

Silkeborg Kommune

Proof-of-Concept: Gudenå Machine Learning

Rapport

4. maj 2022

(revision 1)

Udarbejdet til:
Silkeborg Kommune
Søvej 3
8600 Silkeborg

Udarbejdet af:
EnviDan A/S
David Getreuer Jensen, Ørjan Heggdal,
Jonas Folke Nielsen, Simon Reuss Rahbek
E-mail: dgj@envidan.dk
Direkte tlf.: 46473137
Projektnavn: Proof-of-Concept: Gudenå Machine
Learning Prognose
Projektnr.: 1223114

EnviDan

Indholdsfortegnelse

1.	Indledning og konklusion	3
2.	Kort beskrivelse af Gudenå og datagrundlag	4
3.	Opsætning af ML-model	5
3.1	Datagrundlag.....	5
3.2	Opbygning af model.....	10
4.	Resultater.....	11
4.1	Metrikker	12
4.2	Performance	12
4.3	Opsummering	39
5.	Diskussion.....	39
5.1	Fordele og ulemper med ML-model kontra numerisk model	39
5.2	Beregning af vandstand nedstrøms Tangeværket eller andre opstemninger	40
5.3	Tidsserier fra målestationer	40
5.4	Silkeborg Langsø, Ringvejsbroen	40
5.5	Hybridmodellering	41
5.6	Driftsomkostning for produktionsmoden ML-prognosemodel.....	42
6.	Afklaring af spørgsmål	42
1.	Bredstenbro	45
2.	Langsø, Ringvejsbroen	46
3.	Kongensbro	47
4.	Ulstrup	49
5.	Peaks	52
5.1	Bredstenbro	52
5.2	Klostermølle	53
5.3	Langsø	54
5.4	Kongensbro.....	55

Bilagsfortegnelse

Bilag 1	Resultater fra ML-modeller
Bilag 2	Resultater for peaks

1. Indledning og konklusion

EnviDan har for Silkeborg Kommune undersøgt muligheden for at benytte Machine Learning (ML) til at forudsige vandstanden i Gudenå op til 72 timer frem i tiden.

Opgavens hovedformål er at:

- Redegøre for muligheden for at opstille en automatisk ML model for dele eller hele Gudenå med løbende 72 timers forecast af vandstanden.
- Foretage en performance sammenligning mellem en Machine Learning baseret prognose og fysisk baseret hydrodynamiske prognosemodel (DHI-model)
- Afsøgning/diskussion af, hvilke alternative muligheder der foreligger for forudsigelse af vandstand/flow (herunder "rene" Machine Learning modeller samt hybrid-modellering)

ML-modellen er opstillet for fem steder i Gudenå; Bredstenbro, Klostermølle, Langsø ved Silkeborg, Kongensbro og Ulstrup.

Vandstanden kan forudsiges ved Klostermølle med en 72 timers horisont med gennemsnitlig 1,8 cm afvigelse (bedste R^2 -værdi på 0,953). På peak-værdierne (de højeste vandstande, som er vigtige for Silkeborg Kommune at kunne forudsige) er afvigelsen gennemsnitlig på 1,9 cm ved Klostermølle. Ved Kongensbro er performance også tilfredsstillende og ML modellen kan tydeligt vise tendenserne i vandstanden (om den stiger eller falder meget eller lidt). På peak-værdierne er afvigelsen gennemsnitlig på 5,1 cm.

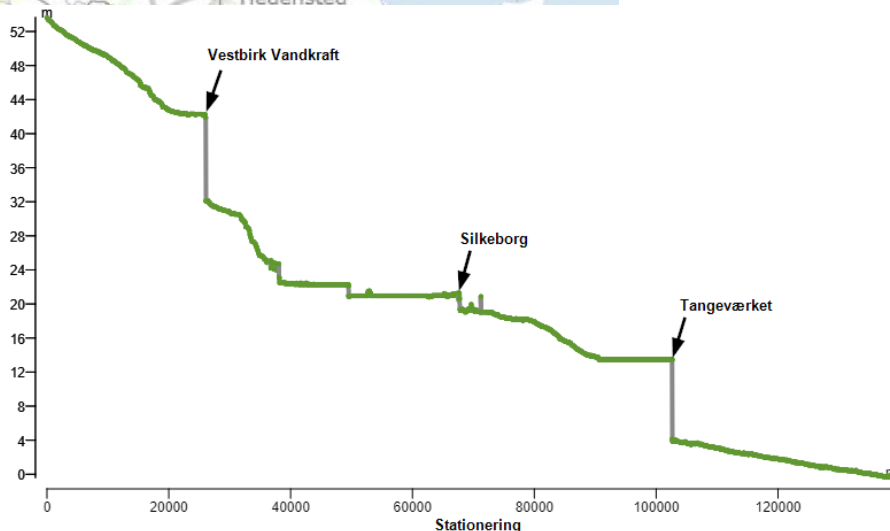
Det har vist sig vanskeligt at forudsige vandstanden nedstrøms opstemninger fx Tangeværket, da der ikke er en logisk styring af vandføringen gennem værket. Det bør undersøges nærmere, om der er andet datainput, som kan danne grundlag for en bedre prædiktion af vandstanden på denne strækning.

Også nedstrøms Silkeborgsøerne er prognosen fra ML-modellen fejlbehæftet, hvilket til dels skyldes manglende loggerdata opstrøms i systemet, samt et hydraulisk kompliceret system ved Silkeborg, som består af bl.a. flere søer, stemmeværk og sluse.

EnviDan vurderer, at en ML-model på størstedelen af Gudenå vil kunne levere en tilstrækkelig god performance til varslingsformål med en prognosehorisont på 72 timer. Nogle steder i Gudenå vurderes det, at man med fordel kunne anvende en hybridmodel mellem en fysisk model og ML-model. Det kunne være ved målestationer, der kun har begrænset mængde data eller manglende opstrøms informationer. En ML-model vil dog kontinuerligt blive mere og mere retvisende i takt med større datahistorik - dette er i øvrigt gældende generelt.

2. Kort beskrivelse af Gudenå og datagrundlag

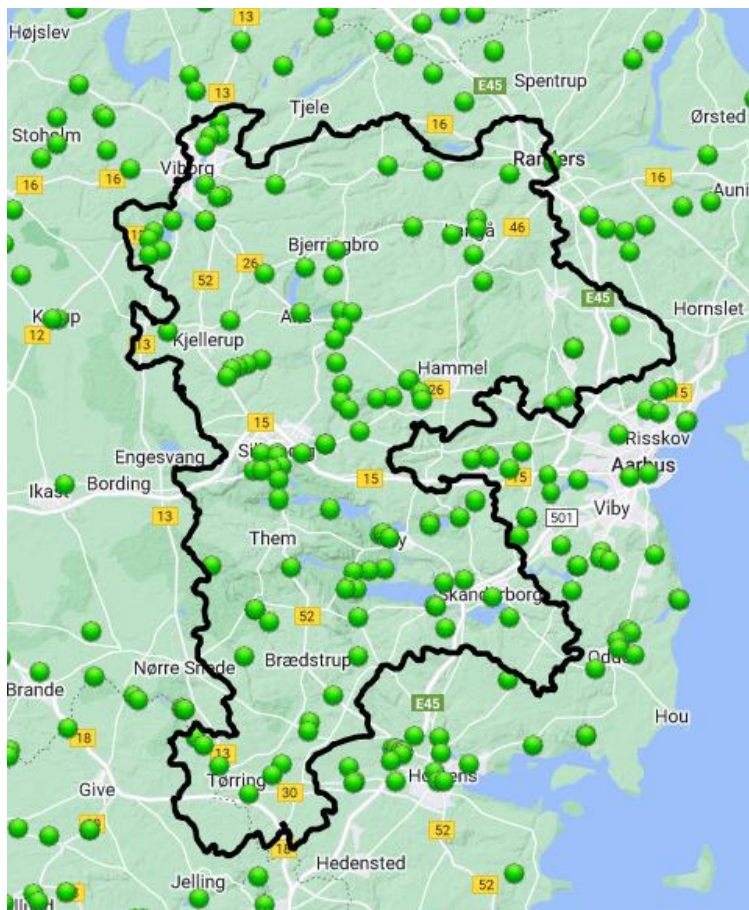
Gudenåen er Danmarks længste vandløb på 149 km og afvander 2600 km², se Figur 1. Udspringet er ved Tørring og udløbet sker til Randers Fjord. Den hydrauliske dynamik ændrer sig meget ned gennem vandløbet, hvor nedbør og grundvandet bidrager forskelligt til de enkelte strækninger. Derudover, så passerer vandet mange større søer, talvise opstemninger og andre reguleringer, som påvirker vandets naturlige hydraulik.



Figur 1 Gudenåens opland og største strømningsveje, samt længdeprofil af vandløbet fra Tørring til Randers. (Kilde: Scalgo Live og Styrelsen for Dataforsyning og Effektivisering (SDFE))

I Gudenåen findes mange loggere til registrering af vandstand og vandføring. Loggerne giver godt datagrundlag for forståelsen af dynamikken og for opbygning af en model, hvadenten det er en numerisk - eller ML-model. Figur 2 viser alle loggere i området, dog er der stor forskel på, hvor lang tid loggeren har målt. En del af loggerne blev nedlagt i 2007 og flere er først genstartet eller

opstartet i 2018. Det skal altså undersøges nærmere, hvor meget data der helt præcis er til rådighed i de forskellige dele af vandløbet.



Figur 2 Markering af vandstandsloggere i Gudenåens opland. (Kilde: Vandportalen.dk og SDFE)

3. Opsætning af ML-model

Vandstandsprognoserne, der udvikles i nærværende projekt, er baseret på machine learning. Machine learning består overordnet af flere grene, hvor der i dette projekt anvendes *Supervised Learning*. Supervised Learning er defineret ved, at kendte historiske værdier anvendes til at træne en algoritme til at forudsige ukendte fremtidige værdier. Dette gøres ved at tilpasse modellens vægte (parametre), indtil modellen er blevet tilpasset i så tilstrækkelig grad, at den kan trække nok information ud af tilgængelige data, til at prædiktere utilgængelige data.

3.1 Datagrundlag

Ved anvendelse af Supervised Learning er det essentielt at have et godt datasæt til at træne sin model. Et godt datasæt i konteksten af vandstandsprognoser i Gudenåen, indebærer for det første, at det indeholder de rigtige features (variable) og for det andet at datamængden er tilstrækkelig.

Det er vigtigt at datasættet indeholder de rigtige features, da det er disse, der giver signalerne til modellen om, hvordan vandstanden varierer. Datamængden er essentiel i den forstand, at modellen lærer af historikken, det er derfor afgørende for modellens performance, at den har nok historiske data, til at modellen bliver repræsentativ for et stort spektrum af hændelser. Modellen skal så at sige lære at 'kende' kraftige sommer- og vinterhændelser, hændelser, hvor det ikke regner, og alle

andre hændelser derimellem. Jo længere tidsserie med historiske data, jo bedre kan modellen forudsige fremtiden.

I det følgende forklares de benyttede data og evt. databehandling inden input i ML modellen.

3.1.1 Nedbørsdata

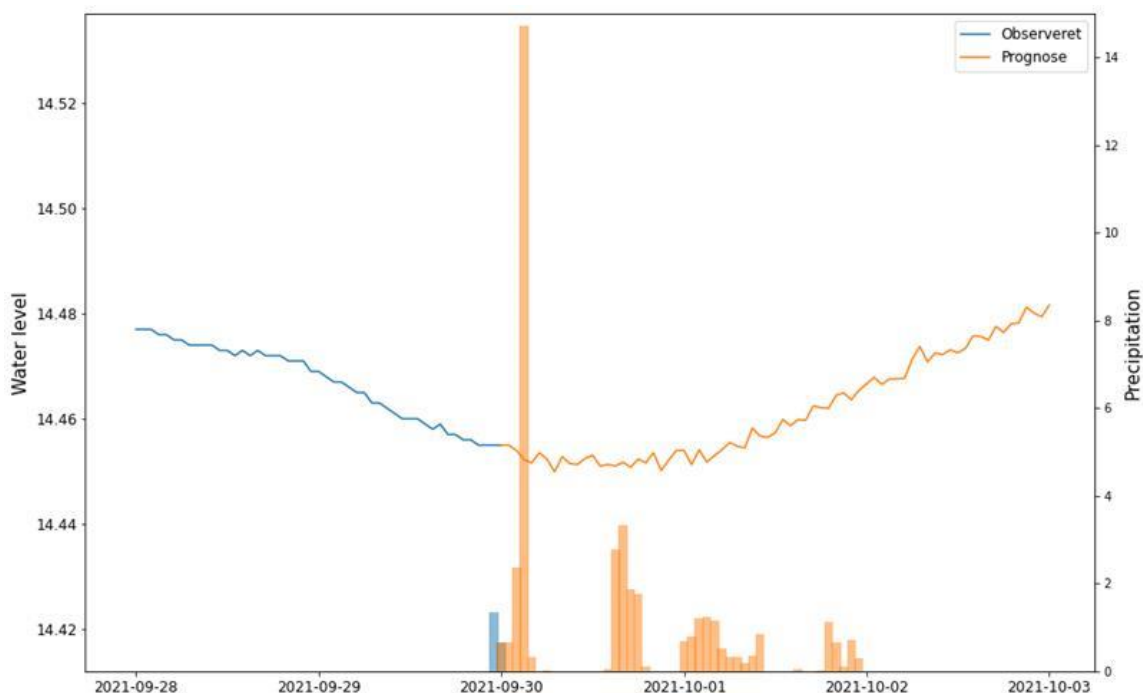
Formålet med modellen, er at forudsige vandstanden/vandstandsændringen op til 72 timer frem i tiden. Når der udvælges features (input data), tages udgangspunkt i forståelse af hydrologien i forhold til vandstandsændringer i Gudenåen. Her ses nedbørsforholdene som værende den primære forklaring.

Qua Gudenåens størrelse varierer dynamikken og responstiderne i systemet ned gennem vandløbet og netop responstiden er af væsentlig karakter, når det gælder prognoser for vandstand. Opstrøms i vandløbet er der fx kort tid fra nedbøren falder i oplandet til vandstanden stiger i vandløbet.

Punkter med kort responstid er afhængig af nyligt faldet nedbør, mens punkter med længere responstid først reagerer på nedbør efter en periode. Dette betyder, at en 72 timers prognose kan være afhængig af viden om nedbør, der endnu ikke er faldet på prognosticerings tidspunktet, ligesom den kan være afhængig af viden om nedbør der er faldet inden prognosticerings tidspunktet. Forklaret i tabellen:

Nedbørsobservationer	Nedbør, som er faldet inden tidspunktet for modelkørsel. Altså er der mere end 72 timers responstid fra nedbør til reaktion i vandløbet (vandstandsændring)
Nedbørsprognoser	Der er mindre end 72 timers responstid fra nedbør til reaktion i vandløbet. Den faktiske nedbør kendes ikke, og derfor bruges prognoser

Af Figur 3 fremgår et eksempel på en nedbørsprognose og de tilhørende nedbørsprognoser og -observationer.



Figur 3 Eksempel på nedbørs påvirkning af vandstand. Prognosen viser vandstanden 72 timer frem i tiden.

Foruden nedbørsdata benyttes også observerede vandstande fra før prognosticerings tidspunktet. For prognosen i et givet punkt i vandløbet anvendes både historiske observationer fra dette punkt, men også for et punkt opstrøms. For at forudsige vandstanden ved fx Bredstenbro anvendes historiske data (vanstandslogning) fra både Bredstenbro og Åstedbro målestation.

Til vandstandsprognoser anvendes observeret nedbør 48 timer før prognosticerings tidspunktet og for nedbørsprognoserne anvendes data for 48 timer ud i fremtiden.

3.1.2 Feature Engineering

For at give modellen de bedste betingelser for at lære mønstre i vandstanden på baggrund af input features, udføres en yderligere behandling af disse. Dette kaldes feature engineering, og betyder at der skabes nye features, ved at lave en forhåndsbearbejdning af data.

Forhåndsbearbejdningen er typisk statistisk behandling, så som f.eks. løbende middelværdi over et givent vindue. Et eksempel på den information denne bearbejdning kan give, kan ses i relationen mellem prognosticeret og observeret nedbør. På Figur 4 på næste side fremgår en sammenligning af observeret og prognosticeret nedbør henholdsvis med og uden forhåndsprocessering (løbende sum over 48 timer). Som det fremgår af figuren, er der en tydelig korrelation mellem de data, hvor sumfilteret er påført, hvorimod data uden filter ser nærmest tilfældig ud. Dette viser, at vi ved statistisk bearbejdning af data, kan udtrække information, der kan hjælpe modellen til bedre forudsigelser.

