/12-2015 til 06/01-2018

Det betyder, at hvis fire timers akkumuleret nedbør skal udgøre det eneste input, så er det bedst at forsinke inputtet 240 minutter fra target. Det er dog ikke en brugbar model, fordi vandstanden i høj grad også afhænger af, hvad der er sket i perioden forud for de seneste fire timer. Det illustreres bedst ved et eksempel, hvor formålet er at forudsige vandstanden ved Ny Grejsdalsvej (ST. 32.22) en halv time frem. I det tilfælde vil det være oplagt at anvende den målte vandstand som input, fordi den netop fortæller noget om historikken i måleserien. Her vil formålet med nedbørsinputtet ikke være at beskrive niveauet af vandstanden, men derimod at justere timingen, og derfor skal forsinkelsen på nedbørsinputtet være langt kortere, som vist på figur 5.4.



Figur 5.4: Den optimale forsinkelse på nedbørsinputtet er afhængig af øvrige input og formålet med modellen. Her ses udviklingen i RMSE ved ændring i forsinkelsen på fire timers akkumuleret nedbør, hvor den målte vandstand også indgår som et input.

Det ses, at når fire timers akkumuleret nedbør fungerer som input sammen med den aktuelle vandstand, så skal det forsinkes med kortere tid for at give en bedre model. Et eksempel herpå er en hændelse med forhøjede vandstande 18/11-2016, som vist på figur 5.5.



Figur 5.5: Generelt er der tydelige forbedringer ved at bruge nedbørsinput i kombination med den målte vandstand. Her ses et eksempel på en 30 minutters prognose ved en hændelse med forhøjede vandstande 18/11-2016.

Denne hændelse illustrerer, at nedbørsinputtet på fire timers akkumuleret nedbør var tilstrækkeligt til at forbedre netværket ved netop denne hændelse. Det er ikke det generelle billede, men det er et eksempel på, hvordan der fejlfindes i modellerne, og at fremgangsmåden er at søge efter bedre input og en bedre kombination af disse. Da der nu er udviklet en metode til at finde de bedste input, er næste skridt at optimere netværkskonfigurationerne.

5.2 Optimering af netværkskonfigurationer til modellering af vandstande

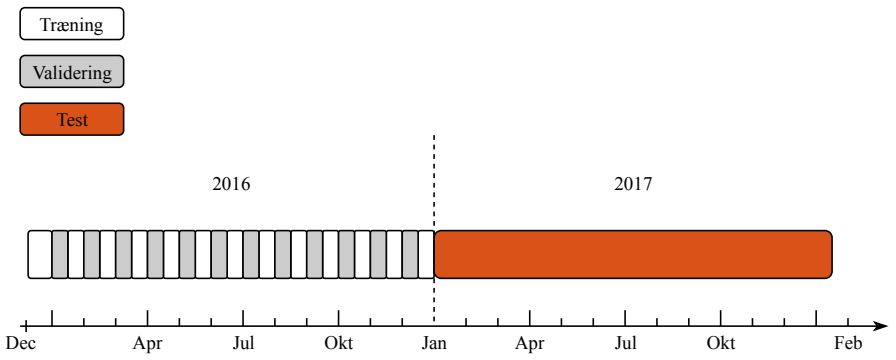
5.2.1 Inddeling af target i trænings-, valideringsdata og testperioder

I eksemplet i kapitel 4 inddeles target tilfældigt i trænings-, valideringsdata og testdata. I det følgende vises et eksempel på, hvorfor det ikke er en god idé, når det ønskes at modellere vandstande, og der er anvendt et netværk, som er blevet optimeret til at forudsige vandstanden ved Grejsdalens Planteskole (ST. 32.06) om 3 timer på baggrund af en sammensætning af vandstands- og nedbørsinput. Det er gjort, fordi betydningen af inddelingen af target er mere tydeligt på en god model. Det forklares først i afsnit 7.2, hvordan netværket er blevet optimeret, da det ikke er interessant i forbindelse med denne analyse. Netværkskonfigurationerne ses i tabel 5.3.

Tabel 5.3: Netværkskonfiguration til 3 timers forudsigelse af vandstanden ved Grejsdalens Planteskole (ST. 32.06) på baggrund af målt vandstand og optimeret nedbørsinput.

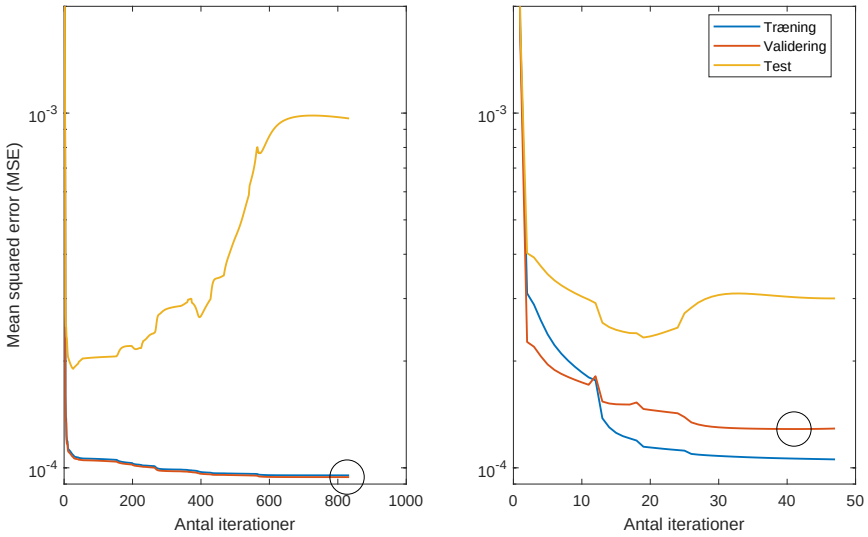
Parameter	Konfiguration
Input	Grejsdalens Planteskole (ST. 32.06) Akkumuleret nedbør, TDL (SVK 5230): 1 min; 2 min ; 5 min; 10 min; 30 min; 1 time; 2 timer; 5 timer; 12 timer; 1 dag
Target	Grejsdalens Planteskole (ST. 32.06)
Forudsigelsestid:	3 timer
Antal neuroner i det skjulte lag	10
Træning/validering	10-Dec-2015 → 31-Dec-2016: 50% træning og 50% validering
Test	01-Jan-2017 → 06-Jan-2018: 100% test

Da ønsket er et netværk, der skal kunne operere på uafhængige data, vælges perioden 01/01-2017 til 06/01-2018 som testværdier. Hvis modellen performer godt i et helt sammenhængende år, og ikke bare på tilfældigt udvalgte værdier mellem træning- og valideringspunkterne, giver det et bedre billede af netværkets evne til at operere på uafhængige data. Perioden 10/12-2015 til 31/12-2016 er afsat til træning/validering.



Figur 5.6: Illustration af inddeling af træning, validering og test i perioder.

Figur 5.7 viser resultaterne af to netværk, hvor det overordnede datagrundlag har været ens, mens trænings- og valideringsdata er udvalgt forskelligt. Til venstre er de inddelt tilfældigt med 50% til træning og 50% til validering, mens de til højre er inddelt efter samme procentsats, men i sammenhængende perioder. Således vil valideringen i januar 2016 udgøre perioden 01-Jan til 14-Jan, mens træningen udgør perioden 15-Jan til 31-Jan og så fremdeles, som illustreret på figur 5.6.



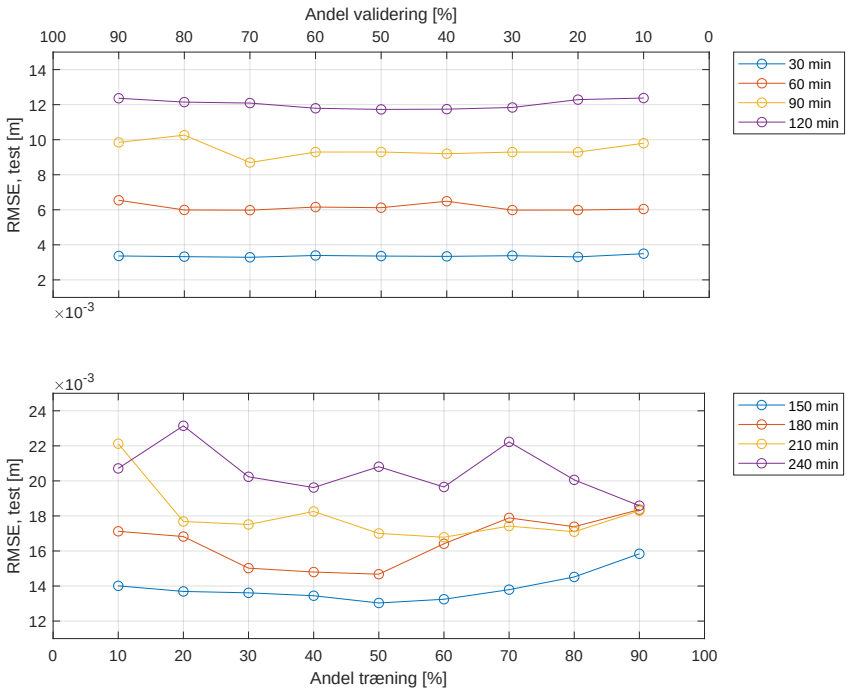
Figur 5.7: Når trænings- og valideringsdata udvælges tilfældigt, er der generelt en stor risiko for overfitting. Tilfældig inddeling (til venstre). Inddelt i perioder (til højre).

Det ses tydeligt, at der i dette tilfælde er en voldsom overfitting på netværket til venstre. Det er forårsaget af den tilfældige inddeling mellem træning og validering, fordi disse datasæt vil ligne hinanden til forveksling. Derfor fejler træningen aldrig på valideringstjekket, hvilket kommer til udtryk ved antallet af iterationer, og at træningen/validering stort set ligger oven i hinanden. Den risiko er ikke til stede i samme grad, når target inddeles i perioder. Her bliver træningen/valideringen ikke så god, men til gengæld ligger testen i nogenlunde samme størrelsesorden. Næste spørgsmål er så, om en ligelig inddeling af træning og validering giver det bedste netværk, eller om der er en fordel i at justere i procentsatserne.

5.2.2 Trænings-/valideringsandel

Det er fastlagt, at det har en fordel at inddele target i perioder fremfor tilfældigt, når det ønskes at forudsige vandstanden ved Grejsdalens Planteskolen (ST. 32.06) om 3 timer. I det tilfælde hvor træning/validering deles 50 %/50 %, giver det et generaliseret netværk, der kan beskrive mønstre i alle tre dele af target med en præcision indenfor samme størrelsesorden. I det følgende undersøges det, hvorvidt det også er tendensen, når forudsigelsestiden øges, og om der er en fordel ved at ændre inddelingen. Basis for undersøgelsen er samme netværk, som beskrevet i tabel 5.3.

Figur 5.8 viser, at forholdet mellem de to perioder har større betydning, som forudsigelseshorisonten øges.



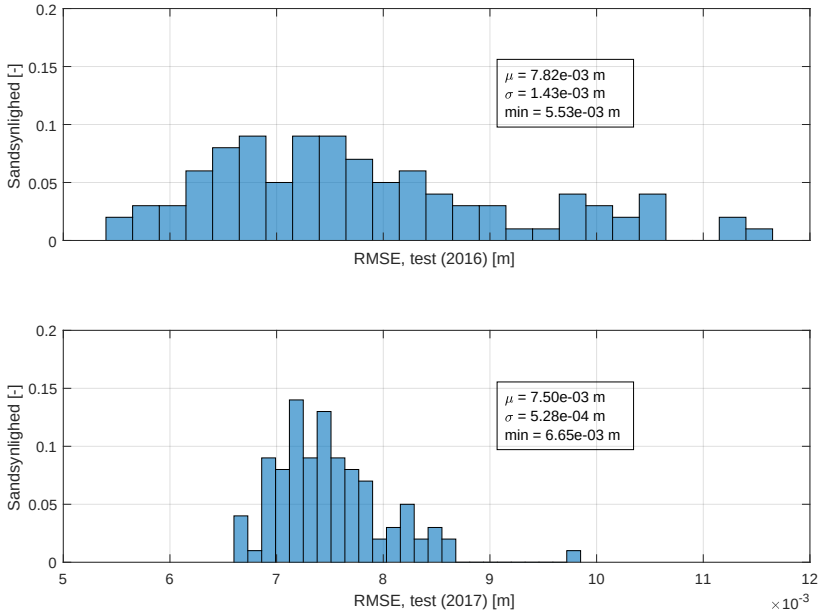
Figur 5.8: Den procentvise fordeling på trænings- og valideringsperioderne har en betydning for kvaliteten i testperioden.

Det er konklusionen, at de bedste netværk forekommer ved en fordeling på 50% træning og 50% validering ved samtlige forudsigelsestider. Derfor vælges det at anvende den inddeling fremover, når det ønskes at finde de bedste netværk.

5.2.3 Ombytning af perioder

Figur 5.8 viser RMSE-værdierne i testen, når den strækker sig i perioden 01/01-2017 til 06/01-2018, mens træningen/validering strækker sig i perioden 10/12-2015 til 31/12-2017. Da der i hele optimeringsprocessen er fokus på at vælge input, der giver gode netværk i den periode, er der risiko for, at de valgte input er specifikke for den periode, og at der på den måde er sket en manuel overfitting. Derfor undersøges det, hvorvidt det samme netværk også giver gode RMSE-værdier i testen, når de to perioder ombyttes. På den måde fås et netværk, der formodes også at kunne præstere godt, hvis det skulle testes fremadrette i eksempelvis den resterende del af 2018 eller 2019. Basis for analysen er igen netværket i tabel 5.3, men denne gang ved en forudsigelsestid på 60 minutter. Figur 5.9 viser resultatet af 1000 netværk, og det er primært gjort for at undgå at præsentere et tilfældigt resultat.

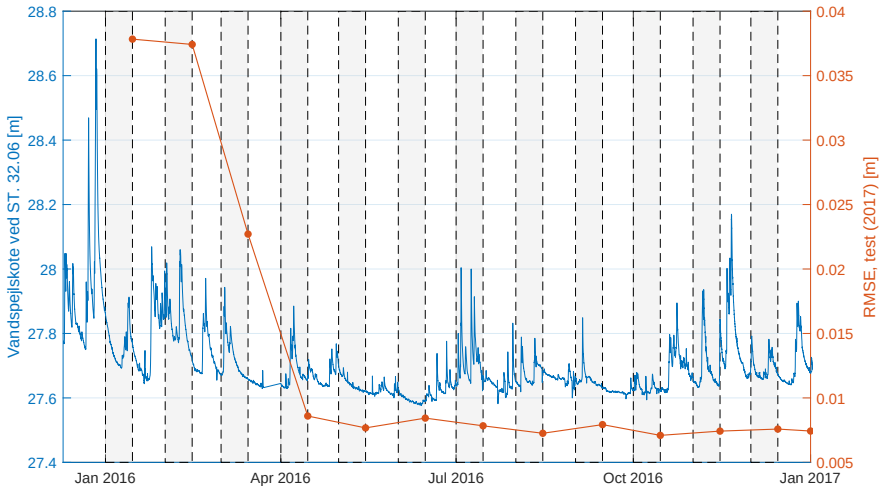
Det ses, at spredningen på RMSE i de to testperioder er noget forskellig, mens middelværdierne er nogenlunde tilsvarende. En af grundende til det kan være, at de højeste vandstande måles ved en oversvømmelseshændelse i december 2015, som altså ikke indgår i træningen på den øverste figur. Fordi netværket ikke har været præsenteret for de høje vandstande før, er det derfor til en vis grad tilfældigt, hvorvidt netværket kan gengive disse vandstande. Det bidrager også til ideen om, at kvaliteten af denne type modeller kun bliver bedre, efterhånden som datagrundlaget øges. Omvendt illustrerer det også begrænsningerne, fordi kvaliteten af forudsigelser af oversvømmelser afhænger af, om det er gået galt tidligere.



Figur 5.9: Øverst haves fordelingen af RMSE i testperioden 01/01-2017 til 06/01-2018, og nederst haves den tilsvarende RMSE i perioden 10/12-2015 til 31/12-2017.

5.2.4 Betydning af datagrundlaget

Tidligere analyser har vist, at inddelingen af trænings- og valideringsdata har en betydning for, hvor godt netværket performer i testperioden. Det er undersøgt under en forudsætning af, at der forefindes data for samtlige to år. I denne analyse undersøges det, hvor godt netværket præsterer i perioden 01/01-2017 til 06/01-2018, efterhånden som trænings- og valideringsdata øges. Figur 5.10 viser resultaten af den analyse, hvor datamængden øges med én måned af gangen. Netværkskonfigurationerne er de samme som i tabel 5.3, men ved en forudsigelsestid på 60 minutter.



Figur 5.10: RMSE i testperioden (01/01-2017 til 06/01-2018) er afhængig af, hvilke vandstande der trænes og valideres på. De hvide og grå markeringer symboliserer henholdsvis træningsperioder og valideringsperioder.

Det ses, at kvaliteten af netværket forbedres betragteligt over de første fire måneder, efterhånden som det præsenteres for nye mønstre og kombinationer mellem input og target. Fra april 2016 til januar 2017 er der også en generel tendens til forbedring, men i form af finjusteringer.

5.3 Sammenfatning på optimering af netværkskonfigurationer

Analyserne i dette kapitel har vist, at korrelationsanalyser er et godt værktøj til at udvælge smarte input til de neurale netværk - i hvert fald når opstrøms vandstand ved Grejsdalens Planteskole (ST. 32.06) skal bruges som input til en model, der skal beskrive vandstanden ved Ny Grejsdalsvej (ST. 32.22). Når nedbør derimod skal anvendes til vandstandsmodellering, har det vist sig at være vanskeligt at vælge en bestemt måde at fortolke nedbøren på. Det giver god mening, fordi det i kapitel 3 også er fundet vanskeligt at finde en simpel relation mellem nedbør og vandstand. Derfor skal nedbør fortolkes forskelligt fra netværk til netværk alt efter, hvilke øvrige input der indgår i netværket.

De bedste netværkskonfigurationer har vist sig at forekomme, når target ikke inddeles tilfældigt, men efter perioder, fordi overfitting for det meste undgås. Det smarteste er at inddele træning og validering i perioder på månedsbasis med 50 % til hver del, som det illustreres på figur 5.6. Ved den inddeling har det vist sig muligt at ombytte test-perioderne, uden det gav betragtelige ændringer modelkvaliteten.

Det er fundet, at det i dette enkeltstående tilfælde er tilstrækkeligt med fire månedes data til at træne et brugbart netværk, som vist på figur 5.10. Det vurderes ikke at være en grundlæggende regel, og det afhænger uforbeholdent af målingerne i perioden.

På den baggrund vurderes det, at forudsætningerne for at modellere vandstande med neurale netværk er på plads, og derfor er næste skridt at dykke ned i de egentlige prognoser på korte og lange forudsigelsestider.

Del III - Vandstandsprognoser i Vejle Midtby

6 Korte forudsigelser 67

6.1 Formålet med korte forudsigelser

6.2 Korte forudsigelser af vandstanden i Grejs Å ved Grejsdalens Planteskole

6.3 Korte forudsigelser af vandstanden i Grejs Å ved Ny Grejsdalsvej

6.4 Korte forudsigelser af vandstanden i Omløbsåen ved Aagade

6.5 Anbefaling til styring

7 Lange forudsigelser 87

7.1 Formål med lange forudsigelser

7.2 Nedbør som input til neurale netværk

7.3 Forkastning af Manningtal og potentiel fordampning som input til neurale netværk

7.4 Anvendelighed af radardata som input til neurale netværk

7.5 Modelkvalitet ved lange forudsigelser

6. Korte forudsigelser

Erfaringer har vist, at betydningen af nedbør som input til det neurale netværk øges med forudsigelseshorisonten. Ved kortere forudsigelseshorisonter er der derfor ikke nogen betydelig forskel på at benytte en model, der kun udnytter information indlejret i vandstandsdataene, fremfor en model der også tager højde for forudgående nedbørsmængder. I dette kapitel beskrives metoden til forudsigelser med korte forudsigelseshorisonter og begrænsningerne ved denne metode.

6.1 Formålet med korte forudsigelser

Formålet med de korte forudsigelser af vandstanden i Vejle Midtby er, at de skal kunne anvendes til at styre fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen. Der er især fokus på at udvikle et prognoseværktøj, der kan mindske sluseanlæggets lukketid, så fisk frit kan vandre i Grejs Å-systemet.

Strategien er først at udvikle en metode, der kan beskrive vandstandsændringer ved de naturlige afstrømningsforhold i Grejs Å ved Grejsdalens Planteskole, hvorefter det undersøges, om disse vandstandsændringer kan anvendes til at prognosticere vandstanden nedstrøms.

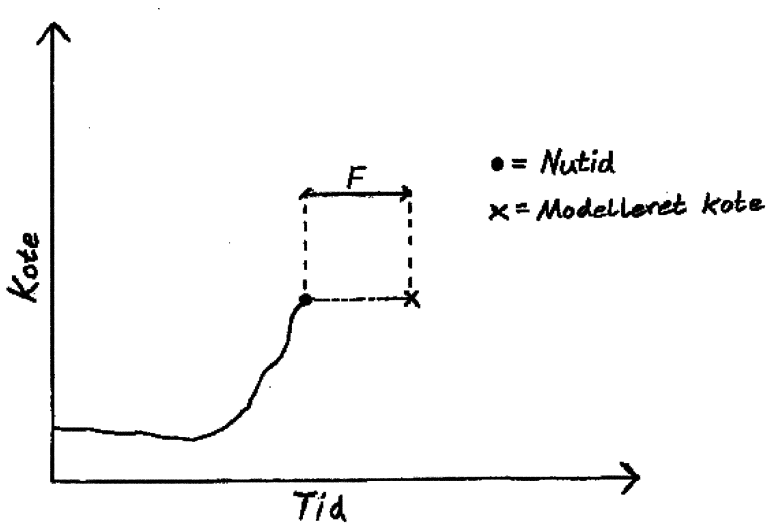
Således bevæges der sig ned i systemet, og det bedste bud på modeller præsenteres for korte forudsigelser i tre stationer:

- Grejsdalens Planteskole, Grejs Å (ST. 32.06).
- Ny Grejsdalsvej, Grejs Å (ST. 32.22).
- Aagade, Omløbsåen (ST. 32.21).

De fulde serier over modellerede vandspejlskoter for træning og valideringsperioden samt testperioden ved de tre stationer er vist ved forudsigelsestiderne 30, 60, 120 og 240 minutter i appendiks G.

6.2 Korte forudsigelser af vandstanden i Grejs Å ved Grejsdalens Planteskole

Hvis der skal laves en simpel forudsigelse af vandstanden i et vandløb indenfor en kort tidshorisont, er det oplagt at anvende den aktuelle vandstand. En sådan måling vil med stor nøjagtighed kunne sige noget om vandstanden kort tid efter målingen er foretaget, mens nøjagtigheden vil forringes jo længere forudsigelsestiden bliver. Det er valgt, at en forudsigelse udelukkende på baggrund af den aktuelle vandstand kaldes en *dummy*-model, og den vil i resten af rapporten anvendes som reference til andre modeller. Dummy-modellen er illustreret på figur 6.1.

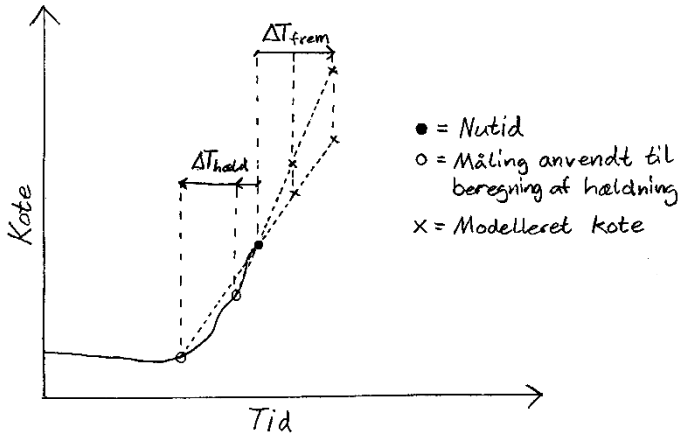


Figur 6.1: Illustration af dummy-modellen, som er en forudsigelse på baggrund af vandstanden nu, hvor F betegner forudsigelsestiden. Dummy-modellen fungerer som referencemodel gennem den resterende del af rapporten.

Prognosen vil i nogle tilfælde kunne forbedres, hvis der samtidig tages højde for hældningen på kurven, altså hvorvidt vandstanden har været stigende eller faldende op til det aktuelle tidspunkt.

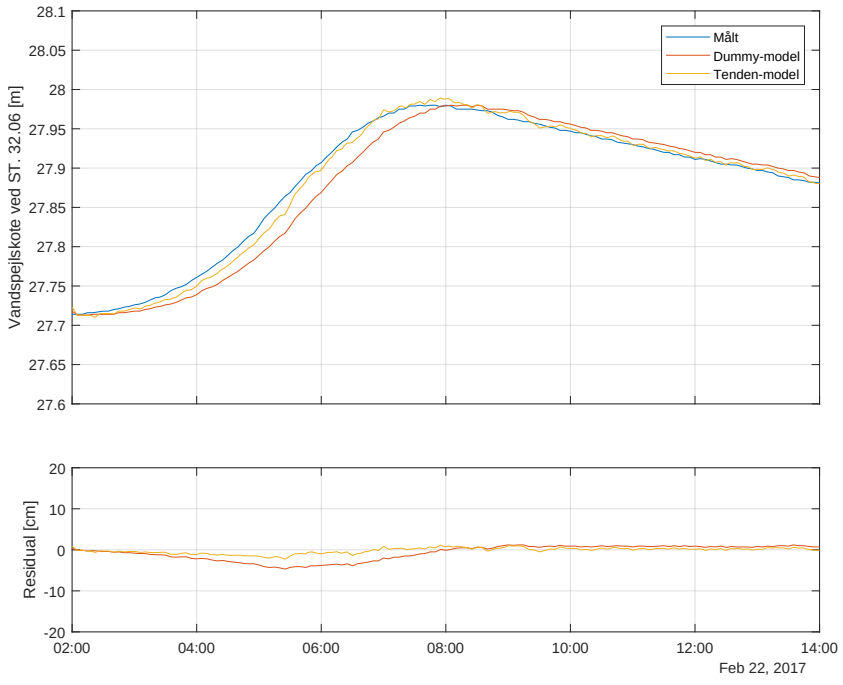
Figur 6.2 viser, hvordan der kan laves forudsigelser på baggrund af hældningen på kurven. Det er valgt, at en forudsigelse på baggrund af hældningen på kurven kaldes en *tendens*-model. For at optimere tendens-modellen skal det vælges hvor lang tid ($\Delta T_{hæld}$), der skal bevæges tilbage på kurven for at beregne hældningen, og hvor lang tid (ΔT_{frem}) det forventes, at denne hældning fortsætter.

Ved at bestemme RMSE-værdien ved anvendelse af forskellige værdier for $\Delta T_{hæld}$ og ΔT_{frem} er værdierne, der giver den mindste RMSE-værdi, fundet. Den optimale $\Delta T_{hæld}$ er fundet til 22 min. Den optimale ΔT_{frem} er bestemt til at være en funktion af forudsigelsestiden, hvor den ved meget lave forudsigelsestider ligger tæt på forudsigelsestiden, mens forskellen mellem forudsigelsestiden og ΔT_{frem} øges med forudsigelsestiden. Analysen af de optimale værdier ses i appendiks F.

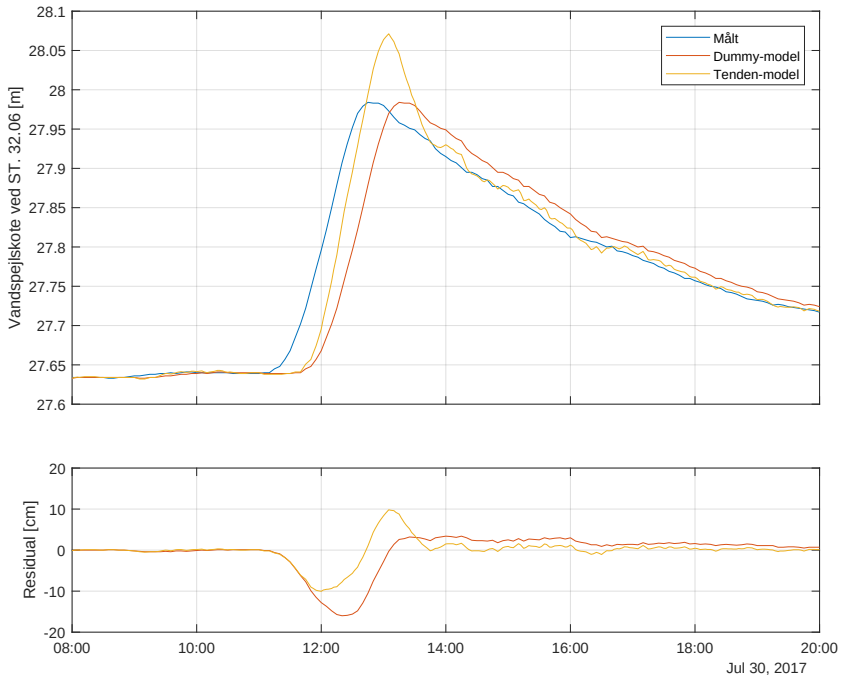


Figur 6.2: Illustration af tendens-modellen, som er en forudsigelse på baggrund af hældningen op til målingen.

På figur 6.3 og 6.4 ses den målte og modellerede vandspejlskote ved Grejsdalens Planteskole (ST. 32.06). Figurerne viser to hændelser, hvor vandspejlskoten ved den første hændelse ændres langsomt, mens vandspejlskoten ved den anden hændelse ændres hurtigt. De modellerede vandspejlskoter er med en forudsigelsestid på 30 minutter. Af de to figurer fremgår det, at forudsigelsen er bedre, når vandspejlskoten ændrer sig langsomt, mens tendens-modellen som regel overestimerer ved hurtige vandstandsændringer.

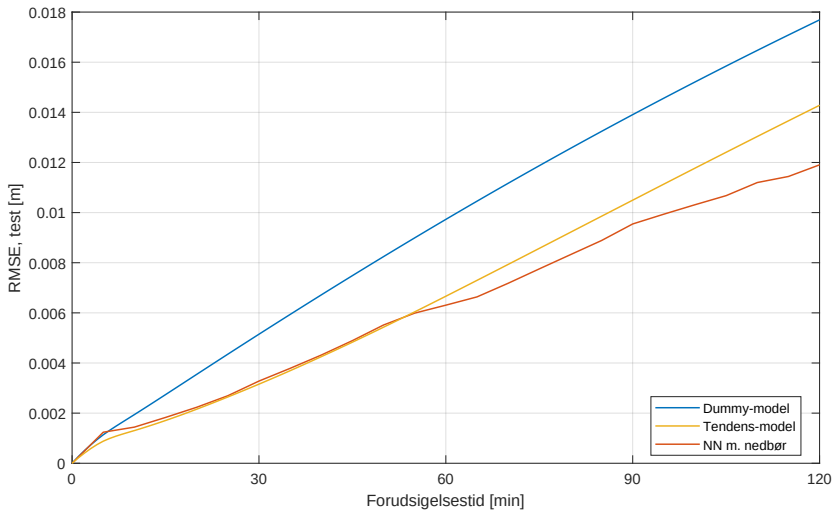


Figur 6.3: Målt og modelleret vandstand ved Grejsdalens Planteskole (ST. 32.06) den 22/02-2017.



Figur 6.4: Målt og modelleret vandstand ved Grejsdalens Planteskole (ST. 32.06) den 30/06-2017.

En hurtig ændring af hældningen på kurven for vandspejlskoten må i de fleste tilfælde formodes at skyldes kortvarig intensiv nedbør. Derfor ville det være naturligt, hvis der eksempelvis i tilfældet vist på figur 6.4 kunne opnås en gevinst ved at anvende neurale netværk med en kombination af tendens-modellen og nedbør som input (nedbør som input til neurale netværk er beskrevet i afsnit 7.2). Dog viser figur 6.5, at der først ses en gevinst ved at anvende neurale netværk frem for tendens-modellen, når der forudsiges mere end én time frem.



Figur 6.5: Forudsigelsestiden skal ud over en time, før anvendelse af neurale netværk for alvor har en gevinst i forhold til tendens-modellen.

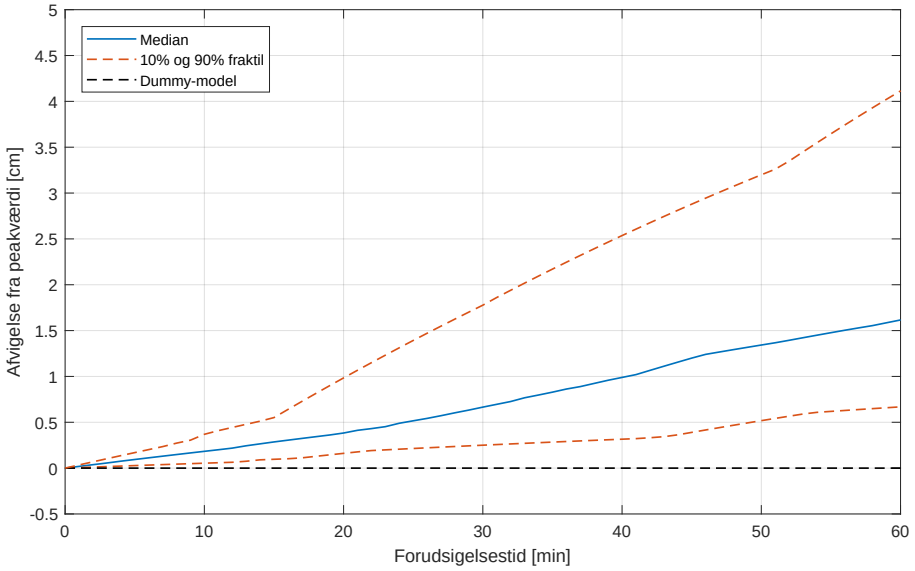
Da det tydeligvis er tendensen i vandstandsændringen, der er mest sigende for vandstanden i den nærmeste fremtid, er det analyseret, om der kan anvendes et neuralt netværk til optimering af en forudsigelse udelukkende på baggrund af de sidst målte værdier for vandstanden. Analysen kan ses i appendiks H. Konklusionen på denne analyse er, at der med et neuralt netværk på baggrund af målte værdier kan opnås et tilsvarende resultat som tendens-modellen. Tendens-modellen vurderes på baggrund af dens simplicitet at være den nemmest anvendelige.

Både tendens-modellen og det neurale netværk er i tilfældet på figur 6.5 optimeret til forhold gældende ved Grejsdalens Planteskole. Ved Grejsdalens Planteskole forekommer der primært naturlig afstrømning, som i henhold til afsnit 3.1 medfører en forsinkelse og udglatning af afstrømningen. Dette er til fordel for tendens-modellen og betyder, at tendens-modellen ikke nødvendigvis er optimal til forudsigelser ved Ny Grejsdalsvej og i Omløbsåen, hvor der kan ske hurtige vandstandsændringer som følge af afstrømning fra befæstede arealer og styring.

Sammenligningsgrundlaget for modellerne er gennemgående RMSE-værdien, men det er også en parameter, som ikke er til at forholde sig til. Der er lavet en peak-analyse på tendens-modellen ved Grejsdalens Planteskole, hvor et peak defineres som følgende:

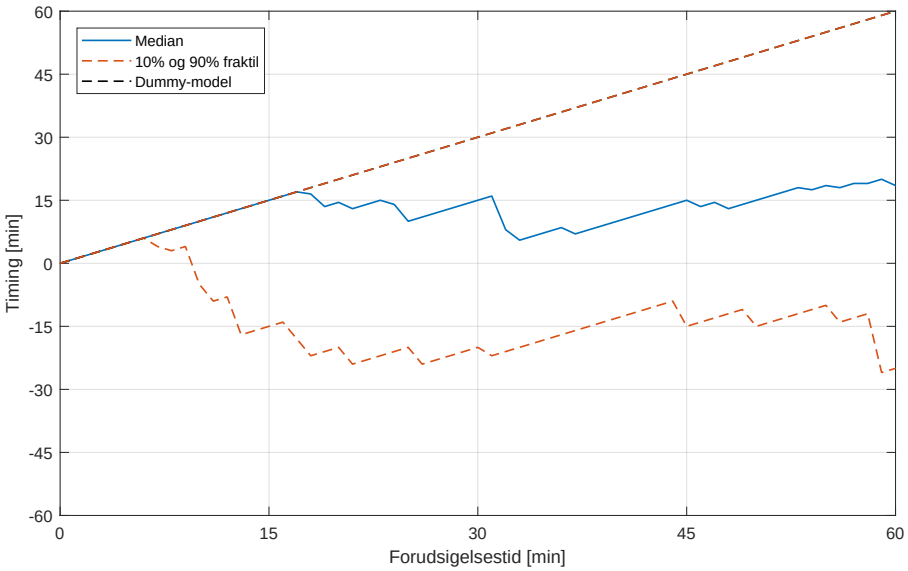
- Der er målt en vandspejlskote over 27,9 m.
- Der er minimum et døgn mellem hvert peak.

Ved den definition haves 52 peaks over hele måleperioden, hvor tendens-modellens timing og afvigelse fra den målte peak-værdi kan undersøges. Afvigelsen fra den målte peak-værdi, ved forøgelse af forudsigelsestiden, ses på figur 6.6. Til sammenligning er dummy-modellen også vist på graferne.



Figur 6.6: Tendens-modellen overestimerer altid peak-værdien, hvilket er en af ulemperne.

Timingen på en model defineres som forskellen i tid mellem det målte peak, og det tidspunkt, hvor modellen peaker. Timingen på tendens-modellen er ved alle forudsigelsestider under 60 minutter, hvilket ses på figur 6.7.



Figur 6.7: Timingen på en model er svær at definere, men tendens-modellen er næsten altid forsinket i forhold det målte peak, hvilket fremgår af en positiv median-værdi ved alle forudsigelsestider under 60 minutter.

På den baggrund vurderes det også, at RMSE er en udmærket performance-parameter, fordi afvigelser fra peak-værdier vokser, når RMSE vokser. Angående timingen, så er der nogle usikkerheder forbundet med den måde, som den defineres. Det vurderes, at selvom der fremgår nogle *"hak"* på kurven, at RMSE også til dels kan beskrive udviklingen i modellens timing.

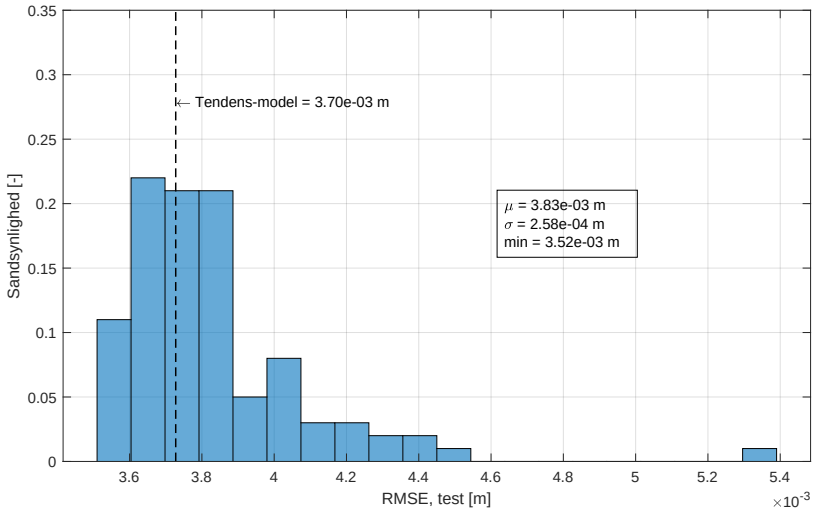
6.3 Korte forudsigelser af vandstanden i Grejs Å ved Ny Grejsdalsvej

Forundersøgelserne i afsnit 3.3 har vist, at der er god korrelation mellem vandstande i Grejs Å ved Grejsdalens Planteskole og Ny Grejsdalsvej. Derfor forventes det, at der er potentiale for at lave et godt neuralt netværk ved Ny Grejsdalsvej, da Grejsdalens Planteskole kan fungere som input.

Da den bedste korrelation mellem vandstanden ved Ny Grejsdalsvej og ved Grejsdalens Planteskole forekommer ved en forsinkelse på 26 minutter, forsøges det at lave en prognose af vandstanden ved Ny Grejsdalsvej ved en forudsigelsestid på 26 minutter. Til det formål opstilles to modeller:

- En tendens-model, der laver en direkte prognose af vandstanden ved Ny Grejsdalsvej om 26 minutter.
- Et neuralt netværk, der både får tendens-modellen og den målte vandstand ved Grejsdalens Planteskole som input.