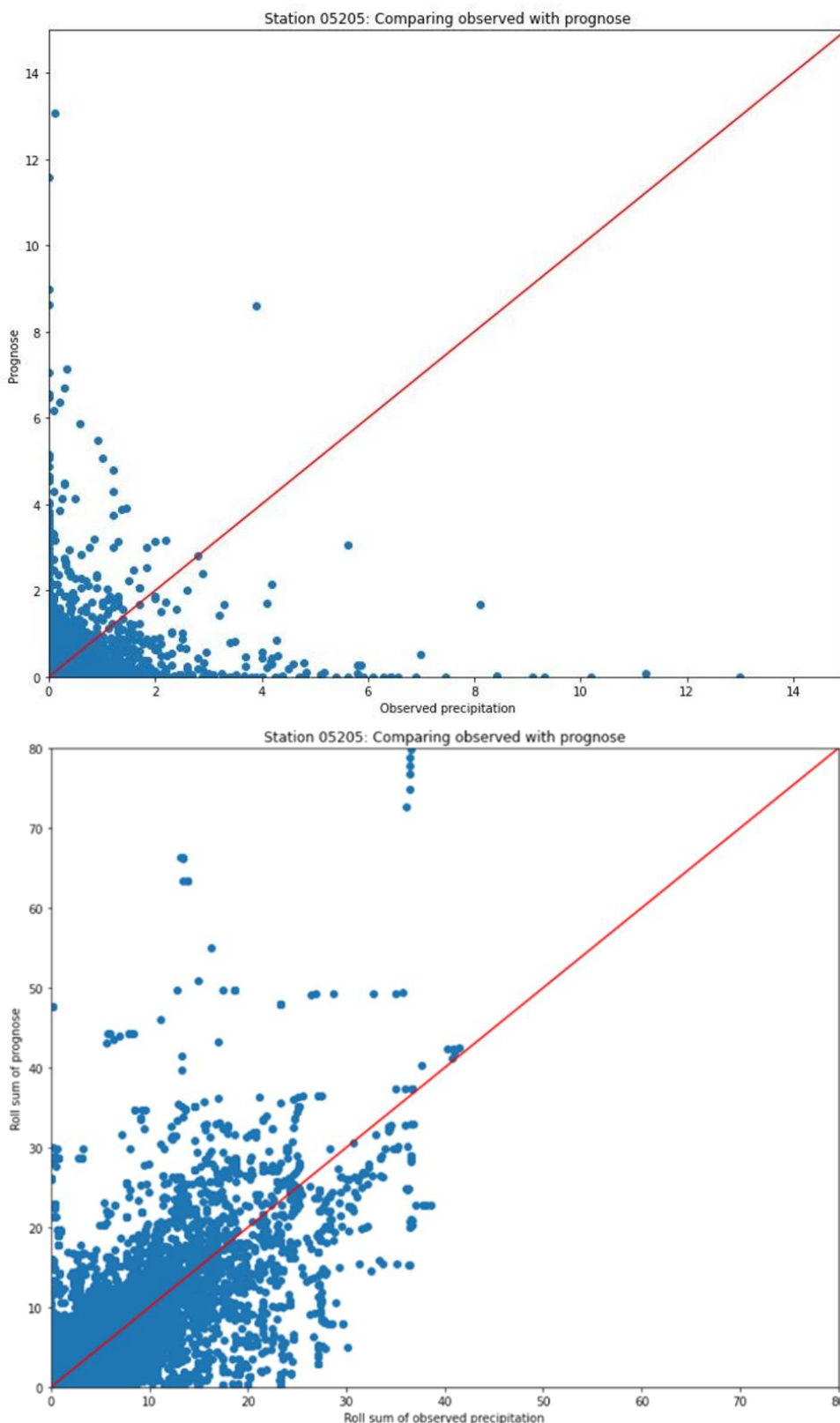
Foruden løbende sum beregnes ligeledes løbende middelværdi for alle input-tidsserier. For vandstand beregnes ligeledes deltaværdier.

Dette resulterer i følgende features som bruges i modellen:

Feature	Periode
Observeret nedbør	48 timer tilbage i tid
Prognosticeret nedbør	48 timer frem i tid
Observeret vandstand for målestation ved punktet	48 timer tilbage i tid
Observeret vandstand for målestation opstrøms punktet	48 timer tilbage i tid
Lufttemperatur	48 timer tilbage i tid

Afledte værdier af ovenstående:

- Løbende sum
- Løbende middelværdi
- Deltaværdier



Figur 4 Sammenligning af observeret og prognosticeret nedbør henholdsvis med (nederst) og uden (øverst) løbende sum af prognose og observationer.

3.1.3 Datakilder

I dette afsnit fremgår hvorfra data stammer, og hvilke data der er brugt i modellen

Vandstand

Vandstandsdata er rekvireret gennem Vandportalen.dk med en opløsning på timebasis. Afhængig af hvilken målestation der udarbejdes prognoser for, anvendes data fra ca. 2015 til 2022. For nogle stationer er vandstandsmålingerne ikke tilgængelig så langt tilbage i tid, eller der er udfald i data, se Tabel 1. Det er en udfordring for modellen at lære tendenser i vandstanden, hvis der ikke er længere, kontinuerte tidsserier.

Tabel 1 Vandstandsdata benyttet i modellen

Model kørt for denne placering, hvor der findes målestation	Periode for data*	Opstrøms målestation	Periode for data*
Bredstenbro	2016-d.d.	21.02 Åstedbro	2017-d.d.
21.54 Klostermølle, Mossø	2016-2018 2021-d.d.	21.142 Bredstenbro	2016-d.d.
21.162 Langsø, Ringvejsbroen	2020-d.d.	21.135 Rye mølle, ns bro	2013-d.d.
21.11 Kongensbro	2010-d.d.	21.01 Tvilumbro	2019-d.d.
21.09 Ulstrup	1986-d.d.	21.07 Bjerringbro	2009-d.d.

*Kun de nyere, sammenhængende dataperioder er angivet.

Nedbørsobservationer

Nedbørsobservationer rekvireres via. DMI's frie data. Der anvendes nedbørsmålinger fra 6 målestationer indenfor Gudenåens opland med en opløsning på timebasis. Nedbørsobservationerne er tilgængelige fra 2015 og frem til i dag.

De 6 målestationer der er anvendt er, markeret på Figur 5:

- Grønbæk
- Flensted
- Vestbirk
- Hesselballe
- Nørre Snede
- Gludsted Plantage Nv



Figur 5 Placering af nedbørsstationer.

Nedbørsprognoser

Nedbørsprognoser rekvireres fra MET.no (det norske meteorologiske institut). Prognoserne stammer fra en numerisk vejrmødel og giver et bud på hvor store nedbørsmængder, der kan forventes ca. 50 timer frem i tid. Vejrmødelens resultater er tidlig diskretiseret på timebasis og stedsligt diskretiseret i 1x 1 km grid. Prognoserne er tilgængelige fra 2019 og frem til i dag. Som input til modellen anvendes ikke hele gridet, men kun de 6 celler der har samme placeringen som de anvendte DMI-målestationer. Således er det muligt at træne modellen på de historiske data som prognoser, men anvende en reel prognose i en driftssituation.

Temperatur

Temperatur rekvireres ligeledes fra DMI's frie data med en opløsning på timebasis.

3.2 Opbygning af model

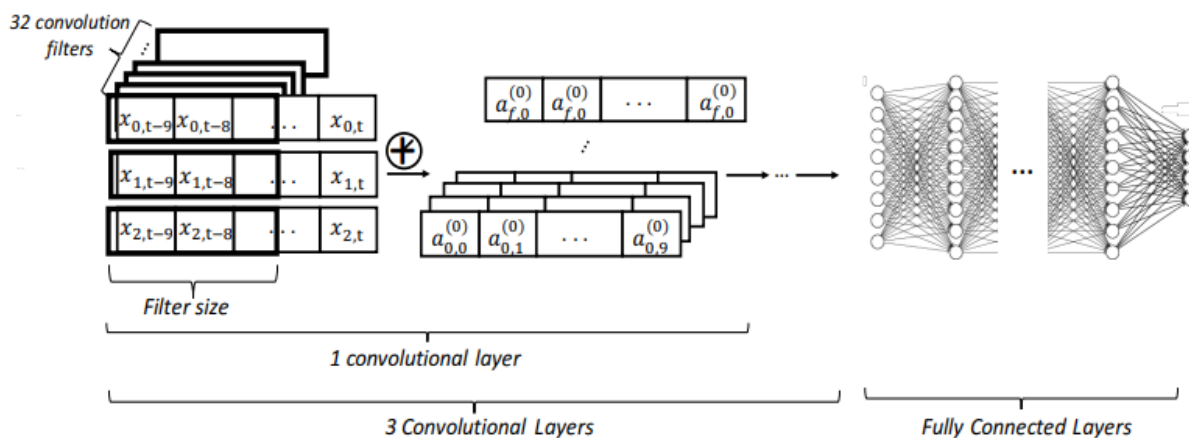
I de følgende afsnit forklares den mere tekniske opbygning af ML-modellen. Der benyttes tre modeller:

- Convolutional Neural Network (CNN)
- CNN med offset
- Persistence (fastholdelse)

3.2.1 CNN-model

Modellen er bygget op som vist i Figur 6 og består af tre 1-dimensionelle Convolutional Neural Network (1D-CNN) lag efterfulgt af et fuldt sammenhængende neuralt netværk. Der henvises til den videnskabelige litteratur for en mere dybdegående gennemgang af modellen opsætningen (Assem et.al, 2017).

Modellen kører i første 1D-CNN lag et filter hen over hver enkelt feature og herved forsøger at lære en sammenhæng ud fra den givende tidsserie. Når inputtet er kommet igennem hele netværket outputter modellen en 3-døgnsprognose på timebasis for vandstanden.



Figur 6 Princip for opbygning af CNN-model. (Assem et.al, 2017)¹

3.2.2 CNN+Offset model

Udover CNN-modellen er der også blevet benyttet en anden model, som kaldes CNN+Offset. Denne model bygger oven på CNN-modellen. Her bruges 3-døgnsprognosen beregnet af CNN-modellen, hvorefter et offset beregnes, forstået ved differencen mellem den observerede vandstand til tiden $t=0$ og den prædikterede vandstand til tiden $t=1$ time. Der korrigeres for denne forskel ved at lægge offset til hele 3-døgnsprognosen og dette er outputtet for CNN+offset modellen. CNN+offset modellen er altså en parallelforskydning af CNN-modellen.

3.2.3 Persistence model

For bedre at kunne evaluere CNN modellerne er der også blevet lavet en baseline model. Modellen tager blot vandstanden til tiden $t=0$ (den senest observerede vandstand) og fastholder den i samme værdi de næste 3 døgn.

Persistence modellen er lavet, for at kunne vise at CNN-modellerne er bedre end bare at fastholde en konstant værdi.

4. Resultater

Afsnittet gennemgår resultaterne af modellen. Det kan være svært at holde styr på resultaterne, da modellerne beregner vandstand for hver time i 3 døgn. Der er altså 72 timer = 72 resultater for hver beregning. Det er naturligvis nemmest for modellen at beregne vandstanden de første timer og vanskeligere at beregne en korrekt værdi på de sidste timer. Derfor er modellens performance (præcision) beregnet både som et gennemsnit over de 72 timers vandstande (hvor alle værdierne har samme vægtning) og kun for den ene værdi der ligger 72 timer frem i tiden (den sværeste værdi at forudsige).

Resultaterne sammenlignes også med resultater fra DHI-modellen.

¹ Assem, Haytham, Salem Ghariba, Gabor Makrai, Paul Johnston, Laurence Gill, and Francesco Pilla. "Urban water flow and water level prediction based on deep learning." In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 317-329. Springer, Cham, 2017.

4.1 Metrikker

Til at beregne performance af modellerne benyttes følgende metrikker:

- Mean Absolute Error (MAE)
 - o Gennemsnitlig afvigelse i meter
- Root Mean Squared Error (RMSE)
 - o RMSE er altid større end MAE
 - o Hvis RMSE er meget større end MAE, så indikerer dette, at der er steder, hvor modellens forudsigelse ligger langt fra den rigtige vandstand
- R^2 -værdi

Derudover har vi beregnet en yderligere metrik defineret ved:

$$R^2_{model} = 1 - \frac{\sum (y_{obs} - y_{model_1})^2}{\sum (y_{obs} - y_{model_2})^2}$$

Denne metrik har det formål, at man direkte kan holde model 1 op imod model 2. Værdien spænder fra -1 til 1. Her skal en positiv værdi forstås som, at model 1's output fitter de observerede værdier bedre end model 2's output og vice versa, hvis værdien er negativ. I det følgende er model 1 en af EnviDans modeller og model 2 DHIs model. Så en positiv værdi betyder, at EnviDans model fitter bedre end DHIs model.

4.2 Performance

Modellerne evalueres på, hvor godt de kan forudsige vandstanden i perioden 1. januar 2021 til 1. marts 2021. Alle resultater herfra findes i Bilag 1.

Modellens performance er også undersøgt i perioden 2021-09-01 til 2022-03-01, hvor der findes resultater fra DHI-modellen. I denne periode er det muligt at sammenligne EnviDans resultater med DHIs resultater.

Resultaterne er inddelt i:

- Hele 3-døgnsprognosen, dvs. alle observationer i prognosen er inkluderet og har samme vægtning
- Kun den sidste observation i 3-døgnsprognosen ($t=72$)

Ydermere er modellens performance også blive undersøgt de steder, hvor vandstanden peaker (top-pene i tidsserien) for at se, hvordan modellen performer, hvor vandstanden er mest kritisk. Det er et specifikt ønske fra Silkeborg Kommune at modellen er retvisende, når vandstanden når kritisk høje koter pga. varslingsanvendelse.

Performance for peak-værdierne er kun beregnet for 72-timers værdien. Dvs. hvis der er et peak i tidsserien fra 3-1-2021 til 4-1-2021 benyttes kun de prognoser, hvis sidste observationer er i dette tidsinterval, altså de prognoser, som er blevet produceret i perioden 1-1-2021 til 1-2-2021.

Bemærk: på nogle figurer blive benyttet navnene *Pred* og *Pred+offset*. Her menes der henholdsvis CNN modellens output og CNN+offset modellens output.

4.2.1 Overordnet performance evaluering

Nærværende afsnit opsamler resultater fra afsnit 4.2.2 til 4.2.6 og sammenligner performance på tværs af modeltype samt målestation. Den overordnede performance evaluering baseres på MAE, se Tabel 2.

Tabel 2 Sammenligning af MAE for de fire modeltyper på alle lokationer for henholdsvis T=72 og 3-døgnsprognozen.

	CNN	CNN+Offset	Persistence	DHI
T=72				
Bredstenbro	0.053	0.063	0.087	0.157
Klostermølle, Mossø	0.022	0.018	0.022	-
Silkeborg Langsø, Ringvejsbroen	0.059	0.047	0.034	0.041
Kongensbro	0.049	0.037	0.037	0.051
Ulstrup	0.150	0.165	0.164	0.247
3-døgnsprognose				
Bredstenbro	0.029	0.035	0.049	0.101
Klostermølle, Mossø	0.016	0.011	0.013	-
Silkeborg Langsø, Ringvejsbroen	0.043	0.025	0.019	0.026
Kongensbro	0.037	0.021	0.021	0.031
Ulstrup	0.108	0.122	0.106	0.177

4.2.2 Bredstenbro

Performance af sidste observation i prognoserne (t=72)

Fra september 2021 og frem har ML-modellerne hver time produceret en 3-døgnsprognose. Kigger man kun på den sidste værdi i hver af disse prognoser, får man en ny tidserie for hele perioden fra september 2021 og til marts 2022.

Det samme er gjort for DHIs model, dog producerer denne kun en prognose hver sjette time, så tidsserie vil i dette tilfælde have en tidlig opløsning på 6 timer.

Det er performance for disse tidsserier, der bliver beregnet. Modellernes performance evalueres kun på de tidspunkter, hvor det er muligt at sammenligne disse. Performance kan aflæses i Tabel 3.