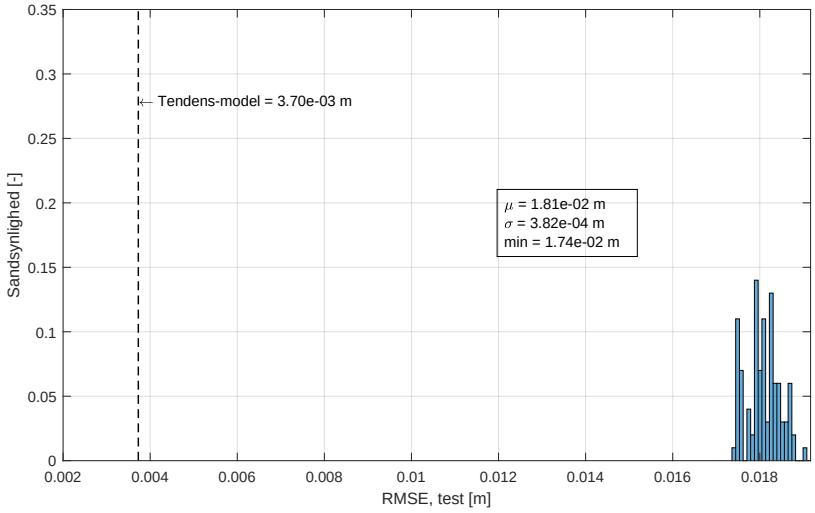
Som tidligere nævnt i kapitel 4 er der risiko for, at en direkte sammenligning kan forstyrres af forskellige minima ved træningen af det neurale netværk. Derfor trænes netværket 100 gange for at undgå et tilfældigt resultat. Resultatet af analysen fremgår af figur 6.8. Det ses, at kvaliteten af tendens-modellen er bedre end middelnetværket, og at det er tilfældigheder der afgør, om et neuralt netværk er bedre.



Figur 6.8: Fordelingen af RMSE-værdier i testperioden ved 100 godkendte netværk. Det ses, at det er tilfældigt, om det neurale netværk har en højere kvalitet end tendens-modellen.

Konklusionen ud fra sammenligning af RMSE-værdierne er, at der ikke er nogen synlig gevinst ved at anvende et neuralt netværk, hvor den aktuelle vandstand målt ved Grejsdalens Planteskole indgår, fremfor tendens-modellen. Det betyder dog ikke, at modellering af vandstanden ved Grejsdalens Planteskole er helt uden betydning, da formålet ikke er at finde den model med den laveste RMSE ved en given forudsigelsestid, men derimod at finde den model, der praktisk set er mest anvendelig.

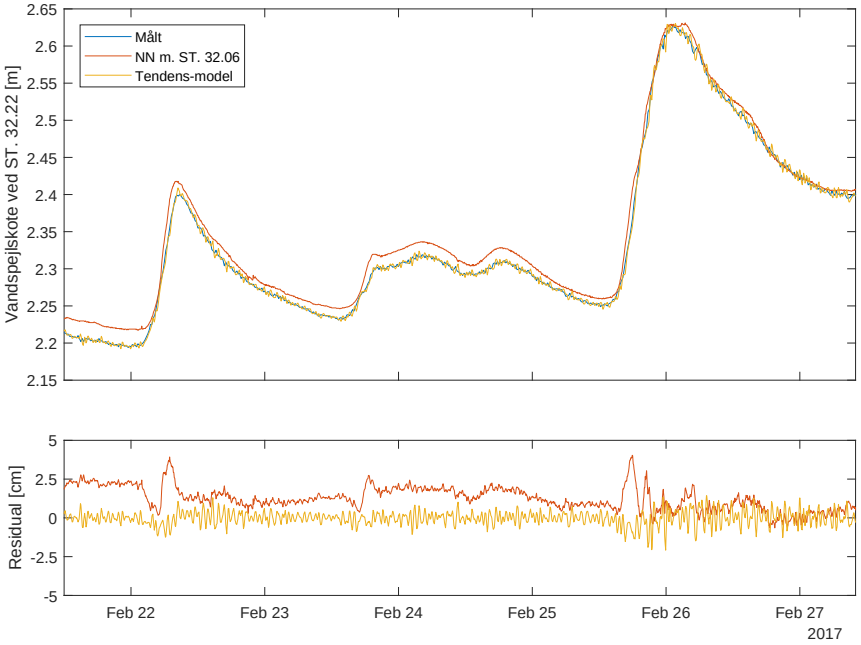
På figur 6.9 vises RMSE-værdier af 100 godkendte netværk, hvor inputtet fra tendens-modellen blev fjernet.



Figur 6.9: Fordelingen af RMSE-værdier i testperioden ved 100 godkendte netværk, hvor inputtet fra tendens-modellen er fjernet. Alle 100 netværk har en meget dårligere RMSE-værdi end tendens-modellen.

Hvis fordelingerne af RMSE på figur 6.9 og 6.8 sammenlignes, ser netværket med både tendens-modellen og Grejsdalens Planteskole som input ud til at være markant bedre. Middelværdien på

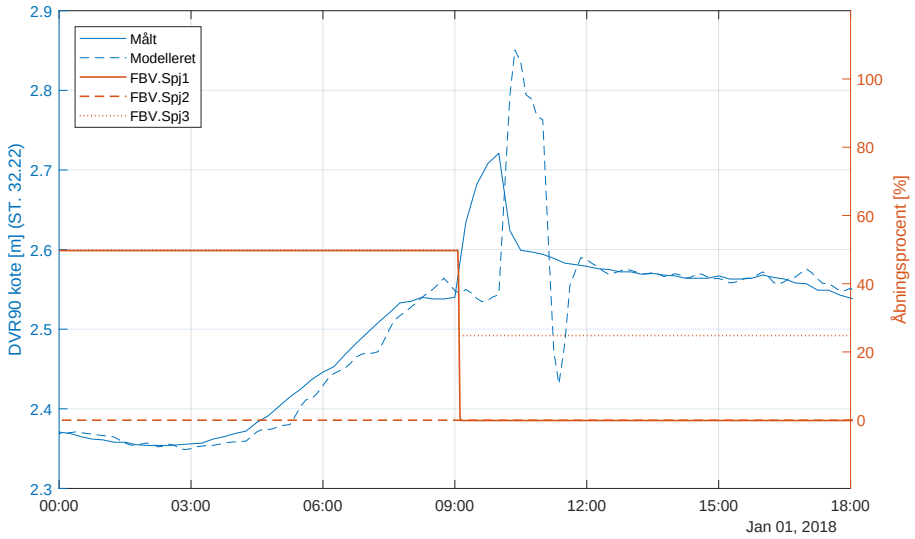
RMSE er en størrelsesorden lavere, men da formålet med modellerne er at forudsige forhøjede vandstande, er det helt afgørende, at modellerne kan forudsige peak-værdier korrekt. På figur 6.10 vises et eksempel fra en periode, hvor der er målt nogle af de højeste vandstande i Grejs Å ved Ny Grejsdalsvej.



Figur 6.10: Det er farligt at stirre sig blindt på RMSE. I denne periode måles nogle af de højeste vandstande i Grejs Å ved Ny Grejsdalsvej, og det neurale netværk med Grejsdalens Planteskole som input kan modellere peaks ganske fint.

Eksemplet viser, at selvom tendens-modellen slår et neuralt netværk, med Grejsdalens Planteskole som input, på RMSE, så er forskellen ikke altid så markant, som RMSE-værdien kan antyde. De største forskelle mellem modellen med det neurale netværk og den målte værdi forekommer ved de lavere vandspejlskoter, hvor modelkvaliteten egentlig er helt uden interesse, fordi formålet er at forudsige peak-vandstande. I perioder, hvor vandstanden ved Ny Grejsdalsvej påvirkes af eksempelvis styring på fordelerbygværket, hvor korrelationen mellem vandstanden ved Ny Grejsdalsvej og Grejsdalens Planteskole ikke er god, har det neurale netværk, med Grejsdalens Planteskole som input, ikke forudsætningerne til at kunne forudsige vandstandene godt - ligesom tendens-modellen heller ikke har det. Den overordnede vurdering er dog stadig, at tendens-modellen performer bedre set over hele modelleringsperioden.

Dykkes der ned i de modellerede vandspejlskoter med tendens-modellen i testperioden ses, at den største forskel mellem modelleret og målt vandspejlskote haves 01/01-2018. Dette er med en forudsigelsestid på 60 minutter. Figur 6.11 viser, at forskellen mellem modelleret og målt vandspejlskote var forårsaget af opstuvning til målestationen ved Ny Grejsdalsvej som følge af lukning af spjældene ved fordelerbygværket klokken 09:00. I dette tilfælde går der tre timer, fra spjældene er blevet lukket, til der igen modelleres korrekt.



Figur 6.11: En lukning af spjældene ved fordelerbygværket forårsagede pludselig vandstandsstigning ved Ny Grejsdalsvej.

Modelkvaliteten af forudsigelserne ved Ny Grejsdalsvej vil eventuelt kunne forbedres ved at inddrage data for åbningsprocenter på spjældene ved fordelerbygværket ved Abelones Plads. I forbindelse med forudsigelser ved Ny Grejsdalsvej er det dog valgt ikke at gå yderligere i dybden med åbningsprocenter på spjældene. Dette skyldes delvist, at kvaliteten af modellen er god, indtil det besluttes at ændre åbningsprocenten af spjældene. Derudover er datagrundlaget fra fordelerbygværket i henhold til afsnit 2.2 begrænset til en relativ kort periode i forhold til vandstandsserierne, hvilket skaber udfordringer i forbindelse med anvendelse af dataene i neurale netværk.

I Omløbsåen har åbningsprocenten på spjældene en væsentlig større betydning end ved Ny Grejsdalsvej, og det er derfor fundet naturligt, at fokus på anvendelsen af åbningsprocenter på spjældene som input til neurale netværk sker ved modelleringen af vandspejlskoten i Omløbsåen.

6.4 Korte forudsigelser af vandstanden i Omløbsåen ved Aagade

ST. 32.21 er placeret i Omløbsåen ved Aagade centralt i risikoområdet, og derfor har prognoser af vandstanden ved denne station en stor betydning. Hvor lang tid det reelt set er nødvendigt at kunne forudsige vandspejlskoten i Omløbsåen for at kunne sætte sluseanlægget i kraft og ændre på spjæld ved fordelerbygværket tids nok til at forhindre en kritisk vandspejlskote i Omløbsåen er afhængigt af, hvor lang tid der går, fra der bliver taget en beslutning om at styre til at vandstanden i Omløbsåen falder. Ved visualisering af data er det fundet, at vandstanden ved Aagade falder indenfor 10 minutter fra opstart af pumper ved sluseanlægget eller ændring af spjæld.

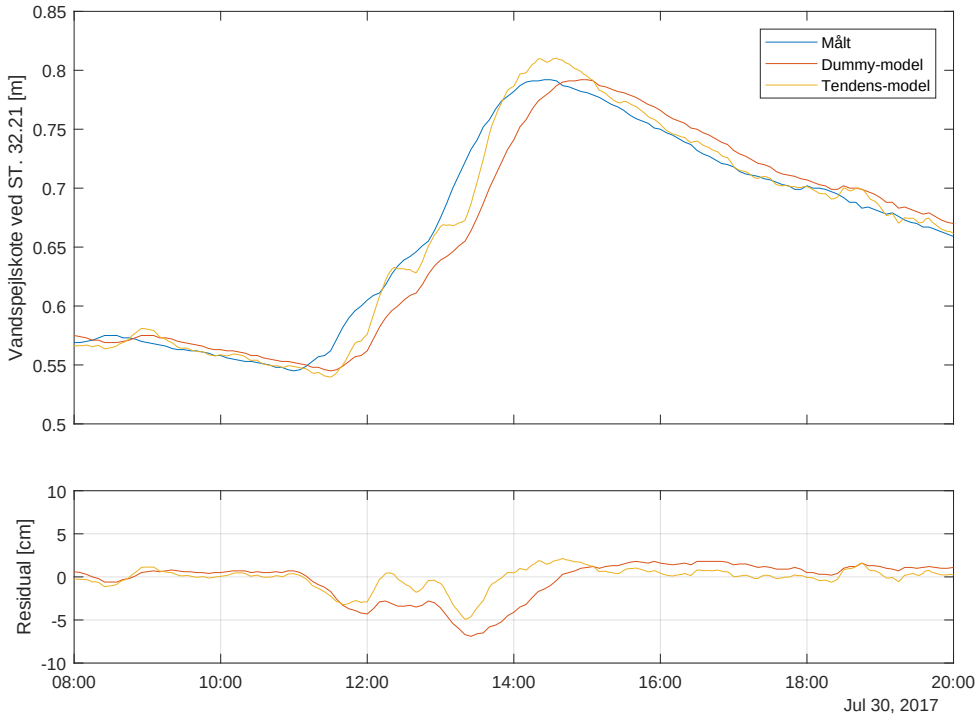
I afsnit 3.2 er det beskrevet, hvordan vandspejlskoten i Omløbsåen ved Aagade afhænger af de naturlige parametre som afstrømningen fra oplandet og vandstanden nedstrøms. Den er derudover styret af fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen. Resultaterne på analyserne i opstrøms stationer ved Grejsdalens Planteskole og Ny Grejsdalsvej (afsnit 6.2 og 6.3) viser dog, at usikkerheden på andre input end vandspejlskoten ved stationen selv er stor nok til, at disse input ikke giver en forbedring i modelresultatet før forudsigelsestiden overstiger 60 minutter. Det er derfor oplagt at forsøge at forudsige vandspejlskoten ved Aagade på baggrund af den tidligere beskrevne tendens-model.

For at tydeliggøre de neurale netværks anvendelighed til forudsigelse af vandspejlskoten i Omløbsåen, beskrives der i de efterfølgende delafsnit 6.4.1 til 6.4.3 begrænsningerne ved anvendelse af tendens-modellen til forudsigelse af vandspejlskoten i Omløbsåen.

6.4.1 Anvendelse af tendens-model til forudsigelse af vandspejlskote under høj afstrømning fra Grejs Å

Det er tidligere illustreret på figur 6.4, at der kan være udfordringer ved at anvende tendens-modellen til at forudsige vandspejlskoten, hvis der sker hurtige ændringer på hældningen af kurven for vandspejlskoten. Afstrømningshændelsen 30/07-2017 er et eksempel på en hændelse, hvor der grundet nedbør og naturlig afstrømning er sket en hurtig ændring på hældningen af kurven for vandspejlskoten ved Grejsdalens Planteskole. For at vurdere tendens-modellens evne til at forudsige vandspejlskoten i Omløbsåen ved Aagade, når der sker en hurtig ændring grundet afstrømning fra oplandet, er det forsøgt at anvende tendens-modellen på hændelsen den 30/07-2017 ved Aagade. Resultatet af en prognose med 30 minutters forudsigelsestid ses på figur 6.12.

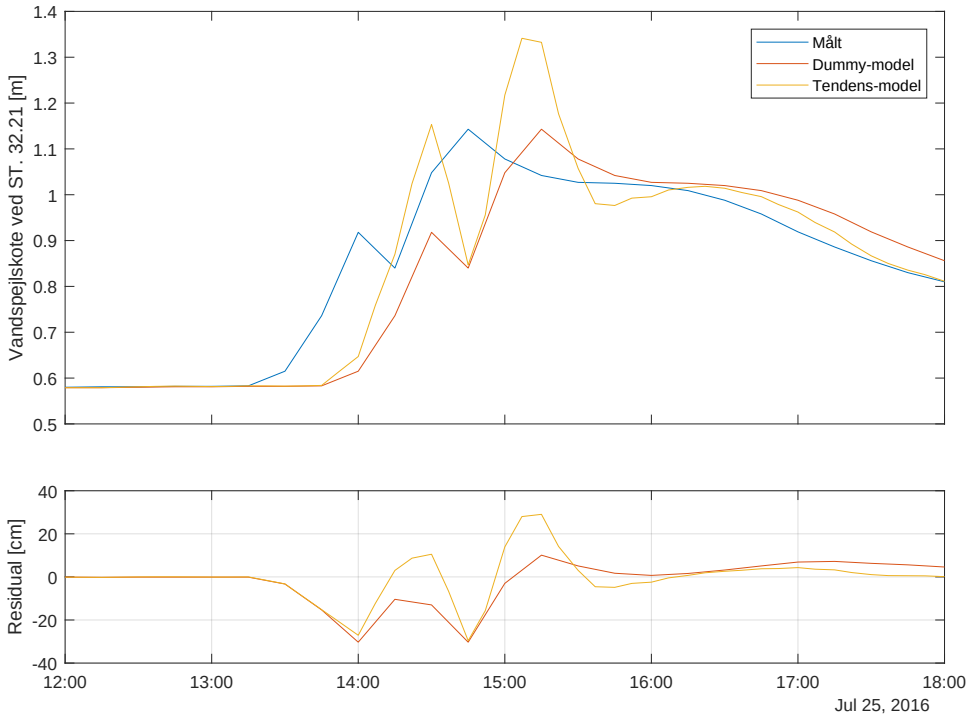
Figur 6.12 viser, at når tendens-modellen anvendes, vil den maksimale afvigelse mellem målt og modelleret vandspejlskote ved Aagade 30/07-2017 kun være 5 centimeter. Det er vurderet som en generel tendens, at tendens-modellen også ved Aagade forudsiger vandspejlskoten uden store afvigelser ved korte forudsigelseshorisonter, hvis ændringen på vandspejlskoten skyldes afstrømningen fra oplandet.



Figur 6.12: Målt og modelleret vandspejlskote i Omløbsåen ved Aagade (ST. 32.21) under en hændelse, hvor ændringen af vandspejlskoten skyldes afstrømning fra Grejs Å.

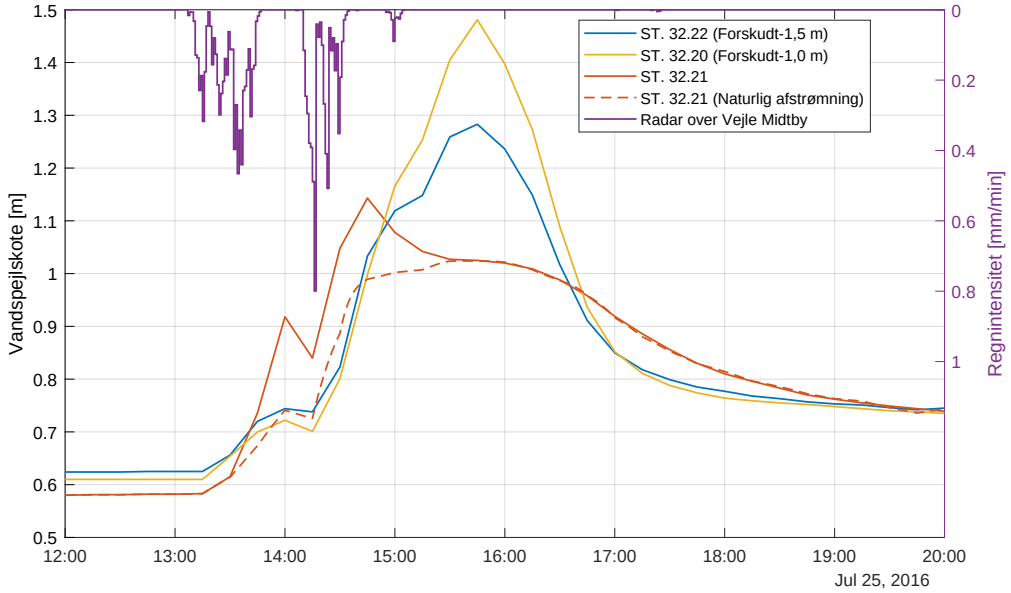
6.4.2 Afstrømning fra Vejle Midtby

Hvis der forekommer intensiv nedbør over Vejle Midtby, kan det resultere i, at vandspejlskoten ændrer sig hurtigere end det normalt er tilfældet ved afstrømning fra oplandet til Grejs Å. Hændelsen 25/07-2016, som beskrives i afsnit 3.3, er en hændelse med stor stedlig fordeling af nedbør (illustreret på figur 3.14). Radarbilleder viser, at nedbøren stort set kun falder i området omkring den nedre del af Grejs Å og over Vejle Midtby. Det ses på figur 6.13, at afvigelsen mellem målt og modelleret vandspejlskote klokken 14:00 er 27 centimeter grundet en utraditionel hurtig stigning i vandspejlskoten mellem klokken 13:30 og 14:00.



Figur 6.13: Målt og modelleret vandspejlskote i Omløbsåen ved Aagade (ST. 32.21) under en hændelse, hvor ændringen af vandspejlskoten skyldes stor afstrømning fra Vejle Midtby.

Radarbilleder viser, at nedbørsintensiteten er høj over Vejle Midtby umiddelbart før de hurtige vandstandsstigninger i Omløbsåen, hvilket indikerer, at den utraditionelle hurtige stigning i vandspejlskoten ved Aagade skyldes hurtig afstrømning/overløb direkte til Omløbsåen fra Vejle Midtby. Det verificeres også af, at vandspejlskoten ved Ny Grejsdalsvej (ST.32.22) og i Mølleåen (ST. 32.20) ikke stiger i tilsvarende tempo i samme periode. Dette er illustreret på figur 6.14. Det er også vist, hvad vandspejlskoten ved Omløbsåen burde have været, hvis der ikke var afstrømning fra Midtbyen. Dette er fundet på baggrund af vandspejlskoten ved Ny Grejsdalsvej og korrelationen mellem vandspejlskoten ved Ny Grejsdalsvej og Omløbsåen efter klokken 15:30.



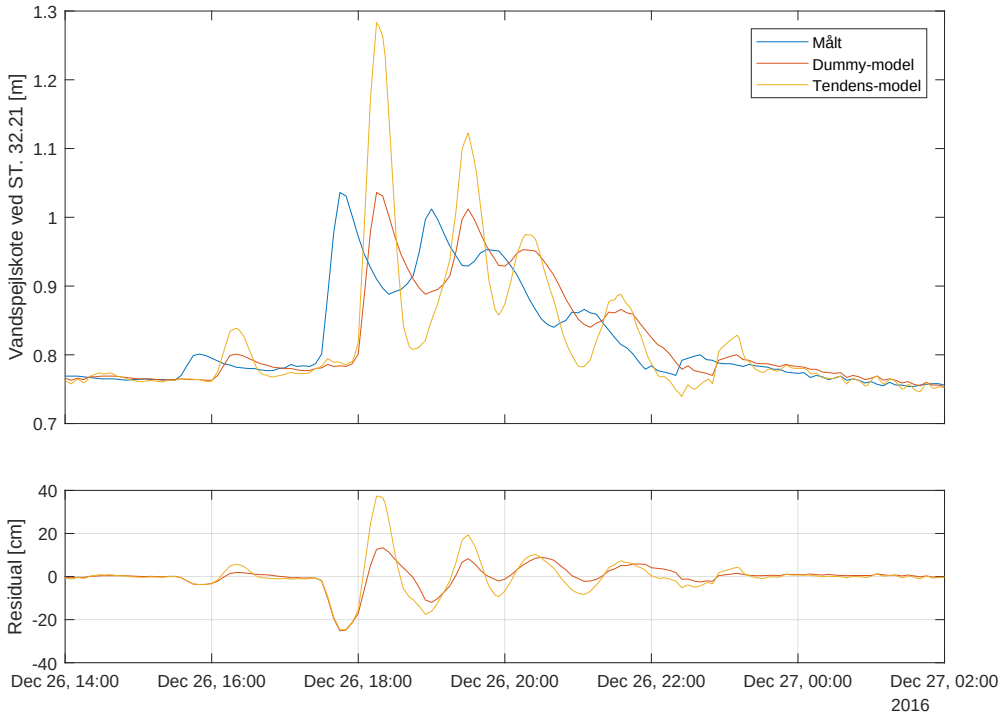
Figur 6.14: Den direkte afstrømning fra Vejle Midtby til Omløbsåen (ST. 32.21) er tydelig, når der sammenlignes med vandspejlskoten ved Ny Grejsdalsvej (ST. 32.22) og Mølleåen (ST. 32.20).

Selvom tendens-modellen i henhold til afsnit 6.4.1 egner sig godt til naturlig afstrømning, er der store fejl mellem den forudsagte værdi og den målte, når der er meget hurtige afstrømninger fra Vejle Midtby. Det er derfor forsøgt, om radardata på lokal nedbør over Vejle Midtby som et input til et neuralt netværk sammen med tendens-modellen kan forudsige vandspejlskoten både ved naturlig afstrømning og de hurtige afstrømninger fra Vejle Midtby.

Inddragelsen af lokal nedbør over Vejle Midtby som input til et neuralt netværk viser sig, på trods af de tydelige påvirkninger ved lokal nedbør, ikke at bidrage til en bedre forudsigelse. Dette vurderes at skyldes en stor usikkerhed på de anvendte radardata. Antallet af hændelser med intensiv lokal nedbør over Vejle Midtby er naturligvis begrænset, hvilket sætter et stort krav til kvaliteten af dataene ved de begrænsede antal hændelser, der optræder i perioden med data. Der er fundet eksempler på, at datakvaliteten på radardataene ved nogle af disse hændelser er tvivlsom, hvilket forklarer den manglende forbedring ved inddragelse af lokal nedbør som input til et neuralt netværk. Eksemplerne på den tvivlsomme datakvalitet er beskrevet i bilag I. For at inddragelsen af lokal nedbør som input til et neuralt netværk skal bidrage til en bedre forudsigelse af vandspejlskoter, vurderes det, at det er nødvendigt med et større datasæt, hvor datakvaliteten ligeledes er bedre.

6.4.3 Anvendelse af tendens-model til forudsigelse af vandspejlskote under høj vandstand i Vejle Havn

Vandstanden i Omløbsåen er i henhold til afsnit 3.2 modsat de andre modellerede stationer også påvirket af vandstanden i Vejle Havn. For at illustrere tendens-modellens anvendelighed, når vandstandsændringerne i Aagade er forårsaget af opstuvning fra Vejle Havn, er der udvalgt en hændelse 26/12-2016, hvor vandspejlkoten i Vejle Havn er høj og samtidig ændrer sig hurtigt. På figur 6.15 ses resultatet af en prognose med 30 minutters forudsigelsestid.

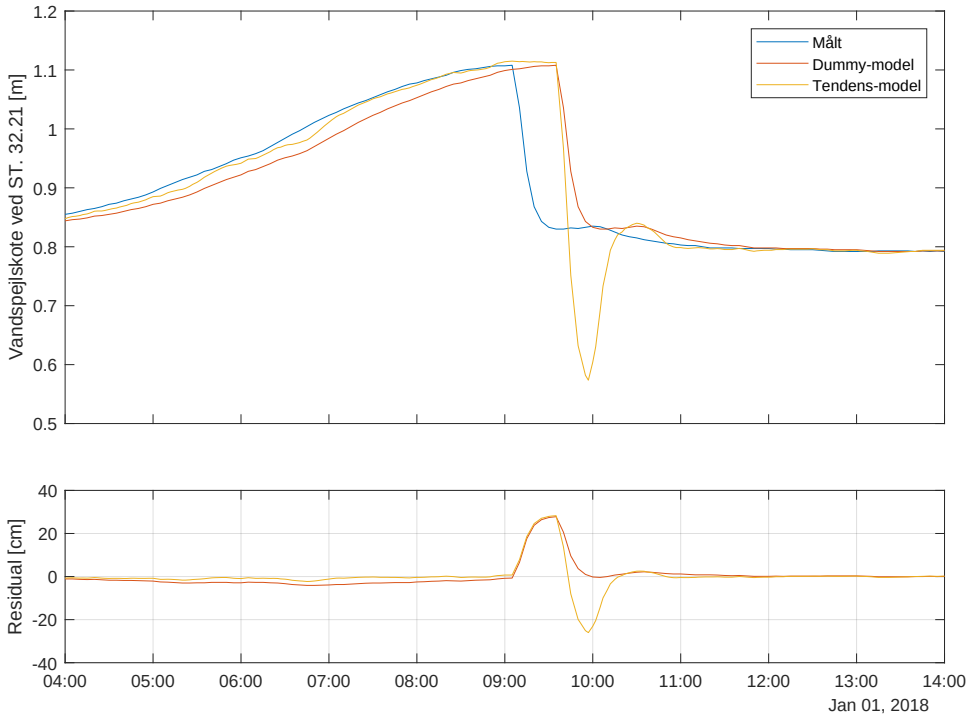


Figur 6.15: Målt og modelleret vandspejlskote ved Aagade (ST. 32.21) under en hændelse, hvor ændringen af vandspejlskoten skyldes opstuvning fra Vejle Havn.

Det er tydeligt, at tendens-modellen i tilfældet 26/12-2017 ikke egner sig til at modellere vandspejlkoten ved Aagade. Der er dog også tilfælde, hvor vandspejlskoten i Vejle Havn stiger over en længere tidsperiode. Tendens-modellen klarer sig i disse tilfælde fint. I tilfælde som hændelsen 26/12-2017 er forudsigelsen af vandspejlskoten i Vejle Havn altafgørende for, om vandspejlskoten ved Aagade kan forudsiges. Vandspejlskoten i Vejle Havn forudsiges allerede på nuværende tidspunkt af DMI (Dansk Meteorologisk Institut 2018). Der er derfor i dette projekt ikke fokuseret på at forudsige vandspejlskoten i Vejle Havn. Det er dog værd at lægge mærke til, at der er en naturlig forsinkelse i systemet, så en forhøjet vandstand i Vejle Havn vil først forplante sig til Aagade cirka et kvarter senere.

6.4.4 Anvendelse af tendens-model til forudsigelse af vandspejlskote under styring

Vandspejlskoten ved Aagade er som beskrevet i afsnit 3.2 påvirket af sluseanlægget ved Bryggen såvel som fordelerbygværket ved Abelones Plads. Når der ændres på spjældene ved fordelerbygværket eller sluseportene lukker eller åbner, sker der er en hurtig påvirkning af vandspejlskoten ved Aagade. Det medfører, at tendensmodellen ikke forudsiger korrekt, hvilket er illustreret på figur 6.16.



Figur 6.16: Målt og modelleret vandspejlskote ved Aagade (ST. 32.21) under en hændelse, hvor ændringen af vandspejlskoten skyldes en påvirkning fra styring.

På den baggrund vurderes det, at tendens-modellen ikke kan modellere vandspejlskoten i Omløbsåen i en kort periode efter en styring er fundet sted. Tendens-modellen vil derefter modellere den regulerede vandspejlskote. For at vurdere, hvornår det ikke længere er nødvendigt at styre vandspejlkoten er det derfor nødvendigt med en alternativ metode. Til dette kan neurale netværk anvendes.

6.4.5 Anvendelse af neuralt netværk til modellering af Omløbsåen i regulerede perioder

Foregående analyser viser, at brug af tendens-modellen er nyttig i perioder, hvor vandstanden i Omløbsåen (ST. 32.21) ikke reguleres af fordelerbygværket eller sluseanlægget. Derfor danner tendens-modellen et fint grundlag til beslutningstagere om enten at justere på spjældene ved fordelerbygværket, lukke sluseportene eller begge dele. I de regulerede perioder får tendens-modellen forventeligt problemer, og derfor kan den ikke anvendes til at vurdere, hvorvidt det er holdbart at åbne spjæld og sluseporte igen. Da målsætningen er at holde sluseportene åbne så meget som muligt, eksempelvis på grund af bedre vandreforhold for fisk, søges det at lave et neuralt netværk, der kan simulere vandstanden i Omløbsåen, hvis den forblev ureguleret.

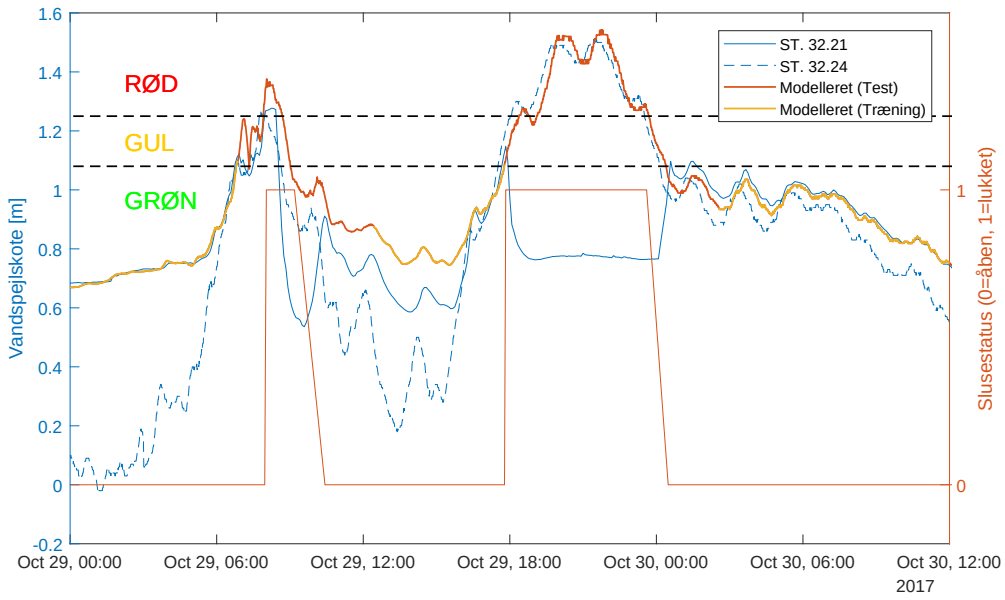
For at identificere uregulerede perioder skal data ved fordelerbygværket og sluseanlægget inkluderes. Af tidslinjen på figur 2.9 ses det, at data ved fordelerbygværket begrænser modellen til perioden 29/03-2017 til 06/01-2018. Hernæst udvælges uregulerede perioder ved at finde ændringstidspunkter på spjældåbninger og lukkestatus ved sluseanlægget, som frasorteres sammen med de efterfølgende 2 timer, da de efterfølgende 2 timer ofte er præget af, at systemet er ved at tilpasse sig en ny ligevægtssituation. Hernæst benyttes vandstanden ved Ny Grejsdalsvej (ST. 32.22) og i Vejle Havn (ST. 32.24) som input til et netværk, der skal trænes ind på vandstanden

i Omløbsåen (ST. 32.21). Det gøres, fordi målestationen ved Ny Grejsdalsvej (ST. 32.22) fint beskriver afstrømningen fra Grejs Å, mens målestationen i Vejle Havn (ST. 32.24) kan beskrive blandt andet tidevandspåvirkning. I henhold til afsnit 3.2 fremgår det, at det er de to væsentligste årsager til vandstandsændringer i Omløbsåen (ST. 32.21), og netværkskonfigurationerne ses af tabel 6.1. Bemærk at der er trænet på alle uregulerede perioder, da modellen er begrænset til en kort periode af dataen fra fordelerbygværket, hvor det vurderes, at det nødvendige grundlagt for både at træne og validere ikke er til stede, blandt andet fordi der ganske sjældent måles kritiske vandspejlskoter i denne periode.

Tabel 6.1: Netværkskonfiguration til modellering af Omløbsåen (ST. 32.21) på baggrund af vandstanden ved Ny Grejsdalsvej (ST. 32.22) og i Vejle Havn (ST. 32.24).

Parameter	Konfiguration
Input	Ny Grejsdalsvej (ST. 32.22) Vejle Havn (ST. 32.24)
Target	Omløbsåen (ST. 32.21)
Antal neuroner i det skjulte lag	10
Træning	Alle uregulerede perioder fra 29/03-2017 → 06/01-2018 : 100% træning
Test	Alle regulerede perioder fra 29/03-2017 → 06/01-2018

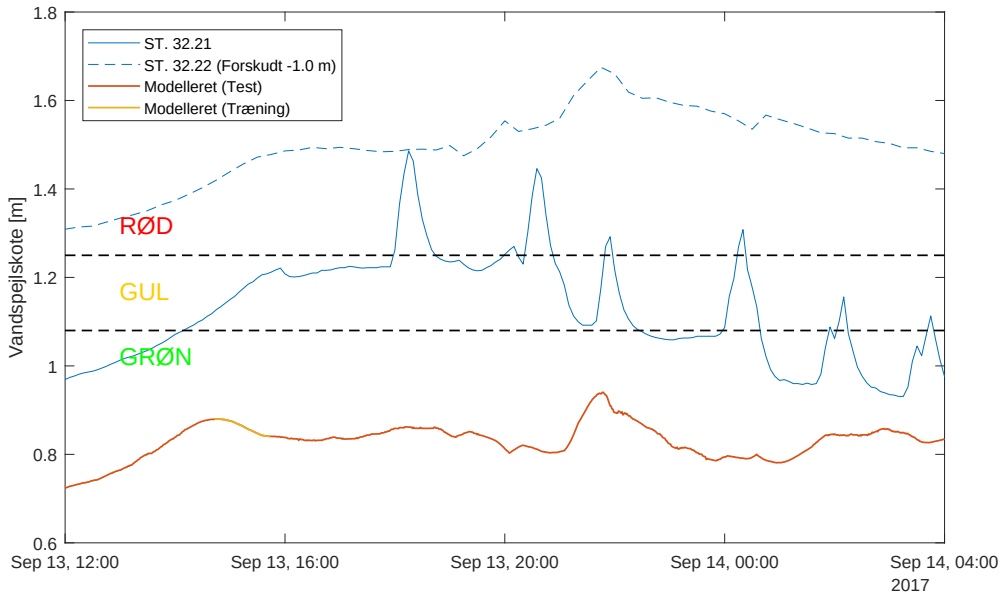
Når netværket trænes på uregulerede perioder, så lærer det en relation mellem vandstanden ved Ny Grejsdalsvej (ST. 32.22) og vandstanden i Vejle Havn (ST. 32.24) under naturlige forhold. Når netværket testes på regulerede perioder er tanken, at netværket skal kunne gengive vandstanden, som den ville have været under naturlige forhold, hvis ikke der var foretaget en styring af sluseanlægget eller fordelerbygværket. Figur 6.17 viser et eksempel, hvor det går godt.



Figur 6.17: Når vandstanden i Omløbsåen (ST. 32.21) i høj grad er styret af vandstanden i Vejle Havn (ST. 32.24), kan det neurale netværk gengive en realistisk potentiel vandstand i de lukkede perioder. Den gule/røde graf er outputtet fra det neurale netværk beskrevet i tabel 6.1.

Vandstanden i Vejle Havn (ST. 32.24) er i dette tilfælde styrende for vandstanden i Omløbsåen (ST. 32.21), og derfor skærmer sluseportene af for opstuvning. Vandspejlskoten i Vejle Havn (ST. 32.24) er en stærk randbetingelse for systemet, når vandstanden er høj i Vejle Havn. Derfor er relationen mellem vandspejlskoten i Vejle Havn og Omløbsåen relativ nem at beskrive.

Ved høj afstrømning fra Grejs Å er spjældindstillingen ved fordelerbygværket væsentlig for vandspejlskoten i Omløbsåen. Det betyder også, at en model uden input fra fordelerbygværket har svært ved at beskrive, hvor stor en indflydelse afstrømningen fra oplandet har, hvilket ses af eksemplet på figur 6.18.



Figur 6.18: Når vandstanden i Omløbsåen (ST. 32.21) afhænger af afstrømningen i Grejs Å (ST. 32.22), kan modellen ikke gengive vandstandene uden input fra fordelerbygværket. Dette er perioden med strømsvigt ved sluseanlægget, som beskrevet i afsnit 3.2.

Det ses, at modellen er forskudt og generelt ligger 30-40 cm under det rette niveau, hvilket skyldes en indstilling på spjældene ved fordelerbygværket, som ikke inkluderes i dette netværk. For at forbedre modellen undersøges det derfor, hvorvidt det er muligt at give et brugbart input fra fordelerbygværket, der kan beskrive niveauforskelle forårsaget af indstillingen på spjældene.

Tolkning af data ved fordelerbygværk

Som beskrevet i kapital 2 er perioden, hvor der haves tilgængeligt data om spjældindstilling, begrænset fra 29/03-2017 til 06/01-2018. De tre spjæld ved fordelerbygværket styres uafhængigt af hinanden, hvilket resulterer i, at der haves tre datasæt over åbningsprocenter. At spjældene er beskrevet hver for sig giver mange kombinationer af spjældindstillinger. De mange kombinationer kræver et stort datagrundlag med mange kombinationer inkluderet. Dette er ikke tilfældet for den begrænsede periode, som der anvendes data fra. Derfor undersøges det, hvorvidt det er muligt at simplificere spjæld-data, så de tre serier for åbningsprocenter reduceres til én serie. En af de store udfordringer ved at simplificere dataene er, at betydningen af en ændring i åbningsprocenten afhænger af, om spjældene befinder sig over eller under vandoverfladen.

Simplificeringen af data for åbningsprocenter forsøges derfor at gøres ved to forskellige metoder:

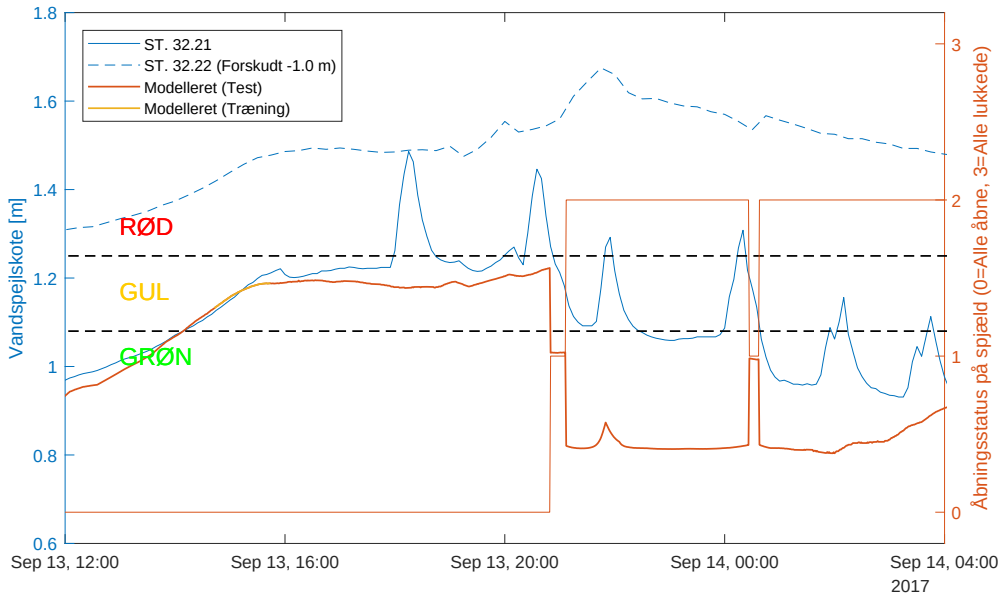
- *Metode 1*: Fastlæggelse af en åbningsprocent, hvor spjældene defineres som åbne, når de er

over denne åbningsprocent og lukkede, når de er under. Det nye input vil således bestå af ét input, som kan antage værdierne 0 til 3 alt efter antal lukkede spjæld. Tanken er, at den fastlagte åbningsprocent kan ses som vandspejlskoten, så spjældene ses at have en betydning for vandstanden, når de er under denne åbningsprocent og ingen betydning, når de er over.