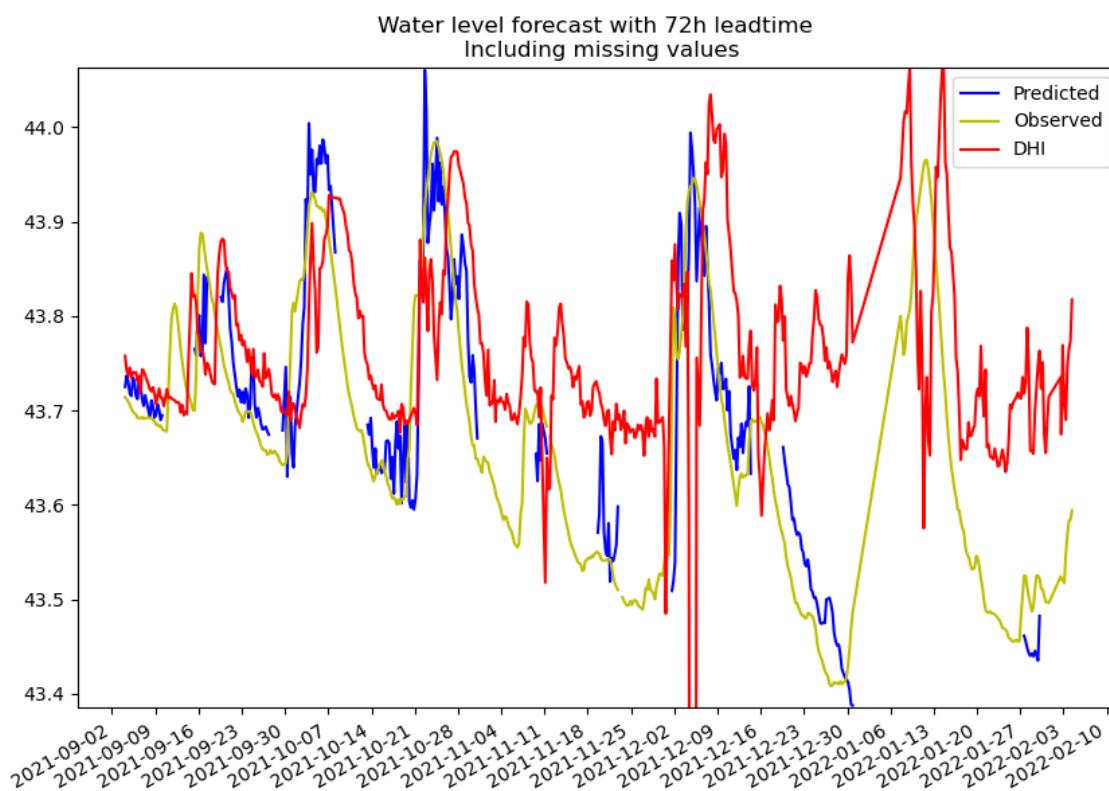
	Model	R^2	MAE	RMSE
Gennemsnitlig performance	CNN	0.801	0.053	0.068
	CNN+offset	0.728	0.063	0.079
	Persistence	0.378	0.087	0.126
	DHI	-1.399	0.157	0.236

Tabel 3 - Bredstenbro: Gennemsnitlig performance for sidste observation i hver prognose i perioden sep. 2021 og frem

Vi ser, at CNN-modellen performer bedst på alle tre parametre med en R^2 -værdi på 0.8 og en gennemsnitlig afvigelse på 5,3 cm, hvorimod DHI-modellen afviger med 15,7 cm i gennemsnit og har en R^2 -værdi på -1.399.

Dette fremgår også af Figur 7, hvor den orange linje viser observeret vandstand, den blå er CNN modellen og den røde er DHIs model. Her er x-aksen er tidspunktet i tidsserien og y-aksen er vandstanden i meter.

Sammenlignes CNN-modellen med DHI-modellen ved brug af R_{model}^2 -metrikken (se afsnit 4.2) fås en værdi på $R_{model}^2 = 0.917$. Dette indikerer, at CNN-modellen ligger betydeligt tættere på den rigtige værdi i forhold til DHI og ikke laver lige så store udsving.

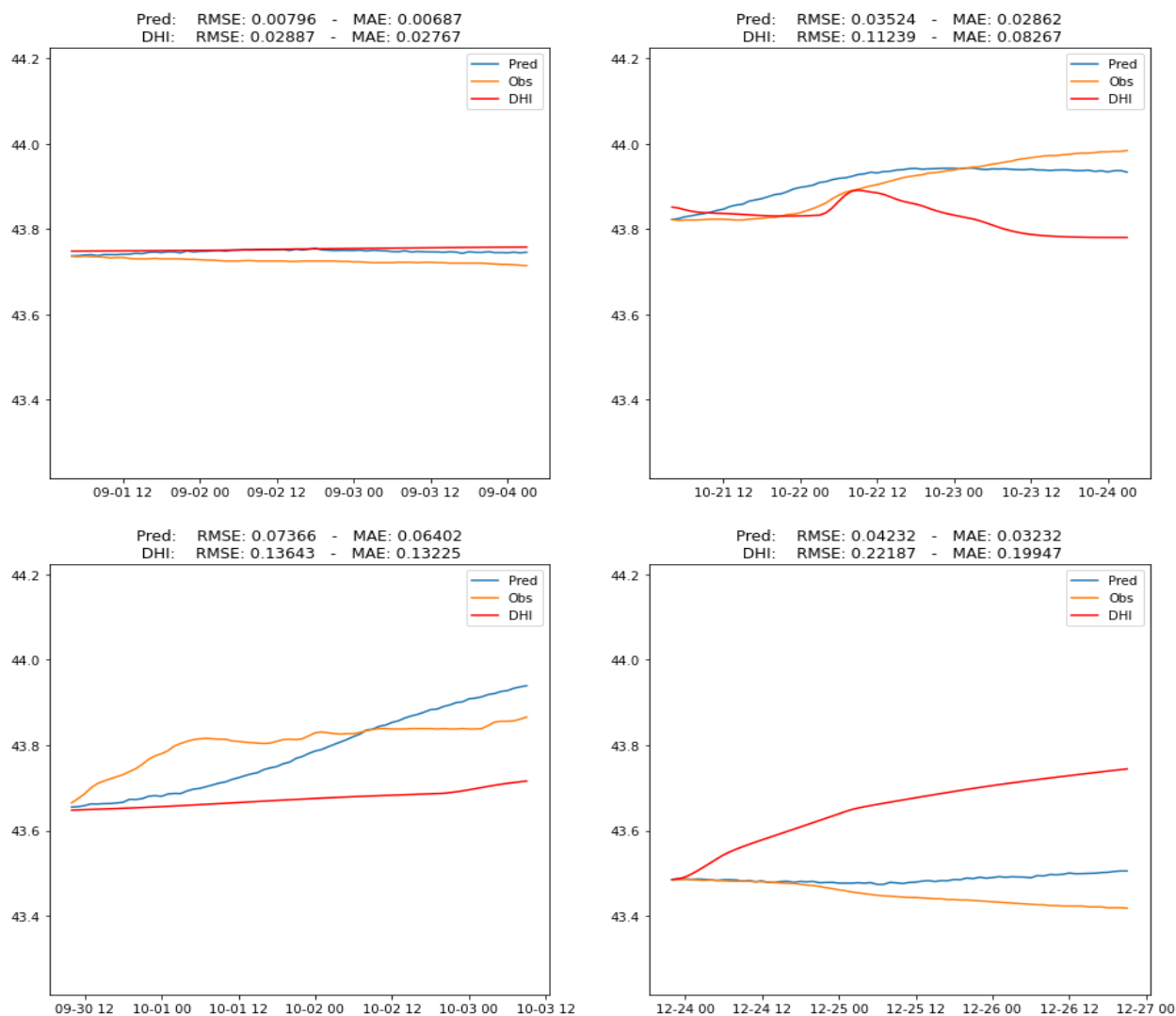


Figur 7 - Bredstenbro - Observerede vandstand, samt ML-modellen og DHI modellens forecast for vandstanden

Hele 3-døgnsprosenen

Som tidligere nævnt outputter ML-modellerne en 72 timers prognose hver time. Vi kigger nu på performance af hele prognosen, hvor alle forudsigelser har samme vægt.

Igen betragtes kun prognoser i perioden september 2021 og frem. Der bliver kun målt performance de steder, hvor vi kan sammenligne med DHI. I Figur 8 vises fire tilfældig udvalgte 72 timer perioder. Her viser x-aksen tidspunktet og y-aksen visser vandstanden i meter.



Figur 8 - Bredstenbro: 3-døgnsprosener for CNN og DHI. Fire tilfældige perioder er vist

I disse fire tilfælde ser vi, at CNN-modellen (blå) ligger tættere på den rigtige vandstand (orange) end DHI-modellen (rød) gør. CNN-modellen ser ud til kunne forudsige tendenser i vandstanden (om den stiger eller falder). Vi ser blandt andet i øverst højre hjørne, at CNN-modellen starter med at skyde lidt over, og så ende med at være lidt lavere end den rigtige vandstand. CNN-modellen har i dette tilfælde en MAE på 0.029, hvilket betyder, at den i gennemsnit ligger 2,9 cm fra den rigtige værdi. Dog er der visse tilfælde, hvor CNN-modellen har stor større afvigelse, hvilket bl.a. ses i nederst venstre hjørne i figuren.

Performance kan aflæses i Tabel 4.

	Model	MAE	RMSE
Gennemsnitlig performance	CNN	0.029	0.035
	CNN+offset	0.035	0.042
	Persistence	0.049	0.058
	DHI	0.101	0.114

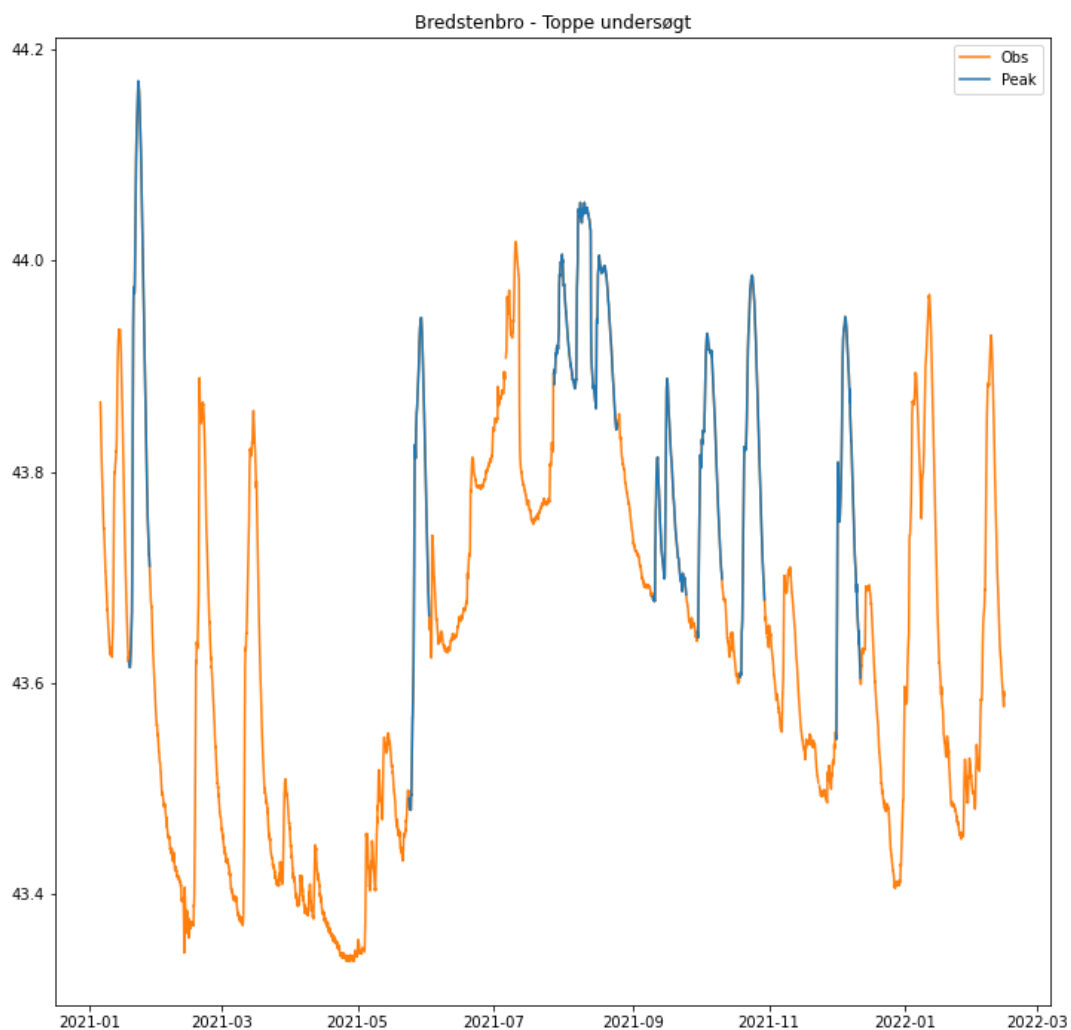
Tabel 4 - Bredstenbro: Gennemsnitlig performance for alle 3-døgnspregninger

Det ses at CNN-modellen er bedre end CNN+offset. Holdes CNN-modellen op mod DHI-modellen fås en $R_{model}^2 = 0.678$.

Peak værdier i 2021

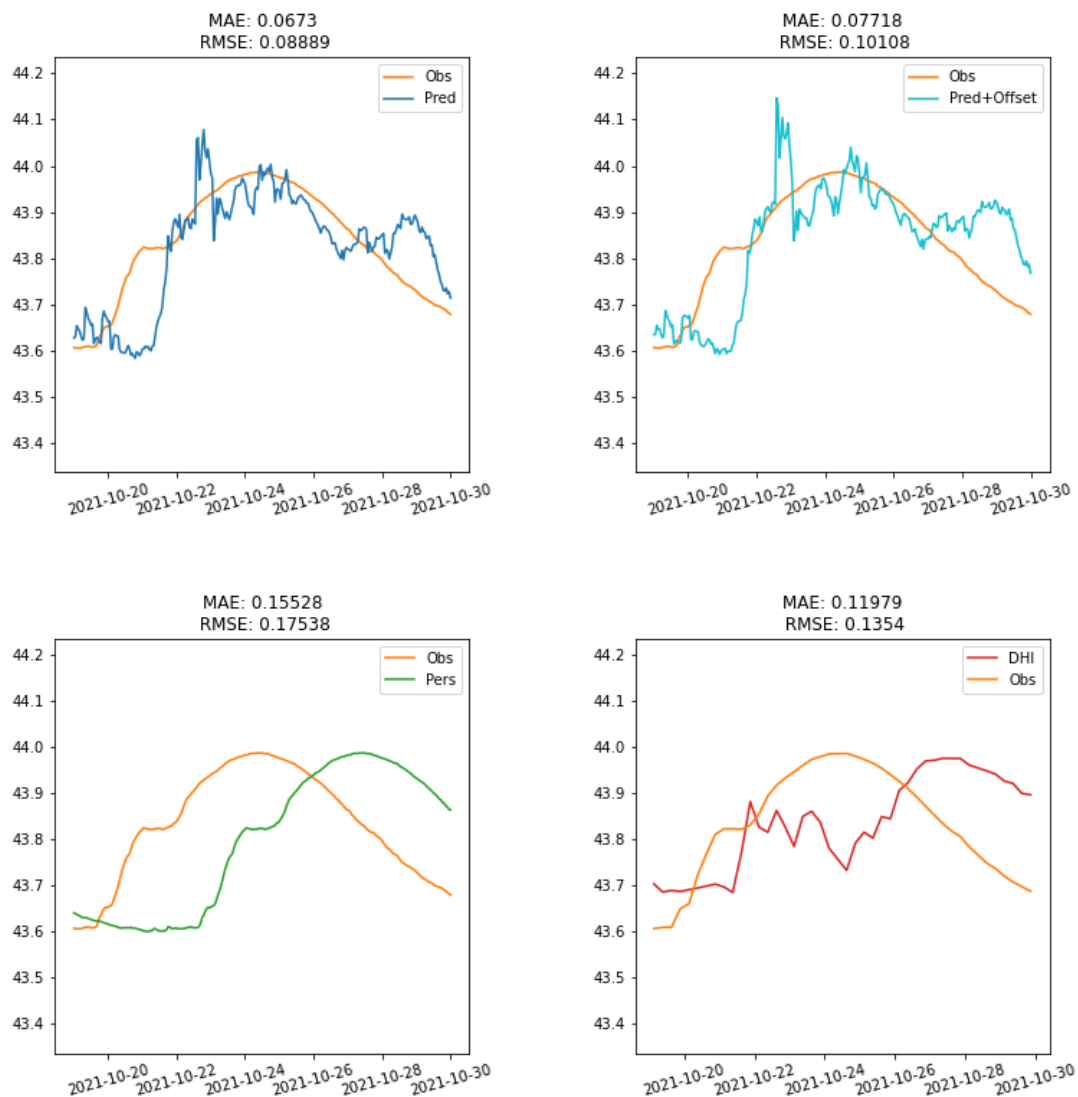
Som tidligere nævnt benyttes kun sidste observationer i prognoserne ($t=72$ timer), når performance for toppene beregnes, dvs. det tidspunkt som er sværest at forudsige (Se afsnit 4.2 for uddybning).

I Figur 9 fremgår det, hvilke toppe i 2021 som modellen har været i stand til at lave tilstrækkelig med prognoser for. De toppe, hvor det ikke har været muligt at lave nok prognoser til skyldes manglende input data, såsom fra nedbørsdata fra enkelte målestationer. Performance for toppene kan findes i Bilag 2.



Figur 9 - Bredstenbro: Toppe i 2021 som modellen har kunnet lave prognoser for (blå)

I Figur 10 er der illustreret, hvordan de fire forskellige modeller performer på et specifikt peak (19. til 30. okt 2021). Vi ser, at Persistence modellen (nederst venstre hjørne) naturligvis først opfanger peaket 3 dage for sent og DHI (nederst højre hjørne) opfanger det også for sent, hvorimod CNN og CNN+offset (øverst) bedre kan vise peaket til den korrekte tid, om end modellen ikke opfanger den allerhøjeste værdi.



Figur 10 - Bredstenbro: Performance af CNN, CNN+offset, Persistence og DHI-modellerne i perioden 19. til 30. okt 2021

Ses på alle toppene observeres i Tabel 5, at CNN-modellen afviger med 7,1 cm i gennemsnit, hvorimod DHIs afviger med 14,8 cm. Persistence modellen er den model, som klarer sig værst målt på MAE, hvilket skyldes, at stigning i vandstanden sker over nogle få dage og Persistence modellen er 3 dage 'bagud'. DHI-modellen har den højeste RMSE, hvilket kunne antyde, at den har visse steder, hvor afvigelsen fra den rigtige vandstand er væsentlig større end dens gennemsnitlige afvigelse.

Alle performance for toppene ved Bredstenbro kan findes i Bilag 2

	Model	MAE	RMSE
Gennemsnit for alle toppene	CNN	0.071	0.092
	CNN+offset	0.083	0.103
	Persistence	0.153	0.175
	DHI	0.148	0.204

Tabel 5 - Bredstenbro: Gennemsnitlig performance ved alle toppene

4.2.3 Klostermølle - Mossø

Performance af sidste observation i prognoserne (t=72):

Ved Klostermølle har det ikke været muligt at sammenligne med DHI, da DHI-modellen ikke har data for denne station. Derfor evaluerer vi modellen på hele perioden fra januar 2021 til marts 2022.

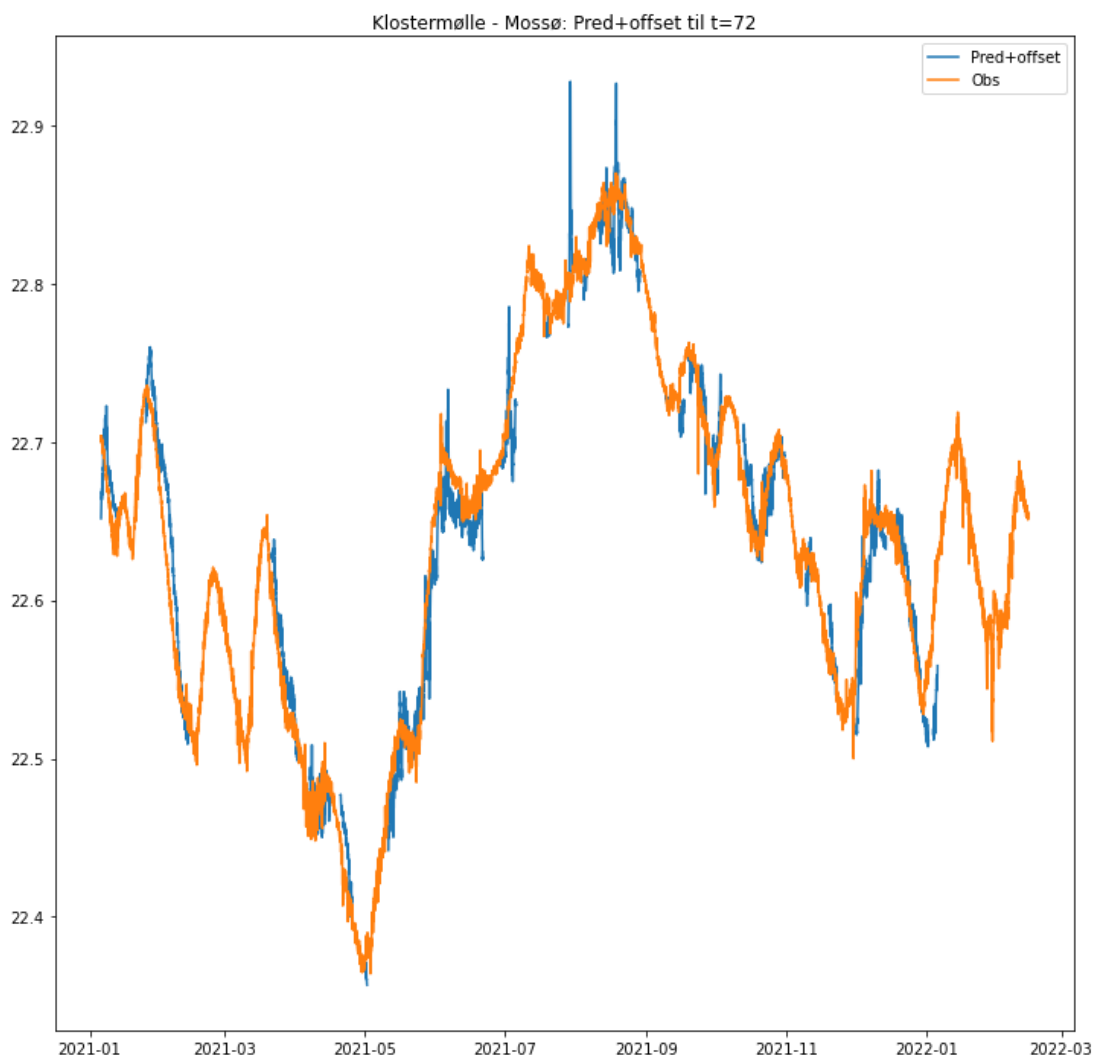
Performance for en 72 timers horisont (for uddybning af dette se afsnit 4.2 og 4.2.2) ses i Tabel 6.

	Model	R^2	MAE	RMSE
Gennemsnitlig performance	CNN	0.934	0.022	0.028
	CNN+offset	0.953	0.018	0.023
	Persistence	0.933	0.022	0.028

Tabel 6 - Klostermølle: Gennemsnitlig performance for sidste observation i hver prognose i perioden sep. 2021 og frem

Her kan det ses, at når CNN+offset modellen afviger med gennemsnit 1,8 cm fra den observerede værdi. Desuden har CNN+offset modellen en R^2 -værdi på 0,953.

Figur 11 understøtter dette, da det ses at CNN+offset modellen følger fint med den observerede værdi, dog med enkelte udfald omkring august måned.



Figur 11 - Klostermølle: CNN+offset modellens prognoser for t=72 ved Klostermølle

Hele 3-døgnsprognozen

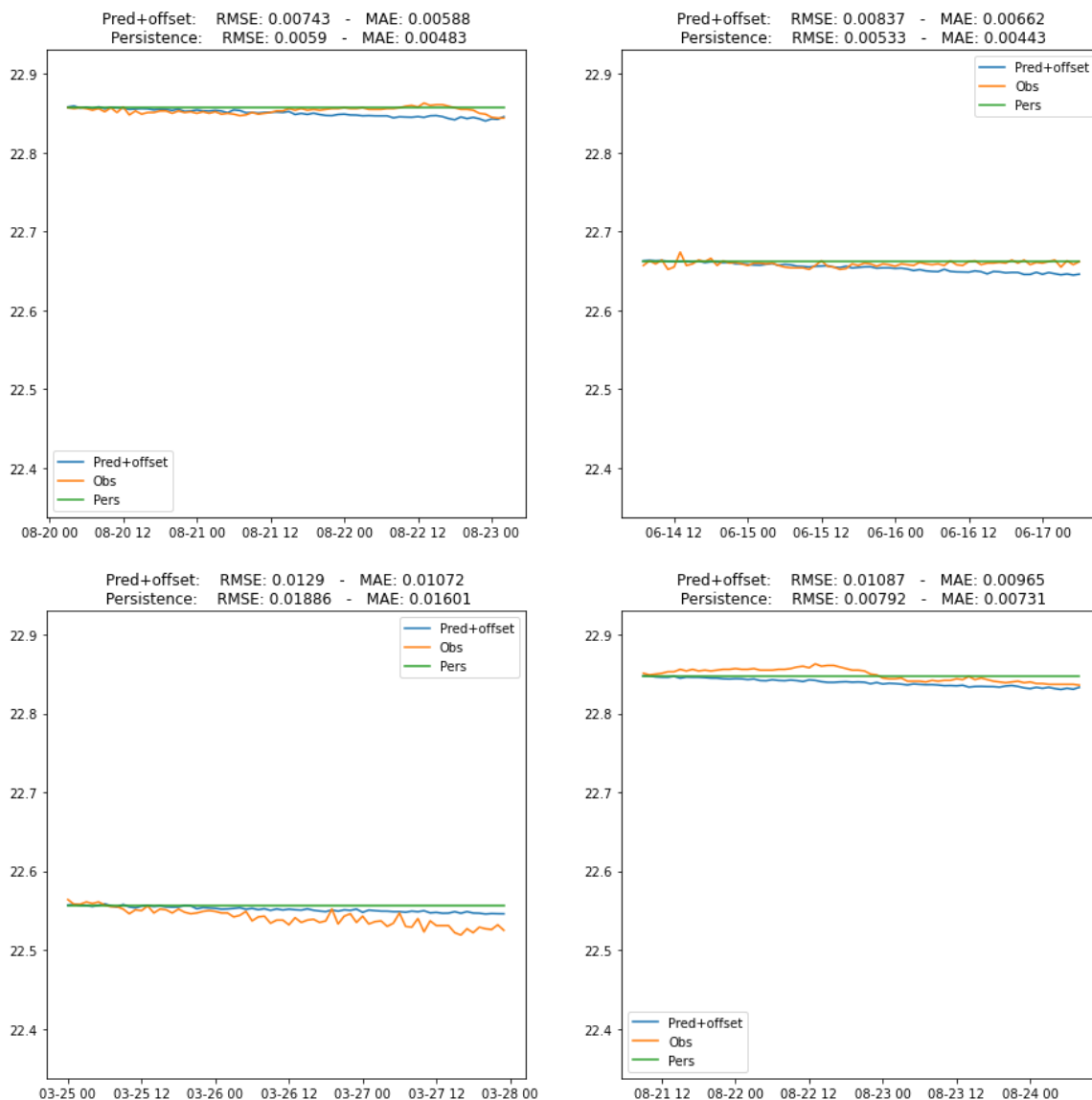
Vi kigger nu på performance af hele prognosen, hvor alle forudsigelser har samme vægt. Vi måler performance på alle prognoser fra januar 2021 til marts 2022.

	Model	MAE	RMSE
Gennemsnitlig performance	CNN	0.016	0.018
	CNN+offset	0.011	0.013
	Persistence	0.013	0.015

Tabel 7 - Klostermølle: Performance for alle værdier i hver 3-døgnsprognoze for perioden jan. 2021 til mar. 2022

I Tabel 7 ser vi, at CNN+offset modellen er bedst i begge kategorier med en MAE på 0.011 og en RMSE på 0.013, hvilket viser, at modellen i gennemsnit afviger med 1,1 cm fra den rigtige vandstand. Derudover antyder den lave RMSE, at modellen generelt ikke har flere større afvigelser.

Dette stemmer overens med de fire tilfældige prognoser for CNN+offset modellen, som er illustreret i Figur 12. Vi kan også se en tendens ved vandstanden ved Klostermølle, nemlig at vandstanden er meget stabil og ændring i vandstand sker langsomt, hvilket også er grunden til Persistence modelens gode performance.



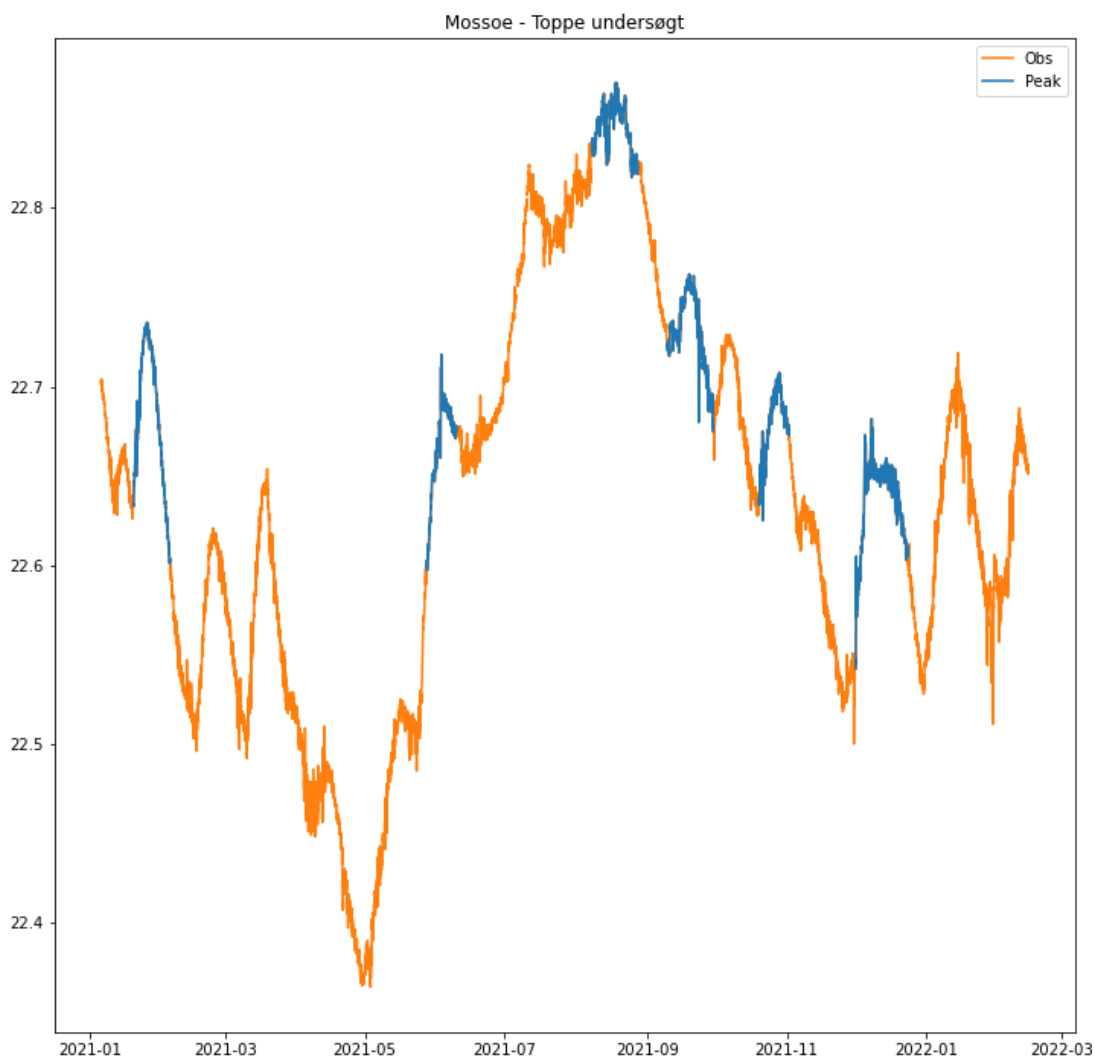
Figur 12 - Klostermølle: Fire tilfældig valgte prognoser fra både CNN+offset modellen (blå) og Persistence modellen (grøn), samt den rigtige vandstand (orange)

Peak værdier i 2021

Som tidligere nævnt benyttes kun sidste observationer i prognoserne (t=72 timer), når performance for toppene beregnes, dvs. det tidspunkt som er sværest at forudsige (Se afsnit 4.2 for uddybning).