- *Metode 2*: Bestemmelse af det tværsnitsareal, som spjældene har under vandspejlskoten. Denne metode vurderes at være mere sigende for spjældenes påvirkning på vandstanden end metode 1. Dog kræver denne metode, at der haves data for vandspejlskoten umiddelbart før spjældene, og at spjældenes åbningsprocenter kan omsættes til en vandspejlskote.

I appendiks J vises en gennemgang af begge metoder. Datagrundlaget for vandspejlskoten umiddelbart før spjældene vurderes at være for dårlig til, at metode 2 kan anvendes. For metode 1 er det fundet, at det giver det bedste resultat, hvis der skelnes mellem, at spjældene er åbne og lukkede ved en åbningsprocent på 33 %.

Ved anvendelse af et neuralt netværk, som også inkluderer det nye simplificerede input, der er bestemt ved metode 1, fås modellen vist på figur 6.19.



Figur 6.19: Når åbningsstatussen på spjældene ved fordelerbygværket inkluderes som input, ses mere realistiske modellerede værdier i første del af perioden ved hændelsen d. 13 september.

Sammenlignes figur 6.19 med 6.18 ses mere realistiske modellerede værdier i første del af perioden, når der inddrages åbningsstatus på spjældene. Når værdien af vandspejlskoten ved Ny Grejsdalsvej (ST. 32.22) stiger over kote 2,4 vurderes nøjagtigheden af den modellerede værdi i Omløbsåen løbende at falde grundet et reduceret datagrundlag ved høje vandspejlskoter.

På trods af udfordringerne ved at inddrage spjældindstillingerne vurderes modelkvaliteten i regulerede perioder at forbedres ved inddragelse af det forsimple input i form af åbningsstatus på spjældene. Der forventes derfor at være potentiale for en klar forbedring af modelkvaliteten, hvis datamængden og datakvaliteten fra fordelerbygværket øges.

6.5 Anbefaling til styring

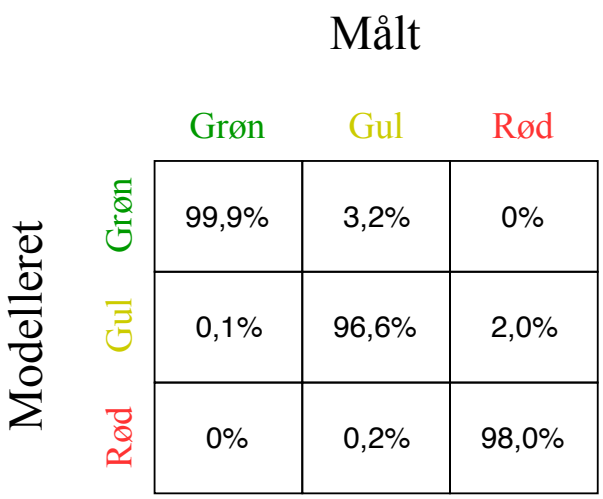
Beslutningsgrundlag

Tendens-modellen er god til at beskrive ændringer af vandstanden i Omløbsåen ved Aagade, indtil der foretages en styring på fordelerbygværket eller sluseanlægget. Det er vanskeligt at lave en analyse af modellens timing og afvigelse fra peak-vandstande, netop fordi der ofte foretages styring ved forhøjede vandstande. Derfor er det undersøgt, hvor ofte modellens output ligger i samme kategori som målingen.

Da der næsten ikke måles nogle vandstande ved ST. 32.21 i rød zone, som beskrevet i afsnit 3.2, justeres kategorierne for at danne et passende statistisk grundlag. Derfor ændres koterne for hver kategori til følgende:

- Grøn: < 0,9 m; 272863 elementer. (Grøn tidligere: <1,08 m)
- Gul: 0,9 m - 1,08 m; 4959 elementer. (Gul tidligere: 1,08 m-1,25 m)
- Rød: > 1,08 m; 510 elementer. (Rød tidligere: >1,25 m)

På det grundlag haves sammenfaldsprocenter mellem den målte vandstand og en prognose af vandstanden ved en forudsigelsestid på 30 minutter, som vist på figur 6.20.



Figur 6.20: Sammenfaldsprocenter mellem den målte vandstand og en prognose ved en forudsigelsestid på 30 minutter for alle tre kategorier.

Det bemærkes, at der ved ingen tilfælde modelleres grøn og måles rød og omvendt. Det betyder, at tendens-modellen aldrig forudsiger ukritiske tilstande, mens der er risiko for oversvømmelser langs Omløbsåen.

Da effekten ved lukning af sluseporte og opstart af pumper forplanter sig indenfor en tidshorisont på under 10 minutter, kan det diskuteres, hvorvidt der er behov for en prognose ved denne tidshorisont, fordi det i givetfald vil være næsten ligeså godt at anvende den aktuelle vandstand til at styre efter.

Styringsstrategi

Anvendelsen af modeller baseret på neurale netværk vurderes at kunne forbedre beslutningsgrundlaget for styringen af fordelerbygværket og sluseanlægget i følgende situationer:

- Før der tages en beslutning om at styre vandstanden i Omløbsåen, kan det med neurale

netværk modelleres, hvad vandspejlskoten vil blive ved forskellige indstillinger på fordelerbygværket, og hvis sluseanlægget tages i brug.
- Den nuværende største gevinst ved at anvende neurale netværk i forbindelse med styringsstrategien vurderes at være muligheden for i regulerede perioder at modellere, hvad den uregulerede vandspejlskote i Omløbsåen vil være. Hermed haves et beslutningsgrundlag mens sluseanlægget er aktivt for, hvornår sluseportene igen kan åbnes.
- Modelleringen af en ureguleret vandstand i Omløbsåen under regulerede perioder kan ligeledes anvendes til at evaluere, hvorvidt den anvendte styringsstrategi under en hændelse var optimal.

Forbedring af datakvalitet

For at opnå et bedre beslutningsgrundlag bør det overvejes, om spjældene ved fordelerebygværket ved Abelones Plads kan styres på en smartere måde. Det anbefales at spjældstyringen ændres, så der er en klar systematik i styringen. Spjældene kan for eksempel styres, så de som udgangspunkt altid har samme lukningsgrad. Spjældindstilling af de tre spjæld vil dermed reduceres fra tre til ét input, når det skal anvendes i et neuralt netværk. Dette vil reducere kombinationer af spjældinstillinger og dermed datamængden, som er nødvendig for at træne gode neurale netværk.

Målinger af vandspejlskoter før og efter ristene ved fordelerbygværket har vist sig at være væsentligt fejlbehæftede. Vandspejlskoterne ved ristene er væsentlige for beskrivelsen af, hvor tilstoppede ristene er og dermed også hvor meget vand, der løber til Omløbsåen. Denne betragtning er ikke medtaget i de foregående analyser grundet det ringe datagrundlag. En bedre datakvalitet vil her gavne beskrivelsen af ristenes betydning.

En bedre beskrivelse af nedbøren over Vejle Midtby kan formentlig bidrage til en bedre forudsigelse af påvirkningen, som den direkte afstrømning til Omløbsåen fra Vejle Midtby har på vandstanden i Omløbsåen. Flere regnmålere tæt ved Vejle Midtby vil kunne bidrage til et bedre kalibreringsgrundlag til radardata over Vejle Midtby.

7. Lange forudsigelser

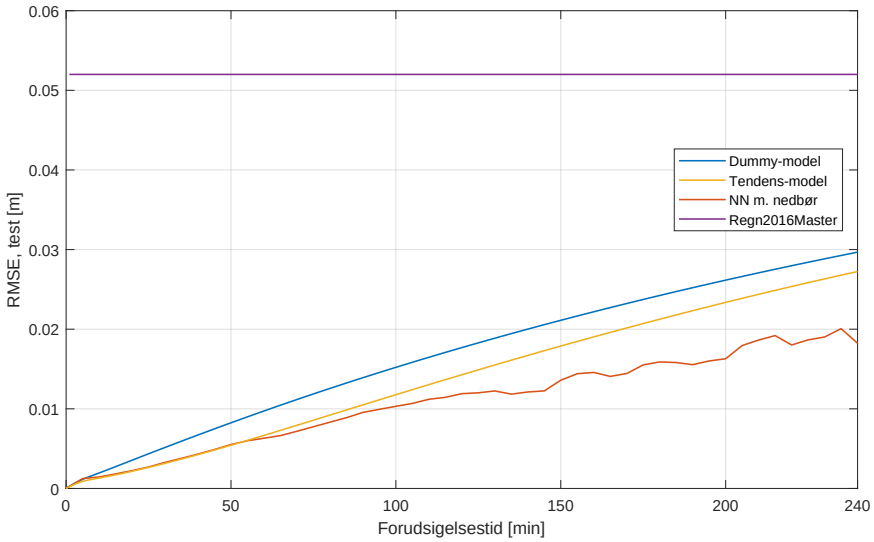
Korte forudsigelser er anvendelige til styring af fordelerbygværket og sluseanlægget. Lange forudsigelser er i dette projekt defineret som modeller, der kan forudsige op til og med 4 timer. Formålet med disse modeller er at danne bedre grundlag for varsling af høj afstrømning fra Grejs Å. I dette kapitel undersøges det, hvordan nedbør kan inddrages i neurale netværk for på den måde at øge kvaliteten af de lange forudsigelser.

7.1 Formål med lange forudsigelser

For at kunne klargøre beredskab og borgere i Vejle Midtby på en eventuel oversvømmelse fra Omløbsåen, er det en nødvendighed til en vis grad at kunne forudse en potentiel høj afstrømning i Grejs Å med en længere forudsigelsestid. Som beskrevet i afsnit 6.2 viser de udførte analyser, at der, indtil der skal forudsiges længere end én time, ikke er en gevinst i at anvende andre input end information omkring vandspejlskoten ved den modellerede station selv. Når forudsigelsestiden øges, vil informationen, der ligger i vandspejlsmålingerne, dog løbende miste sin værdi, og forudsigelsen vil blive mere afhængig af nedbør. Dette princip er illustreret på figur 7.1.

Modellen, som på figur 7.1 kaldes Regn2016Master, er et resultat fra et tidligere afgangsprojekt, som havde til formål at afdække neurale netværks anvendelighed i forhold til hydraulisk modellering (M. M. Rasmussen m.fl. 2016). Projektlokaliteten var tilsvarende Vejle, men fokus lå i at modellere vandspejlskoter alene på baggrund af nedbøren og dermed uden anvendelse af den aktuelle måling ved stationen selv. Analyserne i projektet viser, at det bedste netværk til formålet, bestod af input i form af nedbør i de foregående 130 dage og af en årstidsvariation, som skulle beskrive Manning-tallet.

Det er gennem analyser fundet, at anvendes målinger for vandspejlskoten ved stationen selv som input sammen med nedbør, så ændrer det optimale nedbørs-input sig fra 130 dage til kun at skulle beskrive nedbør 1-2 døgn op til den modellerede vandspejlskote. Dette giver også god fysisk mening, da informationen omkring mængden af nedbør længere tilbage end 2 døgn primært må formodes at beskrive grundvandstilstrømningen, hvilket implicit er beskrevet i vandspejlsmålingen.



Figur 7.1: Der ses en forbedring af tendens-modellen på forudsigelsestider over én time, når den indføres som input i et neuralt netværk sammen med nedbør.

I følgende afsnit beskrives det i hvilket opfang, tendensmodellen (beskrevet i afsnit 6.2) forbedres ved at anvende nedbør som input til et neuralt netværk. Det gennemgås således også, hvordan nedbør som input er optimeret.

7.2 Nedbør som input til neurale netværk

Problemet ved at anvende tendens-modellen ved forudsigelsestider over én time er, at denne ikke forudser en ændring fra den aktuelle tendens på vandspejlskurven. En sådan ændring vil ved naturlig afstrømning oftest forekomme, hvis der er sket en nævneværdig ændring i nedbøren op til det modellerede tidspunkt. Derfor er det forventeligt, at et input i form af nedbør har en gavnlig effekt. Der er dog ikke en simpel metode til tilføjelse af nedbør som input, da nedbøren forekommer som en tidsserie med målinger i ét minuts intervaller. Hvis nedbøren for et døgn skulle beskrives med ét minuts præcision i et neuralt netværk, ville dette svare til $60\,\text{min} \cdot 24\,\text{timer} = 1440$ input! Et så stort antal input vil udover at øge tiden, som det tager modellen at træne ekstremt, også medføre en større sandsynlighed for overfitting.

Det er derfor forsøgt at akkumulere nedbør over forskellige tidsperioder og anvende de akkumulerede værdier som input. Det er vurderet, at forsinkelsen og udglatningen fra nedbør til vandføring (beskrevet i afsnit 3.1) gør sådan en omregning forsvarlig. Det forsøges at optimere nedbørs-inputtet, så vandstandsændringerne forårsaget af naturlig afstrømning kan beskrives bedre. Derfor er analyserne af nedbør som input til de neurale netværk lavet på modelleringen ved Grejsdalens Planteskole, da denne station i henhold til afsnit 3.4 er den eneste station, som ikke er påvirket af styring.

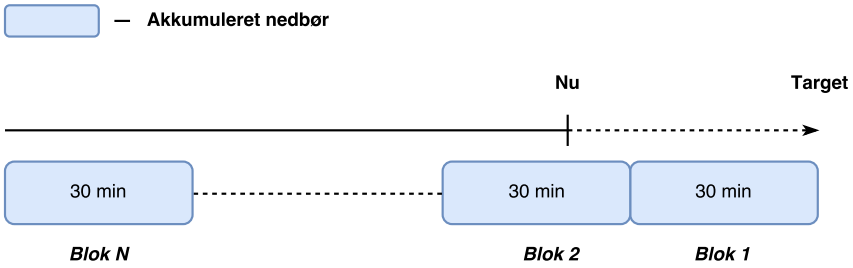
Det er forsøgt at bestemme et godt nedbørs-input ud fra to forskellige strategier:

- **Strategi 1**: Akkumuleringsperioderne er fastholdt på 30 minutter og antallet af input øges indtil en forbedring i netværkskvalitet ikke længere er synlig.
- **Strategi 2**: Akkumuleringsperioderne er varierende, så nedbør, der er faldet tættere på det modellerede tidspunkt, beskrives mere detaljeret. Da der er en naturlig forsinkelse mellem nedbør og afstrømning, optimeres forsinkelsen mellem den modellerede værdi af vandstand og det første nedbørsinput.

At akkumulere signaler på denne måde og sætte dem efter hinanden er hvad, der normalt kaldes en tapped delay line (TDL).

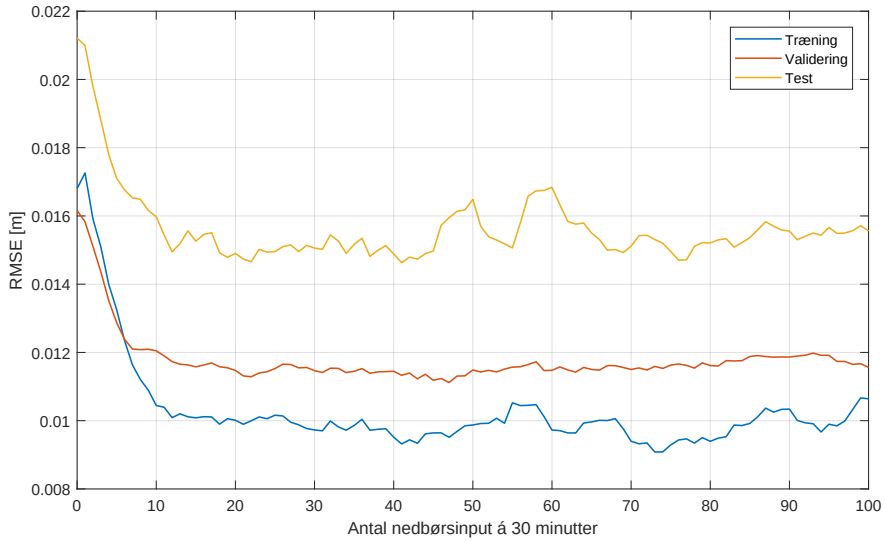
Strategi 1 - Faste akkumuleringsblokke

Det forsøges at akkumulere nedbør fra regnmåleren i Jelling (SVK 5230) over 30 minutters tidsintervaller. Det akkumulerede nedbør anvendes herefter som input sammen med den tidligere omtalte tendens-model til et neuralt netværk til forudsigelse af vandspejlskoten ved Grejsdalens Planteskole 3 timer frem. Der tages ikke højde for forsinkelse mellem nedbør og afstrømning i dette forsøg, hvilket er bevidst, da forsinkelsen er en størrelse, der for det første varierer, men for det andet er vanskelig at beskrive med en nøjagtighed, der gør sig gældende ved alle typer afstrømningshændelser eller på tværs af årstider. Det bemærkes også, at når input indsættes på denne måde, så er det nødvendigt med prognoser af nedbøren, hvorimod der i strategi 2 foræres en del af forudsigelsen gennem forsinkelse i systemet. I følgende analyser antages nedbøren at kunne forudsiges perfekt, eller inden for måleusikkerheden, da der anvendes historisk regn. Inputtet indsættes i form af blokke, som vist på figur 7.2.



Figur 7.2: Nedbørsblokke á 30 minutters akkumuleret nedbør indsættes direkte efter target-vandstanden.

Det er forsøgt at øge antallet af nedbørsblokke, indtil det ikke længere har en synlig positiv effekt på modelresultatet. Resultatet vises på figur 7.3.



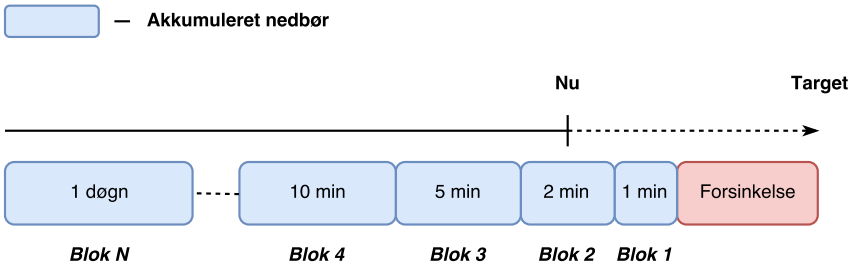
Figur 7.3: Hvis der anvendes over 20 nedbørsinput, falder RMSE-værdien i testperioden ikke yderligere.

Beregningstiden og kompleksiteten af netværket øges kraftigt med antallet af nedbørsinput. RMSE-

værdien i testperioden falder ikke ved tilføjelse af flere input end 20 nedbørsblokke, og dette er derfor vurderet at være det optimale antal. Det svarer til akkumuleret nedbør på ti timer bagudrettet fra target-vandstanden.

Strategi 2 - Varierende akkumuleringsblokke

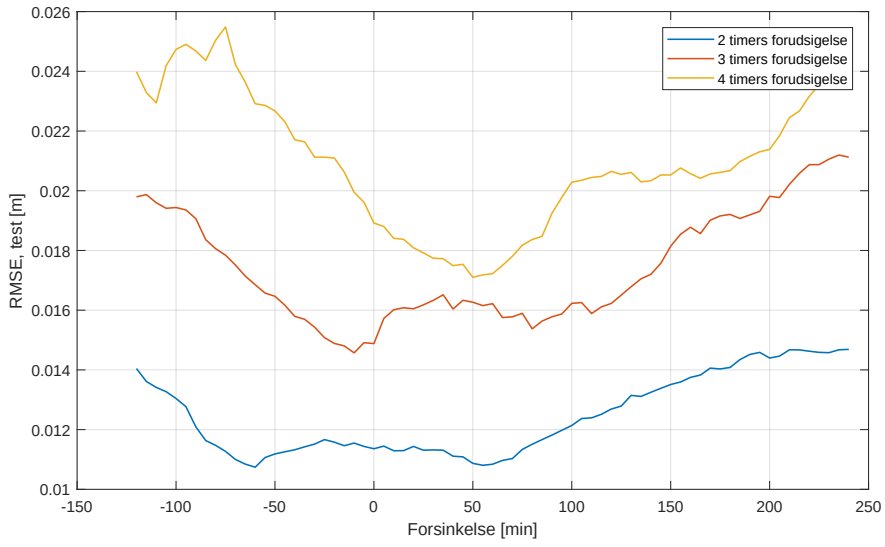
Det forsøges at anvende et forudantaget antal nedbørsblokke og variere akkumuleringstiden fra blok til blok, mens der tages højde for forsinkelsen i systemet. Ved at variere akkumuleringstiden fra kort til lang som vist på figur 7.4, udjævnes nedbøren endnu mere, mens peak-signalet er kraftigere.



Figur 7.4: Akkumuleringstiden varieres fra blok til blok, mens der også tages højde for forsinkelsen mellem målt nedbør i Jelling (SVK 5230) og vandstanden ved Grejsdalens Planteskole (ST. 32.06).

Tabel 7.1: I strategi 2 anvendes ti nedbørsblokke med disse akkumuleringstider.

Blok [-]	1	2	3	4	5	6	7	8	9	10
Akkumuleringstid [min]	1	2	5	10	30	60	120	300	720	1440



Figur 7.5: En forsinkelse på 60 minutter giver et fornuftigt resultat ved de tre analyserede forudsigelsestider.

Figur 7.5 viser RMSE-værdierne ved en ændring i forsinkelsen mellem den modellerede værdi af vandstand og det første nedbørsinput. Det ses, at der er en tendens til, at den optimale forsinkelse stiger med forudsigelsestiden. Dette skyldes, at informationen, der ligger i tendensmodellen som

input, varierer alt efter forudsigelsestiden. Det er ikke valgt at gå yderligere i detaljer med optimering af forsinkelsen, men blot konstateret, at en forsinkelse på 60 minutter giver et fornuftigt resultat ved alle analyserede forudsigelsestider.

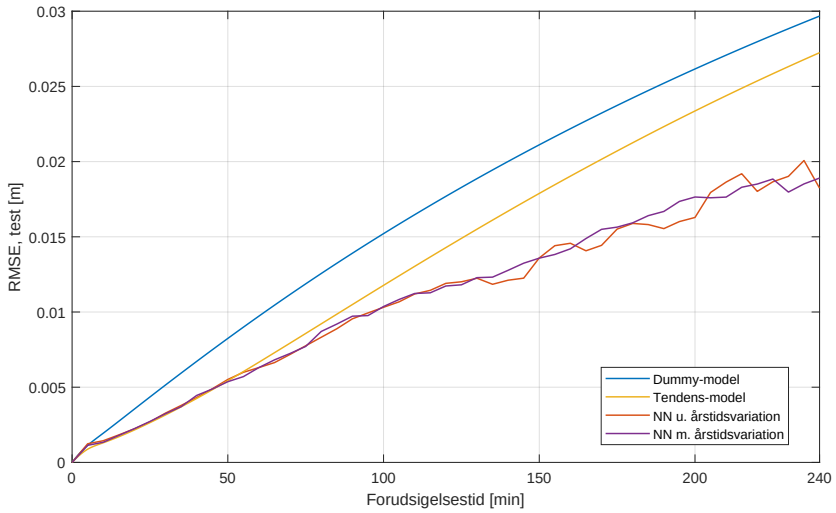
Sammenligning af strategi 1 og 2

Resultaterne er meget lig hinanden ved de to strategier, når begge metoder er optimeret. Dette ses ved, at RMSE-værdien for testperioden på figur 7.3 ligger omkring 0,015 m. Der ses en tilsvarende lav RMSE-værdi for en forudsigelsestid på 3 timer på figur 7.5, og dermed vurderes det, at valget af strategi ikke har en væsentlig betydning.

Da nedbøren ikke er jævnt fordelt over oplandet ved de fleste hændelser, er det forsøgt at inddrage regnserier fra flere regnmålere af gangen som input. Dette har dog ikke nogen nævneværdig effekt på modelkvaliteten, og da det øger den tid, som det tager at træne modellen, er det ikke vurderet at være fordelagtigt at anvende flere regnmålere ad gangen som input.

7.3 Forkastning af Manningtal og potentiel fordampning som input til neurale netværk

I kapitel 3.1 er det beskrevet, hvordan vandstanden i et vandløb er et resultat af mange forskellige parametre. Dermed er det naturligt at forvente, at de neurale netværk til forudsigelse af vandstand potentielt vil have gavn af andre parametre som input sammen med nedbøren. Nogle af de væsentligste styrende parametre formodes at være Manningtal og potentiel fordampning. Som nævnt i afsnit 7.1 er der i et afgangsprojekt fra år 2016 fundet, at den bedste modellering af vandspejlskoter uden brug af målte vandspejlskoter som input forekommer ved anvendelse af en årstidsvariation, som skulle beskrive Manningtallet sammen med data på nedbør (M. M. Rasmussen m.fl. 2016). Afgangsprojektet fra år 2016 viste ikke forbedrede resultater ved inddragelse af potentiel fordampning.



Figur 7.6: Indragelse af en årstidvariation som beskrivelse af Manningtal og potentiel fordampning har ingen synlig påvirkning på modelkvaliteten.

Det formodes, at grunden til, at indragelse af potentiel fordampning kombineret med årstidsvariationen ikke har en positiv effekt på modelkvaliteten i afgangsprojektet fra 2016, skyldes, at årstidsvariationen, som beskriver Manningtallet, samtidig beskriver den potentielle fordampning,

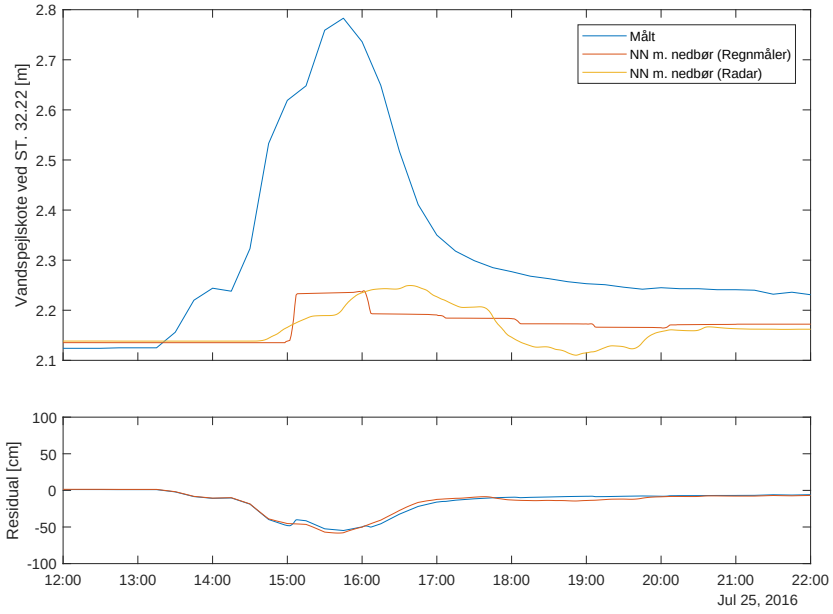
som også varierer hen over året. Ved at anvende en simpel årstidsvariation vurderes det, at sandsynligheden for overfitting reduceres, hvilket også kommer til udtryk i afgangsprojektet fra 2016 ved den laveste forskel i modelkvalitet mellem træning- og testperioden blandt de opsatte modeller. Det er på baggrund af dette valgt at analysere virkningen af indragelse af en årstidsvarition.

På figur 7.6 ses, at inddragelse af en årstidsvariation ingen synlig effekt har på modelkvaliteten. Årstidsvariationen er et input bestående af en sinuskurve med den laveste værdi 1. januar og højeste værdi 1. august.

Da inddragelsen af årstidsvariationen ingen synlig effekt har på modelkvaliteten, vurderes det ikke at være gavnligt at gå i dybden med en bedre beskrivelse af potentiel fordampning.

7.4 Anvendelighed af radardata som input til neurale netværk

Ved hændelsen, som er beskrevet på figur 3.14 i afsnit 3.3, falder der meget nedbør i oplandet, som regnmålerne ikke fanger. Derfor undersøges det, hvorvidt det er muligt at anvende radardata til at forbedre modelkvaliteten. Radardataen er midlet over de topografiske oplande, som beskrevet i kapitel 2, og inddrages som input i de neurale netværk i form af akkumulerede værdier på samme måde som nedbør i regnmålerne fortolkes, eksempelvis som vist på figur 7.4. For at kunne vurdere påvirkningen af anvendelsen af radardata tydeligere, er der valgt at sammenligne to netværk, som ikke får information omkring vandstanden selv som input. Resultatet af den undersøgelse ses på figur 7.7.



Figur 7.7: Der ses ingen forbedringer af det neurale netværk ved afstrømningshændelsen 25/07-2016, hvis der anvendes radardata i stedet.

Der ses ingen gavnlig effekt ved anvendelse af radardata i stedet for regnmålere. Det formodes at skyldes, at radardataene er kalibreret op mod regnmålere, hvilket medfører større usikkerhed på kalibreringen, når der forekommer stedlig fordelt nedbør. Desuden er antallet af hændelser, hvor der er stor variation på nedbøren over oplandet, begrænset, hvilket formodes at medføre manglende grundlag til gøre en model baseret på radardata bedre end en model baseret på nedbørsdata.

7.5 Modelkvalitet ved lange forudsigelser

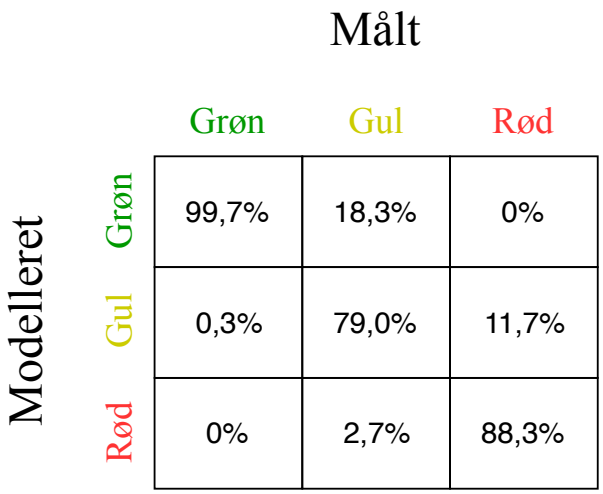
For at illustrere kvaliteten af forudsigelsen af vandspejlkoten vej Grejsdalens Planteskole med en forudsigelsestid på 4 timer, er det undersøgt, hvor ofte forudsigelsen rammer indenfor samme kategori, som måles. Sammenfaldsprocenterne kan på denne måde sammenlignes med de fundne sammenfaldsprocenter for forudsigelsen af vandspejlskoten i Omløbsåen ved Aagade ved en forudsigelsestid på 30 minutter, som er vist i afsnit 6.5.

Der er ikke defineret bestemte kritiske koter ved Grejsdalens Planteskole. Derfor er det valgt at inddele målingerne ved Grejsdalens Planteskole i kategorier, så der haves nogenlunde samme fordeling af elementer i kategorierne som i afsnit 6.5. Koterne for hver kategori inddeles som følgende:

- Grøn: < 28,0 m; 502099 elementer.
- Gul: 28,0 m - 28,2 m; 9268 elementer.
- Rød: > 28,2 m; 851 elementer.

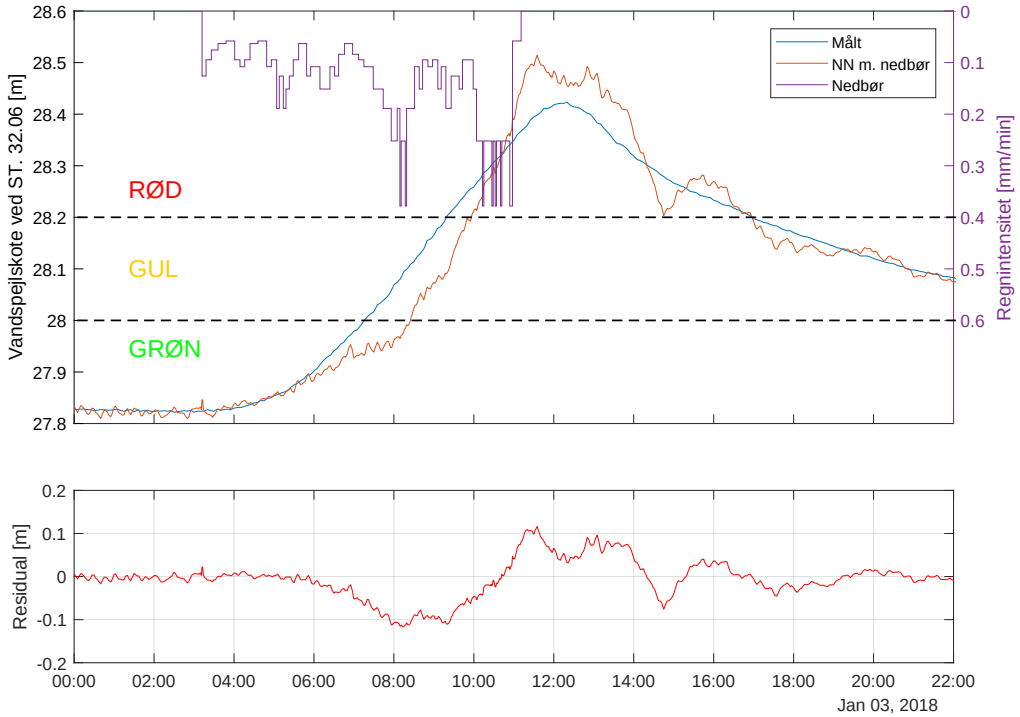
På figur 7.8 ses sammenfaldsprocenter mellem den målte og forudsagte vandspejlskote ved Grejsdalens Planteskole ved en forudsigelsestid på 4 timer.

Sammenfaldsprocenterne på figuren viser, at der er en lavere procentdel af korrekte sammenfald sammenlignet med tidligere resultater ved en forudsigelsestid på 30 minutter i Omløbsåen ved Aagade (figur 6.20). Det giver god mening, da forudsigelsestiden nu er otte gange længere. Derimod bemærkes det, at der fortsat ikke modelleres grøn og måles rød og omvendt, hvilket betyder, at modellen aldrig forudsiger lav vandstand, mens der er risiko for oversvømmelser i Vejle Midtby som følge af høj afstrømning fra Grejs Å.



Figur 7.8: Sammenfaldsprocenter mellem den målte og forudsagte vandspejlskote ved Grejsdalens Planteskole ved en forudsigelsestid på 4 timer for alle tre kategorier.

Afstrømningshændelsen den 03/01-2018 er den hændelse i testperioden, hvor vandstanden har været højest ved Grejsdalens Planteskole, og den er illustreret på figur 7.9.



Figur 7.9: Eksempel på en afstrømningshændelse 03/01-2018, hvor de højeste vandstande ved Grejsdalens Planteskole måles. Når vandstanden er opadgående og i gul zone, så har en model med en forudsigelsestid på 4 timer en tendens til at underestimere. Omvendt er der fin overensstemmelse mellem model og målinger, når vandstanden igen er nedadgående.

Det ses af figur 7.8, at det især er den gule kategori, som viser uregelmæssigheder i sammenfaldsprocenterne. Kategorien beskriver oftest vandspejlskoter, der enten er opadgående eller nedadgående. Når vandspejlskoten er opadgående, og der måles vandstande i gul kategori, så underestimerer modellen en smule, hvilket kommer til udtryk ved 18,3 % sammenfald i kategorien *målt gul og modelleret grøn*. Da en opadgående vandspejlskote forårsages af nedbør, anbefales det derfor også at se på målinger og prognoser af nedbøren, fordi det også er en helt naturligt indikation for, hvor voldsom en afstrømningshændelse, der forventes. Begge disse tilfælde illustreres godt ved hændelsen på figur 7.9, da der er 100 % sammenfald ved nedadgående vandspejlskoter i både gul og rød kategori, mens der underestimeres, når vandspejlskoten er opadgående.

IV

Del IV - Afrunding

8 Konklusion 97

9 Perspektivering 99

8. Konklusion

Det overordnede formål med projektet har været at udvikle datadrevne modeller, som kan forudsige vandstande i Vejle Midtby og opland, så beslutningsgrundlaget for styringen af fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen kan forbedres, og der kan varsles om forhøjet afstrømning fra Grejs Å. Delelementer af den udvikling har været:

- At beskrive de styrende processer for vandstanden i danske vandløb, herunder Grejs Å og Omløbsåen.
- At målrette neurale netværk til modellering og forudsigelser af vandstandene på baggrund af forundersøgelser af de styrende processer.
- At forudsige vandstande i Grejs Å og Omløbsåen på korte og lange forudsigelseshorisonter med datadrevne modeller.
- At præsentere modelresultater på en forståelig måde, så de er praktisk anvendelige og kan forbedre beslutningsgrundlaget for styringen af fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen.

Vandstanden i danske vandløb er primært styret af nedbørsmængder, men flere andre faktorer som eksempelvis fordampning, geologiske forhold og Manningtallet i vandløbet har indvirkning på den resulterende vandstand. I Grejs Å og Omløbsåen gælder særlige forhold som opstuvning i vandløbene fra Vejle Fjord og styring af vandstanden ved fordelerbygværket ved Abelones Plads og sluseanlægget ved Bryggen. Disse forhold har alle betydning for vandstanden, men de er ikke nødvendigvis målbare, hvilket vanskeliggør en simpel omsætning af nedbør til vandstand. Neurale netværk har vist sig brugbare til at beskrive vandstanden ud fra komplicerede sammenhænge i målinger på en hurtig og effektiv måde.

Valget af input til netværkene har vist sig at have stor betydning, og forundersøgelser, såsom korrelationsanalyser, har vist sig at være gode værktøjer til at vælge dem godt. Øvrige netværkskonfigurationer er blevet optimeret til at kunne modellere vandstande, og det er således fundet, at særligt inddelingen af data i træning- og valideringsperioder er vigtig. Den traditionelle inddelingsmetode, hvor target inddeles tilfældigt i træning og validering, har vist sig at forårsage overfitting. Analyser har vist, at den optimale inddelingsmetode til modellering af vandstande er at inddele træningen og valideringen ligeligt i perioder på en halv måned af gangen, fordi det forbedrer netværkenes evne

til at modellere vandstande i sammenhængende perioder med uafhængig data.

På baggrund af kendskabet til de styrende processer for vandstanden i vandløb og de optimale netværkskonfigurationer, er vandspejlskoterne ved målestationerne i Grejs Å ved Grejsdalens Planteskole og Ny Grejsdalsvej og i Omløbsåen ved Aagade blevet modelleret ved forudsigelsestider, der dækker 0-4 timer.

Forudsigelsestiden og de gældende forhold ved den enkelte målestation har en betydning for, hvordan de neurale netværk skal anvendes. Vandstanden i Grejs Å er primært styret af nedbør og naturlig afstrømning. Når vandstandsstigninger er forårsaget af naturlig afstrømning, har det vist sig, at den bedste måde at forudsige vandspejlskoten med en forudsigelsestid op til én time er at beregne en vandspejlskote på baggrund af tendensen i de målte vandstande. Anvendelsen af et neuralt netværk med tidligere målte vandstande som input giver et tilsvarende resultat som en manuel fremskrivning af vandstanden. Der er dermed ikke nogen stor gevinst ved anvendelse af neurale netværk til korte forudsigelser ved naturlig afstrømning, men tilsvarende forringer det ikke kvaliteten af forudsigelsen. Ved længere forudsigelser mere end én time frem viser det sig at være en fordel at kombinere målte vandspejlskoter med nedbør, når de optimale input til et neuralt netværk skal udvælges. Det er ikke vurderet at give en forbedring af vandstandsforudsigelserne at inddrage andre parametre, men dette kunne eventuelt godt ændre sig, hvis forudsigelsestiden blev yderligere forøget, datamængden forøget eller datakvaliteten forbedret.

Påvirkninger fra fordelerbygværket viser sig at have stor betydning for vandstanden i Omløbsåen, og det er på baggrund af det anvendte datagrundlag ikke muligt, at inddrage spjældåbninger på en god og simpel måde for at forbedre modellerne. Datakvaliteten vurderes at kunne forbedres, hvis der foretages en systematisk styring af fordelerbygværket, hvormed kvaliteten af de neurale netværk vil blive forbedret. Det vil betyde, at der på sigt også kan udvikles et neuralt netværk, der kan beskrive den optimale indstilling på fordelerbygværket, så oversvømmelser undgås.

Neurale netværk viser sig særdeles brugbare til at opstille en relation mellem vandstanden i Vejle Havn og afstrømningen fra Grejs Å, der kan beskrive vandstanden i Omløbsåen. Det er anvendeligt i perioder, hvor sluseportene er lukkede, fordi netværket kan beskrive, hvordan det ville have set ud, hvis de forblev åbne, og dermed forbedrer beslutningsgrundlaget for, hvornår sluseportene forsvarligt kan åbnes igen.

9. Perspektivering

I denne perspektivering beskrives tre emner, som vurderes at have relevans for anvendelsen af neurale netværk til modellering indenfor fagområdet "Vand og Miljø" i fremtiden, men som ikke er beskrevet yderligere i nærværende projekt.