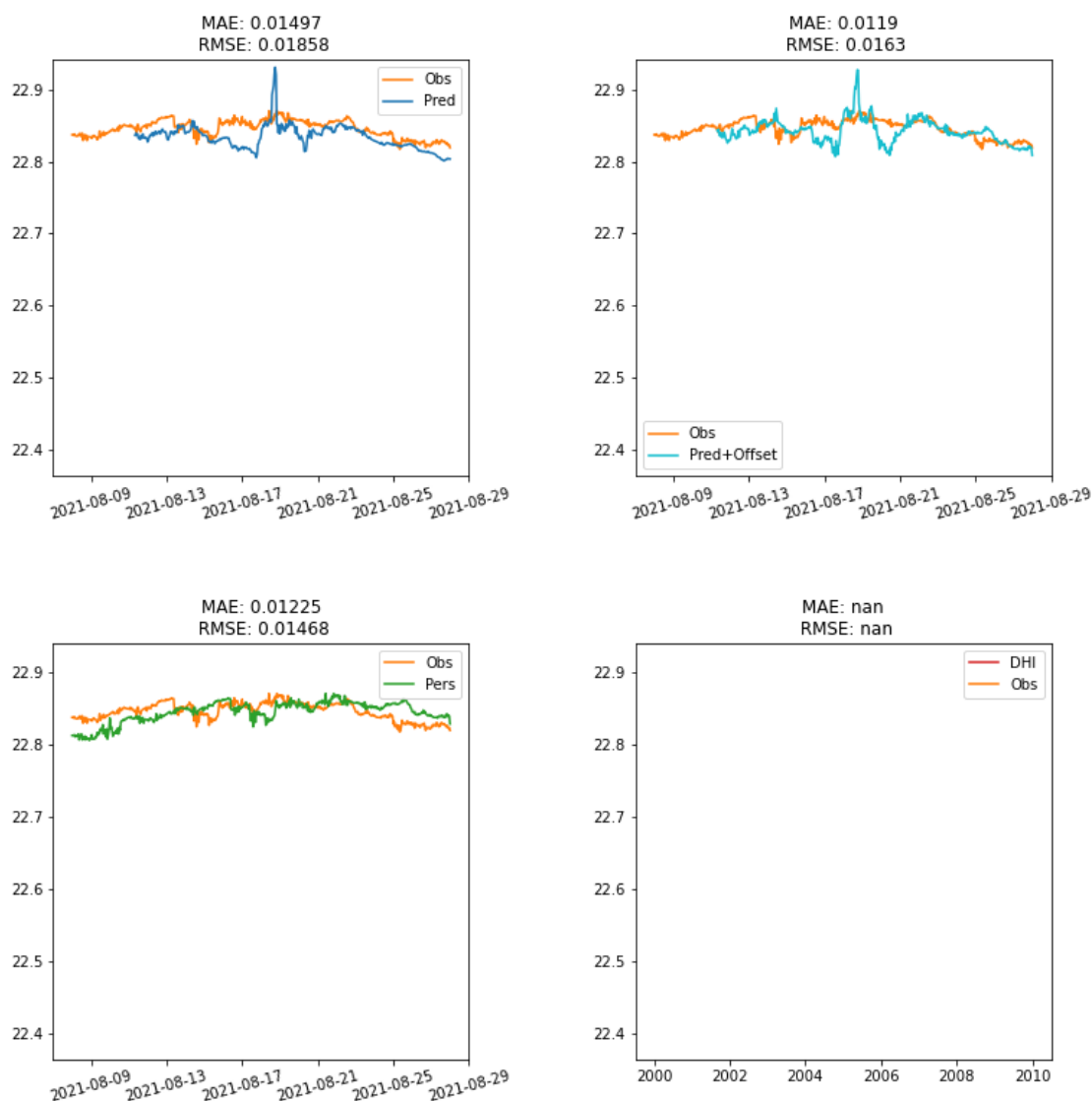
I Figur 13 fremgår det, hvilke toppe i 2021 som modellen har været i stand til at lave tilstrækkelig med prognoser for. De toppe, hvor det ikke har været muligt at lave nok prognoser til skyldes

manglende input data, såsom nedbørsdata fra enkelte målestationer. Performance for alle toppene kan findes i Bilag 2.



Figur 13 - Klostermølle, Mossø: Toppe i 2021, som ML modellerne kunne lave prognose for (blå)

Den højeste vandstand findes i perioden 8. til 28. aug 2021. I denne periode afviger CNN+offset modellen i gennemsnit 1,2 cm fra den rigtige værdi (se Figur 14). Her ser vi, at Persistence-modellen, afviger med ca. samme værdi. Dog har Persistence en lavere RMSE værdi end CNN+offset modellen, hvilket indikerer, at der er enkelte steder, hvor CNN+offset har større afvigelser end Persistence modellen. Dette kommer til udtryk i Figur 14, hvor vi omkring d. 17 august til d. 21 august ser, at CNN+offset modellen springer meget.



Figur 14 - Klostermølle: Peak i perioden 2021-08-08 til 2021-08-28

Persistences gode performance her skyldes, ligesom det generelt er tilfældet her ved Klostermølle, at stigning i vandstanden ved denne top sker langsomt over mange dage. Set over alle toppen afviger CNN+offset modellen med 1,9 cm i gennemsnit med en RMSE på 0.023, hvilket indikerer, at CNN+offset modellen ikke laver store fejl, når den 3 døgns ud i fremtiden forsøger at forudsige disse toppe.

	Model	MAE	RMSE
Gennemsnit for alle top-pene	CNN	0.023	0.027
	CNN+offset	0.019	0.023
	Persistence	0.24	0.028

Tabel 8 - Klostermølle: Gennemsnitlig performance ved alle toppene

Alle performance for toppene ved Klostermølle findes i Bilag 2

4.2.4 Silkeborg Langsø, Ringvejsbroen

Performance af sidste observation i prognoserne (t=72)

Fra september 2021 og frem har ML-modellerne hver time produceret en 3-døgnspåbudsning. Kigger man kun på den sidste værdi i hver af disse påbudsninger, får man en ny tidsrække for hele perioden fra september 2021 og til marts 2022.

Det samme er gjort for DHIs model, dog producerer denne kun en påbudsning hver sjette time, så tidsrække vil i dette tilfælde have en tidslig opløsning på 6 timer.

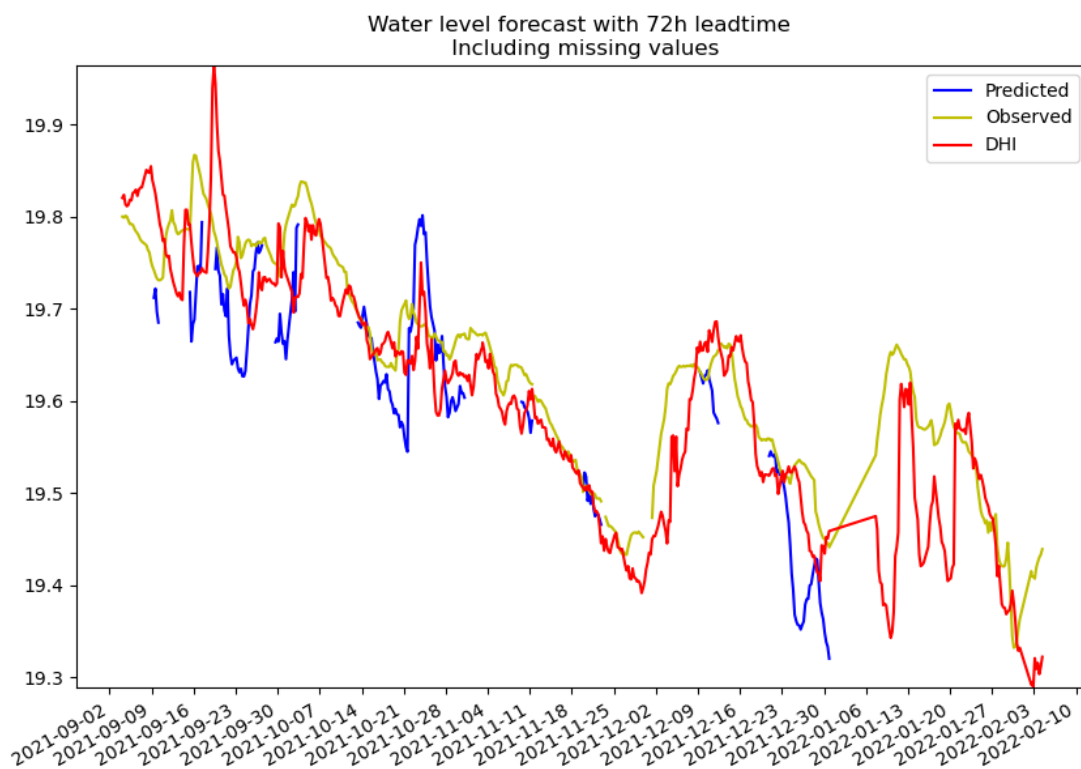
Det er performance for disse tidsrækker, der bliver beregnet. Modellernes performance evalueres kun på de tidspunkter, hvor det er muligt at sammenligne disse. Performance kan aflæses i Tabel 9.

	Model	R^2	MAE	RMSE
Gennemsnitlig performance	CNN	0.44	0.059	0.077
	CNN+offset	0.6252	0.047	0.063
	Persistence	0.842	0.034	0.041
	DHI	0.74	0.041	0.053

Tabel 9 - Langsø: Gennemsnitlig performance for sidste observation i i perioden sep. 2021 og frem

Af Tabel 9 ser vi, at CNN og CNN+offset modellen gennemsnitlig afviger med henholdsvis 5,9 cm og 4,7 cm, imens DHI-modellen kun afviger med 4,1 cm. Dette kommer også til udtryk i Figur 15, hvor vi ser, at CNN-modellen har større udsving end DHI-modellen. Dog er DHI-modellen generelt ikke bedre end Persistence modellen.

Her performer CNN+offset modellen bedre end CNN-modellen. Sammenlignet med DHI får vi $R^2_{model} = -0.31$, hvilket sammen med MAE- og RMSE-værdien indikerer, at den sidste værdi i DHI-modellen generelt ligger tættere på den observerede vandstand end CNN+offset modellen gør. Dette fremgår også af Figur 15.



Figur 15 - Langsø - Observerede vandstand, samt ML-modellen og DHI modellens forecast for vandstanden

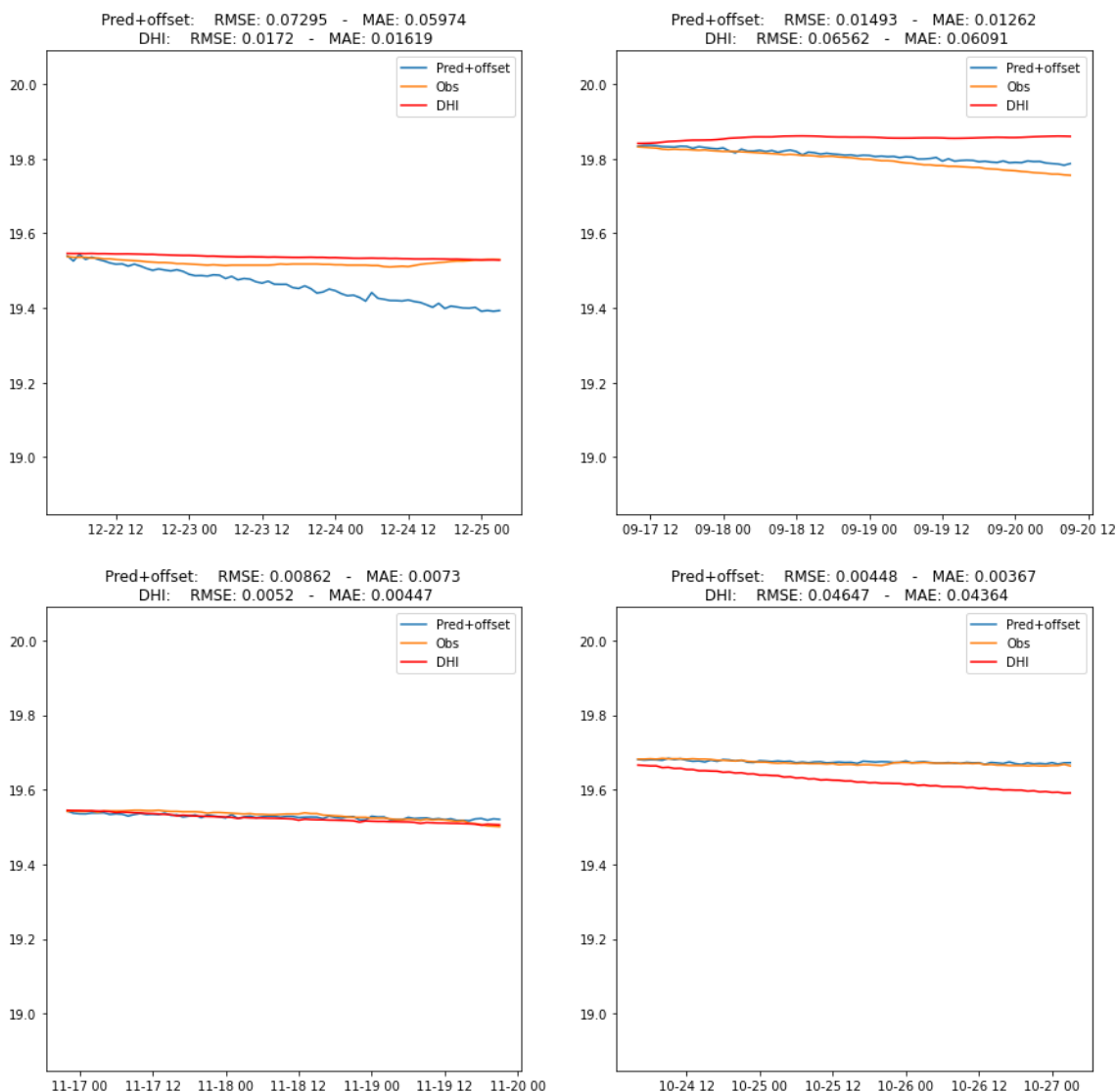
Hele 3-døgnsprosnosen

Modellernes 3-døgnsprosnoser evalueres i perioden september 2021 til marts 2022, hvor det er muligt at sammenligne med DHI-modellen. Tabel 10 nedenfor viser, at CNN+offset modellen er den bedste af de to ML-modeller med en MAE på 0,025, hvilket betyder at alle værdierne i en prognose produceret af CNN+offset modellen i gennemsnit ligger 2,5 cm fra den rigtige værdi. Her ses det også, at DHI-modellens performance er stort set identisk med CNN+offset modellen.

	Model	MAE	RMSE
Gennemsnitlig performance	CNN	0.043	0.046
	CNN+offset	0.025	0.029
	Persistence	0.019	0.023
	DHI	0.026	0.029

Tabel 10 - Langsø: Gennemsnitlige performance af alle 3-døgnsprosnoserne

Dette fremgår også af Figur 16, hvor CNN+offset modellen eksempelvis er bedst i øverst højre hjørne med en gennemsnitlig afvigelse på 1,2 cm, imens DHI-modellen er bedst i fx øverst venstre hjørne med en gennemsnitlig afvigelse på 1,6 cm.



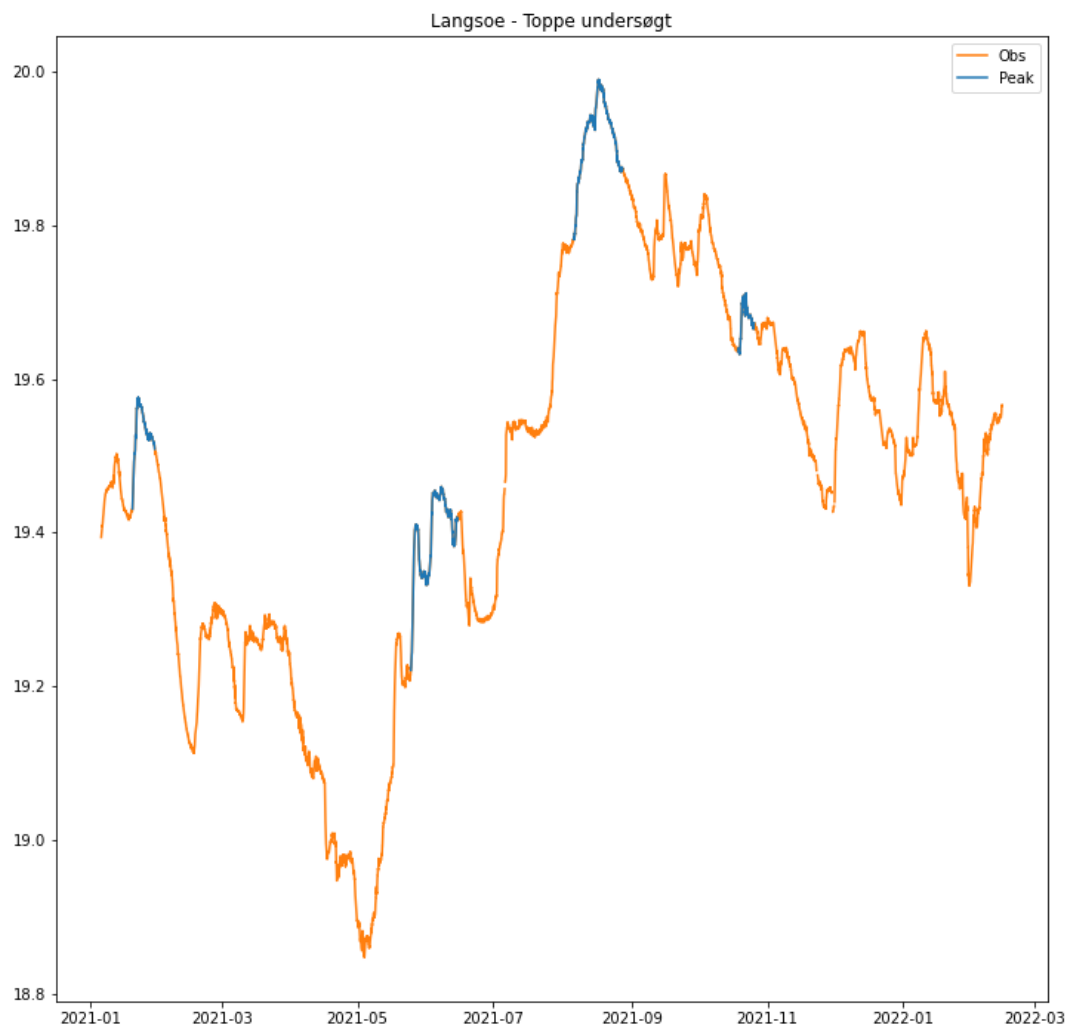
Figur 16 - Langsø: 3-døgnsprognoser for CNN og DHI. Fire tilfældige perioder er vist

Sammenlignes CNN+offset modellen med DHI-modellen er $R^2_{model} = 0.043$. Sammenholdt med MAE- og RMSE-værdier for både CNN+offset og DHI-modellen ser vi, at i dette tilfælde er prognoser fra CNN+offset modellen og prognoser fra DHI-modellen stort set lige gode.