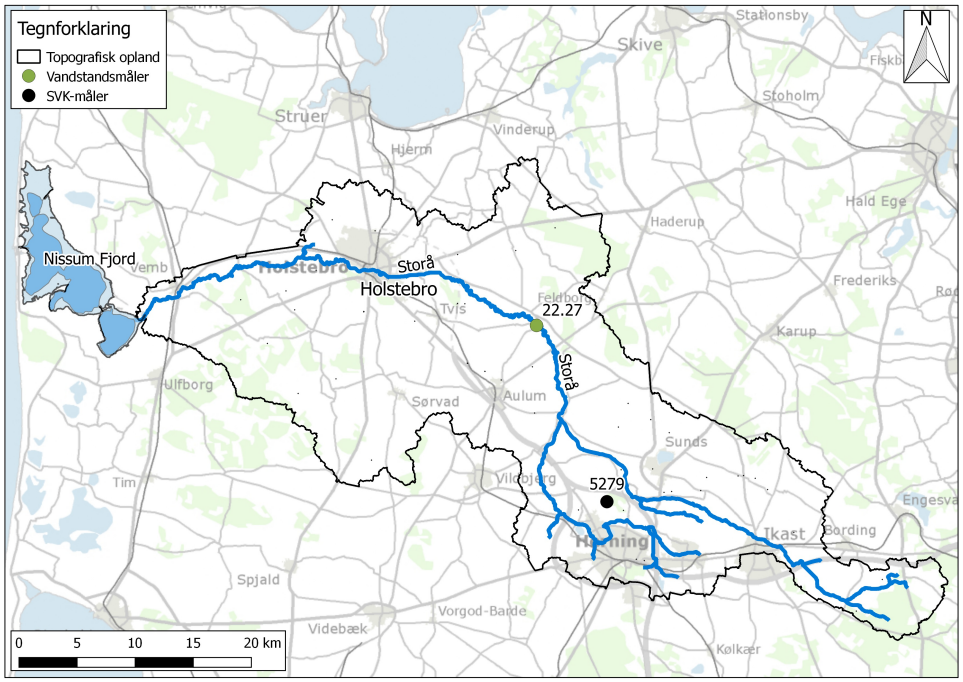
Sikring af bedst mulige forudsætninger for anvendelse af neurale netværk i fremtiden

Igennem projektarbejdet er det erfaret, at data, som har været yderst relevant for beskrivelsen af vandstande i Vejle Midtby, ikke har været tilgængelig i tilstrækkelig mængde og kvalitet. Hvis det i fremtiden ønskes at anvende neurale netværk til modellering, anbefales det derfor at der gøres en ekstra indsats for at forbedre datagrundlaget. Dette vurderes at kunne ske på følgende måder:

- Opsætning af flere målere, eksempelvis regnmålere som vil kunne anvendes til at forbedre kalibrering af radaren, hvilket øger kvaliteten af nedbørsforudsigelsen.
- Bedre datavalidering og udbedring af fejl på målere vil sikre, at flest mulige perioder er tilgængelige at træne de neurale netværk på.
- Arbejde på at skabe flest mulige situationer/kombinationer for netværkene at træne sig ind på. I den specifikke case omhandlende Vejle omfatter dette en løbende test af forskellige spjældindstillinger ved både lave og høje vandstande.

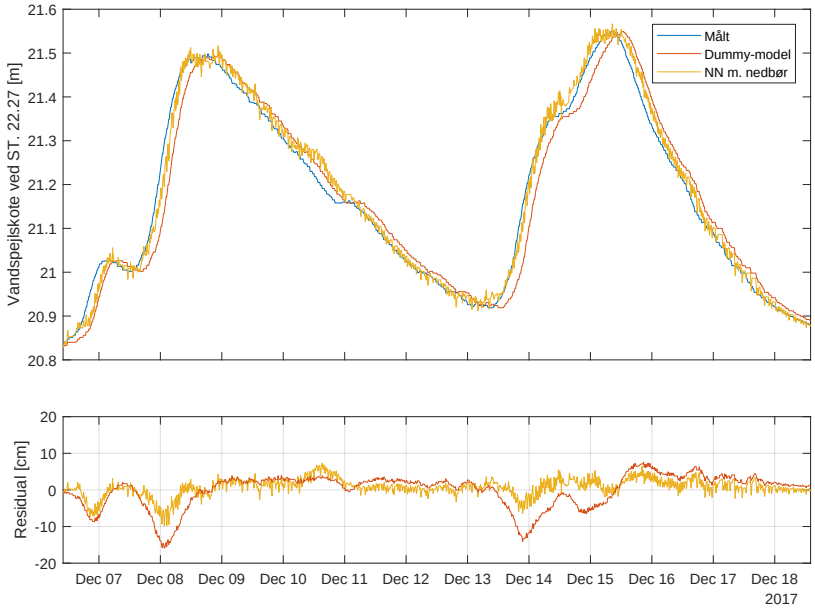
Implementering på en anden lokalitet

I projektets indledning er det vist, hvordan det uden videre er muligt at modellere døgnmiddelvandstanden i Storå, som løber gennem Holstebro. Igennem dette projekt er vandstanden ved Grejsdalens Planteskole prognosticeret med en forudsigelsestid på 4 timer. Det forsøges at anvende den bedste model til formålet på data fra Storå for at illustrere modellens simple implementering og anvendelighed. På figur 9.1 vises Storå med dertilhørende topografiske opland.



Figur 9.1: På kortet vises Storå, opland og vandstandsmåleren, som det ønskes at modellere, og den SVK-måler, der anvendes som input.

Det formodes, at vandstanden ved station 22.27 i Storå ved Rotvig Bro domineres af afstrømningsforhold, der ligner forholdene ved Grejsdalens Planteskole. Derfor er der anvendt et neuralt netværk med input i form af tendens-modellen og nedbør målt i SVK-måleren 5279 til modellering af vandspejlskoten ved Rotvig bro. På figur 9.2 vises et eksempel fra en testperiode, hvor der er to på hinanden følgende afstrømningshændelser.



Figur 9.2: Det bedste netværk fra dette projekt kan uden videre implementeres på en anden lokation, og give en tydeligt bedre forudsigelse af vandstanden om 4 timer end dummy-modellen.

Det neurale netværk giver en forudsigelse af vandspejlskoten ved en forudsigelsestid på 4 timer, som er væsentlig bedre end dummy-modellen, hvilket indikerer, at erfaringerne gjort i nærværende projekt simpelt kan implementeres ved andre lokaliteter.

Andre typer af neurale netværk

I projektet er der anvendt den mest simple form for neurale netværk i form af et feedforward-netværk med ét enkelt skjult lag. Dette er bevidst gjort for at have fokus indenfor fagområdet "Vand og Miljø". Der findes andre variationer af neurale netværk, som har forskellige styrker og svagheder. Ved at tilføje flere skjulte lag kan netværket kaldes et *Deep Feedforward-netværk*. Netværket får på denne måde flere frihedsgrader og mulighed for at beskrive andre og mere komplekse sammenhæng mellem data end et feedforward-netværk med kun et enkelt skjult lag. Et andet eksempel på et neuralt netværk, som potentielt kunne være relevant til anvendelse på tidsserier, og forudsigelser af vandstande, er et *Recurrent Neural Network (RNN)*. Her sendes output tilbage og anvendes som input. Der findes ligeledes mange forskellige typer RNN, som også har både styrker og svagheder. Der formodes derfor at være en god sandsynlighed for at finde andre typer af neurale netværk, som potentielt kunne give en bedre forudsigelse af vandspejlskoterne.

weather radar test facility.

5.3.1 The exponential LAWR data calibration approach

The applied DRO-R relation in Paper IV (Nielsen et al., 2013b) is purely empirically similar to previous LAWR data calibration efforts. Closer inspection of the data revealed that the validity of the linear assumption was doubtful, as illustrated in figure 5.2. If the radar and gauge data are accumulated to rainfall event values, the linear assumption seems plausible. However, if the data is not integrated over time, a nonlinear relation appears.

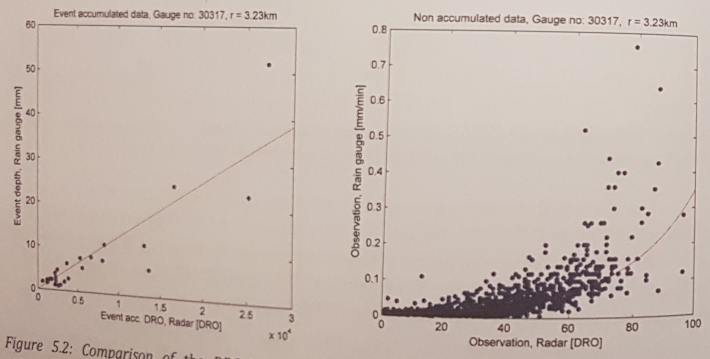


Figure 5.2: Comparison of the DRO-R relationship [illegible] accumulated data (five-minute rainfall intensities) (Nielsen et al., 2013b).

The nonlinear empiric[illegible] exponential relation presented o[illegible] eq. 5.7.

$$b(r) = C_{1b} \cdot \exp(r \cdot C_{2b}) \qquad \text{eq. 5.9}$$

Generally, the exponential LAWR data calibration approach improved the LAWR performance compared to the linear regression approach. The linear regression approach had the tendency to overestimate periods with low-intense precipitation and underestimate periods with moderate or more severe rainfall. The switch between under and overestimation occurred around a mean three-hour intensity of 3.33 mm/hr (10 mm in three hours). The exponential LAWR data calibration resulted in more accurate LAWR QPE most of the time, but it still had difficulties in estimating the more severe rainfall. The most significant improvement was gained for the light rainfall (less than 10 mm in three hours), because the exponential method removed the systematic overestimation as illustrated in figure 5.3. However, as concluded in Paper IV (Nielsen et al., 2013b), the proposed DRO-R relation and LAWR data calibration method had similar difficulties in estimating the more heavy rainfall.

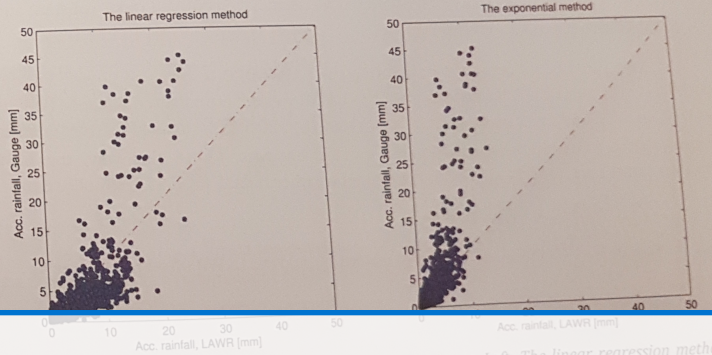


Figure 5.3: Scatter plots of accumulated rainfall over three hours. Left: The linear regression method. Right: The exponential method. (Nielsen et al., 2013b)

Litteraturliste

Achton-Boel, A., B. Refsgaard, C. B. Andersen, K. Jess og T. A. Jensen (2017). *Storå og Nissum Fjord - Klimasikring af Holstebro*. Tek. rap. Institut for Byggeri og Anlæg.

Benzer, R. (2015). „Population dynamics forecasting using artificial neural networks“. I: 24, s. 460–466.

Cappelen, J. (1993). *Kvalitetsmarkering af automatiske nedbørsregistreringer*. Tek. rap. Danish Meteorological Institute.

Chang, F.-J., P.-A. Chen, Y.-R. Lu, E. Huang og K.-Y. Chang (2014). „Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control“. I: *Journal of Hydrology* 517, s. 836–846. URL: http://www.sciencedirect.com/science/article/pii/S0022169414004739.

Dansk Meteorologisk Institut (2018). *Vandstand i Vejle Havn*. URL: http://www.dmi.dk/hav/maalinger/vandstand/.

Faarup, C. (2016). „Nyt sluseanlæg skal beskytte Vejles midtby mod stigende vandmasser“. I: *MIKROPOL*.

Karlby, H., I. Sørensen, L. Ramsay, J. O. Nielsen, M. Jørgensen, W. Bai, A. H. Kristensen, E. Karlsen, E. D. Jensen, K. Brinck, P. Roslev, S. Schmidt, K. Ellsemose og P. Jacobsen (2014). *Vandforsyning*. 3. udgave. Praxis Nyt Teknisk Forlag.

Kasiviswanathan, K., J. He, K. Sudheer og J.-H. Tay (2016). „Potential application of wavelet neural network ensemble to forecast streamflow for flood management“. I: *Journal of Hydrology* 536, s. 161–173. URL: http://www.sciencedirect.com/science/article/pii/S0022169416300786.

Kortforsyningen (2016). *Kortforsyningen*. URL: https://download.kortforsyningen.dk/.

Larsen, T. (2016). *Kronik: Har vi styr på vandføringen i de danske vandløb?* URL: https://ing.dk/artikel/kronik-har-vi-styr-paa-vandfoeringen-de-danske-vandloeb-189143.

Levenberg, K. (1944). „A method for the solution of certain non-linear problems in least squares". I: *Quarterly of Applied Mathematics* 2.2, s. 164–168. URL: http://www.jstor.org/stable/43633451.

Liaison Technologies (2017). *The Importance of Data Quality in Data-Driven Decision Making*. URL: https://www.liaison.com/blog/2017/06/29/importance-data-quality-data-driven-decision-making/.

LIFA A/S Landinspektører (2014). *Tværprofil Omløbsåen*.

Marquardt, D. W. (1963). „An Algorithm for Least-Squares Estimation of Nonlinear Parameters". I: *Journal of the Society for Industrial and Applied Mathematics* 11.2, s. 431–441. URL: https://doi.org/10.1137/0111030.

MathWorks (2017). *MATLAB Neural Net Fitting Toolbox*.

MATLAB (2006). *Levenberg-Marquardt backpropagation*. URL: https://se.mathworks.com/help/nnet/ref/trainlm.html#f8-671992.

McCulloch, W. S. og W. Pitts (1943). „A logical calculus of the ideas immanent in nervous activity". I: *The bulletin of mathematical biophysics* 5.4, s. 115–133. URL: https://doi.org/10.1007/BF02478259.

Miljøstyrelsen (2018). „Udvikling af automatiseret realtids varslings- og styringssystem af Vejle Å, Grejs Å og Omløbsåen til klimasikring af Vejle By". Ikke-publiceret udkast til slutrapport af MUDP projekt J.nr. NST-404-00268.

Naturstyrelsen (2006). *Natura 2000 basisanalyse - Øvre Grejs Ådal*. Tek. rap. Naturstyrelsen.

Naturstyrelsen og Kystdirektoratet (2011). *Endelig udpegning af risikoområder for oversvømmelse fra vandløb, søer, havet og fjorde*. Tek. rap. Naturstyrelsen og Kystdirektoratet.

Okkan, U. (2011). „Application of Levenberg-Marquardt Optimization Algorithm Based Multilayer Neural Networks for Hydrological Time Series Modeling". I: *An International Journal of Optimization and Control: Theories Applications (IJOCTA)*.

Rasmussen, M. M. og K. S. Schneidelbach (2016). „Hydraulisk modellering og modelprædiktiv styring med Neurale Netværk". Speciale. Aalborg Universitet.

Rasmussen, M. R. (1992). *Neurale netværk - Anvendelse af neurale netværk inden for det miljøtekniske område*. Tek. rap.

Refsgaard, J. C., C. Kern-Hansen, F. Plauborg, N. B. Ovesen og P. Rasmussen (2003). *Ferskvandets kredsløb - Kapitel 6. Ferksvandets kredsløb og tidslige variationer*. Tek. rap.

Rojas, R. (1996). *Neural Networks: A Systematic Introduction*. New York, NY, USA: Springer-Verlag New York, Inc.

Samuel, A. L. (1959). „Some Studies in Machine Learning Using the Game of Checkers". I: *IBM J. Res. Dev.* 3.3, s. 210–229. URL: http://dx.doi.org/10.1147/rd.33.0210.

Singh, P. og P. Dwivedi (2018). „Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem". I: *Applied Energy* 217, s. 537–549. URL: http://www.sciencedirect.com/science/article/pii/S0306261918302654.

Straeter, T. A. (1971). *On the Extension of the Davidon-Broyden Class of Rank One, Quasi-Newton Minimization Methods to an Infinite Dimensional Hilbert Space with Applications to Optimal Control Problems*. Tek. rap. Department of Mathematics, Raleigh.

Trap, J. P. (1904). *Kongeriget Danmark, bind 5*. third. Gad.

Vejle Amts Folkeblad (2017). *Video: Slusen svigtede - Aagade var fyldt med vand*. URL: https://vafo.dk/vejle/Video-Slusen-svigtede-AAgade-var-fyldt-med-vand/artikel/492348.

Vejle Byråd (2014). *Risikostyringsplan for oversvømmelse*. Tek. rap. Vejle Kommune.

Vejle Kommune (2014). *Klimatilpasningsplan for Vejle Kommune*. Tek. rap. Vejle Kommune.

— (2015a). *Fuld fart på anlægsarbejdet med ny faunapassage i Grejs Å ved Abelones Plads*. URL: http://apps.vocast.com/r/pressframe/g7q2e5/view/02b3b02da7773d74.

— (2015b). *Sluse- og pumpeanlæg ved Omløbsåen*. URL: https://www.vejle.dk/om-kommunen/fakta-om-os/planer-politikker-og-strategier/.

Vejle Spildevand A/S (2017). *Nyt prognoseværktøj kan forudsige oversvømmelser i Vejle*. URL: https://www.vejlespildevand.dk/nyheder/nyt-prognosevaerktoej-kan-forudsige-oversvoemmelser-i-vejle.

Vejle Sportsfiskerforening (2011). *Grejs Å, Zoneoversigt*. URL: http://www.vsf-vejle.dk/Files/Kortmateriale/Grejs/Grejs_Oversigt_1040x0750.jpg.

Vejle Stadsarkiv (2018). *Vejle Stadsarkiv*. Tek. rap. Vejle Stadsarkiv.

www.112alarm.dk (2015). *Oversvømmelse i Vejle*. URL: http://www.112alarm.dk/oversvoemmelse-i-vejle/.

V

Appendices

A Udtræk af radardata 109

B Skift mellem sommer- og vintertid . . . 115

C Forsinkelse mellem vandspejlskoter . 117

D Eksempel på gengivelse af sinuskurve 123

E Opsætning af neurale netværk i MATLAB 131

F Forudsigelse på baggrund af kurveforløb 149

G Modelresultater 153

H NAR - et alternativ til tendens-modellen 167

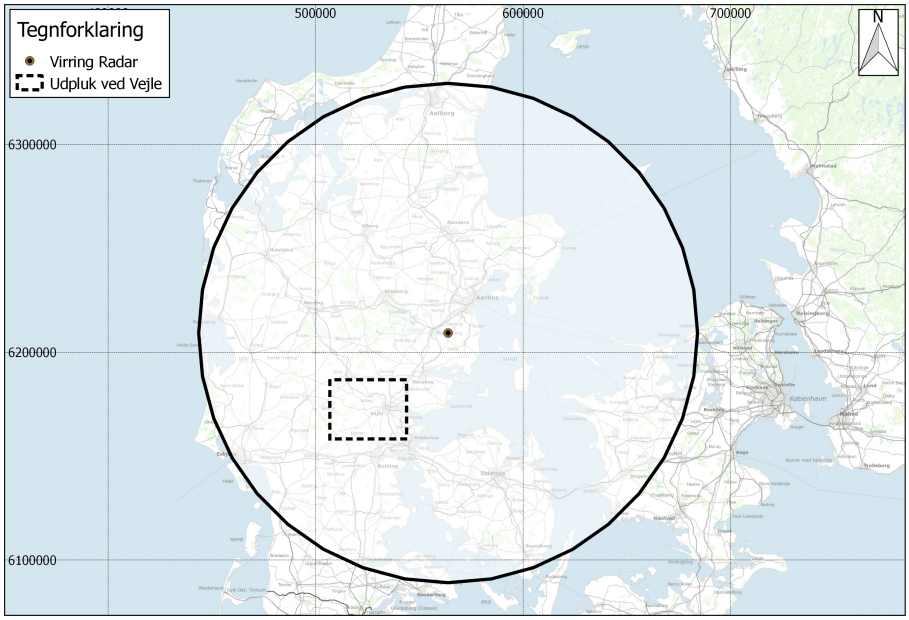
I Radardata over Vejle Midtby 169

J Tolkning af data fra fordelerbygværket 173

A. Udtræk af radardata

Formålet med dette appendiks er at beskrive fremgangsmåden for udtræk af radardata i et specifikt område ved brug af GIS-værktøjer og Matlab.

Der udplukkes signaler fra Virring radaren i et område ved Vejle, som vist på figur A.1.

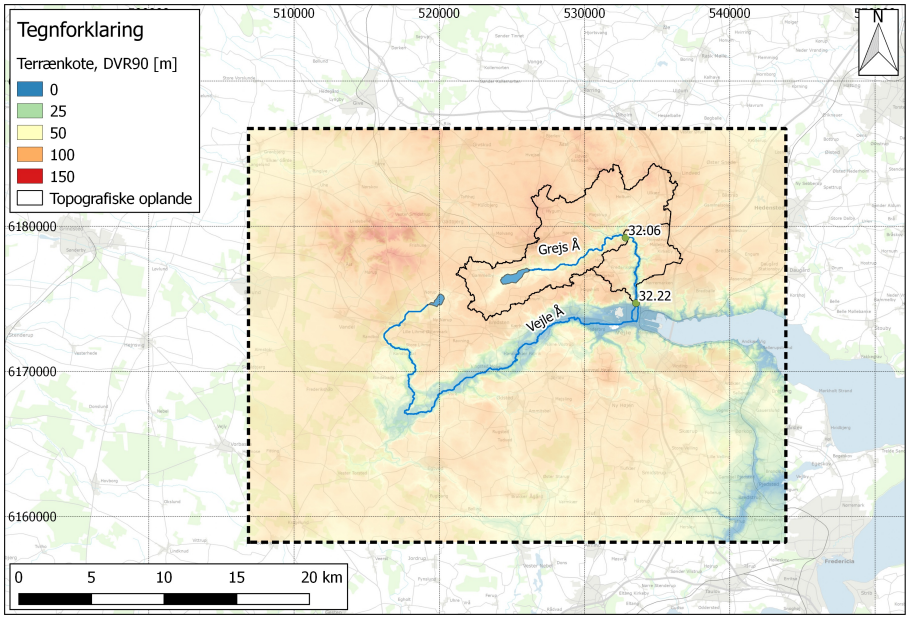


Figur A.1: Radaren i Virring dækker et stort område af landet, men det er kun i området ved Vejle, hvor signalerne er brugbare.

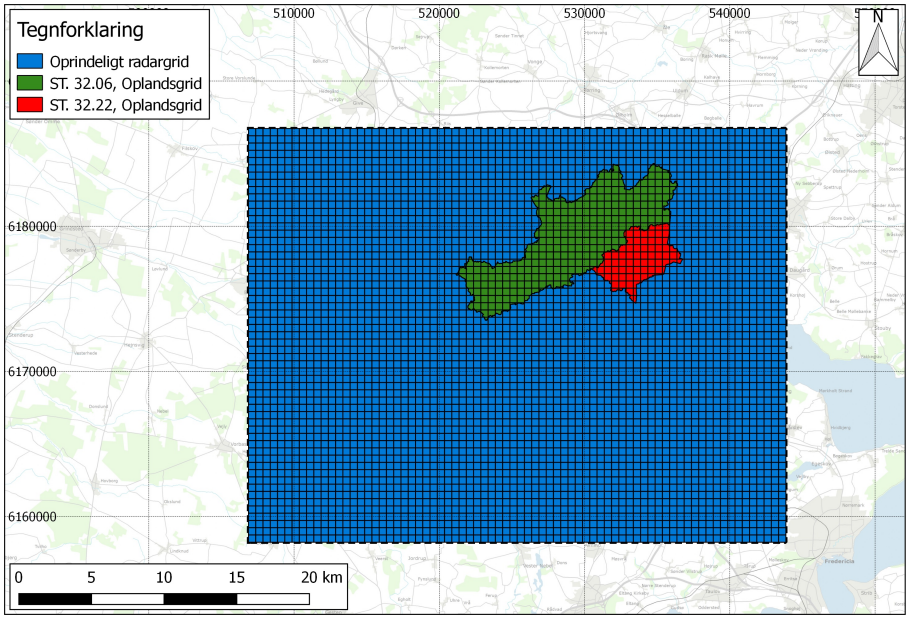
Hele bevæggrunden for at anvendes radarsignalerne er, at der ses voldsomme afstrømninger i Grejs Å ved Ny Grejsdalsvej (ST. 32.22), hvor der ikke er målt nogen nedbør i regnmålerne.

Ud fra terrænkort kan det topografiske opland til Grejs Å bestemmes, og det anvendes til at udvælge

celler i radargriddet, som vist på figur A.2 og figur A.3.



Figur A.2: Det topografiske opland til Grejs Å er bestemt ved en GIS-analyse ud fra et terrænkort med en gridstørrelse på 1,6 m.



Figur A.3: Cellerne i det oprindelige radargrid beskæres med det topografiske opland. Gridstørrelsen er 500x500 meter.

Radardata haves i .df2-fil formatet, og fremgangsmåden for midlingen i oplandet ses af nedenstående.

1. Åbn .dfs2-fil i Mike Zero og klik på *Edit*-fanen. Her findes *Map Projection Coordinates*, hvor koordinaterne til midten af cellen i nederste venstre hjørne står i koordinatsystemet UTM zone 32.

2. Aflæs *Grid spacing* og antal rækker/søjler på illustrationen til venstre. I dette tilfælde $dx = 500\,\text{m}$ og $dy = 500\,\text{m}$.
3. Opret et vektorgitter i QGIS, der ligger oven i dataene i .dfs2-filen. Klik på fanen *Vektor*, find *Forskningsværktøjer* og åbn *Vektorgitter*. Her skal *Gitteromridset* (x_{min}, x_{max}, y_{min}, y_{max}) og cellestørrelsen defineres. QGIS regner fra nederste venstre hjørne af cellen i nederste venstre hjørne, så $dx/2$ og $dy/2$ skal trækkes fra Map Projection Coordinates fra .dfs2-filen.
4. Opret et nyt vektorgitter ved at beskære vektorgitteret efter det topografiske opland og gem begge gitre som .shp-filer.

```
clear all; close all; clc;

% Indlaeser vektorgitteret og attributtabellen, der passer paa .dfs2-filen
[RadarGrid,RadarAtt] = shaperead('RadarGrid.shp');
RadarAtt = struct2cell(RadarAtt);
RadarAtt = cell2mat(RadarAtt)';

xmin = RadarAtt(:,2);
ymin = RadarAtt(:,4);

% OBS: Der er byttet rundt paa raekker og soejler, saa soejle=i og raekke=j
i = unique(xmin);
numcol = [1:length(i)]';
i = [numcol i];

j = unique(ymin);
numrow = [1:length(j)]';
j = [numrow j];
j(:,1) = flip(j(:,1));

l = length(j);
w = length(i);

for z=1:length(RadarAtt)
    for q=1:length(i)
        if RadarAtt(z,2)==i(q,2)
            RadarAtt(z,6)=i(q,1);
        end
    end
    for g=1:length(j)
        if RadarAtt(z,4)==j(g,2)
            RadarAtt(z,7)=j(g,1);
        end
    end
end

% Allokerer en matrix, der skal bruges til at beskrive oplandet
PlanteskoleOpland = zeros(l,w);

% Indlaeser vektorgitteret og attributtabellen, der er beskaaret til det
    topografiske opland
```

```
[PlanteskoleGrid, PlanteskoleAtt] = shape('PlanteskoleGrid.shp');
PlanteskoleAtt = struct2cell(PlanteskoleAtt);

% Beskaarede vektorgitre kan nogengange opfoere sig sjovt i matlab. Det er
    altid cell-arrays i stedet for struct, men i cell-arrayet kan det baade
     vaere characters eller tal. Foelgende soerger for, at det indlaeste
    altid kan behandles.
if iscell(PlanteskoleAtt(1,1))==1
    if ischar(cell2mat(PlanteskoleAtt(1,1)))==1;
        for i=1:size(PlanteskoleAtt,2)
            PlanteskoleAtt(1,i) = num2cell(str2num(cell2mat(PlanteskoleAtt
    (1,i))));
        end
        PlanteskoleAtt = cell2mat(PlanteskoleAtt)';
    elseif isnumeric(cell2mat(PlanteskoleAtt(1,1)))==1;
        PlanteskoleAtt = cell2mat(PlanteskoleAtt)';
    end
end

% Finder sammenfald i attributter i de to gitre
for z=1:length(RadarAtt)
    for h=1:length(PlanteskoleAtt)
        if RadarAtt(z,1)==PlanteskoleAtt(h,1)
            RadarAtt(z,8)=1;
        end
    end
end

% Laver 1-taller i de sammenfaldende celler. Resten er stadig 0
for z=1:length(RadarAtt)
    if RadarAtt(z,8)==1
      PlanteskoleOpland(RadarAtt(z,7),RadarAtt(z,6)) = 1;
    end
end

RadarAtt(:,8) = [];

% Tjekker koordinaterne en sidste gang for en sikkerhedsskylde
TopLeft = [RadarAtt(1,2) RadarAtt(1,5)];

% Gemmer
save('PlanteskoleOpland.mat','PlanteskoleOpland')
```

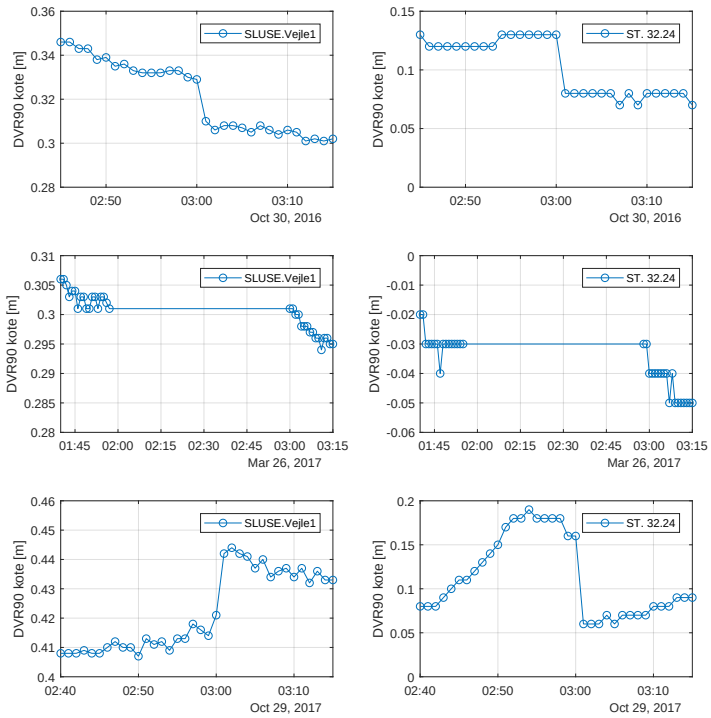
```
clear all; close all; clc;

% Goer det muligt at importere .dfs2-filer i matlab. DHI-toolbox skal
    installeres foerst
NET.addAssembly('DHI.Generic.MikeZero.DFS');
import DHI.Generic.MikeZero.DFS.*;
```

```
% Indlaeser oplandsmatricen fra forrige kode
loadstr = 'PlanteskoleOpland.mat';
load(loadstr);

% Definerer start- og sluttidspunkt
startdate = datenum('01-Feb-2015 00:00');
lastdate = datenum('28-Feb-2018 23:59');

% Gaar ind og henter de rigtige .dfs2-filer ved at aendre stien ud fra
    datoen.
for z=startdate:lastdate;
    time = z;

    year = datestr(time,'yyyy');
    month = datestr(time,'mm');
    day = datestr(time,'dd');

    infile = strcat('dfs2Archive/dfs2Archive/Virring/',year,'/',month,'/','
    dkvir',day,'.dfs2');

    if (~exist(infile,'file'))
      [filename,filepath] = uigetfile('*.dfs2','Select the .dfs2');
      infile = [filepath,filename];
    end

    dfs2 = DfsFileFactory.Dfs2FileOpen(infile);

    saxis = dfs2.SpatialAxis;
    x = saxis.X0 + saxis.Dx*(0.5+(0:(saxis.XCount-1)))';
    y = saxis.Y0 + saxis.Dy*(0.5+(0:(saxis.YCount-1)))';

    itemname = char(dfs2.ItemInfo.Item(0).Name);
    itemunit = char(dfs2.ItemInfo.Item(0).Quantity.UnitAbbreviation);
    nsteps   = dfs2.FileInfo.TimeAxis.NumberOfTimeSteps;

    deleteval = double(dfs2.FileInfo.DeleteValueFloat);

    col = dfs2.SpatialAxis.XCount;
    row = dfs2.SpatialAxis.YCount;

    cells_Opland = sum(sum(PlanteskoleOpland)');
    int = zeros(row,col,1440);

    for i=0:nsteps-1
      h = double(dfs2.ReadItemTimeStep(1,i).To2DArray())';

      I = find(h == deleteval);
      h(I) = NaN;
```

```
      h = flipud(h);    % Dataene skal flippes, da raekkenummerering i Mike
     Zero er omvendt fra matlab.
      int(:,:,i+1) = h.*(1000/60/60);        % Omskriver intensiteterne til
     mm/min
      int(:,:,i+1) = int(:,:,i+1).*PlanteskoleOpland;        % Soerger for
    kun at fange signaler inde i oplandet

      Precip(i+1,z-startdate+1) = sum(sum(int(:,:,i+1))')/cells_Opland;
     % Midler signalet over antallet af celler i oplandet
    end
    dfs2.Close();
end

Precip = Precip(:);
t = [startdate:1/24/60:lastdate]';

Radar_Planteskole = [t Precip];

save('Radar_Planteskole.mat','Radar_Planteskole')
```

B. Skift mellem sommer- og vintertid

Formålet med følgende appendiks er at illustrere skift mellem sommer- og vintertid, så det er muligt at korrigere, da fejlen eller bliver trænet ind i det neurale netværk



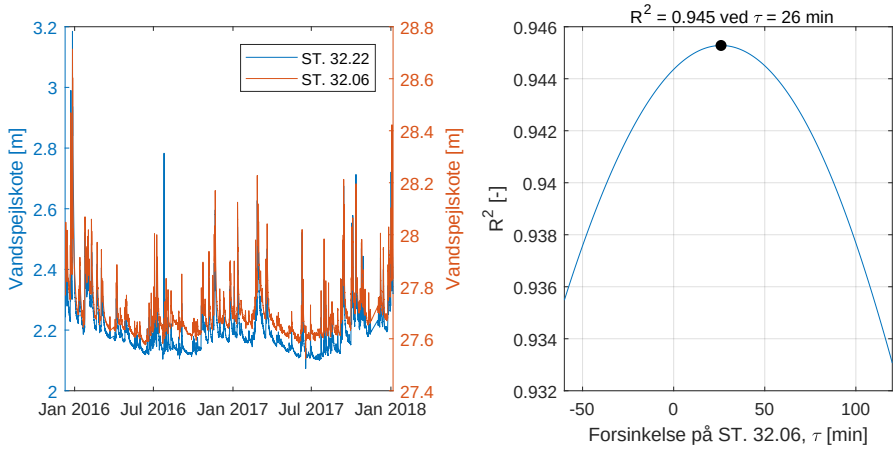
Figur B.1: Der ses tydelige skift mellem sommer- og vintertid i måledata.

Måleperioden er der fire skift mellem sommer- og normaltid. Dette er 27/03-2016 02:00, 30/10-2016 03:00, 26/03-2017 02:00 og 29/10-2017 03:00. Der er dog afbrud i måleserierne på samtlige vandstandsdata den 27. marts. Data fra Vejle Spildevand (data omhandlende spjæld og sluseanlæg) og data fra DMI (vandstand i Vejle Havn) er vurderet at inkludere sommertid. Skiftene mellem sommer- og normaltid for henholdsvis data fra sluseanlægget og data fra Vejle Havn er vist på figur B.1.

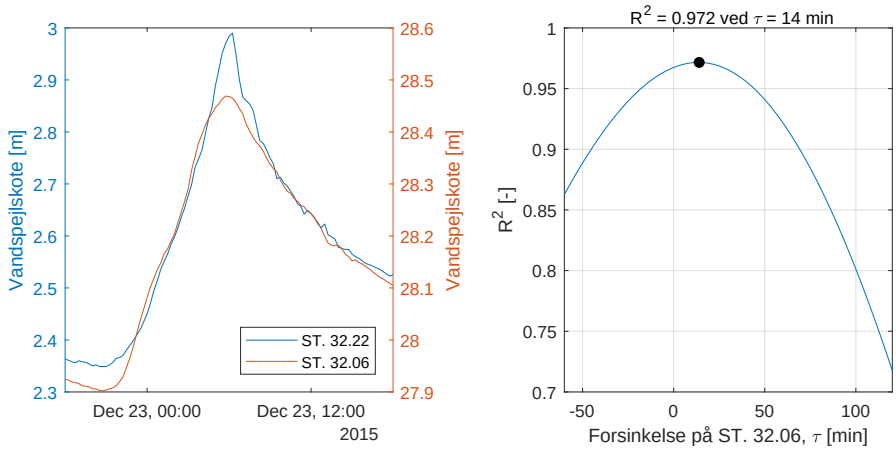
C. Forsinkelse mellem vandspejlskoter

Dette appendiks indeholder resultater fra korrelationsanalysen mellem vandspejlskoten ved Grejsdalens Planteskole og Ny Grejsdalsvej.

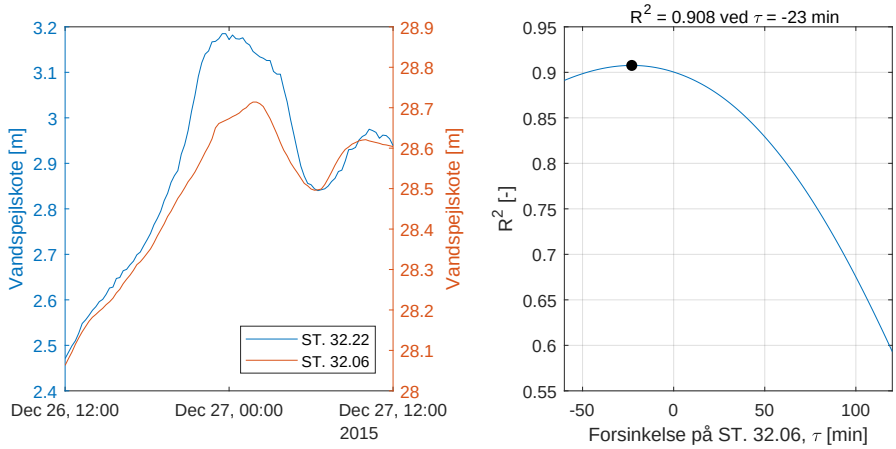
I figurerne i dette appendiks er der til venstre vist vandspejlskoten ved de to stationer uden nogen forskydning og til højre er vist R^2-værdien for den lineære korrelation mellem de to stationer ved forskellige forsinkelsestider på målingerne ved Grejsdalens Planteskole (ST. 32.06). På figur C.1 er resultatet af analysen på hele datasættet vist. De resterende figurer, C.2 til C.11 illustrerer de 10 hændelser, hvor vandspejlskoten har været højest ved Grejsgdalens Planteskole.



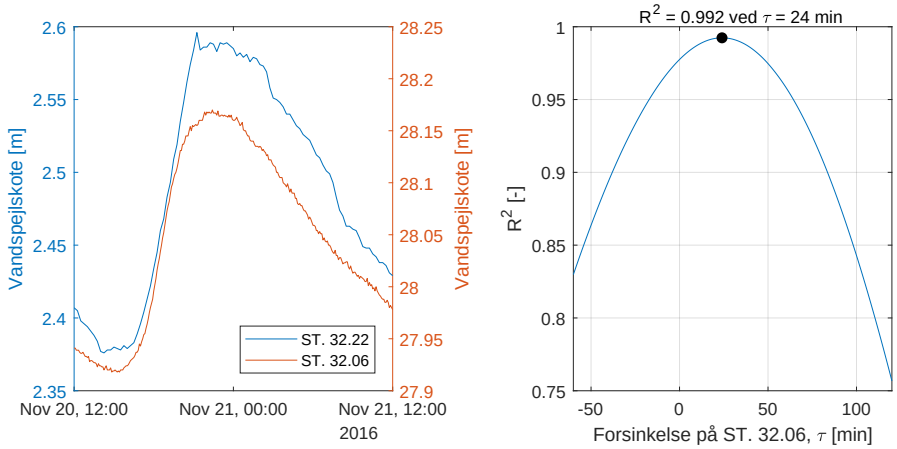
Figur C.1: Den bedste lineære korrelation mellem vandspejlskoten ved Grejsdalens Planteskole og Ny Grejsdalsvej forekommer, hvis målingerne ved Grejsdalens Planteskole forsinkes med 26 minutter.



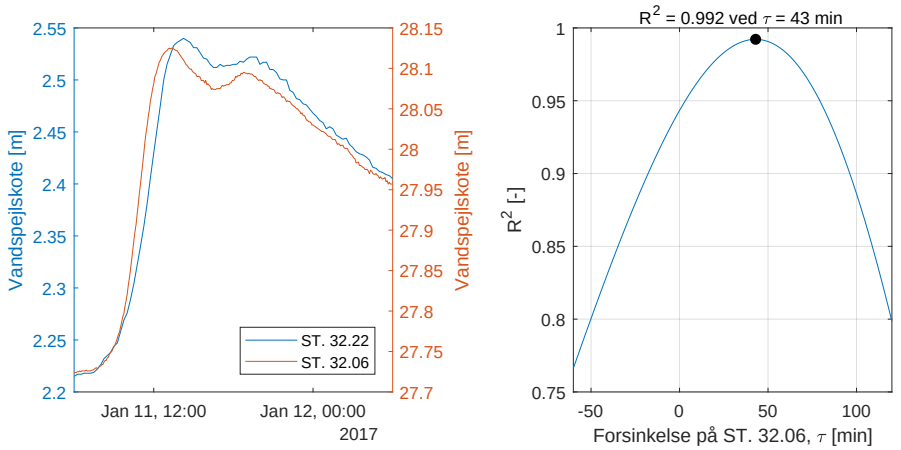
Figur C.2: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på 14 minutter.



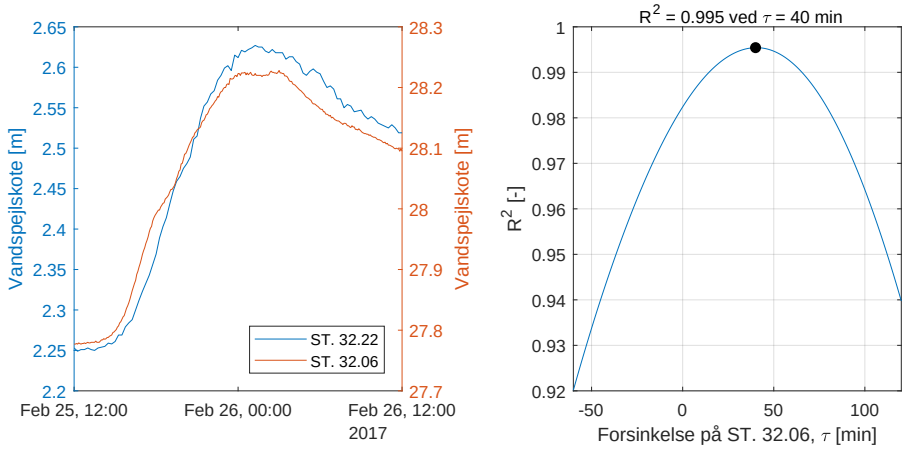
Figur C.3: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på -23 minutter.



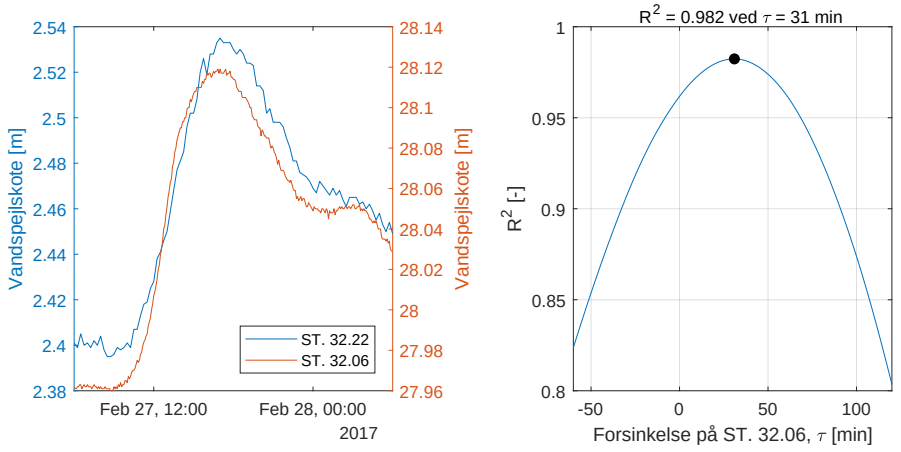
Figur C.4: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på 23 minutter.



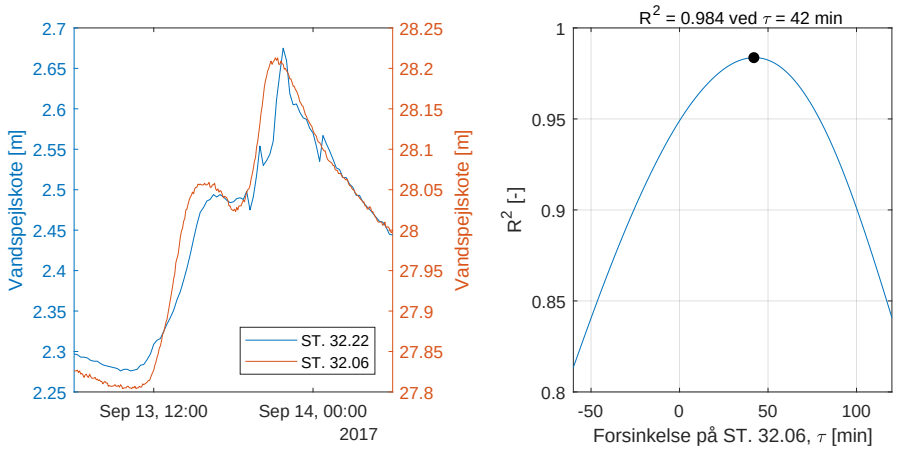
Figur C.5: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på 43 minutter.



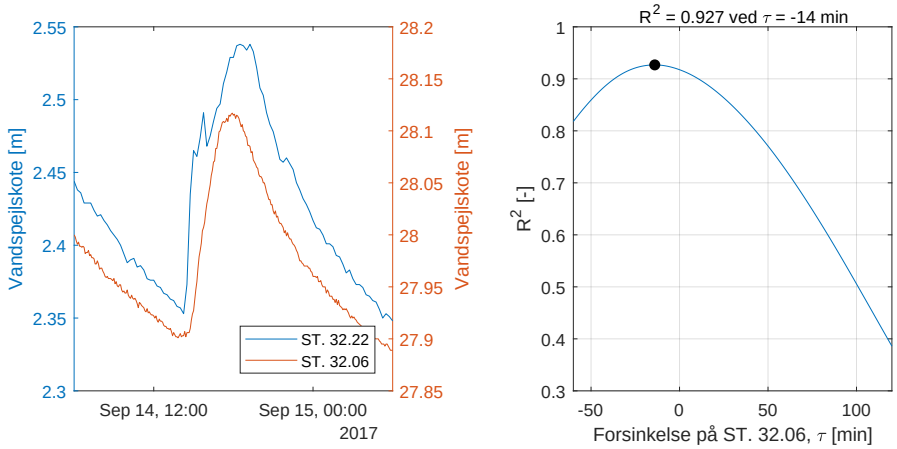
Figur C.6: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på 40 minutter.



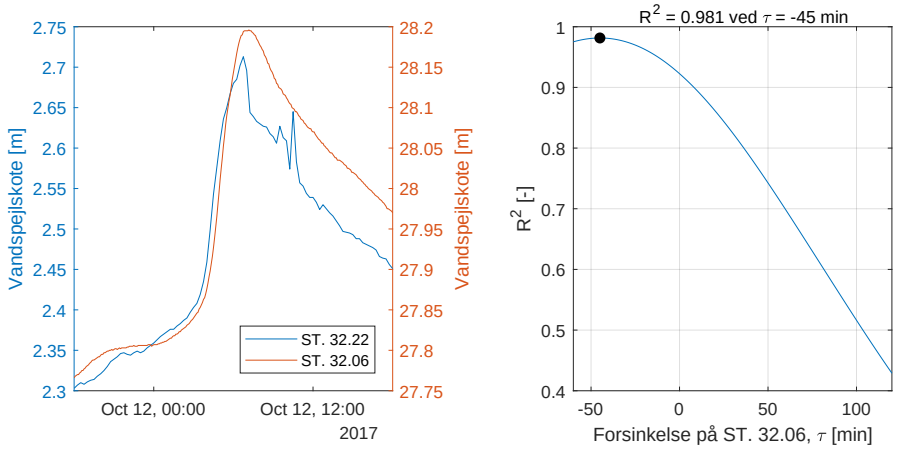
Figur C.7: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på 31 minutter.



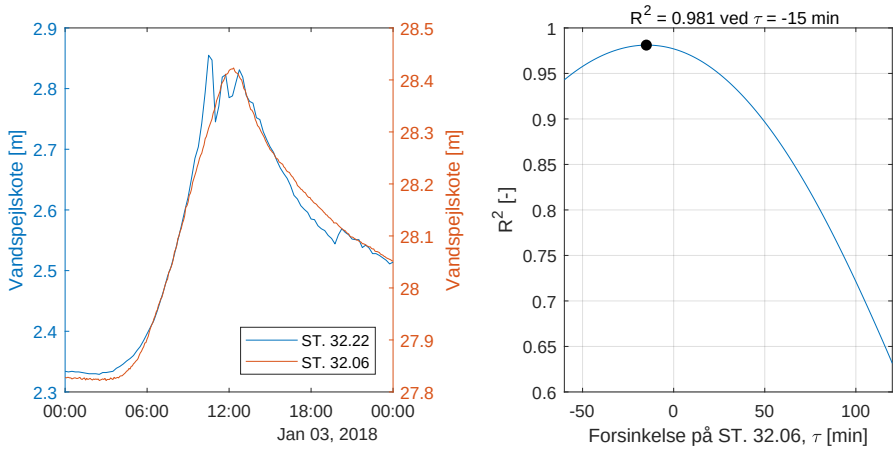
Figur C.8: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på 42 minutter.



Figur C.9: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på -14 minutter.



Figur C.10: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på -45 minutter.



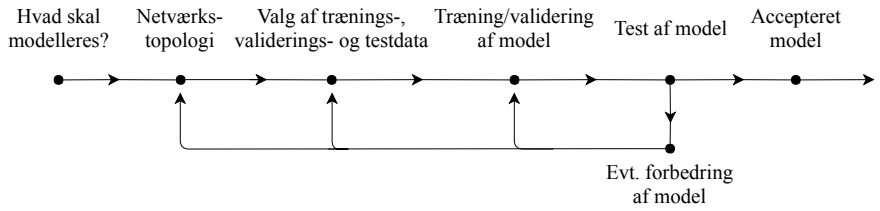
Figur C.11: Den bedste lineære korrelation forekommer for denne hændelse ved en forsinkelse af ST. 32.06 på -15 minutter.

D. Eksempel på gengivelse af sinuskurve

Dette appendiks omhandler de bagvedliggende processer ved modellering af sinus-afhængige data med neurale netværk.

Fremgangsmåde

Fremgangsmåden for modellering med neurale netværk er vist på figur D.1. Figuren er baseres på et overordnet billede af informationerne fundet i bogen *"Neural Networks - A Systematic Introduction"*. (Rojas 1996)



Figur D.1: Fremgangsmåden for modellering med neurale netværk.

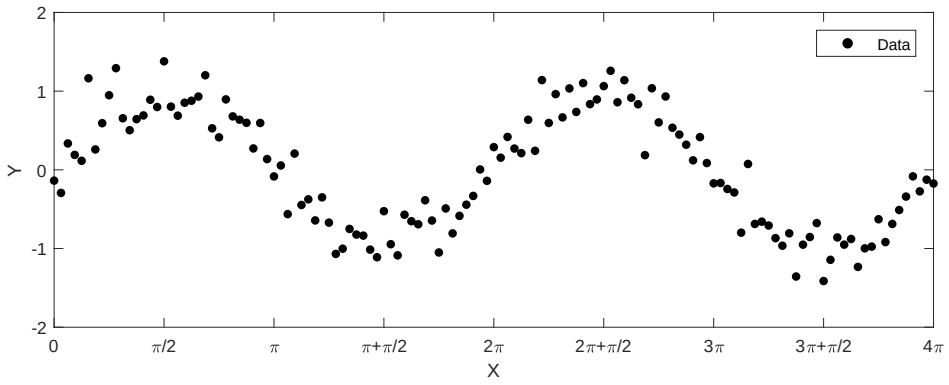
Der er fokus på på især matematikken bag indtræningen af det neurale netværk, og eksemplet er gennemarbejdet i Matlabs *Neural Net Fitting Tool* (MathWorks 2017). Det er ikke et fokuspunkt at præsentere modelresultatet.

D.0.1 Hvad skal modelleres?

Det ønskes at gengive data, som er genereret på følgende måde og illustreret på figur D.2. Det er disse punkter, som i forbindelse med modellering med neurale netværk kaldes ***target***.

```
clc; clear all; close all;
tic
```

```
% Lav tilfaeldig data
N = 3000; % Antal elementer

dx = 32; % Antallet af datapunkter per udsving

for i=0:N-1
    X(i+1,1) = i;
    Y(i+1,1) = sin((pi/dx)*i)+randn(1,1)/4; % Tilfaeldige
end
```



Figur D.2: Datapunkterne viser en tydelig sinusafhængighed, som det forsøges at gengive med et neuralt netværk.

Af koden fremgår det, at der er 32 punkter for hvert udsving.

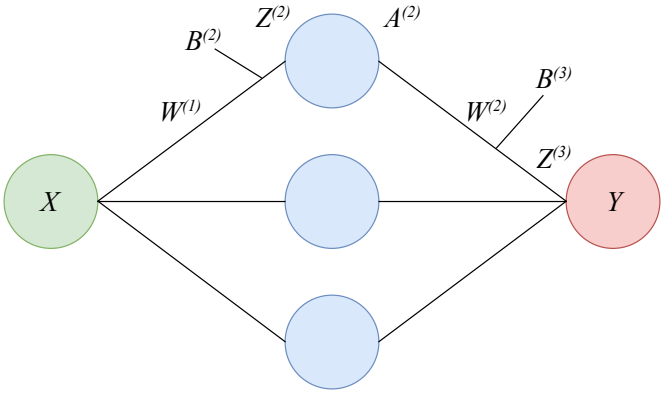
D.0.2 Netværksstruktur

Valg af input

Det ses, at der er tydelig harmonisk afhængighed mellem X og Y, hvorfor X vælges som input.

Antal neuroner i det skjulte lag

Der skal fastsættes et antal neuroner, som senere kan tilpasses, hvis der forekommer voldsom overfitting, eller hvis netværket ikke har et tilstrækkeligt antal frihedsgrader til at gengive data. Som udgangspunkt vælges tre neuroner, og netværksstrukturen ses af figur D.3

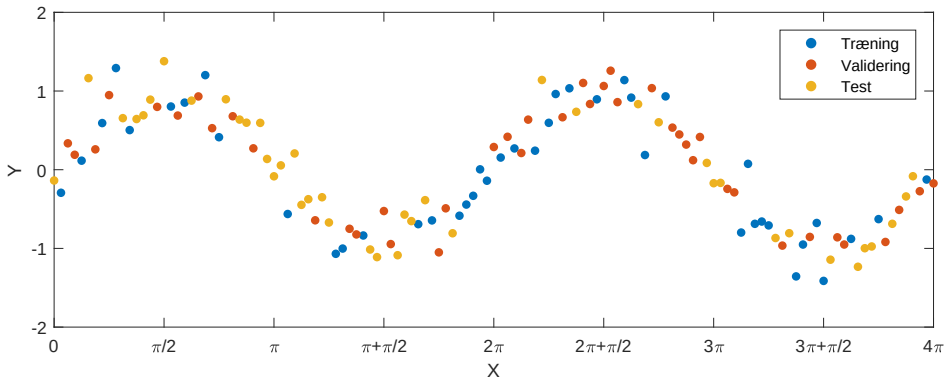


Figur D.3: Strukturen af det tænkte eksempel. Notationen forklares i det følgende.

Et netværk af denne type kaldes et ***feedforward***-netværk.

D.0.3 Valg af træning, validering og test

Af target udvælges tre dele: én til ***træning***, én til ***validering*** og én til ***test***. Træningen er den del af target, som det neurale netværk forsøger at indstille sine frihedsgrader efter. Valideringen er en uafhængig del af target, hvor kvaliteten af netværket afprøves ved hver iteration. Formålet med at validere løbende er at undgå overfitting. Test er ligesom valideringen en uafhængig del af target, men kvaliteten af netværket på denne del af data afprøves først, når frihedsgraderne er justeret helt på plads. Traditionelt udvælges de tre dele tilfældigt, hvilket også gøres i eksemplet, som vist på figur D.4.



Figur D.4: Af target udvælges tilfældige værdier til træning, validering og test, og der afsættes en tredjedel til hvert formål.

Der er afsat en tredjedel til hver del.

D.0.4 Valg af træningsalgoritme

Det vælges at anvende Levenberg-Marquardt træningsalgoritmen, hvor første skridt består af ***forward propagation***, hvorefter fejlen minimeres ved ***backpropagation***.

D.0.5 Træning af netværket

Levenberg-Marquardt træningsalgoritmen anvendes, og første skridt består af forward propagation.

Forward propagation

Forward propagation er en fremadrettet beregning fra input til output, og der skal derfor gættes på et sæt frihedsgrader. Der er tre synapser før og efter det skjulte lag, hvilket giver i alt seks frihedsgrader, samt nogle bias-justeringer som i princippet også er frihedsgrader. Der haves i alt 43 træningspunkter, så inputtet X er en søjlevektor med 43 indgange:

$$X = \begin{pmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{43,1} \end{pmatrix} \tag{D.1}$$

Hernæst introduceres vægtene på de tre synapser før det skjultelag, som rækkevektoren $W^{(1)}$:

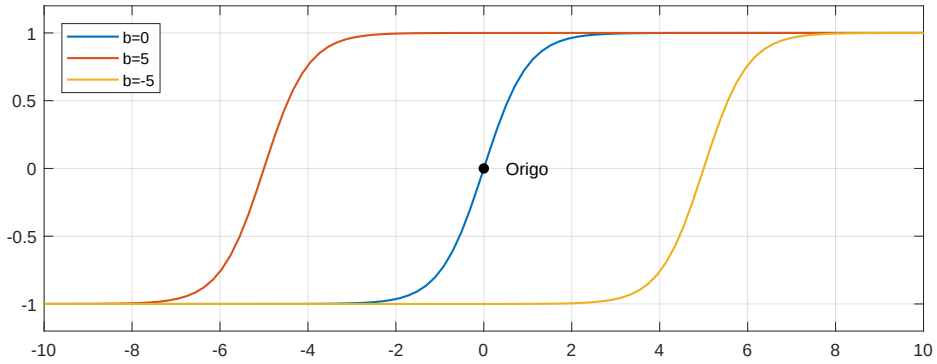
$$W^{(1)} = \begin{pmatrix} w_{1,1}^{(1)} & w_{1,3}^{(1)} & w_{1,3}^{(1)} \end{pmatrix} \tag{D.2}$$

Hvoraf det vægtede input $Z^{(2)}$ findes ved matrix-multiplication mellem X og W(1):

$$\begin{aligned} Z^{(2)} &= XW^{(1)} \\ &= \begin{pmatrix} x_{1,1} \\ x_{2,1} \\ \vdots \\ x_{43,1} \end{pmatrix} \begin{pmatrix} w_{1,1}^{(1)} & w_{1,3}^{(1)} & w_{1,3}^{(1)} \end{pmatrix} \\ &= \begin{pmatrix} x_{1,1}w_{1,1}^{(1)} & x_{1,1}w_{1,2}^{(1)} & x_{1,1}w_{1,3}^{(1)} \\ x_{2,1}w_{1,1}^{(1)} & x_{2,1}w_{1,2}^{(1)} & x_{2,1}w_{1,3}^{(1)} \\ \vdots & \vdots & \vdots \\ x_{43,1}w_{1,1}^{(1)} & x_{43,1}w_{1,2}^{(1)} & x_{43,1}w_{1,3}^{(1)} \end{pmatrix} \end{aligned} \tag{D.3}$$

Det vægtede input $Z^{(2)}$ adderes en bias $B^{(2)}$ for at give transferfunktionen en ekstra, uafhængig, konstant, der sørger for, at transferfunktionen ikke tvinger $Z^{(2)}$ gennem origo. Transferfunktionen er en hyperbolsk tangens (ligning D.4), og betydningen af bias er vist på figur D.5.

$$f(x)^{(2)} = \frac{2}{1+\exp(-2\cdot(x+B^{(2)})} - 1 \tag{D.4}$$



Figur D.5: Bias har en betydning for den horisontale forskydning af transferfunktionen (ligning D.4).

Transferfunktionen anvendes på alle indgange af det vægtede input $Z^{(2)}$, og der haves aktiveringsmatricen $A^{(2)}$:

$$\begin{aligned} A^{(2)} &= f(Z^{(2)} + B^{(2)}) \\ &= \begin{pmatrix} f(z^{(2)}_{1,1} + b^{(2)}_{1,1}) & f(z^{(2)}_{1,2} + b^{(2)}_{1,2}) & f(z^{(2)}_{1,3} + b^{(2)}_{1,3}) \\ f(z^{(2)}_{2,1} + b^{(2)}_{1,1}) & f(z^{(2)}_{2,2} + b^{(2)}_{1,2}) & f(z^{(2)}_{2,3} + b^{(2)}_{1,3}) \\ \vdots & \vdots & \vdots \\ f(z^{(2)}_{43,1} + b^{(2)}_{1,1}) & f(z^{(2)}_{43,2} + b^{(2)}_{1,2}) & f(z^{(2)}_{43,3} + b^{(2)}_{1,3}) \end{pmatrix} \end{aligned} \tag{D.5}$$

$A^{(2)}$ sendes videre af de tre synapser efter det skjulte lag og vægtes af søjlevektoren $W^{(2)}$:

$$W^{(2)} = \begin{pmatrix} w^{(2)}_{1,1} \\ w^{(2)}_{2,1} \\ w^{(2)}_{3,1} \end{pmatrix} \tag{D.6}$$

Af det haves søjlevektoren $Z^{(3)}$ ved matrix-multiplication:

$$
\begin{aligned}
Z^{(3)} &= A^{(2)}W^{(2)} \\
&= \begin{pmatrix} a^{(2)}_{1,1} & a^{(2)}_{1,2} & a^{(2)}_{1,3} \\ a^{(2)}_{2,1} & a^{(2)}_{2,2} & a^{(2)}_{2,3} \\ \vdots & \vdots & \vdots \\ a^{(2)}_{43,1} & a^{(2)}_{43,2} & a^{(2)}_{43,3} \end{pmatrix} \begin{pmatrix} w^{(2)}_{1,1} \\ w^{(2)}_{2,1} \\ w^{(2)}_{3,1} \end{pmatrix} \\
&= \begin{pmatrix} a^{(2)}_{1,1}w^{(2)}_{1,1} + a^{(2)}_{1,2}w^{(2)}_{2,1} + a^{(2)}_{1,3}w^{(2)}_{3,1} \\ a^{(2)}_{2,1}w^{(2)}_{1,1} + a^{(2)}_{2,2}w^{(2)}_{2,1} + a^{(2)}_{2,3}w^{(2)}_{3,1} \\ \vdots \\ a^{(2)}_{43,1}w^{(2)}_{1,1} + a^{(2)}_{43,2}w^{(2)}_{2,1} + a^{(2)}_{43,3}w^{(2)}_{3,1} \end{pmatrix} \\
&= \begin{pmatrix} z^{(3)}_{1,1} \\ z^{(3)}_{2,1} \\ z^{(3)}_{3,1} \end{pmatrix}
\end{aligned} \tag{D.7}
$$

Det næste skridt er at introducere en lineær transferfunktion i output ($f^{(3)}$), som kan skalere $Z^{(3)}$ i forhold til target, såfremt det er nødvendigt. I dette skridt introduceres endvidere endnu en bias $B^{(3)}$, hvis funktion er fuldstændigt det samme som tidligere. Det er en uafhængig konstant, som kan forskyde lineariteten væk fra origo. Af det haves et endeligt output $\hat{Y}$:

$$
\begin{aligned}
\hat{Y} &= f^{(3)}(Z^{(3)} + B^{(3)}) \\
&= \begin{pmatrix} \hat{y}_{1,1} \\ \hat{y}_{2,1} \\ \vdots \\ \hat{y}_{43,1} \end{pmatrix}
\end{aligned} \tag{D.8}
$$

Deraf haves fire ligninger, der tilsammen udgør den fremadrettede beregning:

$$Z^{(2)} = XW^{(1)} \tag{D.9}$$

$$A^{(2)} = f^{(2)}(Z^{(2)} + B^{(2)}) \tag{D.10}$$

$$Z^{(3)} = A^{(2)}W^{(2)} \tag{D.11}$$

$$\hat{Y} = f^{(3)}(Z^{(3)} + B^{(3)}) \tag{D.12}$$

Backpropagation

Backpropagation er en bagudrettet beregning, som søger at minimere fejlen mellem outputtet $\hat{Y}$ og target Y. Fejlen er det, som i realiteten er netværkets evalueringsparameter. Det er i dette tilfælde mean squared error, noteres J, og den beregnes som vist nedenfor:

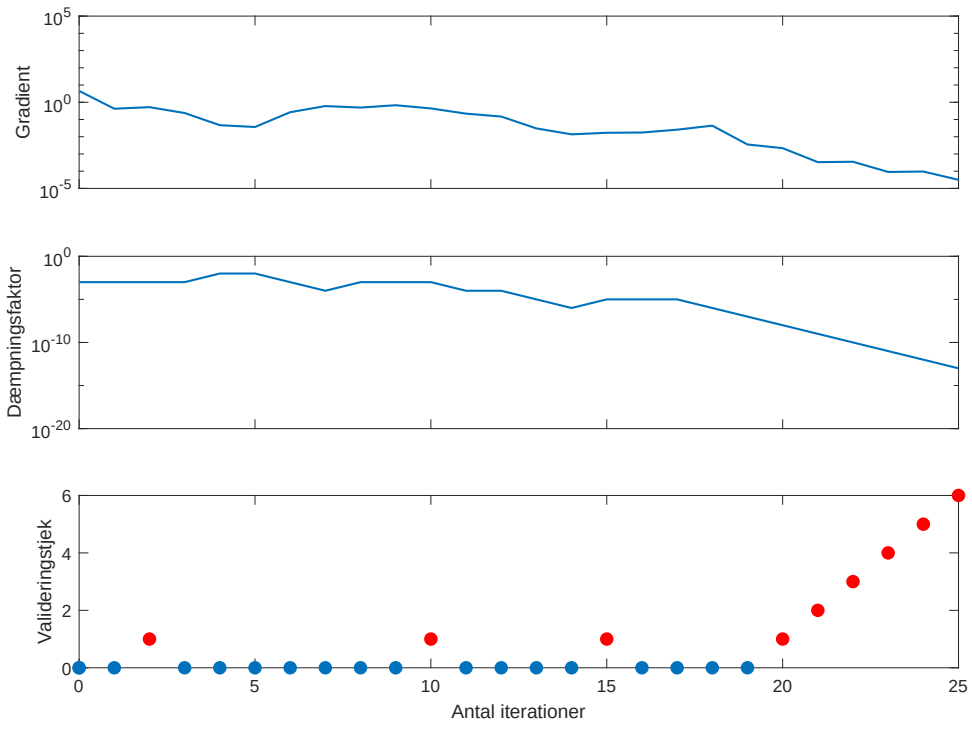
$$J = \frac{\sum_{i=1}^{43}(\hat{Y}_i - Y_i)^2}{43} \tag{D.13}$$

Herefter søges det at finde de to vektorer $W^{(2)}$ og $W^{(3)}$, der giver den laveste værdi J. Levenberg-Marquardt algoritmen træner efter ***gradientmetoden***, som egentlig blot går ud på at løse ligningen

$dJ/dW = 0$ med hensyn til W ved iterering, og sætte W ind i de fire ligninger i den fremadrettede beregning (Straeter 1971). Det er her computerkraften får en betydning, da det ofte er umuligt at løse den ligning. Her er der eksempelvis seks forskellige frihedsgrader, og løsningen ligger derfor i et 6-dimensionalt vektorrum. Derfor introduceres stopkriterierne, som gør, at der accepteres andre løsninger end det globale minimum. Det reducerer beregningstiden betragteligt.

- Det maksimale antal iterationer er nået (standard er 1000).
- Den maksimale tidsvarighed overskrides (standard er ∞).
- Evalueringsparameteren er konvergeret mod en målsat MSE (standard er 0).
- Gradienten på evalueringsparameteren er faldet under en minimumsgrænse (standard er 10^{-7}).
- Dæmpningsfaktoren, μ overskrider den maksimale dæmpningsfaktor (standard er $\mu_{max} = 10^{-10}$).
- Hvis MSE i valideringen bliver større ved et antal yderligere iterationer ($\max_{fail}$), selvom MSE i træningen mindskes. (standard er $\max_{fail} = 6$).

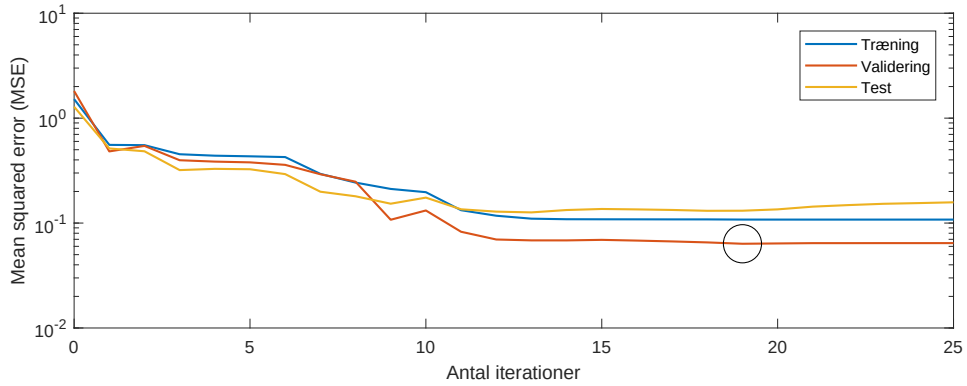
Træningen stoppes ved iteration 19, som vist på figur D.6. I dette tilfælde stoppes træningen af stopkriterie 6, da yderligere iterationer gør MSE større i seks forsøg. Efter seks fejlede valideringstjek antages det, at forringelsen i valideringen ikke længere er tilfældigt. Iteration 3, 10, og 15 giver også en dårligere MSE i valideringen, men her er det tilfældigt, fordi iteration 4, 11 og 16 giver en bedre MSE igen.



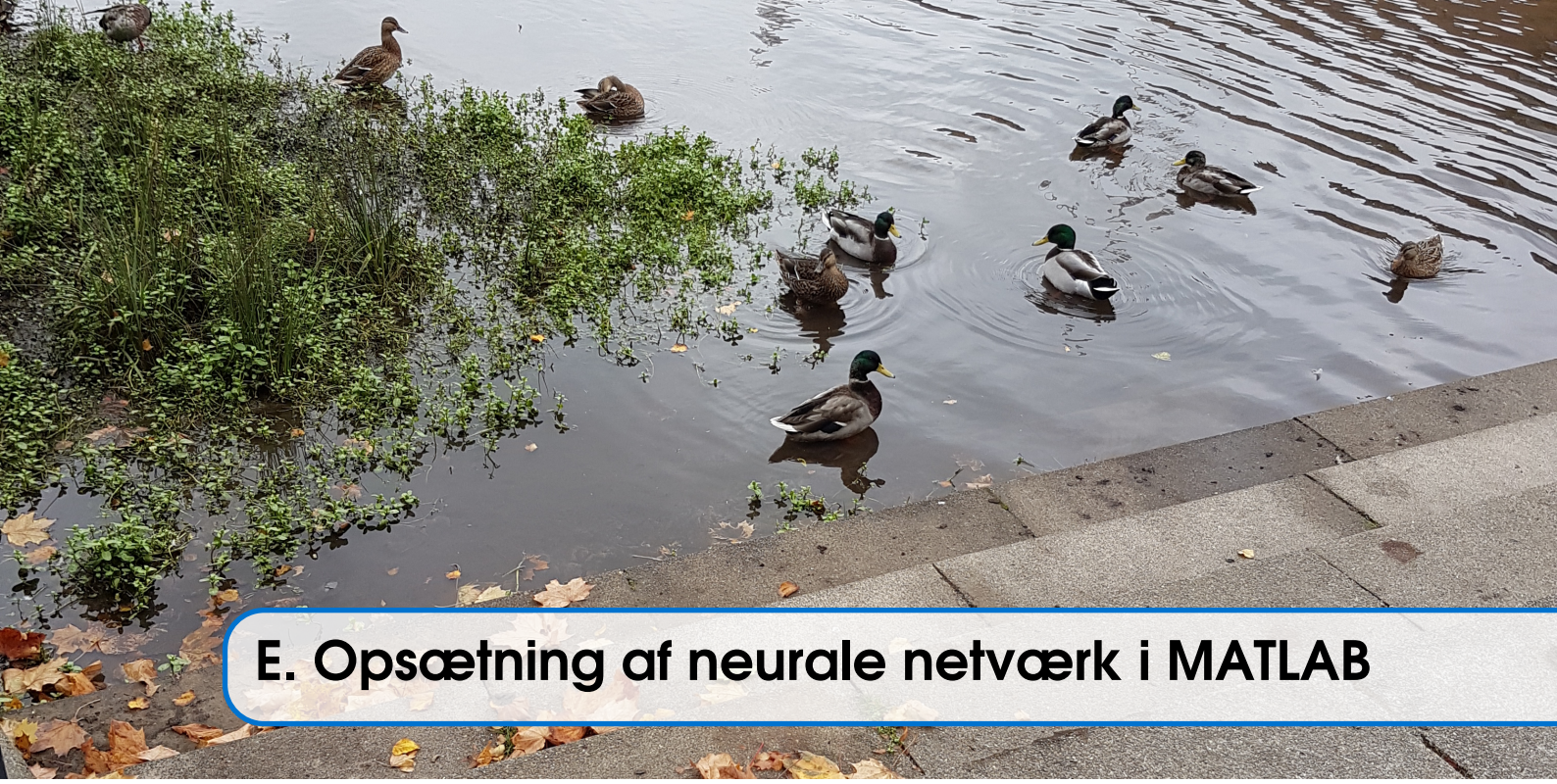
Figur D.6: I dette tilfælde stoppes træningen som følge af seks fejlede valideringstjek.

Resultatet på træningen er en model med en MSE i størrelsesordenen 10^{-1}, hvilket ses på figur D.7. Her vises udviklingen i MSE i træningen, valideringen og testen som funktion af antallet af iterationer, og de tre grafer skal helst følge hinanden, da modellen i det tilfælde vil virke lige godt på alle perioder. Det ses også, at testen er bedre end træningen og valideringen mellem iteration 0 og 8,

hvilket erfares ved flere tilfælde senere. Testen er i eksemplet udvalgt tilfældigt, men værdierne er uafhængige af træningen. Derfor kan det tolkes som om, at modellen er bedre end det umiddelbart kunne forventes på baggrund af træningen/valideringen.



Figur D.7: Træningen stopper efter 19 iterationer, og resultatet er en model med en MSE i størrelsesordenen 10^{-1}.



E. Opsætning af neurale netværk i MATLAB

Dette appendiks viser den vigtigste kode i hele projektet, og den er udviklet til nemt at kunne definere input, forudsigelsestider og i det hele taget teste forskellige netværkskonfigurationer.

Opsætning af netværk

Følgende kode er et eksempel til modellering af vandstanden ved Grejsdalens Planteskole (ST. 32.06), og der haves næsten identiske koder for modellering af de resterende stationer. Derfor anses denne kode som standard.

E.0.1 Anvendte data

```
clc; tic
clear all;

%% DataPath
Datapath = 'C:\Users\bent2\OneDrive - Aalborg Universitet\Thomas og
    Benjamin - Speciale\Data';

%% Anvendte data
% Vandstande
load(fullfile(Datapath,'Vandstande\Hobs_Planteskolen_3206_inter.mat'));
Station_3206 = Hobs_Planteskolen_3206_inter;    % Planteskolen (ST. 32.06)

% Regn
% Ud/indkommenter alt efter, om der skal anvendes radar eller regnmaaler
load(fullfile(Datapath,'Regnmaaler\JellingSVK_5230SommerTid.mat'));
load(fullfile(Datapath,'Radar\PlanteRadar_intSommerTid.mat'));
```

```
% Maaler
Regn = JellingSVK_5230SommerTid; % [E-6 m/s]

Regn = [Regn(:,1) Regn(:,2).*(60/1000)]; % [mm/min]

% % Radar
% Regn = PlanteRadar_intSommerTid; % [E-6 m/s]
% Regn = [Regn(:,1) Regn(:,2).*(60/1000)]; % [mm/min]
```

E.0.2 Valg af træning, validering og test

```
%% Targetperiode
TargetStart = '10-Dec-2015 00:00';
TargetSlut = '06-Jan-2018 00:00';

%% Definer traenings-, validerings- og testperioder
% Traening
Starttid = ['10-Dec-2015 00:00';
           % '27-Dec-2015 15:01'; Maaleafbrud
            '15-Jan-2016 00:01';
            '15-Feb-2016 00:01';
            '15-Mar-2016 00:01';
           % '23-Mar-2016 03:01'; Maaleafbrud
            '15-Apr-2016 00:01';
            '15-May-2016 00:01';
            '15-Jun-2016 00:01';
            '15-Jul-2016 00:01';
            '15-Aug-2016 00:01';
            '15-Sep-2016 00:01';
            '15-Oct-2016 00:01';
            '15-Nov-2016 00:01';
            '15-Dec-2016 00:01'
            ];


Sluttid = ['27-Dec-2015 15:00';
          % '01-Jan-2016 00:00'; Maaleafbrud
           '01-Feb-2016 00:00';
           '01-Mar-2016 00:00';
           '23-Mar-2016 03:00';
          % '01-Apr-2016 01:00'; Maaleafbrud
           '01-May-2016 00:00';
           '01-Jun-2016 00:00';
           '01-Jul-2016 00:00';
           '01-Aug-2016 00:00';
           '01-Sep-2016 00:00';
           '01-Oct-2016 00:00';
           '01-Nov-2016 00:00';
           '01-Dec-2016 00:00';
```

```
            '01-Jan-2017 00:00'
           ];

TrainPeriod = [datenum(Starttid) datenum(Sluttid)];

% Validering
Starttid = ['01-Jan-2016 00:01';
            '01-Feb-2016 00:01';
            '01-Mar-2016 00:01';
            '01-Apr-2016 01:01'; % Forskudt en time pga afbrud
            '01-May-2016 00:01';
            '01-Jun-2016 00:01';
            '01-Jul-2016 00:01';
            '01-Aug-2016 00:01';
            '01-Sep-2016 00:01';
            '01-Oct-2016 00:01';
            '01-Nov-2016 00:01';
            '01-Dec-2016 00:01'
            ];

Sluttid = ['15-Jan-2016 00:00';
           '15-Feb-2016 00:00';
           '15-Mar-2016 00:00';
           '15-Apr-2016 00:00';
           '15-May-2016 00:00';
           '15-Jun-2016 00:00';
           '15-Jul-2016 00:00';
           '15-Aug-2016 00:00';
           '15-Sep-2016 00:00';
           '15-Oct-2016 00:00';
           '15-Nov-2016 00:00';
           '15-Dec-2016 00:00'
           ];

ValidationPeriod = [datenum(Starttid) datenum(Sluttid)];

% Test
Starttid = ['01-Jan-2017 00:01';
            '01-Dec-2017 10:46'];

Sluttid = ['17-Nov-2017 03:43'; % Maaleafbrud sidste halvdel af november
    2017
           '06-Jan-2018 00:00'];

TestPeriod = [datenum(Starttid) datenum(Sluttid)];
```

E.0.3 Antal neuroner og træningsalgoritme

```
%% Antal neuroner i det skjulte lag
```

```
hiddenLayerSize = 10;

%% Traeningsalgoritme
TrainFcn = 'trainlm';
```