Peak værdier i 2021

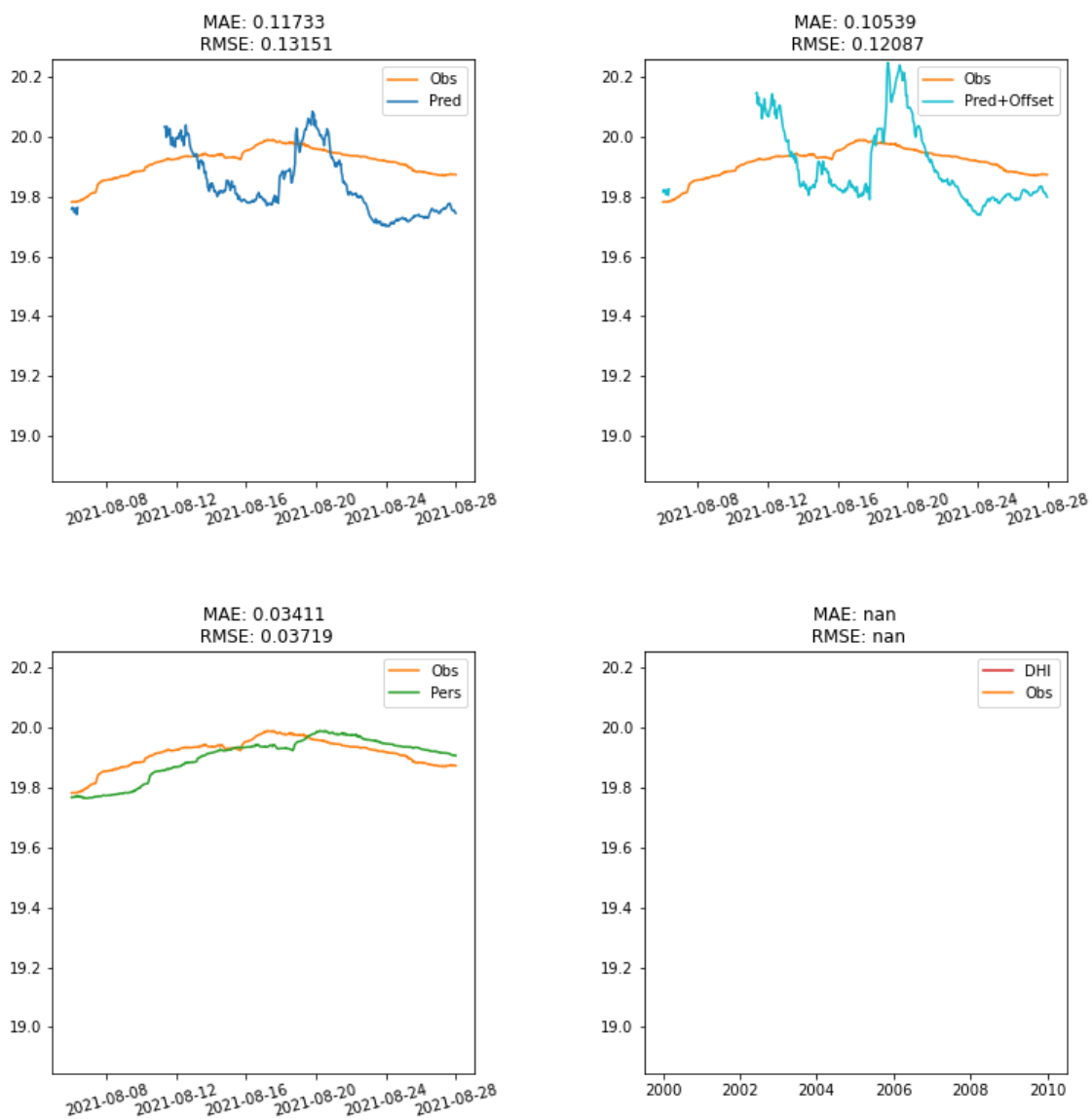
Som tidligere nævnt benyttes kun sidste observationer i prognoserne (t=72 timer), når performance for toppene beregnes, dvs. det tidspunkt som er sværest at forudsige (Se afsnit 4.2 for uddybning).

I Figur 17 fremgår det, hvilke toppe i 2021 som modellen har været i stand til at lave tilstrækkelig med prognoser for. De toppe, hvor det ikke har været muligt at lave nok prognoser til skyldes manglende input data, såsom nedbørsdata fra enkelte målestationer. Performance for alle toppene kan findes i Bilag 2.



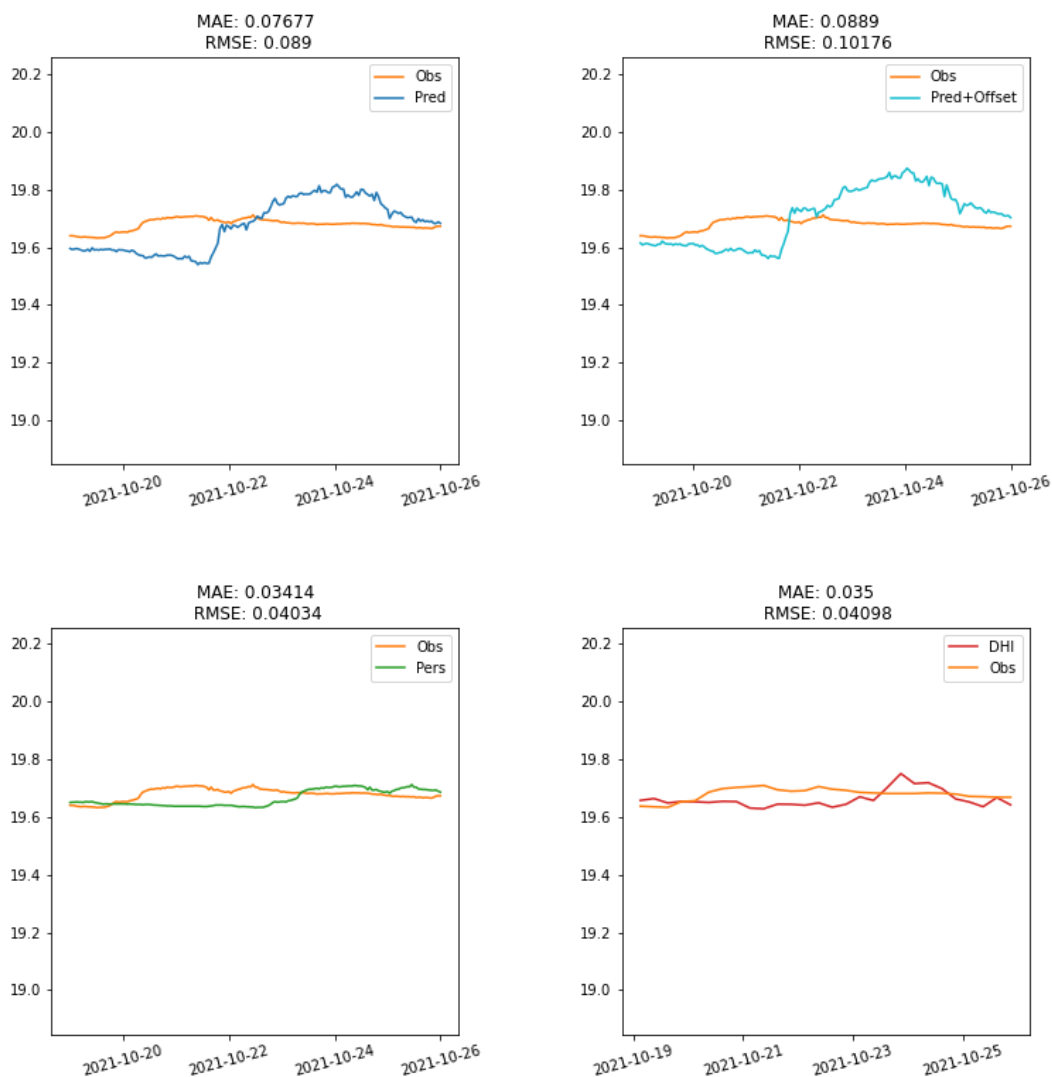
Figur 17 - Langsø: Toppe i 2021 som modellen har kunnet lave prognoser for (blå)

Den højeste værdi forekom i perioden 6. - 28. august 2021. Som det fremgår fra Figur 18, så ligger de to ML-modeller i gennemsnit 11,7 cm og 10,5 cm fra observeret værdi, hvorimod Persistence-modellen kun ligger 3,4 cm fra i gennemsnit. Persistence-modellens gode performance skyldes, at vandstanden ved dette peak stiger langsomt over en længere periode.



Figur 18 - Langsø: Performance ved højeste peak i 2021

I Figur 19 betragtes den top ved Langsø, hvor det var muligt at sammenligne ML modellen med DHI-modellen. Dette er i perioden fra 19. til 26. oktober 2021.



Figur 19 - Langsø: Performance af CNN, CNN+offset, Persistence og DHI-modellerne i perioden 19. til 26. okt 2021

Af Tabel 11 ses at ML-modellerne har udfordringer ved Langsø, hvor CNN-modellen afviger med 7,4 cm i gennemsnit. Bemærk, at den gennemsnitlige performance for DHI kun viser værdien for toppen i Figur 19, da dette var den eneste top, hvor der var prognoser nok fra både ML-modellerne og DHI-modellen til at foretage en sammenligning.

	Model	MAE	RMSE
Gennemsnit for alle toppene	CNN	0.074	0.09
	CNN+offset	0.075	0.092
	Persistence	0.045	0.057
	DHI*	0.035	0.041

Tabel 11 - Langsø: Gennemsnitlig performance ved alle toppene. *DHIs performance kun for 19. - 26. okt 2021

Alle performance for toppene ved Langsø kan findes i Bilag 2.

4.2.5 Kongensbro

Performance af sidste observation i prognoserne (t=72)

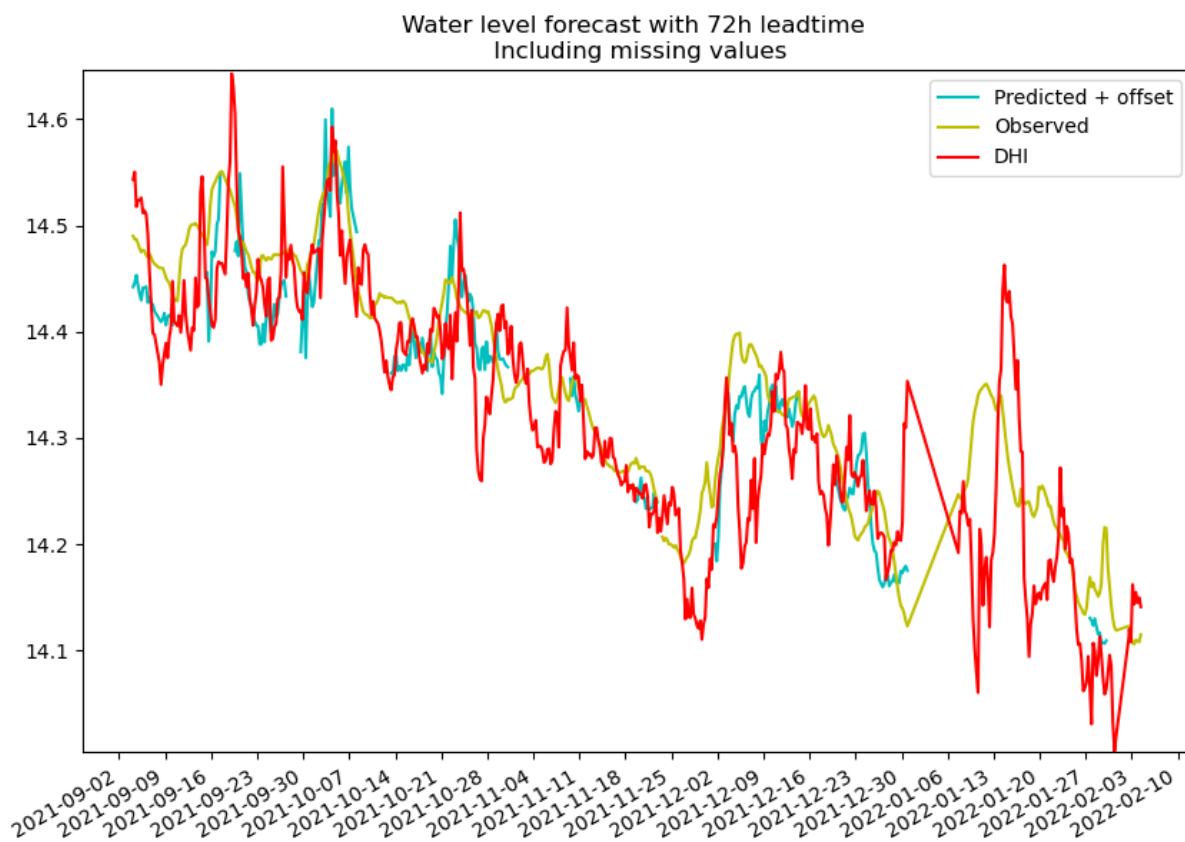
Som tidligere er det perioden fra september 2021 og frem, som bliver betragtet. Der benyttes kun prognoser i denne tidsperiode, hvor det er muligt at sammenligne med DHI-modellen.

	Model	R^2	MAE	RMSE
Gennemsnitlig performance	CNN	0.733	0.049	0.056
	CNN+offset	0.838	0.037	0.043
	Persistence	0.803	0.037	0.048
	DHI	0.619	0.051	0.067

Tabel 12 - Kongensbro: Gennemsnitlig performance for sidste observation i hver prognose i perioden sep. 2021 og frem

I Tabel 12 ser vi, at ved Kongensbro performer CNN+offset modellen bedst på alle tre metrikker med en R^2 -værdi på 0.84 og en gennemsnitlig afvigelse på 3,7 cm. Holdt op imod DHI-modellen får vi at $R_{model}^2 = 0.574$, hvilket stemmer godt overens med Figur 20.

Dette kommer især til udtryk omkring 15. december, hvor CNN+offset modellen forudser en for lav vandstand, men DHI modellen afviger væsentlig mere. Bemærk, at ML-modellerne næsten ingen prognoser har lavet i 2022, så DHI-modellens prognoser er i dette tilfælde ikke medregnet i performance.



Figur 20 - Kongensbro: Sidste observation i hver prognose produceret af CNN+offset modellen og DHI-modellen

Hele 3-døgnsprognozen

Som tidligere nævnt outputter ML-modellerne en 72 timers prognose hver time. Vi kigger nu på performance af hele prognosen, hvor alle forudsigelser har samme vægt.