E.0.4 Valg af input

```
%% Antal traeninger
N = 1;

%% Forudsigelsestid
Forudsigelsestid = 120; %[min]

%% Forsinkelse paa regn
ForsinkelseRegn = 86;  % [min]

%% Target
Starttid = datenum(TargetStart); %Target
Sluttid = datenum(TargetSlut); %Target
StarttidIndgang = find(Station_3206<(Starttid+0.0001) & Station_3206>(
    Starttid-0.0001));
SluttidIndgang = find(Station_3206<(Sluttid+0.0001) & Station_3206>(Sluttid
    -0.0001));
id = [StarttidIndgang; SluttidIndgang];
Station_3206_Target = Station_3206(id(1):id(2),:);

TargetNeural = Station_3206_Target(:,2);
Dates = Station_3206_Target(:,1);

% Target
save('TargetNeural.mat','TargetNeural')
save('Dates.mat','Dates')

%% Definering af input
% Loops
counter = 1;

for F=Forudsigelsestid %[min]
% Vandstandsmaaler input 1
    Starttid = datenum(TargetStart)-datenum(minutes(F)); %Target-
    Forudsigelsestid
    Sluttid = datenum(TargetSlut)-datenum(minutes(F)); %Target-
    Forudsigelsestid
    StarttidIndgang = find(Station_3206<(Starttid+0.0001) & Station_3206>(
    Starttid-0.0001));
    SluttidIndgang = find(Station_3206<(Sluttid+0.0001) & Station_3206>(
    Sluttid-0.0001));
    id = [StarttidIndgang; SluttidIndgang];
```

```
    Station_3206_Input = Station_3206(id(1):id(2),:);

% Gradient input 1
    Starttid = datenum(TargetStart)-datenum(minutes(F)); %Target-
    Forudsigelsestid
    Sluttid = datenum(TargetSlut)-datenum(minutes(F)); %Target-
    Forudsigelsestid
    StarttidIndgang = find(Station_3206<(Starttid+0.0001) &  Station_3206>(
    Starttid-0.0001));
    SluttidIndgang = find(Station_3206<(Sluttid+0.0001) &  Station_3206>(
    Sluttid-0.0001));
    id = [StarttidIndgang; SluttidIndgang];

    vsTid = 0.0000075738*(F^3)-0.0043946950*(F^2)+0.8485778114*F;   %
    Empirisk bestemt ved gradientoptimering af kurveforloeb.
    Grad = 22;
    Station_3206_Koter = Station_3206(:,2);

    InputGradient = Station_3206_Koter(id(1):id(2),:)+(Station_3206_Koter(
    id(1):id(2),:)-Station_3206_Koter(id(1)-Grad:id(2)-Grad,:))/Grad*vsTid;

%% Regninputs som Tapped Delay Line
% Regninput 1
    for For1 = ForsinkelseRegn
        Akkumuleringstid = 60; %[min]
        ForsinkelseFraTarget = For1; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn1 = Regn(id(1):id(2),:);

        Regninput1(:,1) = Regn1((1+Akkumuleringstid):end,1);

        for i=1:length(Regn1)-Akkumuleringstid
            Regninput1(i,2)=sum(Regn1((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 2
    for For2 = For1+Akkumuleringstid
        Akkumuleringstid = 60; %[min]
        ForsinkelseFraTarget = For2; %[min]
```

```
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn2 = Regn(id(1):id(2),:);

        Regninput2(:,1) = Regn2((1+Akkumuleringstid):end,1);

        for i=1:length(Regn2)-Akkumuleringstid
            Regninput2(i,2)=sum(Regn2((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 3
    for For3 = For2+Akkumuleringstid
        Akkumuleringstid = 60; %[min]
        ForsinkelseFraTarget = For3; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn3 = Regn(id(1):id(2),:);

        Regninput3(:,1) = Regn3((1+Akkumuleringstid):end,1);

        for i=1:length(Regn3)-Akkumuleringstid
            Regninput3(i,2)=sum(Regn3((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 4
    for For4 = For3+Akkumuleringstid
        Akkumuleringstid = 60; %[min]
        ForsinkelseFraTarget = For4; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
```

```
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn4 = Regn(id(1):id(2),:);

        Regninput4(:,1) = Regn4((1+Akkumuleringstid):end,1);

        for i=1:length(Regn4)-Akkumuleringstid
            Regninput4(i,2)=sum(Regn4((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 5
    for For5 = For4+Akkumuleringstid
        Akkumuleringstid = 60; %[min]
        ForsinkelseFraTarget = For5; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn5 = Regn(id(1):id(2),:);

        Regninput5(:,1) = Regn5((1+Akkumuleringstid):end,1);

        for i=1:length(Regn5)-Akkumuleringstid
            Regninput5(i,2)=sum(Regn5((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 6
    for For6 = For5+Akkumuleringstid
        Akkumuleringstid = 1440; %[min]
        ForsinkelseFraTarget = For6; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
```

```
        Regn6 = Regn(id(1):id(2),:);

        Regninput6(:,1) = Regn6((1+Akkumuleringstid):end,1);

        for i=1:length(Regn6)-Akkumuleringstid
            Regninput6(i,2)=sum(Regn6((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 7
    for For7 = For6+Akkumuleringstid
        Akkumuleringstid = 1440; %[min]
        ForsinkelseFraTarget = For7; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn7 = Regn(id(1):id(2),:);

        Regninput7(:,1) = Regn7((1+Akkumuleringstid):end,1);

        for i=1:length(Regn7)-Akkumuleringstid
            Regninput7(i,2)=sum(Regn7((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 8
    for For8 = For7+Akkumuleringstid
        Akkumuleringstid = 1440; %[min]
        ForsinkelseFraTarget = For8; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn8 = Regn(id(1):id(2),:);

        Regninput8(:,1) = Regn8((1+Akkumuleringstid):end,1);

        for i=1:length(Regn8)-Akkumuleringstid
```

```
            Regninput8(i,2)=sum(Regn8((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 9
    for For9 = For8+Akkumuleringstid
        Akkumuleringstid = 1440; %[min]
        ForsinkelseFraTarget = For9; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn9 = Regn(id(1):id(2),:);

        Regninput9(:,1) = Regn9((1+Akkumuleringstid):end,1);

        for i=1:length(Regn9)-Akkumuleringstid
            Regninput9(i,2)=sum(Regn9((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 10
    for For10 = For9+Akkumuleringstid
        Akkumuleringstid = 10080; %[min]
        ForsinkelseFraTarget = For10; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn10 = Regn(id(1):id(2),:);

        Regninput10(:,1) = Regn10((1+Akkumuleringstid):end,1);

        for i=1:length(Regn10)-Akkumuleringstid
            Regninput10(i,2)=sum(Regn10((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 11
    for For11 = For10+Akkumuleringstid
```

```
        Akkumuleringstid = 20160; %[min]
        ForsinkelseFraTarget = For11; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn11 = Regn(id(1):id(2),:);

        Regninput11(:,1) = Regn11((1+Akkumuleringstid):end,1);

        for i=1:length(Regn11)-Akkumuleringstid
            Regninput11(i,2)=sum(Regn11((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 12
    for For12 = For11+Akkumuleringstid
        Akkumuleringstid = 30240; %[min]
        ForsinkelseFraTarget = For12; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn12 = Regn(id(1):id(2),:);

        Regninput12(:,1) = Regn12((1+Akkumuleringstid):end,1);

        for i=1:length(Regn12)-Akkumuleringstid
            Regninput12(i,2)=sum(Regn12((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 13
    for For13 = For12+Akkumuleringstid
        Akkumuleringstid = 40320; %[min]
        ForsinkelseFraTarget = For10; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
```

```
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn13 = Regn(id(1):id(2),:);

        Regninput13(:,1) = Regn13((1+Akkumuleringstid):end,1);

        for i=1:length(Regn13)-Akkumuleringstid
            Regninput13(i,2)=sum(Regn13((i+1:(i+Akkumuleringstid)),2));
        end

% Regninput 14
    for For14 = For13+Akkumuleringstid
        Akkumuleringstid = 80640; %[min]
        ForsinkelseFraTarget = For14; %[min]
        Starttid = datenum(TargetStart)-datenum(minutes(Akkumuleringstid))-
    datenum(minutes(ForsinkelseFraTarget)); %Target-Akkumuleringstid-
    ForsinkelseFraTarget
        Sluttid = datenum(TargetSlut)-datenum(minutes(ForsinkelseFraTarget)
    ); %Target-ForsinkelseFraTarget
        StarttidIndgang = find(Regn(:,1)<(Starttid+0.0001) & Regn(:,1)>(
    Starttid-0.0001));
        SluttidIndgang = find(Regn(:,1)<(Sluttid+0.0001) & Regn(:,1)>(
    Sluttid-0.0001));
        id = [StarttidIndgang; SluttidIndgang];
        Regn14 = Regn(id(1):id(2),:);

        Regninput14(:,1) = Regn14((1+Akkumuleringstid):end,1);

        for i=1:length(Regn14)-Akkumuleringstid
            Regninput14(i,2)=sum(Regn14((i+1:(i+Akkumuleringstid)),2));
        end

% Saml input i InputNeural-matrix
InputNeural = [Station_3206 Regninput1(:,2) Regninput2(:,2) Regninput3(:,2)
     Regninput4(:,2) Regninput5(:,2) Regninput6(:,2) Regninput7(:,2)
    Regninput8(:,2) Regninput9(:,2) Regninput10(:,2) Regninput11(:,2)
    Regninput12(:,2) Regninput13(:,2) Regninput14(:,2)];

% Input der ikke er i brug:
% InputGradient

% Gem Input
save('InputNeural.mat','InputNeural')
```

```
%% Koer netvaerkskoden
run('BenjaminNetvaerk_3206.m')

%% Performance gemmes
Resultat(counter,1) = ForsinkelseRegn(1,counter);
Resultat(counter,2) = RMSE_Peak;
Resultat(counter,3) = median(rmse_tperf(:,1));
Resultat(counter,4) = median(rmse_vperf(:,1));
Resultat(counter,5) = median(rmse_perf(:,1));
Resultat(counter,6) = counter;

counter = counter+1
    end
    end
    end
    end
    end
    end
    end
    end
    end
    end
    end
    end
    end
    end
end

save('Planteskolen/Resultat.mat','Resultat')

toc
```

Evaluering af netværk

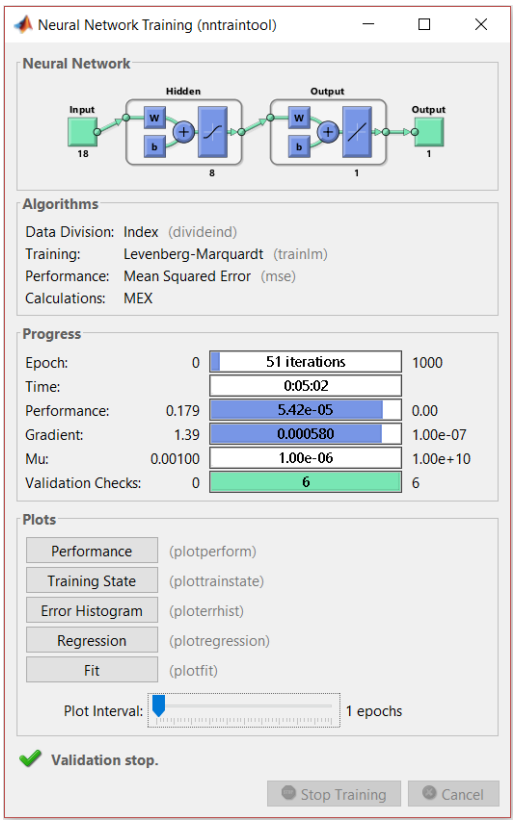
E.0.5 Træning

```
clc; tic

%% Indlaes Input, target og tidsstempler
load('InputNeural.mat')
load('TargetNeural.mat')
load('Dates.mat')

%% Opsaet netvaerket
for j=1:N
X = InputNeural';
T = TargetNeural';

if TrainFcn=='trainlm'
```

```
    trainFcn = 'trainlm';  % Levenberg-Marquardt backpropagation.
elseif TrainFcn=='trainbr'
    trainFcn = 'trainbr';  % Bayesian Regularization.
elseif TrainFcn=='trainscg'
    trainFcn = 'trainscg'; % Scaled Conjugate Gradient.
else isempty(TrainFcn)==1
    trainFcn = 'trainlm';  % Saetter Levenberg-Marquardt backpropagation
    som default.
end

% Laver et Fitting Network
net = fitnet(hiddenLayerSize,trainFcn);

% Choose Input and Output Pre/Post-Processing Functions
net.input.processFcns = {'removeconstantrows','mapminmax'};
net.output.processFcns = {'removeconstantrows','mapminmax'};

net.divideFcn = 'divideind';  % Inddeler data efter index

% Index for traening
[m,n] = size(TrainPeriod);
ip = zeros(length(Dates),m);


for i=1:size(TrainPeriod,1)
    StarttidIndgang = find(Dates<(TrainPeriod(i,1)+0.0001) & Dates>(
    TrainPeriod(i,1)-0.0001));
    SluttidIndgang = find(Dates<(TrainPeriod(i,2)+0.0001) & Dates>(
    TrainPeriod(i,2)-0.0001));
    id = [StarttidIndgang; SluttidIndgang];
    ip(id(1):id(2),i) = id(1):id(2);
end

v = ip(:);
id = find(v==0);
v(id)=[];
v=v';

net.divideParam.trainInd = v;

clear ip v

% Index for validering
[m,n] = size(ValidationPeriod);
ip = zeros(length(Dates),m);

for i=1:size(ValidationPeriod,1)
    StarttidIndgang = find(Dates<(ValidationPeriod(i,1)+0.0001) & Dates>(
    ValidationPeriod(i,1)-0.0001));
```

```
    SluttidIndgang = find(Dates<(ValidationPeriod(i,2)+0.0001) & Dates>(
    ValidationPeriod(i,2)-0.0001));
    id = [StarttidIndgang; SluttidIndgang];
    ip(id(1):id(2),i) = id(1):id(2);
end

v = ip(:);
id = find(v==0);
v(id)=[];
v=v';

net.divideParam.valInd = v;

% Index for test
clear ip v

[m,n] = size(TestPeriod);
ip = zeros(length(Dates),m);

for i=1:size(TestPeriod,1)
    StarttidIndgang = find(Dates<(TestPeriod(i,1)+0.0001) & Dates>(
    TestPeriod(i,1)-0.0001));
    SluttidIndgang = find(Dates<(TestPeriod(i,2)+0.0001) & Dates>(
    TestPeriod(i,2)-0.0001));
    id = [StarttidIndgang; SluttidIndgang];
    ip(id(1):id(2),i) = id(1):id(2);
end

v = ip(:);
id = find(v==0);
v(id)=[];
TestInd = v;
v=v';

net.divideParam.testInd = v;

clear ip v

% Evalueringsparameter
net.performFcn = 'mse';  % Mean Squared Error

% Traen netvaerket
    [net,tr] = train(net,X,T);

rmse_tperf(j,1) = sqrt(tr.best_tperf);
rmse_vperf(j,1) = sqrt(tr.best_vperf);
rmse_perf(j,1) = sqrt(tr.best_perf);

    if rmse_tperf(j,1)<median(rmse_tperf(1:j-1,1))
```

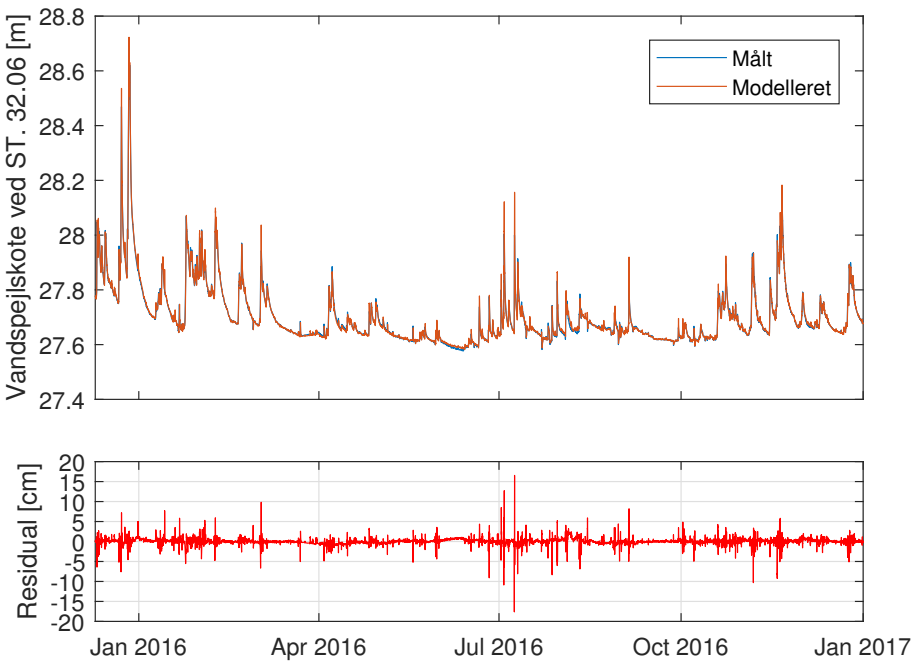
```
        BedsteNet = net;
        BedsteTr = tr;
    elseif j==1
        BedsteNet = net;
        BedsteTr = tr;
    end
end

y = BedsteNet(X)';
e = y-TargetNeural;
```



E.0.6 Plot træning/validering

```
T = [datetime('01-Dec-2015 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Jan-2016 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Apr-2016 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Jul-2016 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Oct-2016 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Jan-2017 00:00','InputFormat','dd-MMM-yyyy HH:mm')];

fig = figure(1)
fig.Renderer='Painters';
fig.PaperUnits='centimeters';
fig.PaperPosition=[0 0 23 16.2]
fig_pos = fig.PaperPosition;
```

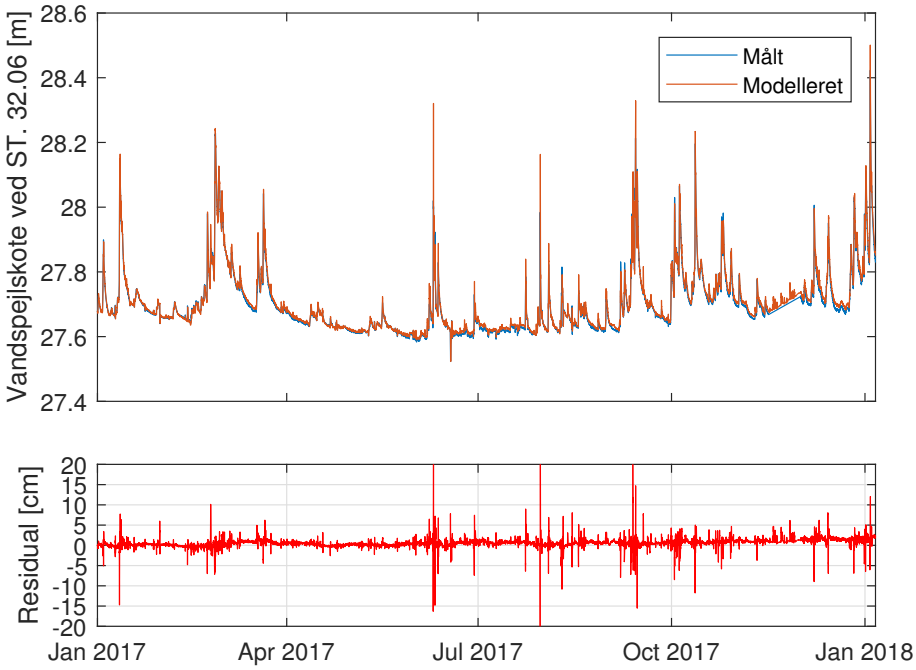
```
fig.PaperSize = [fig_pos(3) fig_pos(4)];

Dateplot = datetime(Dates,'ConvertFrom','datenum');
sub1 = subplot(3,1,1:2)
plot(Dateplot,TargetNeural,'Color',[0    0.4470    0.7410]);
hold on
plot(Dateplot,y,'Color',[0.8500    0.3250    0.0980]);
hold off
xlim([Dateplot(1,1) Dateplot(TestInd(1,1)-1,1)])
legend('Maalt','Modelleret','Location','NorthEast')
sub1.XTick = T;
sub1.XTickLabels = {};
ylabel('Vandspejlskote ved ST. 32.06 [m]')

sub2 = subplot(3,1,3)
s2 = plot(Dateplot,e.*100,'Color','r');
xlim([Dateplot(1,1) Dateplot(TestInd(1,1)-1,1)])
sub2.YTick = [-20 -15 -10 -5 0 5 10 15 20]
ylabel('Residual [cm]')
ylim([-20 20])
grid on
box on
sub2.XTick = T;
```



E.0.7 Plot test

```
T = [datetime('01-Jan-2017 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Apr-2017 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Jul-2017 00:00','InputFormat','dd-MMM-yyyy HH:mm')
    datetime('01-Oct-2017 00:00','InputFormat','dd-MMM-yyyy HH:mm')
```

```
    datetime('01-Jan-2018 00:00','InputFormat','dd-MMM-yyyy HH:mm')];

fig = figure(2)
fig.Renderer='Painters';
fig.PaperUnits='centimeters';
fig.PaperPosition=[0 0 23 16.2]
fig_pos = fig.PaperPosition;
fig.PaperSize = [fig_pos(3) fig_pos(4)];

Dateplot = datetime(Dates,'ConvertFrom','datenum');
sub1 = subplot(3,1,1:2)
plot(Dateplot,TargetNeural,'Color',[0    0.4470    0.7410]);
hold on
plot(Dateplot,y,'Color',[0.8500    0.3250    0.0980]);
hold off
xlim([Dateplot(TestInd(1,1)-1,1) Dateplot(TestInd(end,1),1)])
legend('Maalt','Modelleret','Location','NorthEast')
sub1.XTick = T;
sub1.XTickLabels = {};
ylabel('Vandspejlskote ved ST. 32.06 [m]')

sub2 = subplot(3,1,3)
s2 = plot(Dateplot,e.*100,'Color','r');
xlim([Dateplot(TestInd(1,1)-1,1) Dateplot(TestInd(end,1),1)])
sub2.YTick = [-20 -15 -10 -5 0 5 10 15 20]
ylabel('Residual [cm]')
ylim([-20 20])
grid on
box on
sub2.XTick = T;

toc
```



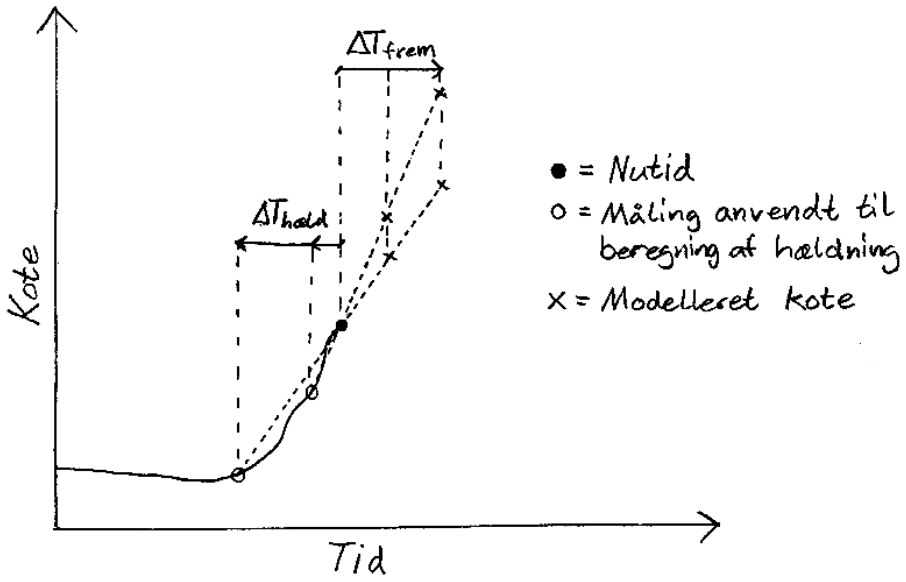
E.0.8 Gem resultater

```
% Opret mappe til resultatfiler og tjekker om mappen allerede eksisterer
filedirOutput = strcat('Planteskolen/KortForudsigelse/Output',num2str(F),'.
    mat');
filedirError = strcat('Planteskolen/KortForudsigelse/Error',num2str(F),'.
    mat');

if ~exist('Planteskolen/KortForudsigelse')==1
    mkdir('Planteskolen/KortForudsigelse')
end

% Gemmer resultatfil i den nye mappe
save(filedirOutput,'y')
save(filedirError,'e')
```

F. Forudsigelse på baggrund af kurveforløb

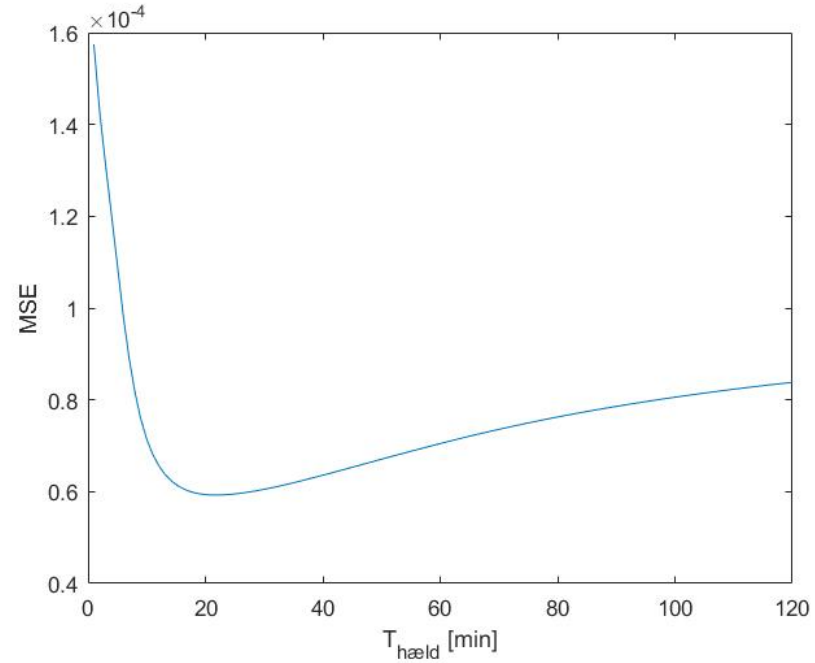
I det følgende appendiks beskrives det, hvordan tendens-modellen er blevet optimeret til at gengive vandstande på bedst mulig vis.

For at forudsige vandspejlskoten på bagrund af hældningen på kurven, skal det vælges hvor lang tid ($\Delta T_{hæld}$), der skal bevæges tilbage på kurven for at beregne hældningen, og hvor lang tid (ΔT_{frem}) det forventes, at denne hældning fortsætter. Dette er illustreret på figur F.1.

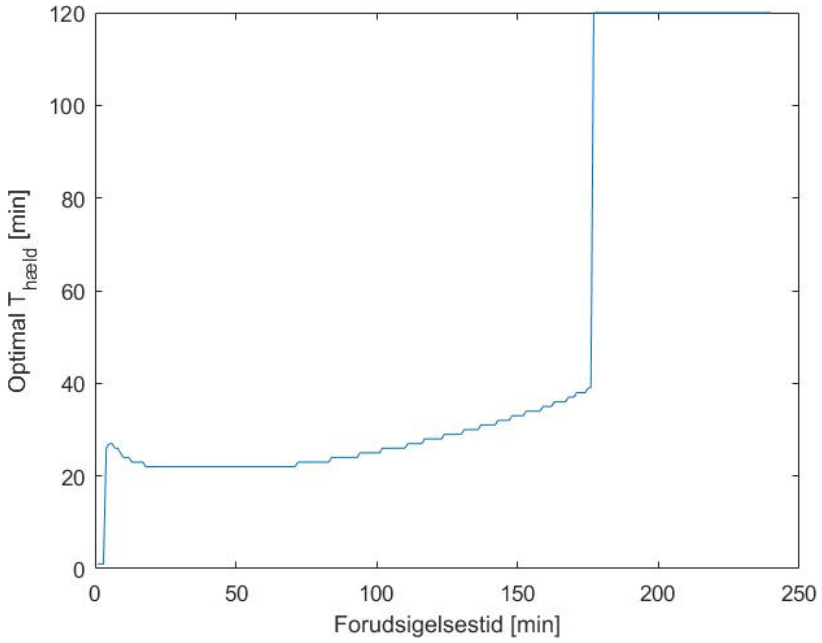


Figur F.1: Illustration af forudsigelse på baggrund af hældning.

Ved at bestemme MSE-værdien ved forskellige værdier af $\Delta T_{hæld}$ (illustreret på figur F.2) er den optimale værdi af $\Delta T_{hæld}$ fundet for forudsigelsestiderne fra 1 til 240 minutter. De optimale værdier af $\Delta T_{hæld}$ er vist på figur F.3.



Figur F.2: MSE-værdien ved forskellige værdier af $\Delta T_{hæld}$ ved en forudsigelsestid på 60 minutter.



Figur F.3: $\Delta T_{hæld}$-værdien, som giver den mindste MSE-værdi ved forskellige forudsigelsestider.

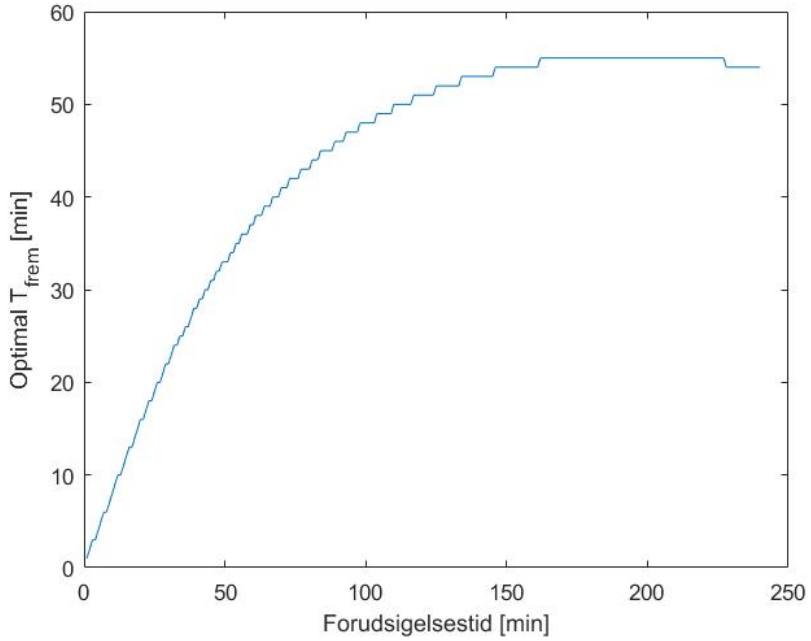
Ved forudsigelsestider over 175 minutter ses en ekstrem stigning i den optimale $\Delta T_{hæld}$. Dette skyldes, at det ikke giver mening at fremskrive vandstanden på baggrund af den aktuelle hældning ved lange forudsigelsestider. Det er derfor valgt, at anvende en værdi for $\Delta T_{hæld}$ på 22 minutter, da dette giver det bedste resultat på de kortere forudsigelsestider.

På samme vis som for $\Delta T_{hæld}$ er den optimale værdi for ΔT_{frem} fundet. De optimale værdier af

ΔT_{frem} er vist på figur F.4. Som den optimale værdi for ΔT_{frem} anvendes en værdi, der er afhængig af forudsigelsestiden:

$$\Delta T_{frem} = 0.0000076 * F^3 - 0.0044 * F^2 + 0.85 * F$$

Funktionen er tilpasset værdiern plottet på figur F.4 og er dermed kun valid i intervallet 0-240 minutter.



Figur F.4: ΔT_{frem}-værdien, som giver den mindste MSE-værdi ved forskellige forudsigelsestider.

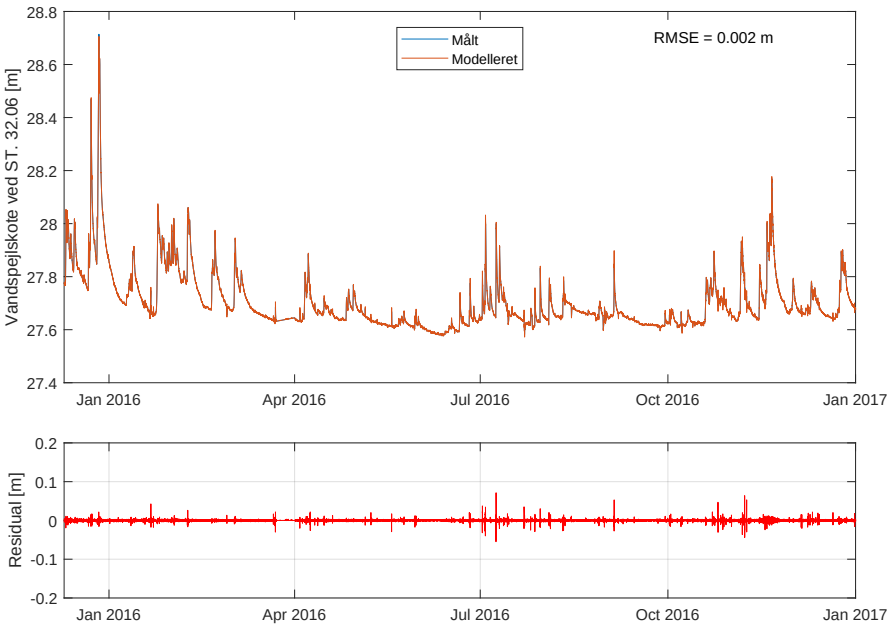
G. Modelresultater

I dette appendiks præsenteres de fulde modellerede serier for henholdsvis stationen ved Grejsdalens Planteskole, Ny Grejsdalsvej og Aagade.

For samtlige stationer er resultatet vist for forudsigelsestider på 30, 60, 120 og 240 minutter. De modellerede vandspejlskoter er et resultat af neurale netværk med input i form af tendens-modellen og nedbør. Inputtene er beskrevet i henholdsvis appendiks F og afsnit 7.2. Der er på de viste modellerede serier ikke frasorteret perioder ved Ny Grejdsdalsvej og Aagade, hvor der har været styring. Derfor er der også naturligt store residualer ved disse stationer i enkelte modelleringsperioder.

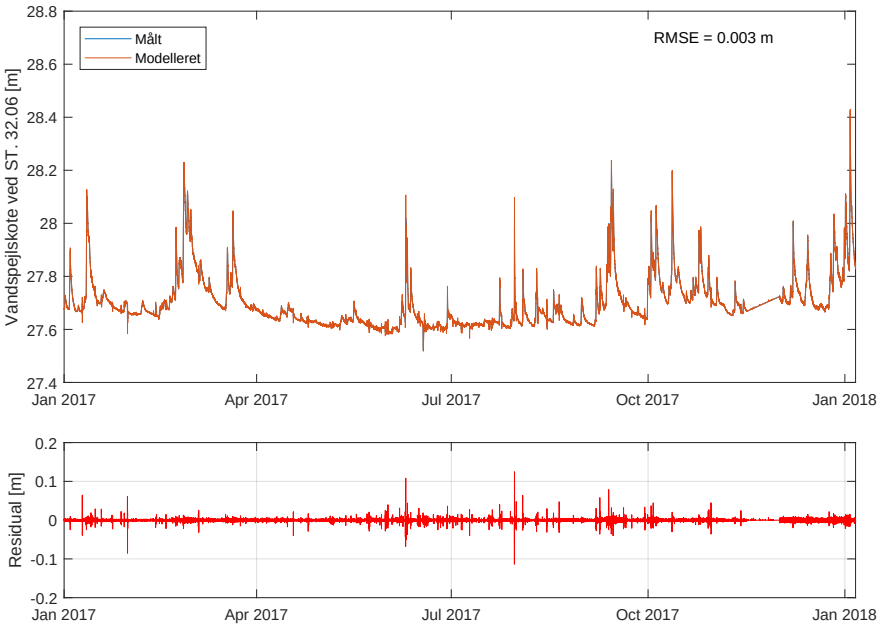
Modelresultater: Grejs Å ved Grejsdalens Planteskole

G.0.1 Træning/validering



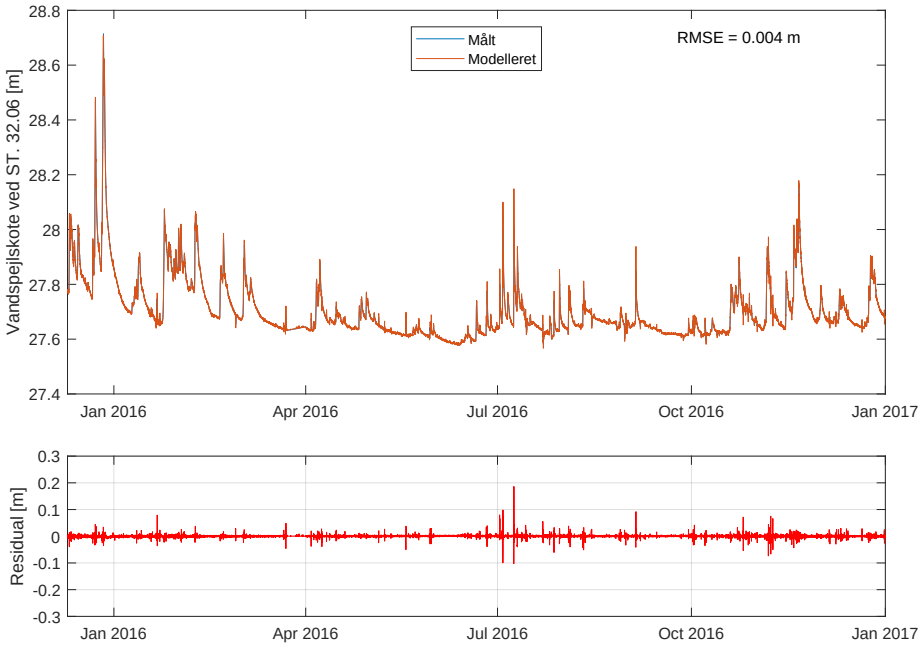
Figur G.1: Trænings-/valideringsperiode for 30 minutters forudsigelse ved Grejsdalens Planteskole.

G.0.2 Test



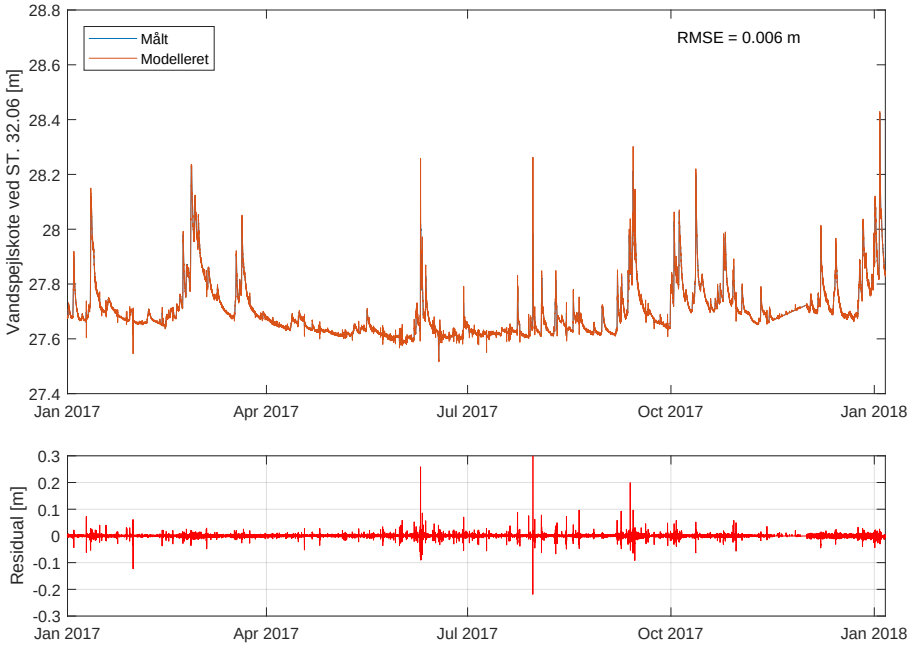
Figur G.2: Testperiode for 30 minutters forudsigelse ved Grejsdalens Planteskole.

G.0.3 Træning/validering



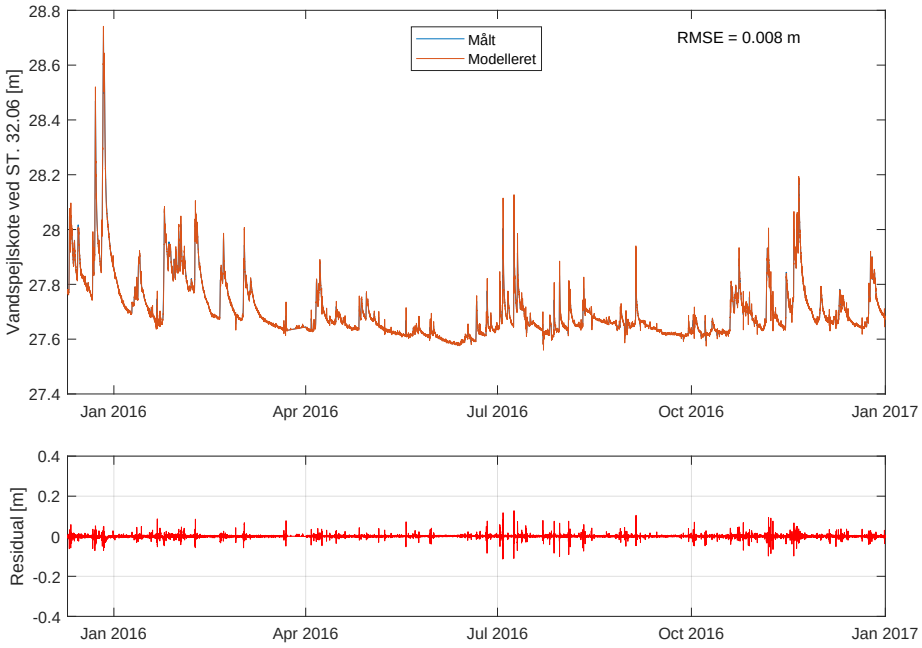
Figur G.3: Trænings-/valideringsperiode for 60 minutters forudsigelse ved Grejsdalens Planteskole.

G.0.4 Test



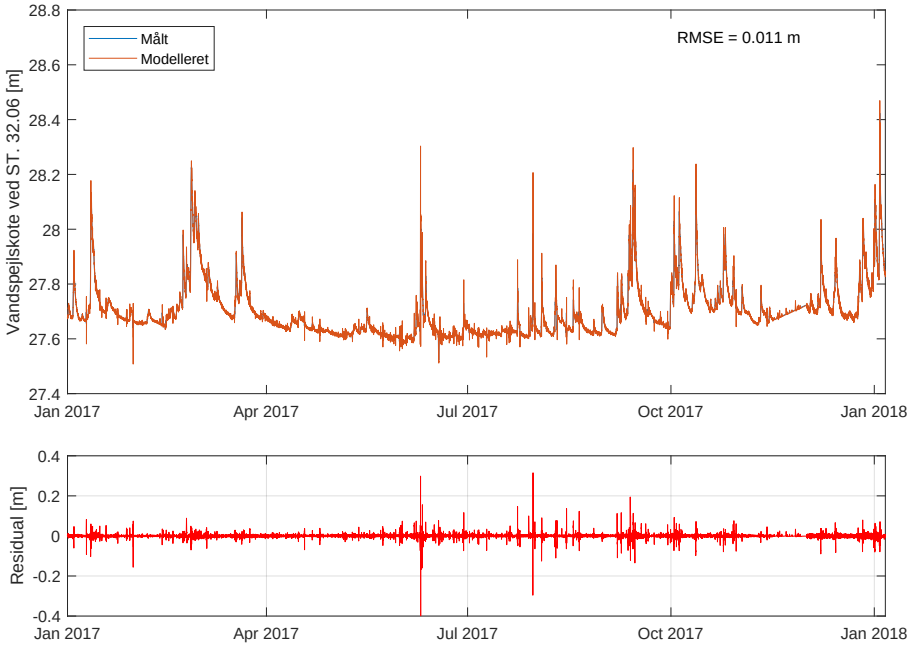
Figur G.4: Testperiode for 60 minutters forudsigelse ved Grejsdalens Planteskole.