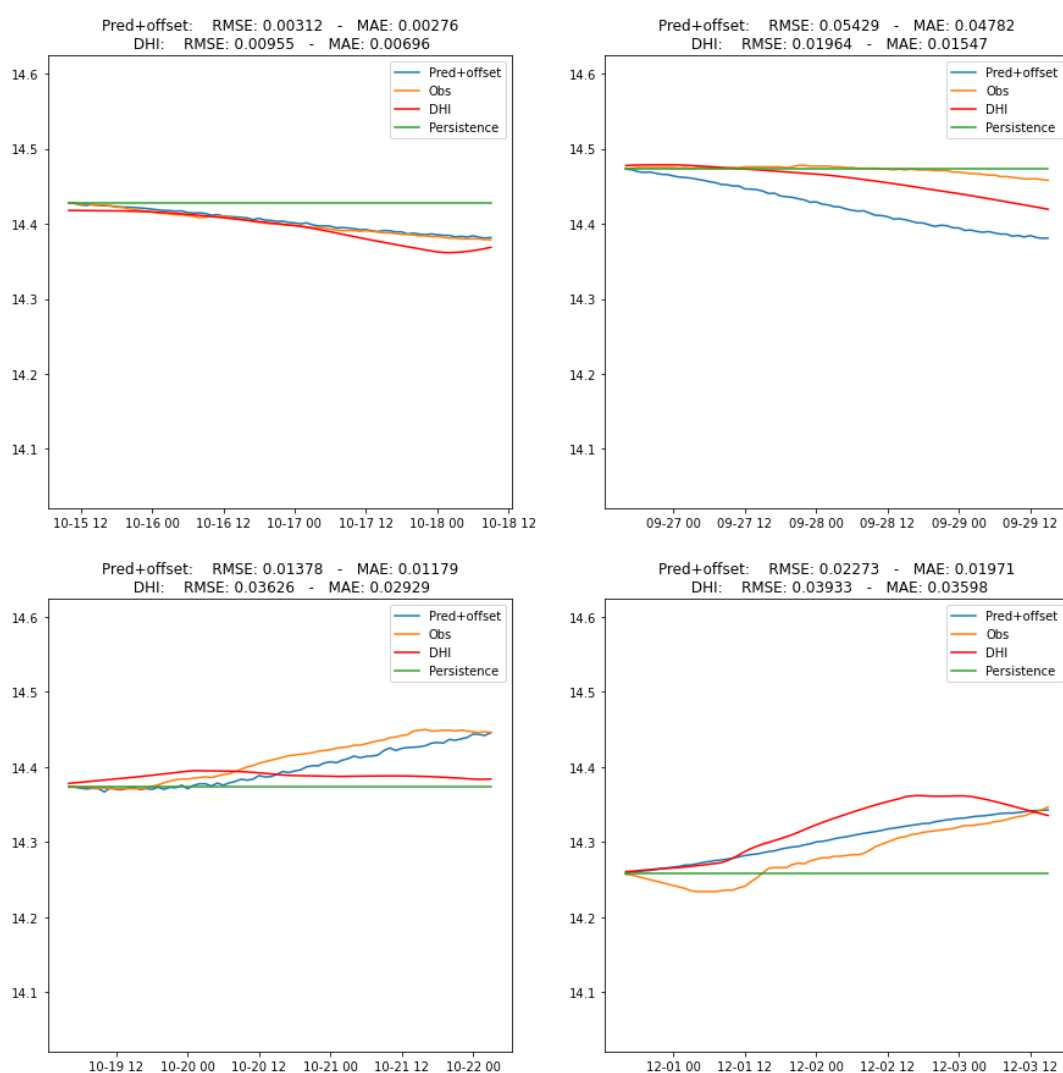
Af Tabel 13 fremgår den gennemsnitlige performance for 3-døgnsprognozer i perioden september 2021 og frem. ML-modellernes prognoser som benyttes i denne tidsperiode, er ligesom tidligere, de der kan sammenlignes direkte med DHI-modellen.

	Model	MAE	RMSE
Gennemsnitlig performance	CNN	0.037	0.038
	CNN+offset	0.021	0.025
	Persistence	0.021	0.024
	DHI	0.031	0.035

Tabel 13 - Kongensbro: Gennemsnitlig performance for hele 3-døgnsprognozen

Det ses at CNN+offset og Persistence modellerne performer bedst med en RMSE på henholdsvis 0,025 og 0,024 og begge med en gennemsnitlig afvigelse på 2,1 cm. I Figur 21 ser vi prognoser fra fire tilfældig valgte tidspunkter. Vi ser, at CNN+offset modellen klarer sig godt i de tilfælde, hvor vandstanden enten stiger eller falder, med en afvigelse der varierer mellem 1,2 cm og 2,8 cm. Faktisk ser vi også, at CNN+offset modellen stort set rammer rigtig til sidst i prognosen. Derimod ser vi i øverst højre hjørne, både ML-modellen og DHI-modellen tror, at vandstanden falder, selvom den forbliver forholdsvis stabil. Dette kommer også til udtryk ved, at CNN+offset modellens prognose i dette tilfælde afviger med 4,7 cm i gennemsnit og til sidst afviger med omkring 7-8 cm.

CNN+offset modellen performer bedst af de to CNN-modeller ved Kongensbro. Her får vi, at $R^2_{model} = 0.23$, når CNN+offset modellen holdes op imod DHI-modellen.



Figur 21 - Kongensbro: 3-døgnsprognoser for CNN og DHI. Fire tilfældige perioder er vist

Peak værdier i 2021

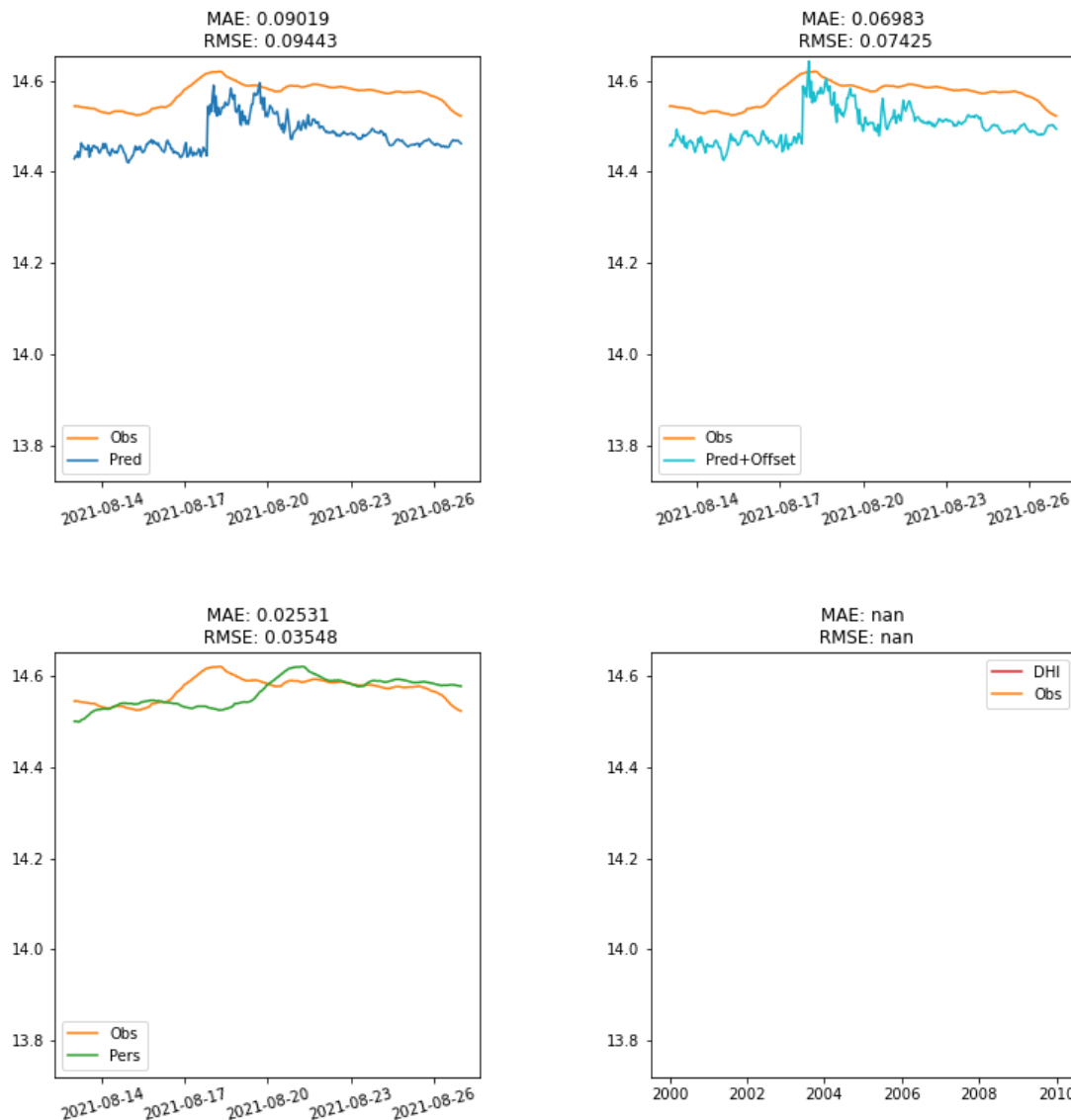
Som tidligere nævnt benyttes kun sidste observationer i prognoserne ($t=72$ timer), når performance for toppene beregnes, dvs. det tidspunkt som er sværest at forudsige (Se afsnit 4.2 for uddybning).

Af Figur 22 fremgår det, hvilke toppe i 2021 som modellen har været i stand til at lave tilstrækkelig med prognoser for. De toppe, hvor det ikke har været muligt at lave nok prognoser til skyldes manglende input data, såsom nedbørsdata fra enkelte målestationer. Performance for alle toppene kan findes i Bilag 2.



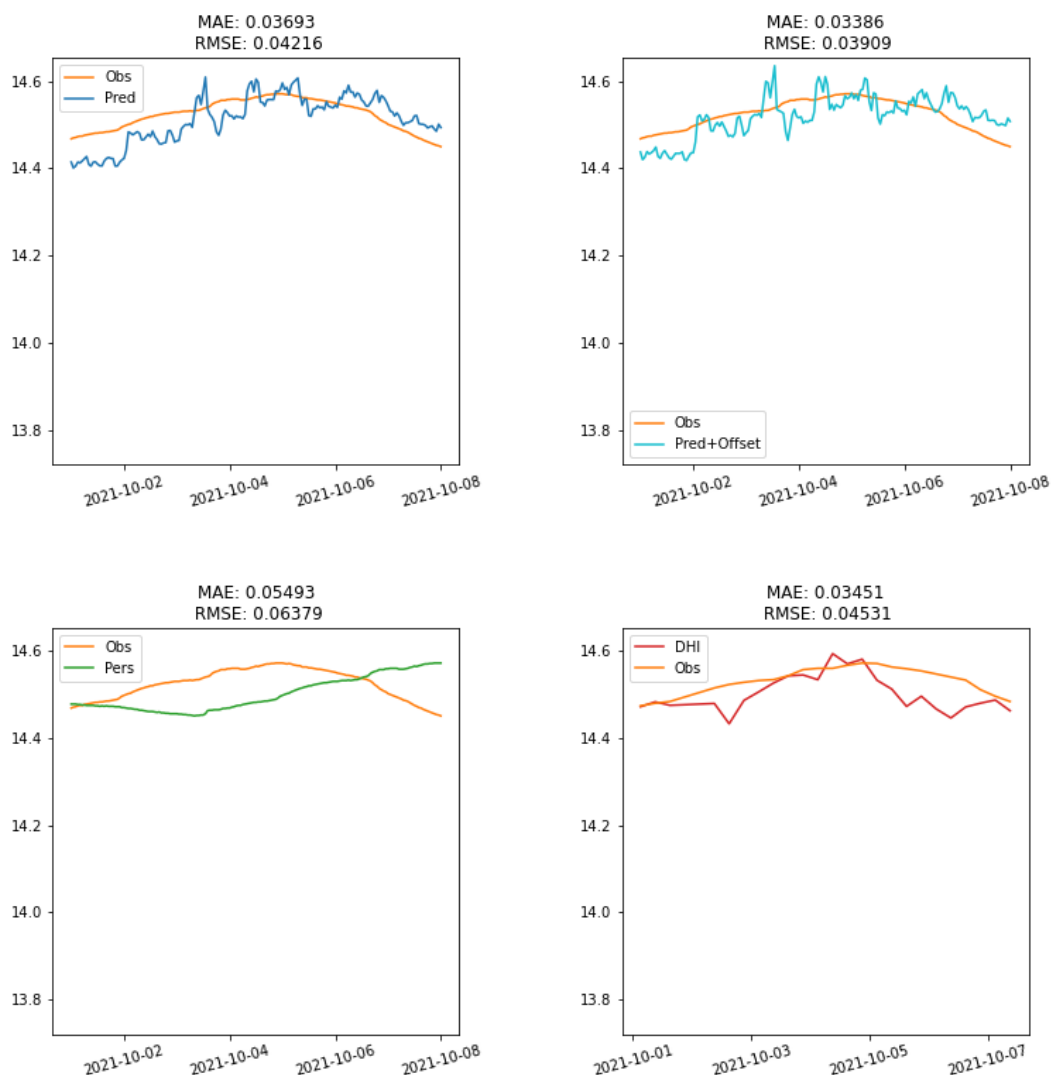
Figur 22 - Kongensbro: Toppe i 2021 som modellen har kunnet lave prognoser for (blå)

Som det ses i ovenstående figur, så forekommer den højeste vandstand omkring perioden 13. - 27. august 2021, se Figur 23. Her fremgår det, at den CNN+offset modellen i gennemsnit afviger med 6,9 cm i denne periode. Generelt er modellens vandstand lavere end den observerede værdi. Persistence modellens performance er generelt bedre, men når ikke at opfange, når der kommer en stigning i vandstanden d. 17. august.



Figur 23 - Kongensbro: Performance for modeller i perioden 13. - 27. august 2021. Ingen data for DHI-modellen

Betragtes perioden 1. - 8. oktober 2021 kan det aflæses i Figur 24, at CNN+offset modellen i gennemsnit ligger 3,3 cm fra den rigtige værdi, imens Persistences modellens gennemsnitlige afvigelse er 5,4 cm. Nederst højre hjørne fremgår det, at DHI-modellen afviger med 3,4 cm i gennemsnit. Dog er DHI-modellens RMSE på 0.045, hvorimod at CNN+offset modellen har en RMSE på 0.039. Dette indikerer, at DHI-modellen har visse steder med større afvigelse, fx. d. 6. oktober.



Figur 24 - Kongensbro: Performance for modeller i perioden 2021-10-01 til 2021-10-08

I Tabel 14 ses den gennemsnitlige performance ved alle toppene. Dette stemmer overens med den tendens, som vi så ved toppen i perioden 1. - 8. oktober 2021. CNN+offset modellen har den bedste MAE. CNN+offset- og DHI-modellen adskiller sig på RMSE-parameteren, hvilket antyder, at DHI-modellen har visse steder hvor dens forudsigelse ligger langt fra den rigtige vandstand.

	Model	MAE	RMSE
Gennemsnit for alle toppene	CNN	0.056	0.064
	CNN+offset	0.051	0.059
	Persistence	0.067	0.078
	DHI	0.056	0.071

Tabel 14 - Kongensbro: Gennemsnitlig performance ved alle toppene

Alle performance for toppene ved Kongensbro kan findes i Bilag 2.

4.2.6 Ulstrup

Performance af sidste observation i prognoserne (t=72)

ML-modellernes, Persistences og DHI-modellens performance for den sidste observation (t=72) i perioden primo september 2021 til primo marts 2022 kan aflæses i Tabel 15.

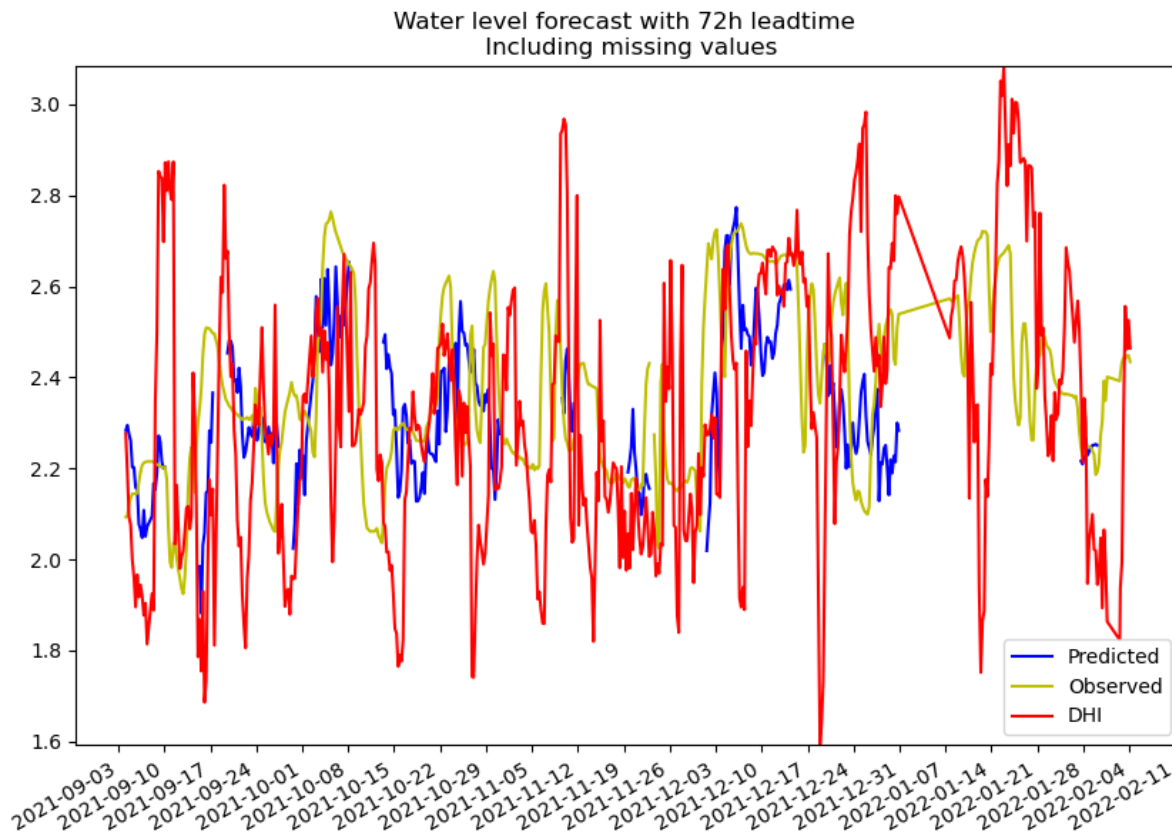
	Model	R^2	MAE	RMSE
Gennemsnitlig performance	CNN	0.05	0.150	0.191
	CNN+offset	-0.151	0.165	0.21
	Persistence	-0.27	0.164	0.221
	DHI	-1.77	0.247	0.326

Tabel 15 - Ulstrup: Gennemsnitlig performance for sidste observation i hver prognose i perioden sep. 2021 og frem

Vi ser, at modellerne performer dårligere her end ved de andre stationer. Det fremgår af de lave R^2 -værdier, at ingen af modellerne formår at finde tendenserne for om vandstanden ved Ulstrup stiger, falder eller er uændret. Det bedste der kan opnås ved Ulstrup er CNN-modellen, der med en MAE på 0.150 i gennemsnit ligger 15 cm fra den rigtige værdi, når den skal forudsige vandstanden om præcis 3 dage ud i fremtiden.