G.0.5 Træning/validering



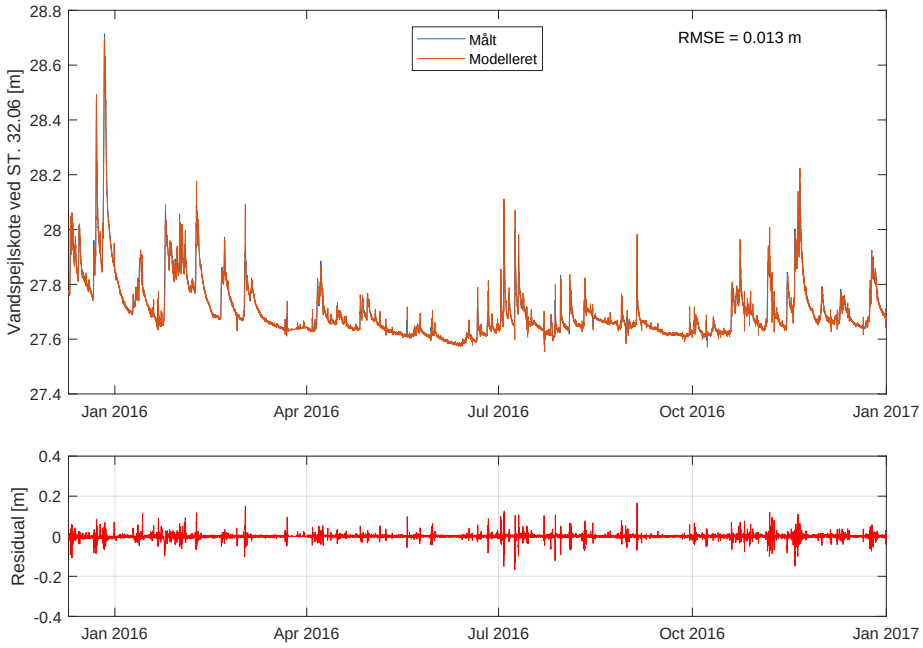
Figur G.5: Trænings-/valideringsperiode for 120 minutters forudsigelse ved Grejsdalens Planteskole.

G.0.6 Test



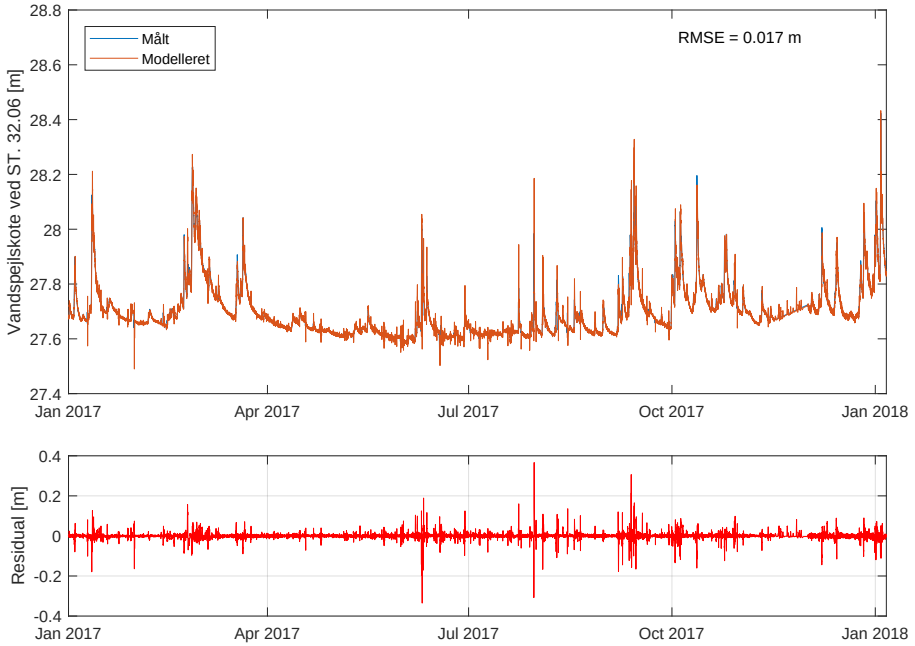
Figur G.6: Testperiode for 120 minutters forudsigelse ved Grejsdalens Planteskole.

G.0.7 Træning/validering



Figur G.7: Trænings-/valideringsperiode for 240 minutters forudsigelse ved Grejsdalens Planteskole.

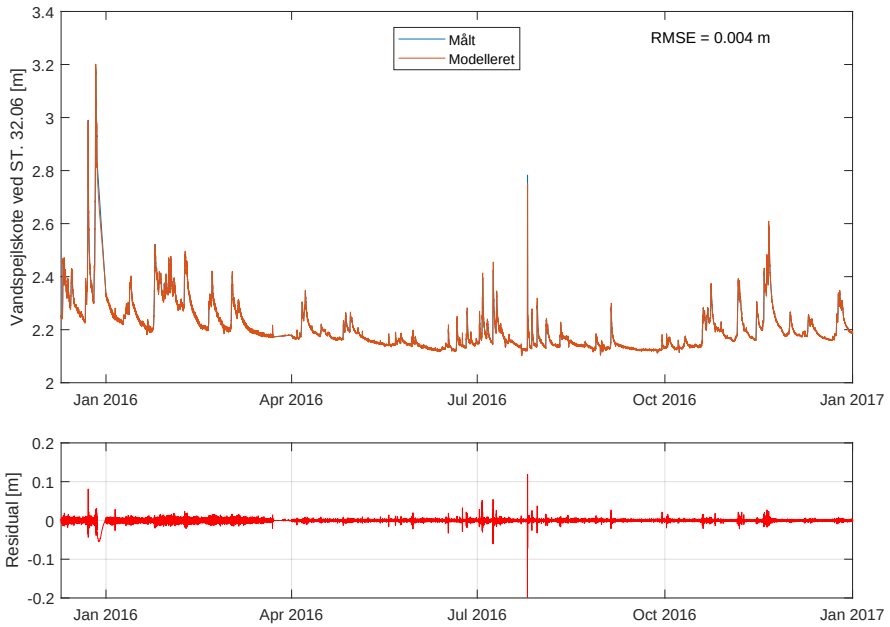
G.0.8 Test



Figur G.8: Testperiode for 240 minutters forudsigelse ved Grejsdalens Planteskole.

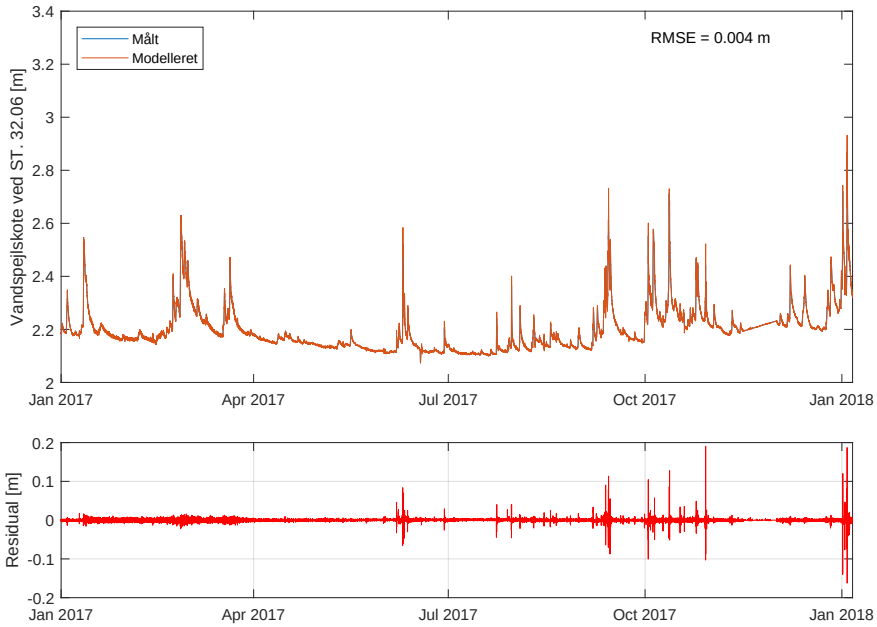
Modelresultater: Grejs Å ved Ny Grejsdalsvej

G.0.9 Træning/validering



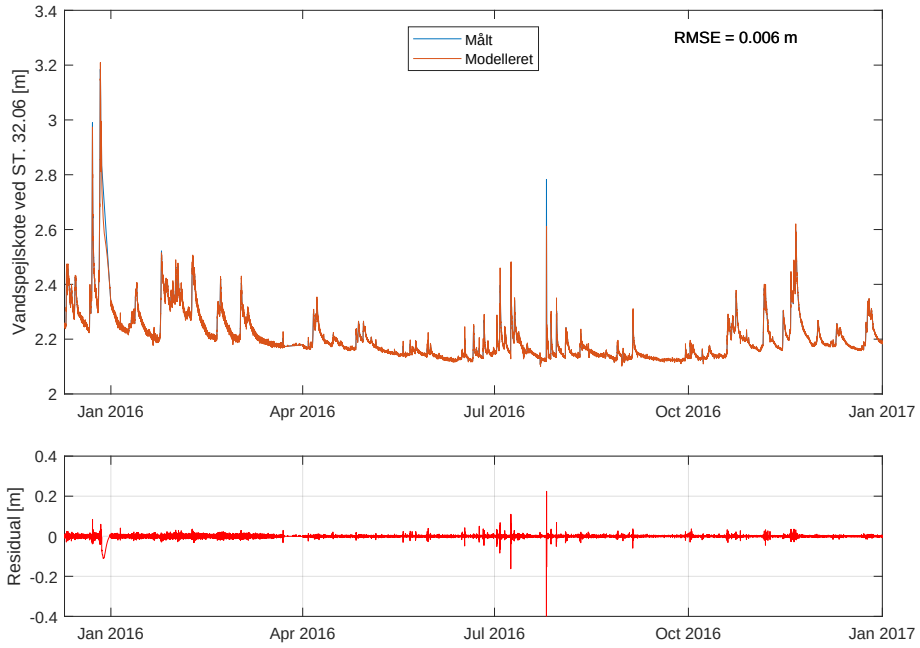
Figur G.9: Trænings-/valideringsperiode for 30 minutters forudsigelse ved Ny Grejsdalsvej.

G.0.10 Test



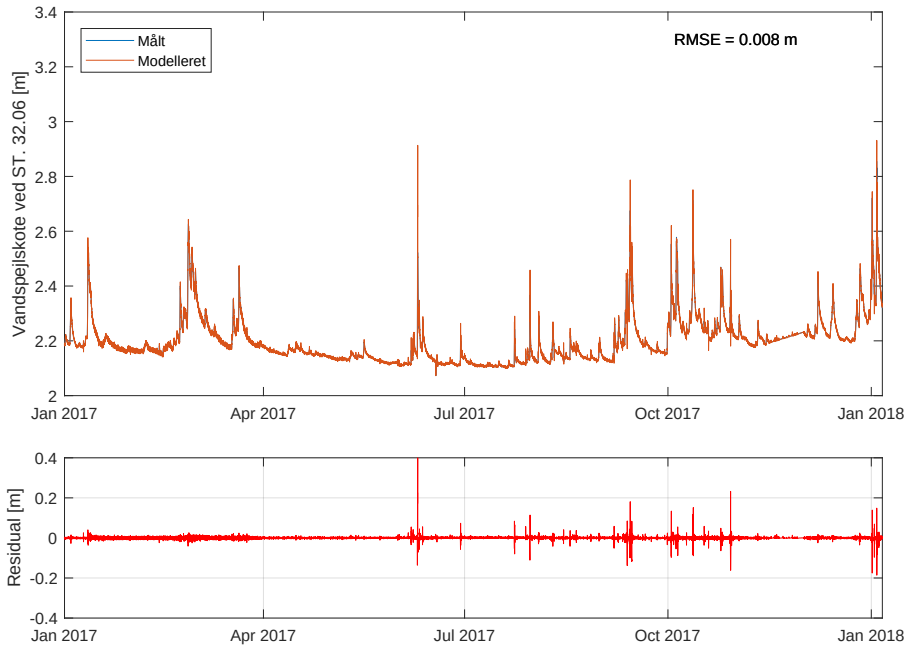
Figur G.10: Testperiode for 30 minutters forudsigelse ved Ny Grejsdalsvej.

G.0.11 Træning/validering



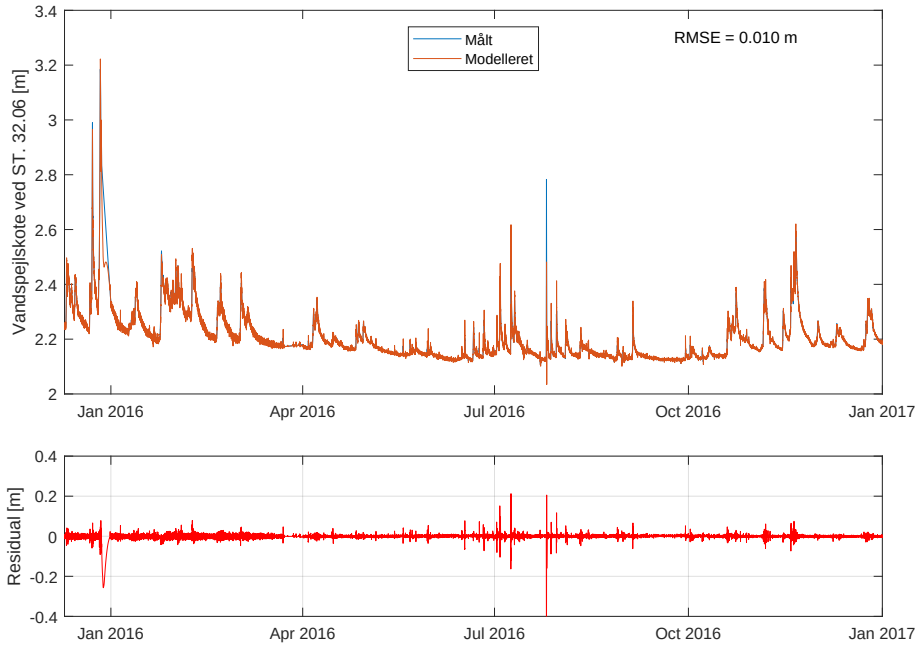
Figur G.11: Trænings-/valideringsperiode for 60 minutters forudsigelse ved Ny Grejsdalsvej.

G.0.12 Test



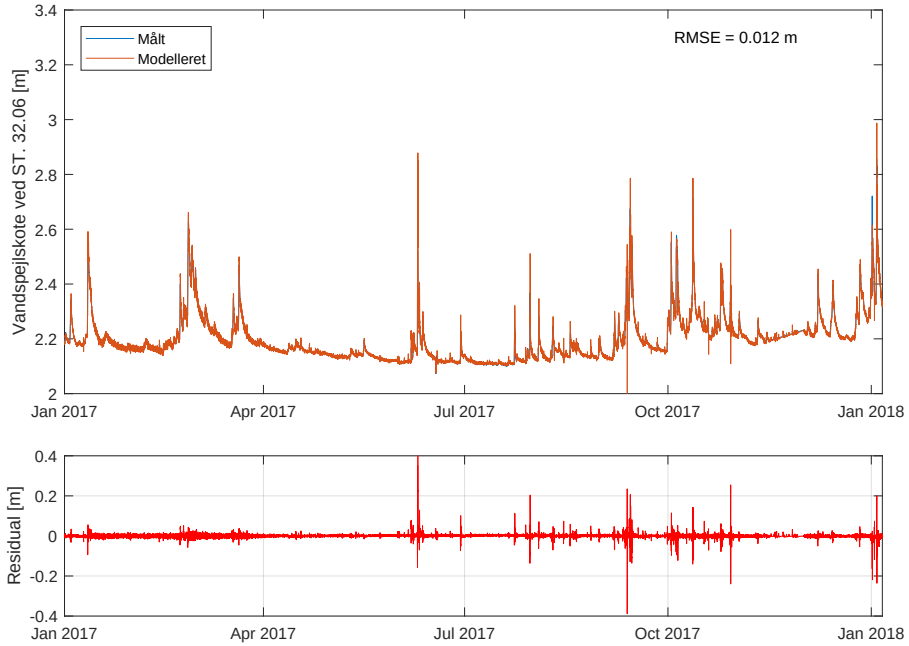
Figur G.12: Testperiode for 60 minutters forudsigelse ved Ny Grejsdalsvej.

G.0.13 Træning/validering



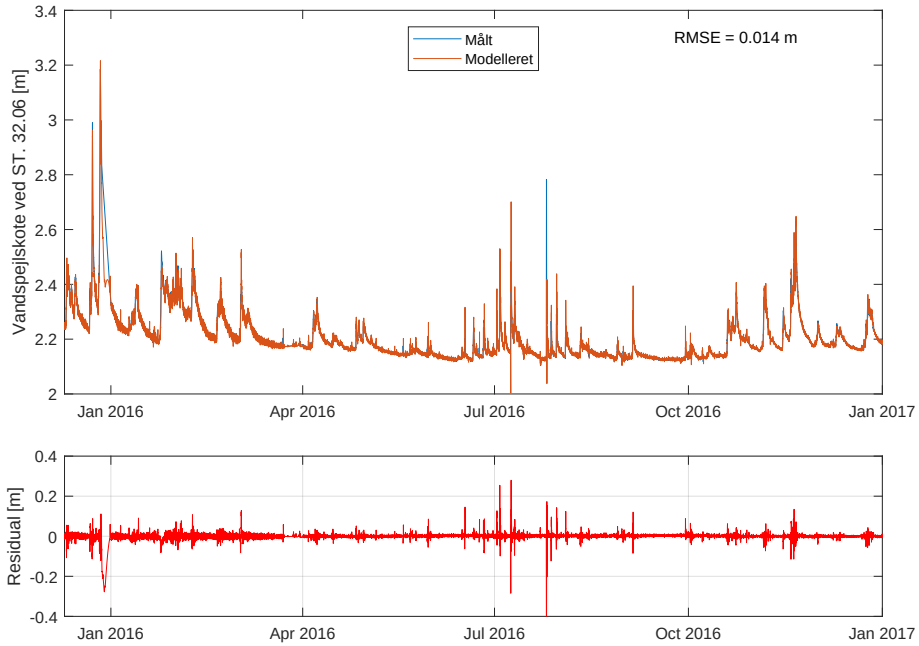
Figur G.13: Trænings-/valideringsperiode for 120 minutters forudsigelse ved Ny Grejsdalsvej.

G.0.14 Test



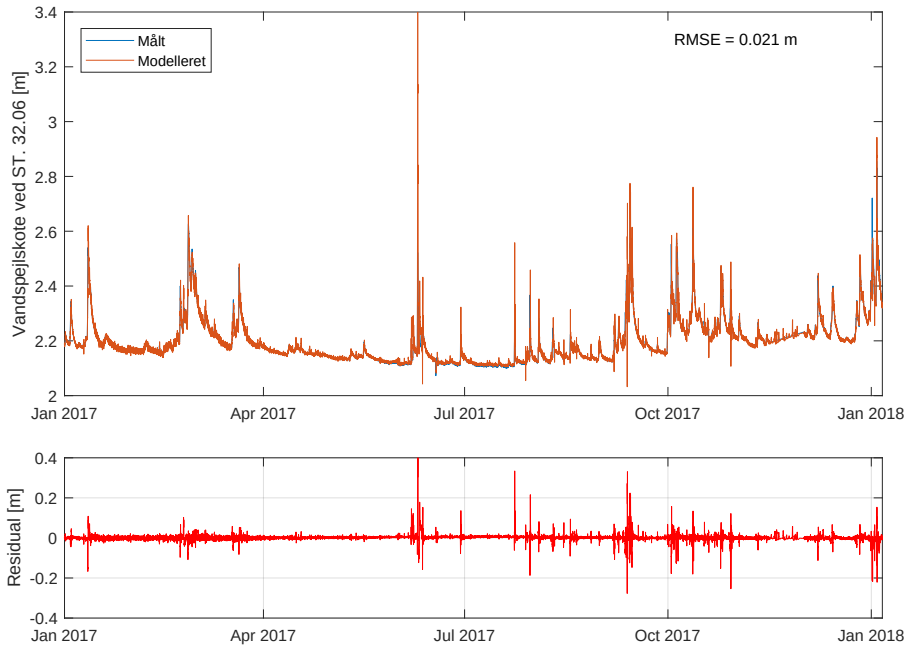
Figur G.14: Testperiode for 120 minutters forudsigelse ved Ny Grejsdalsvej.

G.0.15 Træning/validering



Figur G.15: Trænings-/valideringsperiode for 240 minutters forudsigelse ved Ny Grejsdalsvej.

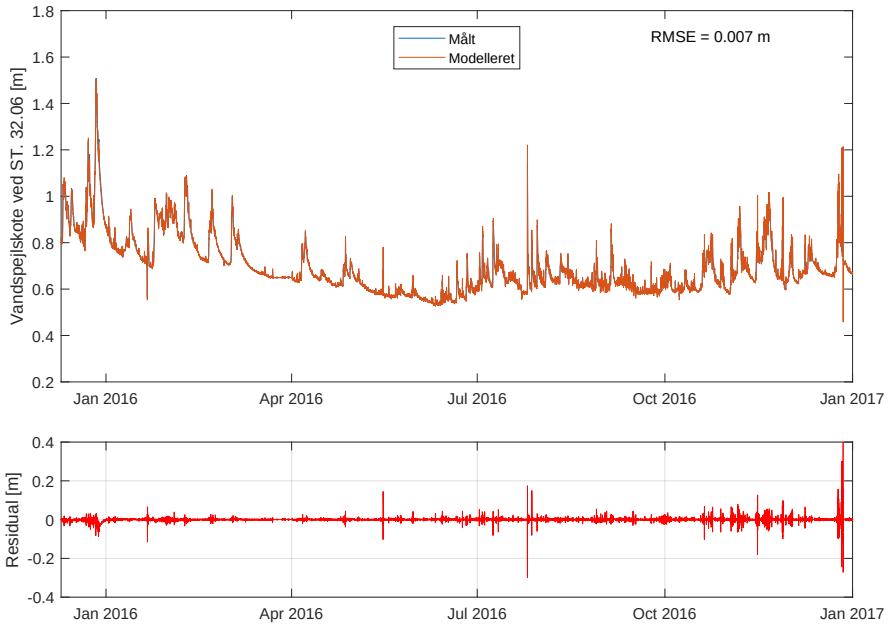
G.0.16 Test



Figur G.16: Testperiode for 240 minutters forudsigelse ved Ny Grejsdalsvej.

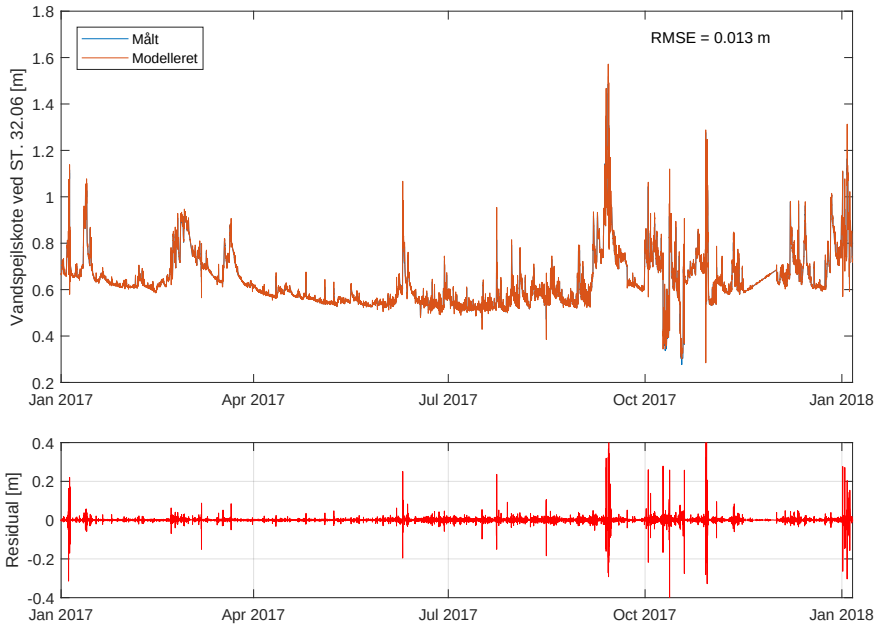
Modelresultater: Omløbsåen ved Aagade

G.0.17 Træning/validering



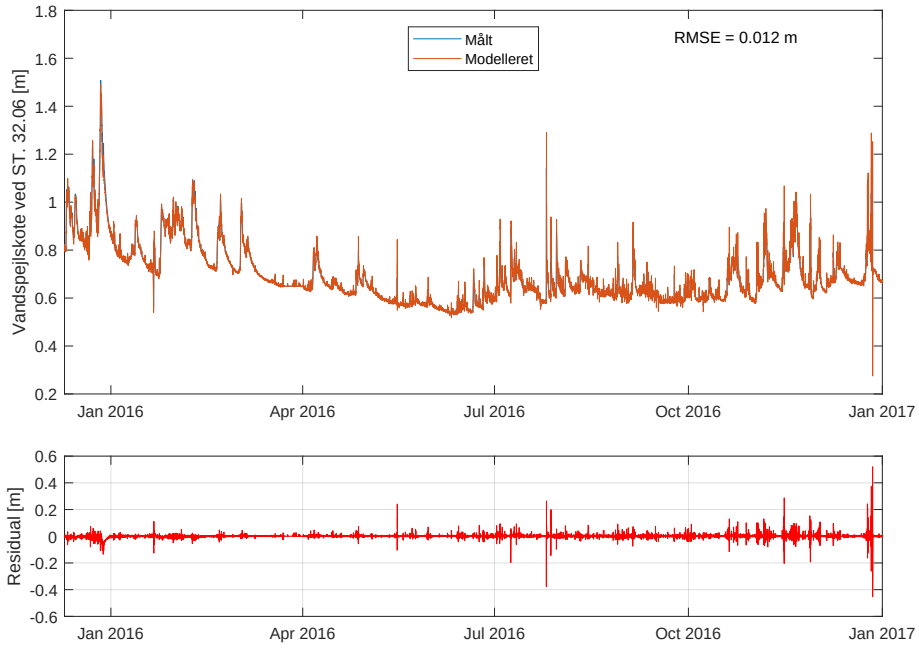
Figur G.17: Trænings-/valideringsperiode for 30 minutters forudsigelse ved Aagade.

G.0.18 Test



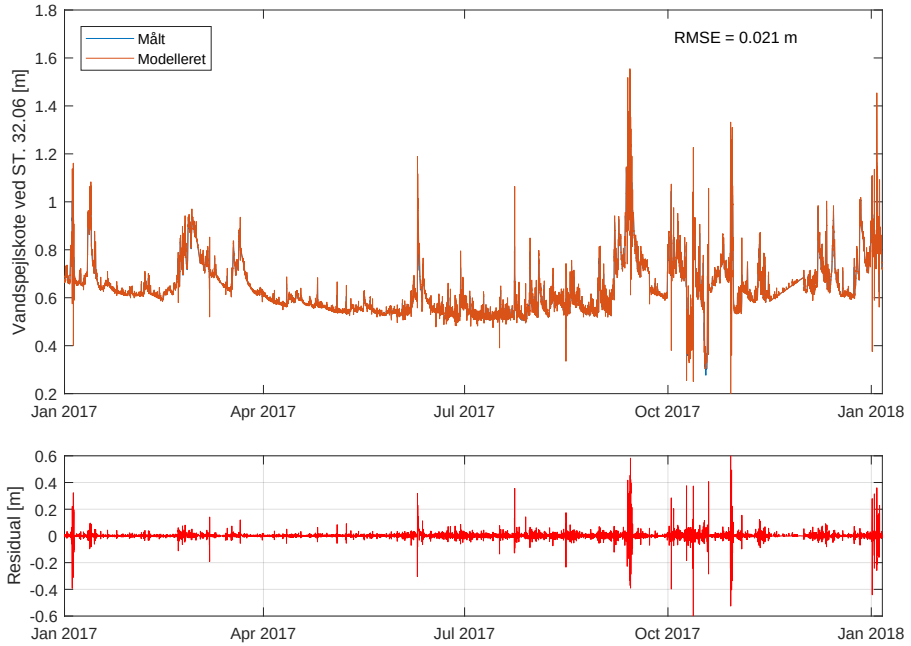
Figur G.18: Testperiode for 30 minutters forudsigelse ved Aagade.

G.0.19 Træning/validering



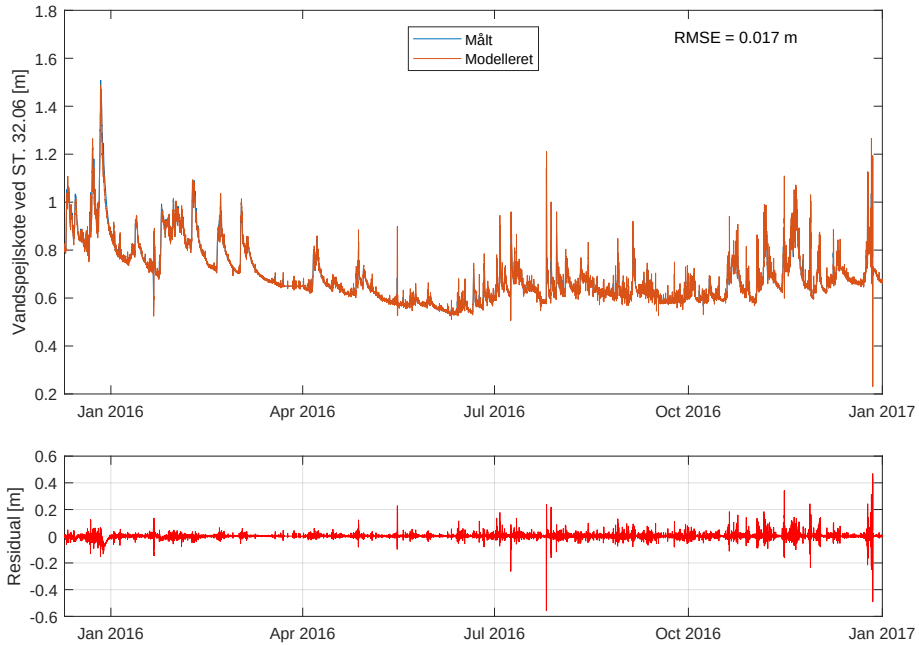
Figur G.19: Trænings-/valideringsperiode for 60 minutters forudsigelse ved Aagade.

G.0.20 Test



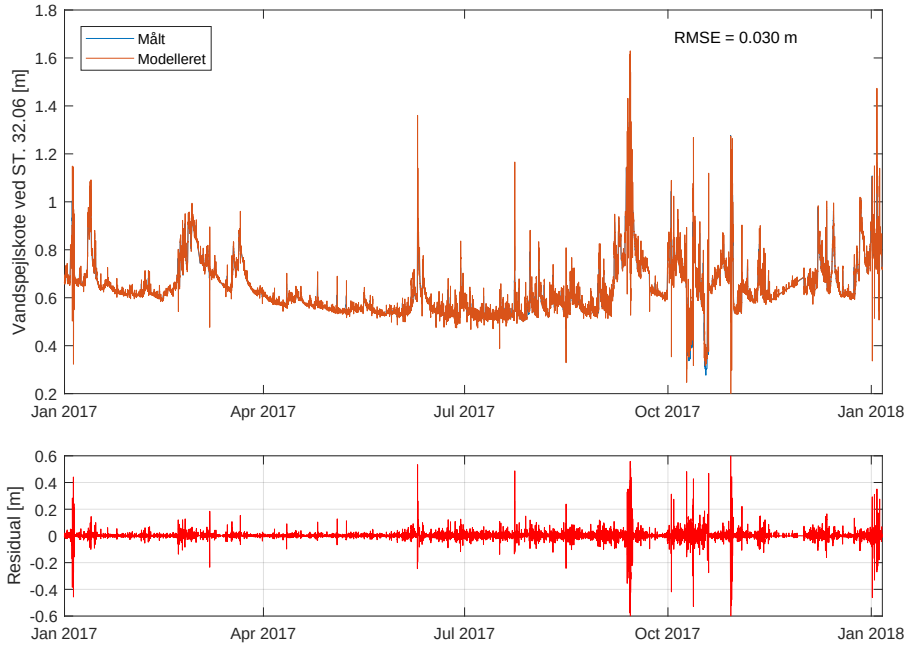
Figur G.20: Testperiode for 60 minutters forudsigelse ved Aagade.

G.0.21 Træning/validering



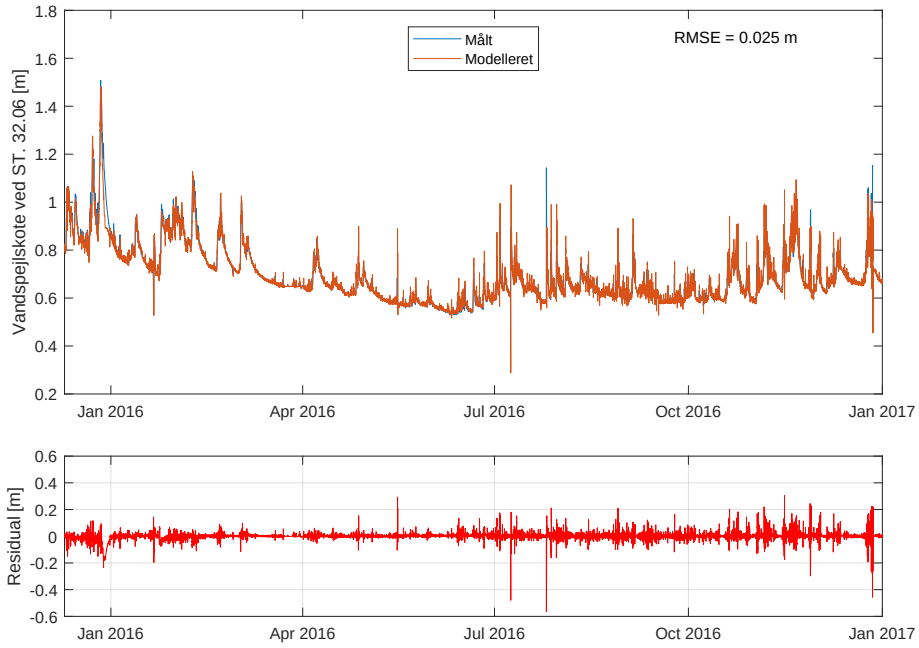
Figur G.21: Trænings-/valideringsperiode for 120 minutters forudsigelse ved Aagade.

G.0.22 Test



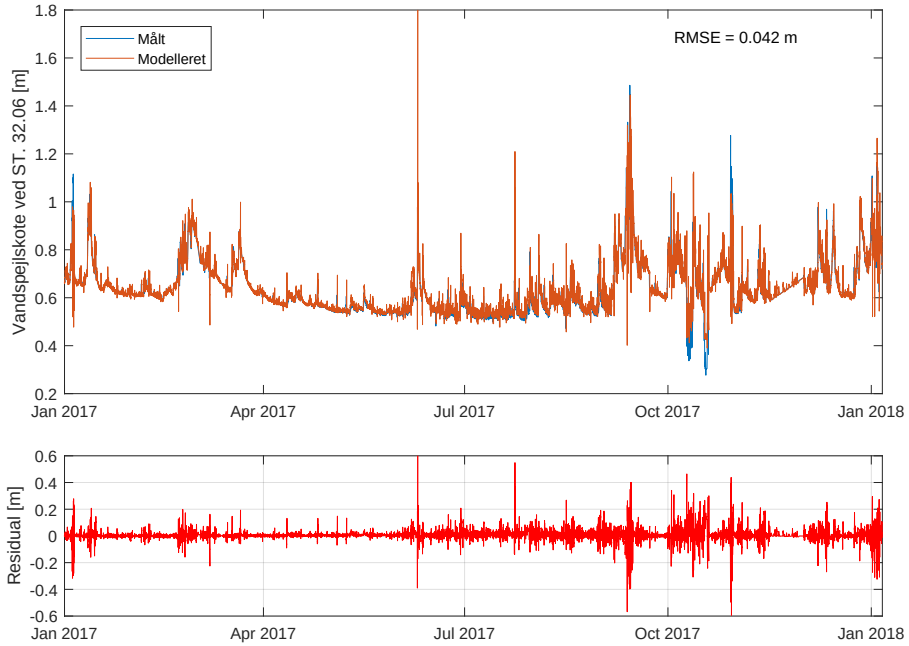
Figur G.22: Testperiode for 120 minutters forudsigelse ved Aagade.

G.0.23 Træning/validering



Figur G.23: Trænings-/valideringsperiode for 240 minutters forudsigelse ved Aagade.

G.0.24 Test

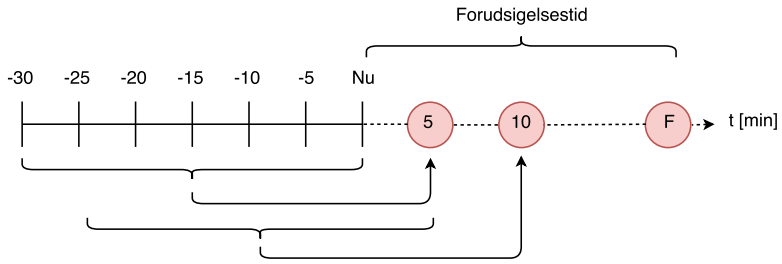


Figur G.24: Testperiode for 240 minutters forudsigelse ved Aagade.

H. NAR - et alternativ til tendens-modellen

Da tendens-modeller giver gode resultater, er det forsøgt at anvende et neuralt netværk, der på samme måde tager højde for den foregående tendens. Et sådan netværk kaldes et "Nonlinear autoregressive (NAR) neural network".

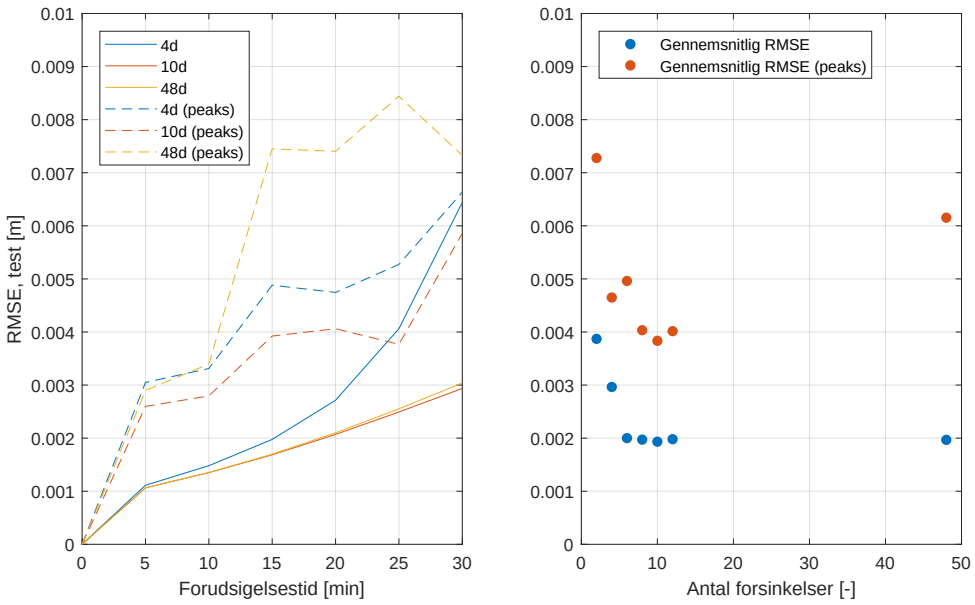
Da tendens-modellen giver det bedste resultat ved de korte forudsigelsestider, er det forsøgt om en forudsigelse udelukkende på baggrund af vandstandsmålinger fra stationen selv kan forbedres ved anvendelse af et neuralt netværk. Typen af netværk, som forudsiger den næste værdi i en tidsserie på baggrund af tidligere værdier, kaldes for et "*Nonlinear autoregressive (NAR) neural network*". Princippet i brugen af NAR-netværket er illustreret på figur H.1. Det er illustreret, at værdien, som måles i det aktuelle tidspunkt og værdierne, som er målt hvert femte minut de sidste 30 minutter anvendes som input til at forudsige målingen 5 minutter senere. Anvendes den forudsagte værdi kan der herefter forudsiges yderligere 5 minutter frem osv.



Figur H.1: Princippet i brugen af NAR-netværket.

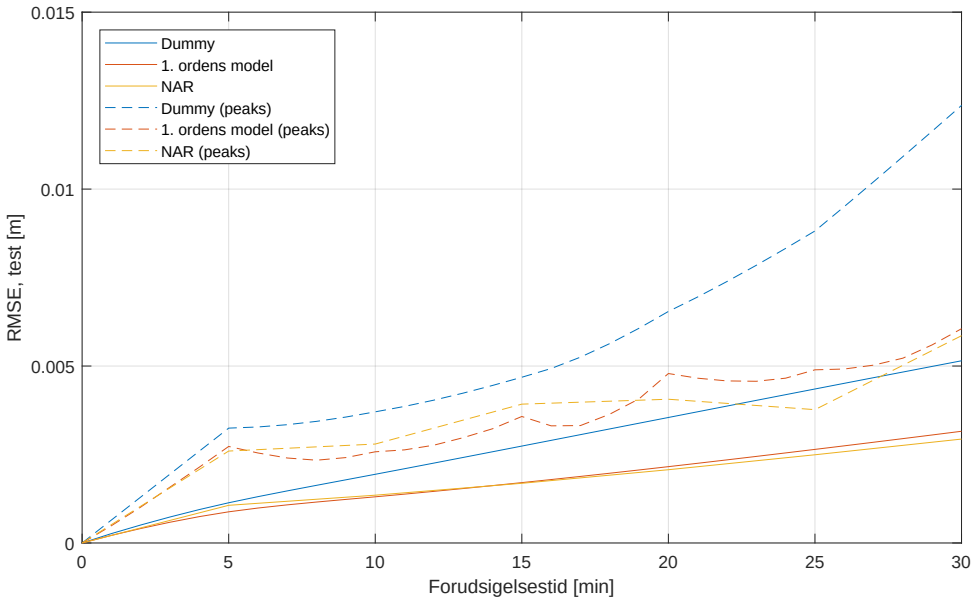
Ligesom ved tendens-modellen er det for NAR-netværket væsentligt at fastlægge hvilken tidsperiode, som medtages i analysen. På figur H.2 er det illustreret, at et antal forsinkelser i spændet 8-12 er passende. Dette svarer til målingerne for de sidste 35-55 minutter. Det er derudover fundet, at 10 neuroner er passende at anvende samt trænings- og valideringsperioder, som er beskrevet i 5.2. RMSE-værdien i peaks er generelt højere end for den fulde testperiode, men især i peaks er det vigtigt ikke at have tage for mange forsinkelser med (ses ved den høje gennemsnitligt RMSE-værdi

ved peaks ved anvendelse af 48 forsinkelser).



Figur H.2: RMSE-værdier ved brug af NAR-netværk med forskellige antal forsinkelser d (til venstre). Gennemsnitlig RMSE i intervallet med forudsigelestid på 0-30 minutter afhængig af antal forsinkelser d (til højre).

På figur H.3 ses, at NAR-netværkets kvalitet er meget lig tendens-modellen både generelt og i peaks. Det er derfor vurderet, at den bedste model er den, som i situationen er nemmest at anvende. Dette må forventes i de fleste tilfælde at være tendens-modellen på grund af dens simplicitet.

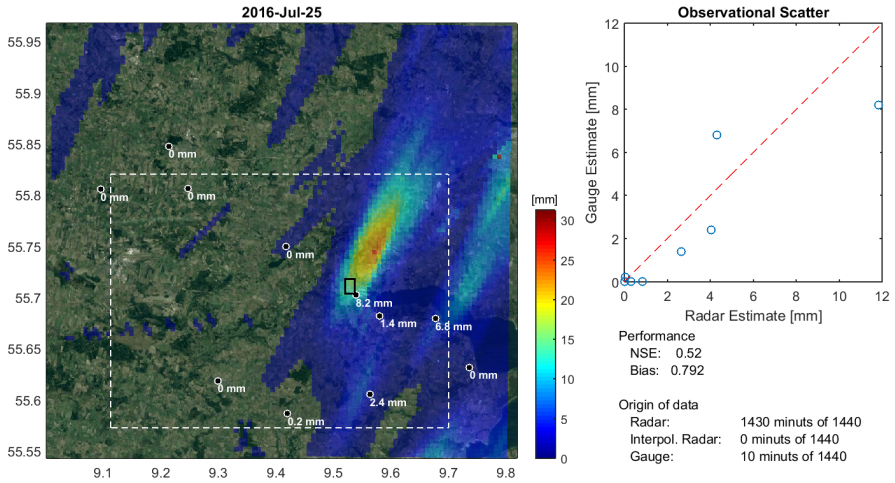


Figur H.3: Udviklingen i RMSE ved forøgelse af forudsigelsestiden, hvor også resultatet med NAR-netværket er vist.

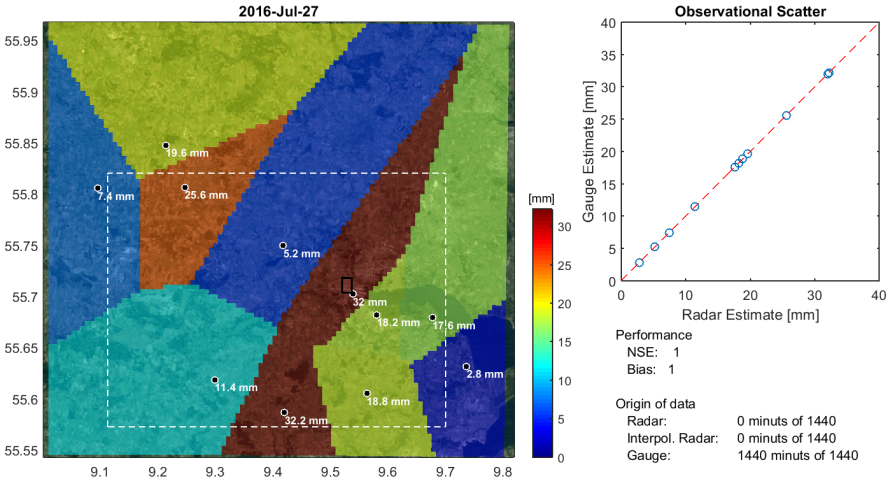
I. Radardata over Vejle Midtby

For at beskrive hurtig afstrømning fra Vejle Midtby til Omløbsåen er det forsøgt at inddrage radardata på den lokale nedbør over Vejle Midtby som input til et neuralt netværk. Dette tyder dog ikke på at forbedre forudsigelsen af vandspejlskoten. Det vurderes at skyldes tvivlsomt data, som beskrives i dette appendiks.

Den omtalte hændelse den 25/07-2016 og en hændelse to dage senere 27/07-2016 viser tydeligt udfordringerne ved anvendelse af radardata. Begge hændelser indikerer hurtig afstrømning fra Vejle Midtby. Hændelsen den 25/07 (figur 6.14) viser stor afstrømning fra Vejle Midtby, mens hændelsen den 27/07 (ikke illustreret) viser mindre afstrømning fra Vejle Midtby. Den akkumulerede nedbør over de to døgn er illustreret på figur I.1 og I.2. Her er billedet et andet, da "radardata"på trods af den lavere afstrømning den 27/07 viser markant mere nedbør over Vejle Midtby end den 25/07.

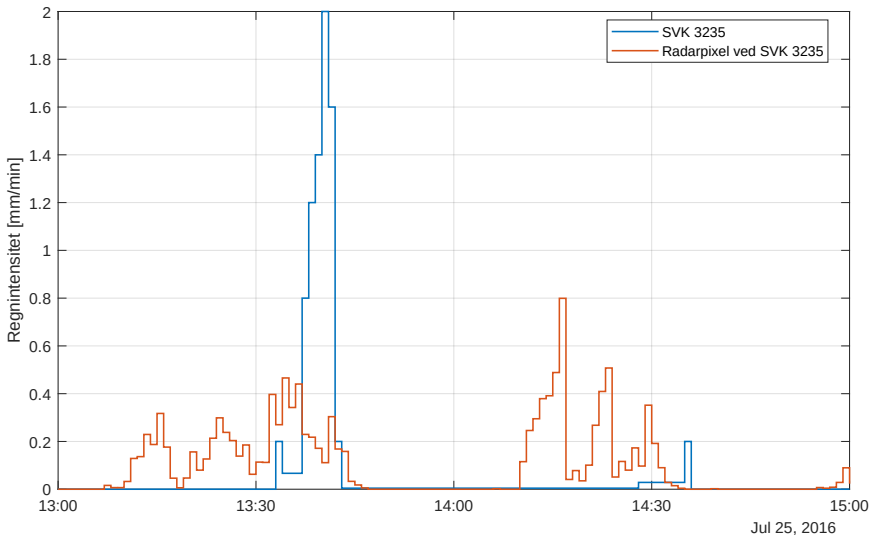


Figur I.1: Radardata fra 25/07-2016. Vejle Midtby er markeret med en sort firkant. (Billede udleveret af Jesper Ellerbæk Nielsen)



Figur I.2: Radardata fra 27/07-2016. Vejle Midtby er markeret med en sort firkant. (Billede udleveret af Jesper Ellerbæk Nielsen)

Det ses på figur I.2, at der ikke forefindes radardata fra den pågældende dato, og at radardataene derfor er baseret på regnmålere. Regnmåleren ved Vejle Centralrenseanlæg (SVK 5235), som er blevet anvendt ved Vejle Midtby, har formodentlig ikke været repræsentativ for Vejle Midtby, da der ikke ses en kraftig respons ved vandspejlskoten i Omløbsåen.



Figur I.3: Der er stor forskel på fordelingen af nedbørsintensiteter målt med regnmåleren og radaren. Det kan derfor diskuteres om en kalibrering af radardata op mod regnmåleren giver mening i dette tilfælde.

Det kunne potentielt give en bedre forudsigelse af vandspejlskoter, hvis perioder, hvor radardataene er baseret på regnmålere, frasorteres. Det vurderes dog på baggrund af vandspejlsstigningerne, at radardataene for hændelsen den 25/07 heller ikke giver et realistisk billede på, hvad regnintensiteten reelt har været over Vejle Midtby, selvom radardataene for denne hændelse ikke er baseret på regnmålere. På figur I.3 sammenholdes den målte regnintensitet i SVK-måleren (SVK 5235) ved

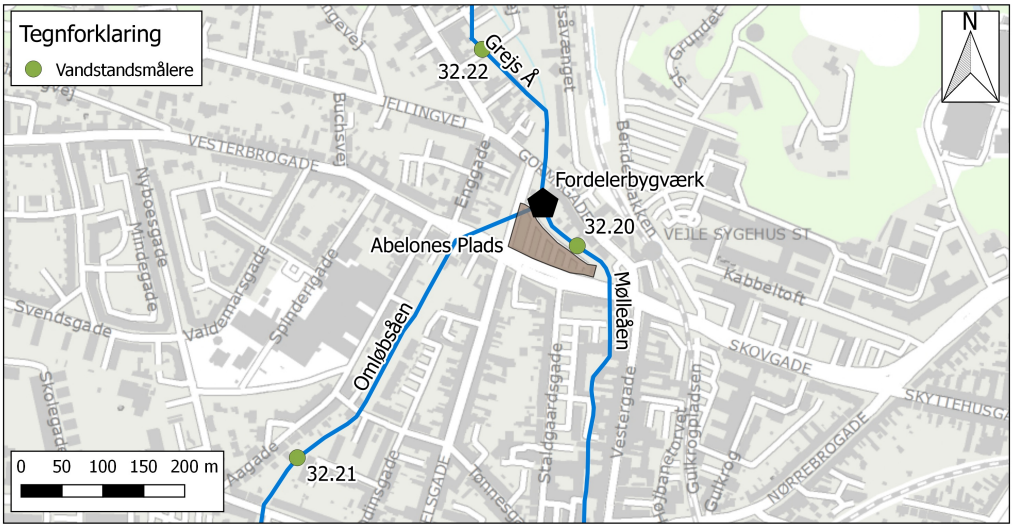
Vejle Centralrenseanlæg med den tilsvarende målte regnintensitet af radaren i samme pixel, hvor SVK-måleren er placeret.

Der ses en stor forskel på fordelingen af nedbørsintensiteter målt med regnmåleren og radaren. Forudsætningen for, at det giver fysisk mening at kalibrere radardata op imod regnmålere er, at de viser nogenlunde samme tendens over en hændelse. Dette er langt fra tilfældet her og kalibreringen kan dermed sagtens medføre, at radardataene giver et misvisende billede af de reelle nedbørsintensiteter.

J. Tolkning af data fra fordelerbygværket

Fordelerbygværket ved Abelones Plads har vist sig at have stor betydning for vandstanden i Omløbsåen. For at anvende data om spjældåbninger ved fordelerbygværket som input til de neurale netværk, forsøges det i dette appendiks at omskrive åbningsprocenter til et mere simpelt input.

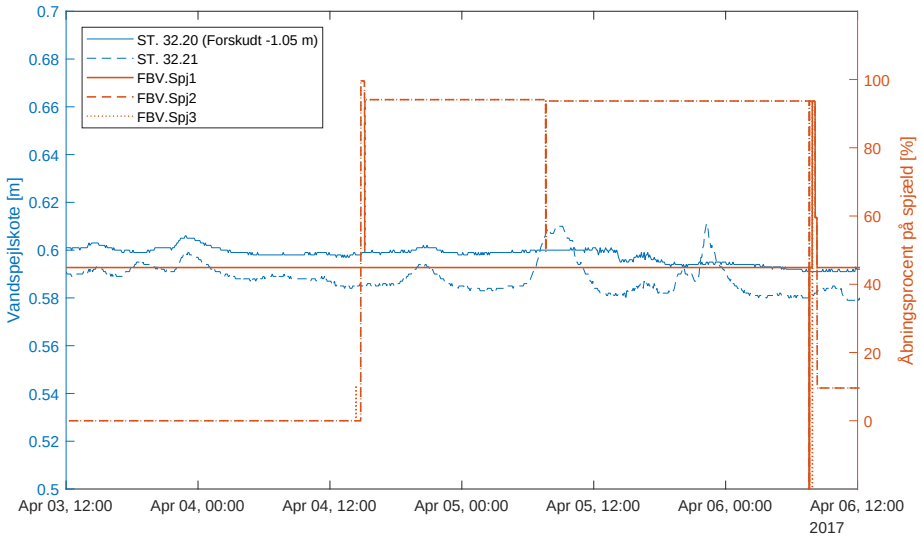
En af målsætningerne i projektet er at forudsige vandstanden i Omløbsåen (ST. 32.21), som ligger midt i risikoområdet. Da vandstanden i Omløbsåen i høj grad afhænger af afstrømningen fra oplandet, kræver det et neuralt netværk, som beskriver den afstrømning bedst muligt. Det er beskrevet i kapitel 3, at vandstandsmåleren i Grejs Å ved Ny Grejsdalsvej (ST. 32.22) dækker afstrømningen fra oplandet, og derfor er den målestation et godt input til modellering af vandstanden i Omløbsåen. Problematikken opstår, fordi denne afstrømning fordeles mellem Omløbsåen og Mølleåen, som vist på figur J.1.



Figur J.1: Grejs Å deles i Omløbsåen og Mølleåen ved Abelones Plads, og der haves vandstandsmålere i alle tre grene.

Fordelingen styres af et fordelerbygværk med tre spjæld, der kan regulere vandstanden i Omløbsåen og Mølleåen. Fordelerbygværket er skitseret på figur 2.4 i kapitel 2, hvor også placeringen af diverse målere er markeret.

Åbningsprocenterne af de tre spjæld kan justeres alt efter det ønskede niveau i Omløbsåen og Mølleåen. Spjældene lukker fra oven og ned, og derfor vil en reducering af åbningsprocenten ikke altid have en effekt på vandføringen i de to vandløb, hvis spjældene ikke har ramt vandoverfladen endnu. Det konstateres ved flere tilfælde, eksempelvis i begyndelsen af april 2017 som vist på figur J.2. Figuren viser også negative åbningsprocenter på spjældene, hvilket forekommer ofte i datasættet. Det antages at være fejlmålinger.



Figur J.2: Når spjældenes åbningsprocenterne reguleres i et interval, hvor spjældene ikke er i berøring med vandoverfladen, har det ingen betydning for den resulterende vandstand i Omløbsåen (ST. 32.21). Det er især tydeligt i situationer med lav vandstand. Ydermere konstateres en afstikker til en negativ åbningsprocent, som må antages at være en fejlmåling.

Når åbningsprocenterne af de tre spjæld ved fordelerbygværket benyttes som input til et neuralt netværk erfares det, at ændringer af åbningsprocenter i et interval, som ikke medfører vandstandsændringer, er ugunstigt for netværket. Ydermere bidrager de til antallet af mulige kombinationer mellem vandstande og åbningsprocenten på spjældene, fordi alle spjæld kan have 1000 (måles med tre betydende cifre) forskellige indstillinger og dermed $1000 \cdot 1000 \cdot 1000$ forskellige kombinationer foruden inkludering af vandstandene, hvilket potentielt kan resultere i en voldsom overfitting. På den baggrund undersøges det, hvorvidt det er muligt at forsimple åbningsprocenterne.

Det ønskes af finde et brugbart input fra fordelerbygværket, hvilket kræver en forsimpling af de åbningsprocenter, der allerede er tilgængelige. Derfor oprettes en ny tidsserie FBV.Status, som består af fire indstillinger:

- 0 = Alle spjæld åbne.
- 1 = 1 spjæld lukket.
- 2 = 2 spjæld lukkede.
- 3 = Alle spjæld lukkede.

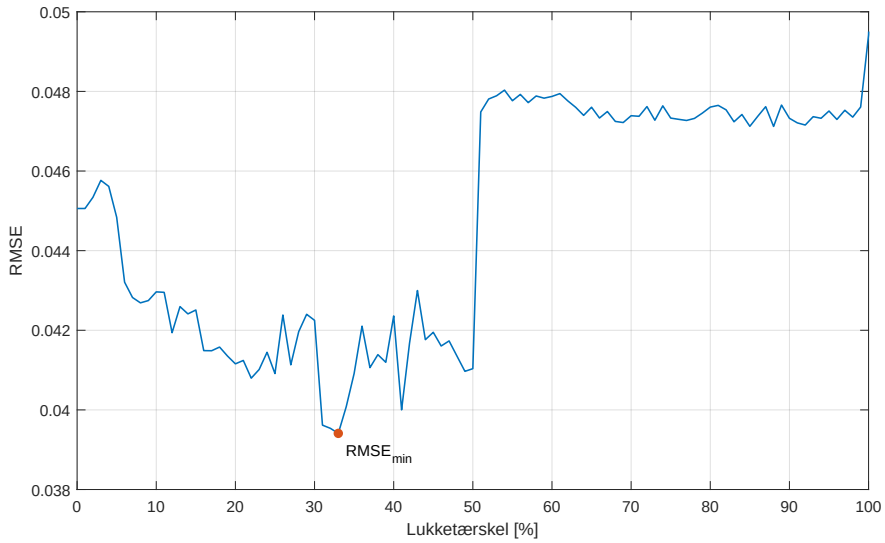
Hernæst undersøges, hvilken åbningsprocent der kræves, før et spjæld antages lukket. Det gøres på baggrund af optimering af et neuralt netværk til modellering af vandstanden i Omløbsåen (ST. 32.21), hvis netværkskonfigurationer fremgår af tabel J.1.

Tabel J.1: Netværkskonfiguration til modellering af Omløbsåen (ST. 32.21) på baggrund af vandstanden ved Ny Grejsdalsvej (ST. 32.22), i Vejle Havn (ST. 32.24) og åbningsstatus af spjældene ved fordelerbygværket.

Parameter	Konfiguration
Input	Ny Grejsdalsvej (ST. 32.22) Vejle Havn (ST. 32.24) Åbningsstatus af fordelerbygværk
Target	Omløbsåen (ST. 32.21)
Antal neuroner i det skjulte lag	10
Træning/validering	Alle uregulerede perioder fra 29/03-2017 → 06/01-2018 : 100% træning
Test	Ingen

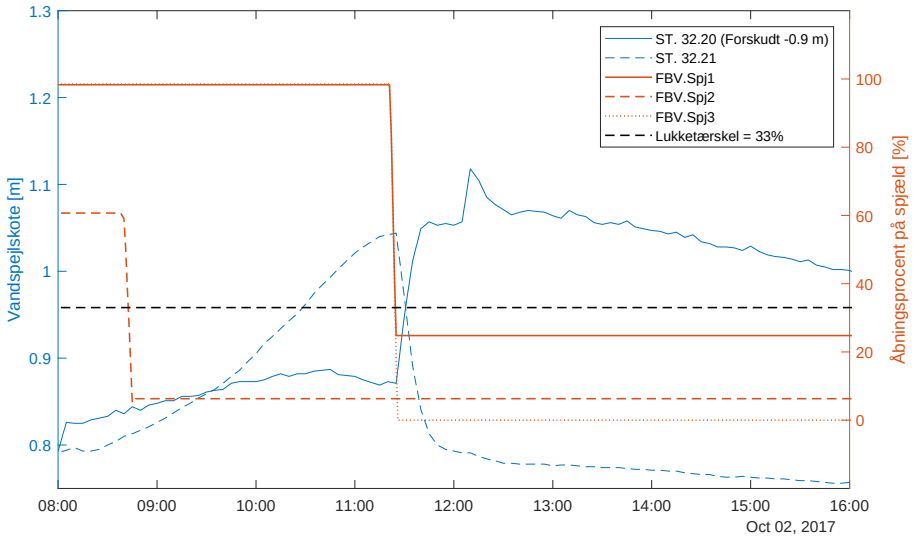
I tabellen fremgår det, at der trænes på alle uregulerede perioder. Med de perioder menes, at systemet har opnået en naturlig ligevægt efter en ændring i åbningsprocenten på spjældene og lukkeperioder ved sluseanlægget.

Netværket trænes på åbningsstatus af fordelerbygværket, som optimeres ved at ændre lukketærsklen mellem 0 og 100. Hernæst haves en optimal lukketærskel på 33%, som vist på figur J.3.



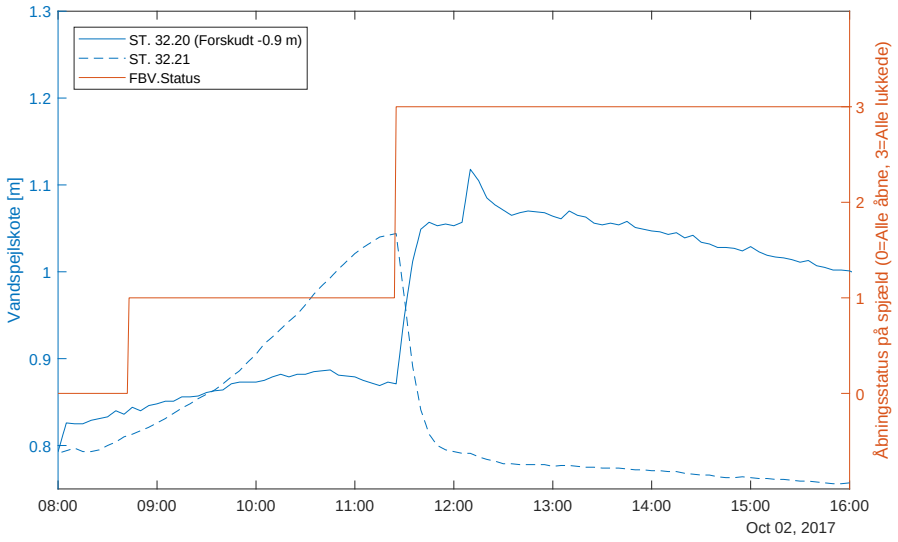
Figur J.3: Lukketærsklen justeres, indtil det neurale netværk giver den laveste RMSE. Det store spring i RMSE ved en åbningsprocent på 50 % skyldes, at spjældene har den indstilling i store dele af måleperioden. Derfor betyder en lukketærskel på 50 %, at alle spjæld er lukkede i 80 % af tilfældene, og alle spjæld er åbne i de resterende 20 %.

Der er et stort spring i RMSE, når lukketærsklen er over 50%. Det skyldes, at åbningsprocenten på spjældene ofte står i det niveau. Derfor antages spjældene lukkede i stort set hele perioden, hvilket svarer til slet ikke at have et input fra fordelerbygværket. Af det udledes også, at det overordnet set er en fordel at kende indstillingen på fordelerbygværket. Figur J.4 viser et eksempel på, at den optimale lukketærskel ikke er urealistisk, da der ses en tydelig ændring i fordelingen af afstrømning til Omløbsåen (ST. 32.21) og Mølleåen (ST 32.20).



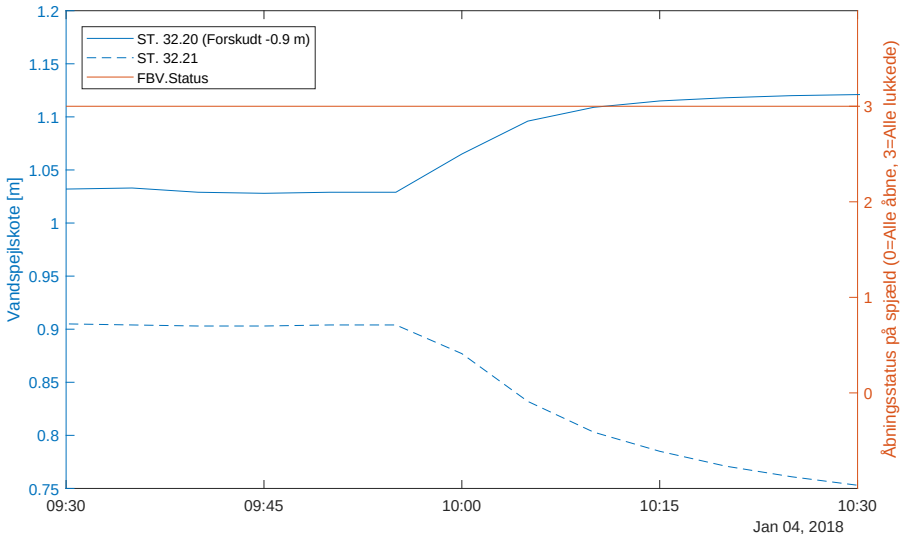
Figur J.4: Den optimale lukketærskel er optimeret på baggrund af modelresultater, men den virker realistisk, da der ses tydelige ændringer i fordelingen af afstrømning til hhv. Omløbsåen (ST. 32.21) og Mølleåen (ST. 32.20), når åbningsprocenterne passerer tærsklen.

Af det haves det nye forsimplede fordelerbygværksinput, som nu består af én serie fremfor tre separate serier med åbningsprocenter. Figur J.5 viser, hvordan det nye input ser ud sammenlignet med det rå signal på figur J.4.



Figur J.5: Her ses fortolkningen af åbningsstatus på baggrund af lukketærsklen.

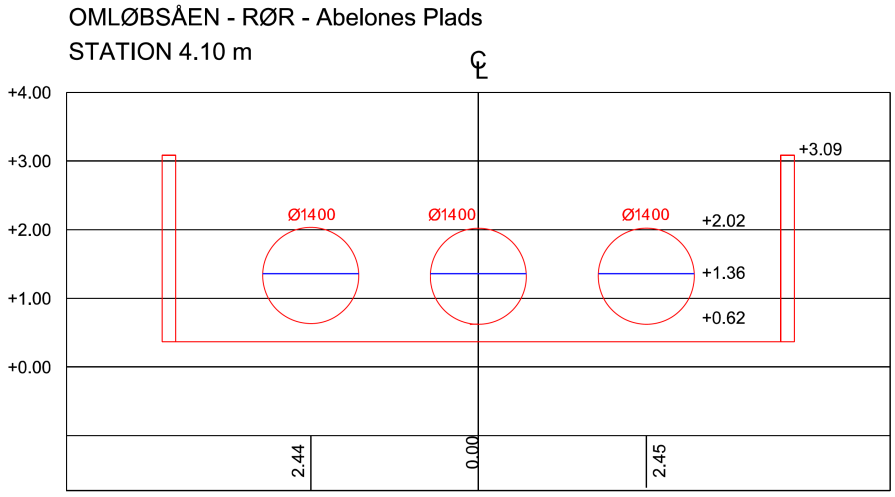
Når åbningsprocenterne forsimples på den måde ved brug af en lukketærskel, har det nogle faldgrupper. Figur J.6 viser, at der i en periode, hvor alle spjæld antages lukkede, stadig sker ændringer i fordelingen mellem Omløbsåen (ST. 32.21) og Mølleåen (ST. 32.20), som er forårsaget af ændringer i åbningsprocenten.



Figur J.6: I denne situation ses det, at lukketærsklen er sat for højt, og at alle spjæld antages lukkede i hele perioden, selvom det ikke er tilfældet i realiteten.

At fordelerbygværket har indflydelse på vandstanden i perioder, hvor åbningsstatus ikke ændres, giver problemer i modellen. Det kan potentielt løses med et mere dynamisk input, som begrænses til perioder, hvor spjældene er under vandoverfladen.

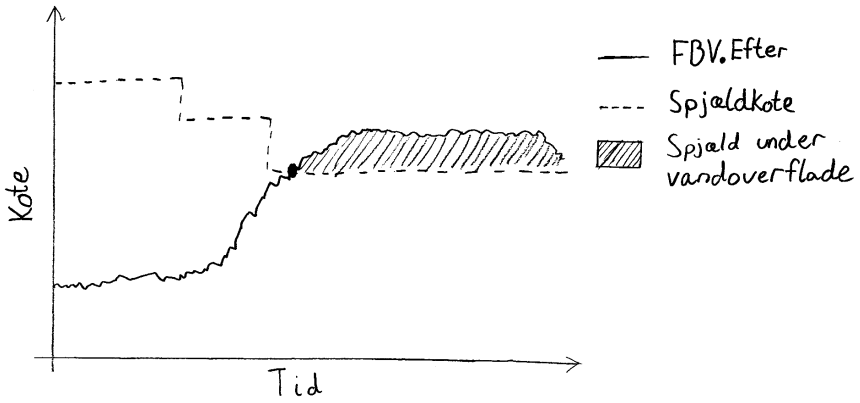
Af figur 2.4 fremgår det, at vandstanden måles ved fordelerbygværket efter ristene (FBV.Efter) umiddelbart før de tre spjæld. Det er en information, der potentielt kan sige noget om, hvorvidt spjældene har været i berøring med vandoverfladen. Tværsnittet på figur J.7 viser top- og bundkoten af de tre ledninger, der fører vandet fra fordelerbygværket og over i Omløbsåen. Det er disse rør, som blokeres af spjældene, når vandstandene i Vejle Midtby skal reguleres.



Figur J.7

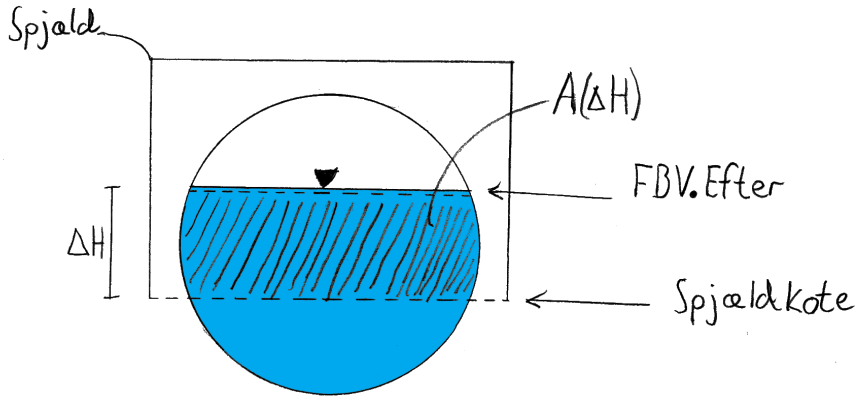
På den baggrund er det muligt at give et skøn på, hvilken kote spjældet har stået i. Det vides ikke med sikkerhed, hvordan åbningsprocenterne måles. Det mest korrekte ville være at måle åbningsprocenterne i forhold til rørene, hvilket vil sige, at spjældene er helt åbne ved kote +2,02 og helt lukkede ved kote +0,62. Hernæst kan åbningsprocenten enten være udtrykt ved det åbne areal,

eller lineært i forhold til top- og bundkoten. I det tilfælde vil krydsningspunktet mellem spjældkoten og vandspejlskoten i FBV.Efter betyde, at spjældet netop er i berøring med vandoverfladen, som skitseret på figur J.8.



Figur J.8: Det skraverede område efter krydsningspunktet mellem spjældkoten og vandspejlskoten (FBV.Efter) kan anvendes til at forbedre inputtet fra fordelerbygværket.

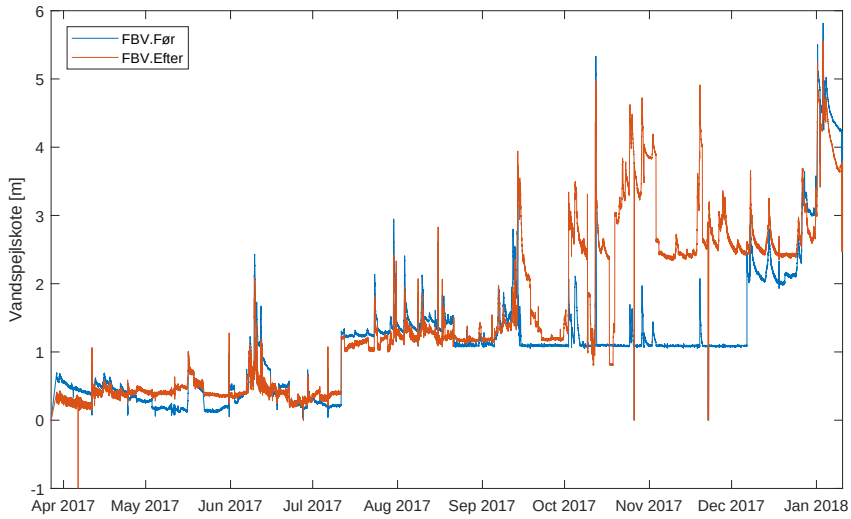
Det skraverede område efter krydsningspunktet fortæller, at spjældkoten er lavere end vandspejlskoten udenfor spjældene. Det er i den periode, at fordelerbygværket har en reel effekt og kan fordele afstrømningen mellem Omløbsåen og Mølleåen, hvilket sker gennem en reduktion af afstrømningen til Omløbsåen. Det kan eksempelvis beskrives ved en reduktion af det gennemstrømmede areal i rørene, som skitseret på figur J.9.



Figur J.9: Når vandstandskoten overstiger spjældkoten, kan inputtet eksempelvis beskrives som et arealinput.

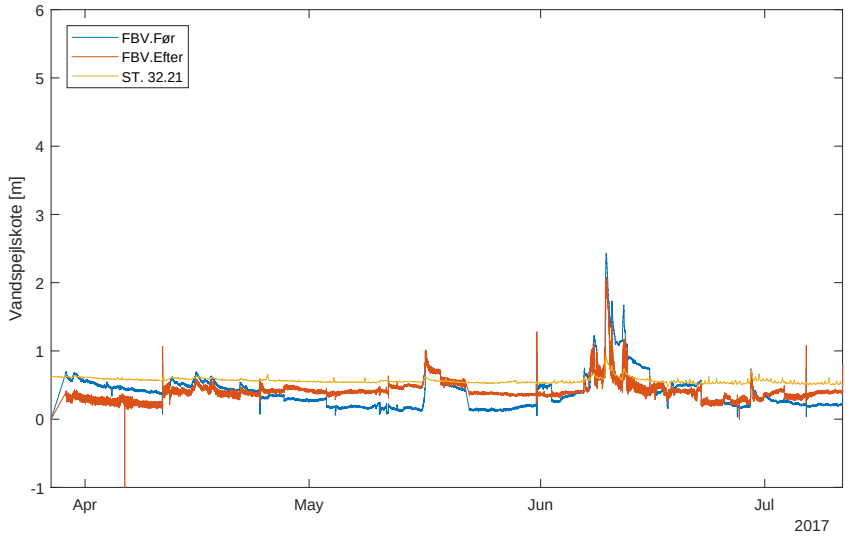
I Manningformlen (ligning 3.1) fremgår det, at et mindre gennemstømmet tværsnitsareal resulterer i en lavere vandføring gennem rørene. Selvom trykniveauet øges som følge af opstuvningseffekt, og energilinjegradienten dermed bliver større, så vurderes det ikke at kunne gøre op for reduktionen af det gennemstrømmede areal. Især fordi det netop er opstuvningseffekten, der leder vandet over i Mølleåen, så Omløbsåen aflastes.

Før analysen påbegyndes, undersøges kvaliteten af vandstandsmålinger ved fordelerbygværket før og efter ristene (FBV.Før og FBV.Efter). Det formodes, at vandstanden før risten bør være højest af de to, fordi den er længere oppe i systemet, men især på grund af enkelttab forårsaget af ristene, som er særligt udtalt, når de tilklogges af flydegods. De to vandstandsserier vises på figur J.10.



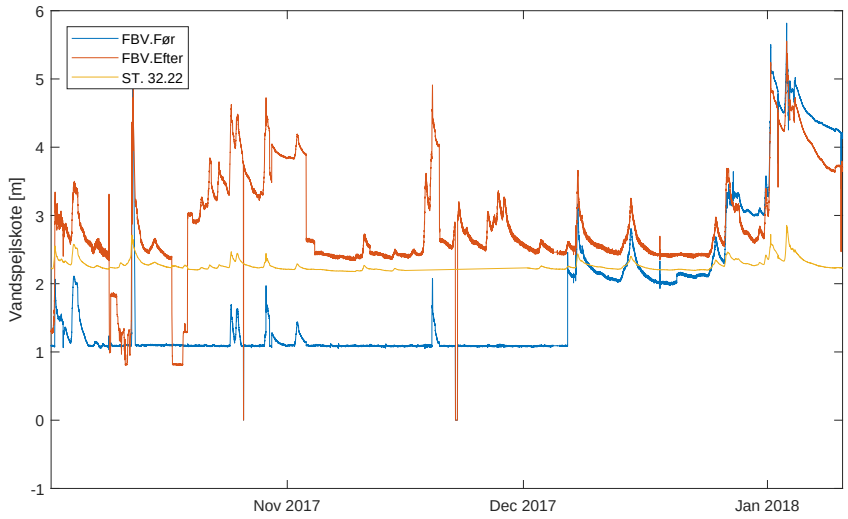
Figur J.10: Der forekommer nogle markante og systematiske spring i data, som ikke er realistiske.

Det ses, at vandstanden har nogle markante forskydninger. Der er en tydelig voksende tendens i vandspejlskoten, som formodes at være fejlagtig. I perioden april til juli 2017 ligger vandspejlskoten under nedstrøms vandstandsmåling i Omløbsåen (ST. 32.21), som vist på figur J.11. Derfor er det usandsynligt, at data i denne periode repræsenterer et realistisk billede af den egentlige vandspejlskote foran de tre spjæld.

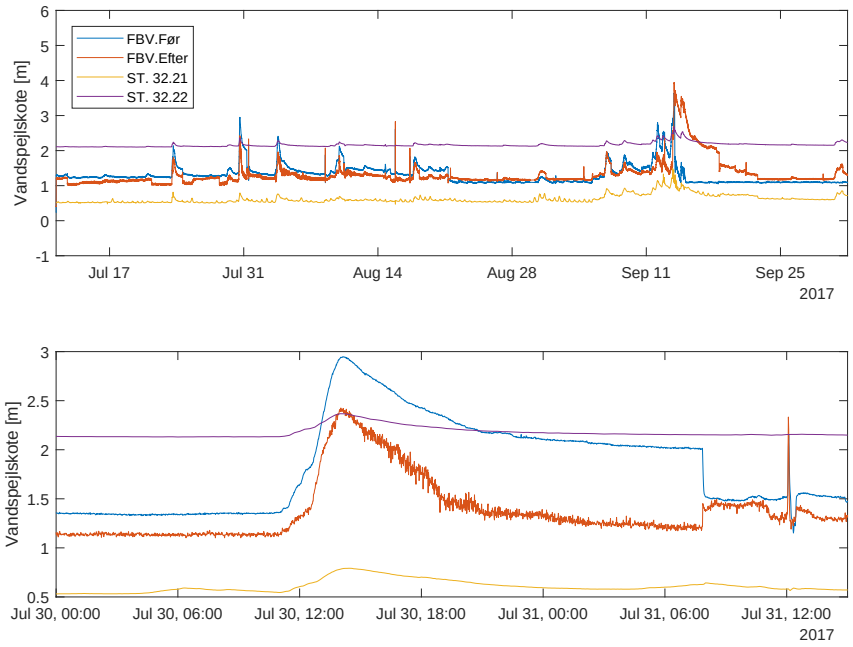


Figur J.11: I perioden april til juli 2017 ligger vandspejlskoten under nedstrøms vandstandsmåling i Omløbsåen (ST. 32.21)

Tilsvarende ses det på figur J.12, at vandspejlskoten ved fordelerbygværket i perioden oktober 2017 til januar 2018 ligger over vandspejlskoten ved Ny Grejsdalsvej (ST. 32.22), som er en opstrøms vandstandsmåler. Ydermere konstateres det, at vandstandsmåleren før risten (FBV.Før) i flere sammenhængende perioder måler noget markant andet end vandstandsmåleren efter risten (FBV.Efter), hvilket ikke kan forklares naturligt.



Figur J.12: Vandstandsmåleren efter risten (FBV.Før) måler en højere vandspejlskote end opstrøms måler ved Ny Grejsdalsvej (ST. 32.22), hvilket er urealistisk. Ydermere konstateres fejlmålinger af vandstanden før risten (FBV.Før) i lange sammenhængende perioder.



Figur J.13: Perioden juli til oktober 2017 viser til dels et realistisk billede af vandstanden før spjældene, og eksemplet d. 30.-31. juli viser et eksempel herpå, og hvordan rensning af riste også påvirker vandstanden i Omløbsåen (ST. 32.21).

Tilbage haves perioden juli til oktober 2017, som potentielt set godt kunne være realistiske målinger. På figur J.13 er hele perioden vist sammen med et eksempel nedenfor. Sammenlignet med andre anvendte data er der relativt meget støj især på måleren, der er placeret efter risten, hvilket formentlig skyldes turbulens. Eksemplet viser en afstrømningshændelse 30. til 31. juli 2017, hvor korrelationen

mellem de to målere er fin. Forskellen mellem dem er mindre før hændelsen, en den er bagefter, hvilket kan skyldes akkumulering af flydegods ved risten som følge af forhøjet vandføring. Om morgenen d. 31 juli ser det ud til, at ristene renses, hvilket faktisk øger vandstanden i Omløbsåen (ST. 32.21) en smule. Det er en anden detalje, som i givet fald vil komme med i et input formuleret ved en reduktion af det gennemstrømmede areal.

Det er tale om et enkeltstående tilfælde og kvaliteten af målingerne vurderes overordnet dårligt, og frasortering af urealistiske målinger vil medføre et meget begrænset datagrundlag til en sommerperiode, hvor vandstandende generelt er ukritiske. På det grundlag vurderes det, at den bedste mulighed for inddragelse af fordelerbygværket i modellerne ligger ved brug af det forsimplede input eksemplificeret på figur J.5, selvom der ligger et potentiale, hvis datagrundlaget ved fordelerbygværket forbedres og forøges.