Af Figur 25 fremgår det, at prognoserne (blå og rød) fluktuerer meget og ligger langt fra den rigtige vandstand (orange).

Vi får, at $R_{model}^2 = 0.658$, når vi sammenligner CNN-model med DHI-modellen. Sammenholdes dette med de tre metrikker i Tabel 15 og Figur 25, så er CNN modellen væsentlig bedre end DHI-modellen, dog performer CNN-modellen stadig dårligt ved Ulstrup.



Figur 25 - Ulstrup: Observerede vandstand, samt ML-modellen og DHI modellens forecast for vandstanden

Hele 3-døgnsprognozen

Modellernes 3-døgnsprognozer evalueres her i perioden september 2021 til marts 2022. Som tidligere benyttes der kun i denne tidsperiode, hvor det er muligt at sammenligne med DHI.

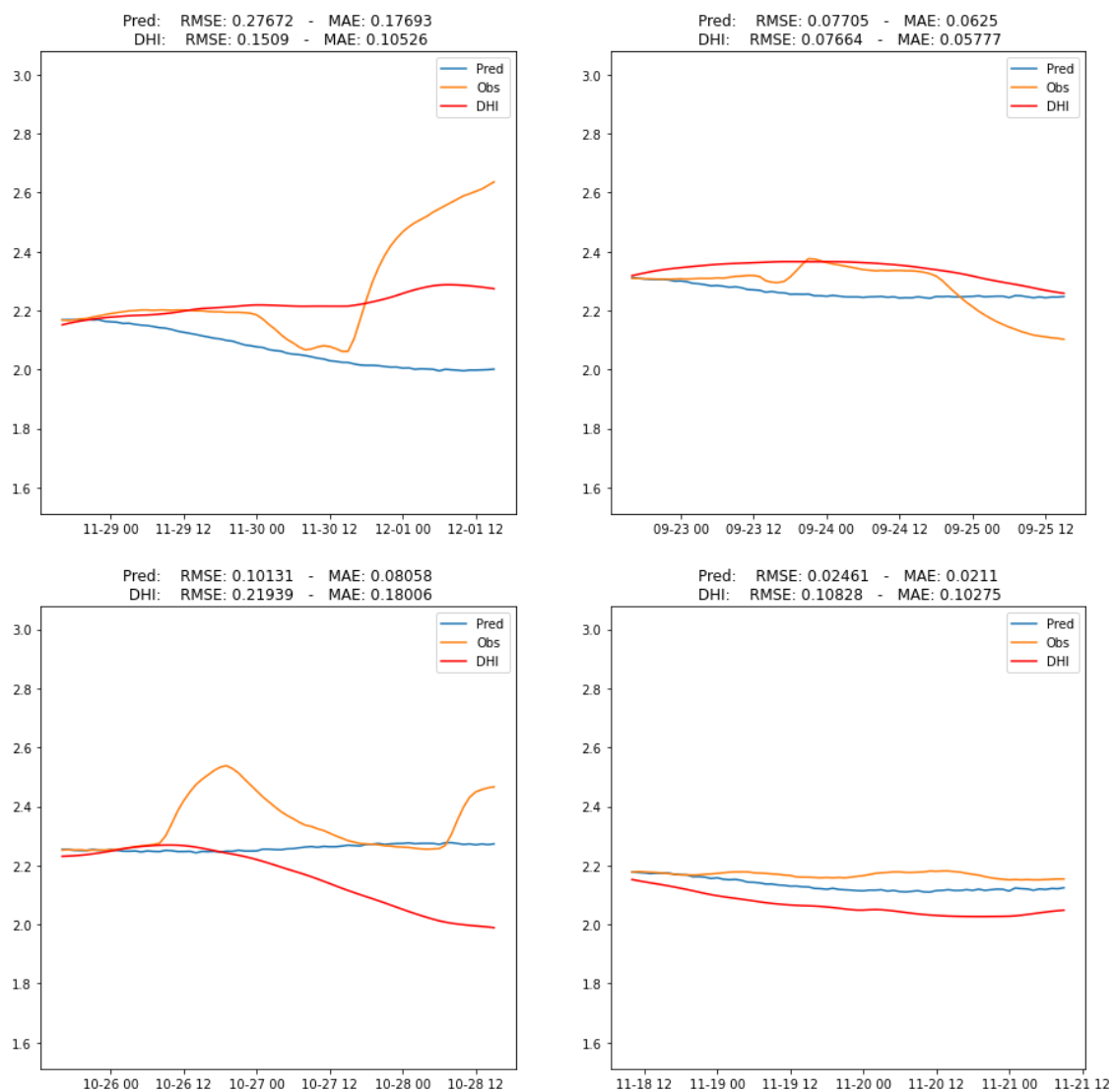
Tabel 16 nedenfor viser, at CNN-modellen er den bedste af de to ML-modeller med en MAE på 0,108, hvilket betyder at alle værdierne i en prognose produceret af CNN- modellen i gennemsnit ligger 10,8 cm fra den rigtige værdi. Her ses det også, at DHIs modellens performance er dårligere med en MAE på 0,177.

	Model	MAE	RMSE
Gennemsnitlig performance	CNN	0.108	0.126
	CNN+offset	0.122	0.141
	Persistence	0.106	0.125
	DHI	0.177	0.202

Tabel 16 - Ulstrup: Gennemsnitlig performance for alle 3-døgnsprognozer

Dette viser igen, at alle modellerne har problemer med at forudsige vandstanden ved Ulstrup, hvilket også fremgår af Figur 26, hvor det tydeligt ses, at hverken DHI (rød) eller CNN-modellen (blå) opfanger de pludselige ændringer i vandstanden (orange linje). Disse pludselige ændringer er forårsaget af Tangeværkets styring og der henvises til diskussionen i afsnit **Fejl! Henvissningskilde ikke fundet.** for nærmere beskrivelse af problematikken.

Igen er CNN bedre end CNN+offset og vi får, at $R^2_{model} = 0.318$, når denne model holdes op i mod DHIs.



Figur 26 - Ulstrup: 3-døgnsprognoser for CNN og DHI. Fire tilfældige perioder er vist

Evaluering af modellens performance i toppene i 2021 er i dette tilfælde udeladt, da modellen i sin nuværende form ikke kan benyttes ved Ulstrup grundet styringen ved Tangeværket.

4.3 Opsummering

- Modellerne kan mest præcis forudsige vandstanden ved Klostermølle med en R^2 -værdi på 0.953 og gennemsnitlig afvigelse på 1,8 cm fra observerede vandstand.
- ML-modellerne (CNN og CNN+offset) er generelt bedre end DHI-modellen (Undtagen for t=72 ved Silkeborg Langsø)
- Alle modellerne har problemer med at forudsige vandstand og tendenser ved Ulstrup grundet Tangeværket.
- ML-modellerne er bedre end DHI-modellen på 4 ud af de 5 steder i Gudenåen, hvor der er lavet sammenligning

5. Diskussion

5.1 Fordele og ulemper med ML-model kontra numerisk model

I nærværende projekt er der opstillet en ML-model for Gudenå og denne sammenholdes med en eksisterende hydraulisk (numerisk) model lavet af DHI. Det er to meget forskellige modeltilgange, som her sammenlignes.

Den hydrauliske model er opbygget med en grundforståelse af vands fysiske egenskaber og hvad disse betyder for hastighed, vandstand osv. For at den hydrauliske model er retvisende er det nødvendigt at indstille forskellige parametre korrekt. Det kan være ruheden/Manningtallet, vandløbets fysiske dimensioner og oplandsstørrelser. Derudover er der en lang række fysiske parametre, hvor standardværdier benyttes. Modellens resultat bliver kun så god, som de parametre, der bruges som input. Når modellen er opstillet laves en kalibrering, som har til formål at 'skruer' på parametrene, så modellen bedre passer med virkeligheden. En parameter, som ofte ændres ved kalibrering er Manningtallet, da denne er meget usikker (det er svært at vurdere den præcise modstand i et vandløb, og dennes konstante ændring i løbet af året). En stor fordel ved den hydrauliske model er, at man kan lægge en 'forståelse' af vandløbssystemet ind: man kan fx beskrive, hvor der er søer eller opstemninger og dette kan gøre resultatet mere retvisende nedstrøms disse fysiske 'forandringer'.

ML-modellen bruger nogle af de samme input (nedbør og vandføring), men finder resultatet ved selv at lære en sammenhæng mellem input og output. Modellen så at sige kalibrerer sig selv løbende, indtil den kan 'regne ud', hvad resultatet bliver, ud fra de input, den er præsenteret for. Ved ML-modellen er der ingen forhåndsviden om fx vandløbsdimensioner eller søer og opstemninger. Konsekvensen af disse bør modellen selv kunne lære, når den bare bliver præsenteret for tilstrækkeligt data, hvor alle mange situationer forekommer. Det er derfor vigtigt med et stort lærings-datasæt, altså en lang tidsserie af nedbør og observerede vandstande.

ML-modellen kan altså selv finde sammenhænge mellem data, som kan være svære eller tidskrævende at beskrive i en hydraulisk model. Men ML-modellen skal have store mængder data for at kunne se den korrekte sammenhæng.

Udfordringen ved ML-modellen, som vi også ser i projektets resultater, er at finde sammenhænge, som er ulogiske. Når Tangeværket oplyser, at de ikke har en egentlig styring af, hvor meget vand de leder gennem værket, så er det svært for ML-modellen at finde en sammenhæng, som den kan bruge til at forudsige vandstanden i fremtiden. Det kan derfor vise sig meget svært at forudsige

vandstanden visse steder i Gudenå med kun en ML-model. Her kan en hybrid af hydraulisk- og ML-model måske vise sig bedre.

5.2 Beregning af vandstand nedstrøms Tangeværket eller andre opstemninger

De flere fysiske opstemninger i vandløbet har stor betydning for den vandstand og vandføring, som der er nedstrøms. Det største eksempel er Tangeværket, som med styringen af turbiner kan ændre markant på vandføring og dermed den vandstand, der er i Gudenåen langt nede i vandløbet. Efter at have kontaktet direktør Rasmus Lambert, står det klart, at der ikke er en entydig, 'gennemskuelig' styring, som er nødvendig for ML-modellen kan lære sammenhængen. Det er primært de ansatte, som bestemmer, hvor mange turbiner, der er åbne og hvor meget vand der ledes gennem værket. Deres vurdering kan bygge på vejruddsigten, udfordringer på elnettet, et ønske om en minimums og maksimums kote i søen eller andre parametre.

I andre tilfælde (fx opstemning ved Ry) er det endnu mere manuel styring, hvor der fjernes eller opsættes stemmebrædder efter lyst og behov.

I en fuld model for hele Gudenåen skal det undersøges, hvordan styringen af hver enkelt opstemning fungerer. Måske kan en hybrid af ML- og hydraulisk model være løsningen. I den hydrauliske model er det nemmere at beskrive de fysiske forhold ved fx stemmekanter m.v. Men resultaterne fra DHI-modellen, som vi har sammenlignet med, viser også, at denne har svært ved at forudsige vandstanden nedstrøms Tangeværket.

5.3 Tidsserier fra målestationer

For en del af målestationerne er der kun en kort periode med data, se afsnit 3.1.3. Det er ikke undersøgt, hvor meget data der er tilgængelig på de resterende målestationer i Gudenå, men stationerne i dette projekter er udvalgt, da de er nogle af dem med længst data. Så andre steder i Gudenå kan det vise sig, at der er endnu dårligere data til rådighed.

Der bør dannes et overblik over al tilgængelige data, før vi endelig kan vurdere, hvor godt ML-modellen vil performe alle steder i vandløbet. Det kan blive en hæmsko for os, hvis der er ringe grundlag på de strækninger, som vi ikke har haft fokus på indtil nu.

5.4 Silkeborg Langsø, Ringvejsbroen

ML modellernes ringe performance ved Silkeborg Langsø, Ringvejsbroen kan skyldes mængden af data som er tilgængelig. Ved Silkeborg Langsø er der kun 2 års data og som nævnt i afsnit 3 og 5.1, så skal ML-modellerne bruge en vis mængde data for at lære sammengænge. For at kunne lave prognoser til hele 2021 har modellen kun haft 1 år at træne på. Havde modellen haft mere data kunne det tænkes, at dens performance ville være bedre, da den ville have bedre grundlag for at lære generelle tendenser for vandstanden ved Langsø.

Desuden har der været problemer med at finde en god opstrøms målestation (som også bruges som input til ML-modellen). Der er valgt en station ved Rye Mølle, men søerne mellem Ry og Silkeborg har en dæpende effekt på dynamikken i vandløbet.

Man kunne have valgt en opstrøms station tættere på, men vi valgte denne, da stationen skal være langt nok oppe i systemet, så der er en vis forsinkelse/responstid mellem vandstanden ved Langsø og vandstanden opstrøms. Desuden var tidsserierne meget korte for nogle af de andre stationer mellem Langsø og Ry og derfor ikke brugbare.

ML-modellen har desuden svært ved at finde sammenhænge mellem data, da der er en opstemning ved Silkeborg Papirfabrik. Der var ikke tid i projektet til at dykke nærmere ned i sammenhænge

mellem vandstand opstrøms & nedstrøms opstemningen. En sammenhæng kan sikkert findes, og med indarbejdelse af denne i ML-modellen er det sandsynligt at performance bliver bedre. Dette er en opgave, hvis projektet fortsættes.

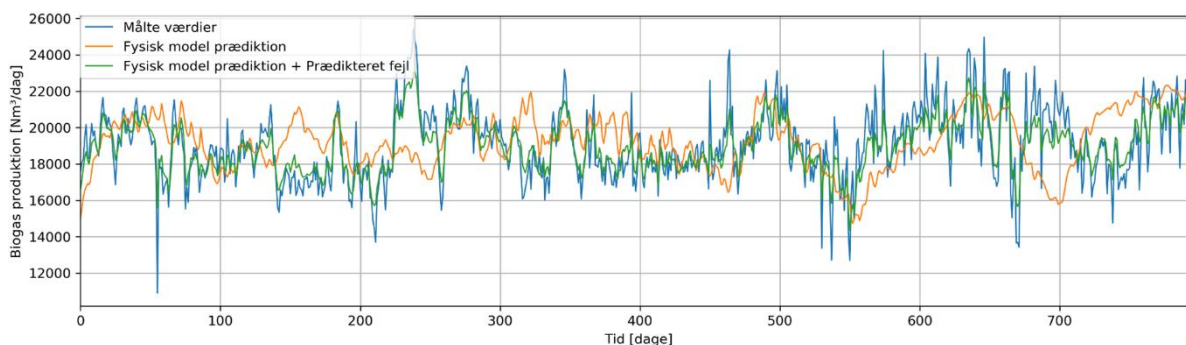
5.5 Hybridmodellering

Der er fordele og ulemper ved numerisk modellering (fysisk HD-model) og data-drevne modeller (Machine Learning modeller), hvilket er en direkte årsag til, at der er blevet og forsat, bliver forsket meget i, hvordan man bedst muligt kan kombinere disse to tilgange. Overordnet set kan disse opdeles i tre metodikker.

- (A) Del-udskiftning af numerisk model med ML-model
- (B) Korrektion af numerisk model resultater
- (C) Optimering af numerisk skema for hastighedsforøgelse (se f.eks. [Li et al. 2020](#))

Af (A) menes med en del-udskiftning, at dele af den numeriske model udskiftes med en ML-model. I denne kontekst kunne det være en del-strækning af Gudenå, hvor de numerisk beregnede resultater (vandføring/vandstand) ikke er interessante, men udfordringer på denne strækning påvirker modelresultaterne længere nedstrøms. En ML-model kan hermed benyttes til at erstatte denne del og resultaterne herfra assimileres direkte i den numeriske model for at levere en samlet set bedre prognose nedstrøms i modellen.

Ved metode (B) forudsiges fejlen mellem den numeriske model og observation af en ML-model. Denne prædikterede fejl adderes med resultaterne fra den numeriske model og herved korrigeres, hvilket samlet set giver en bedre prognose. Sagt med andre ord estimeres en delta-værdi, der benyttes til at korrigere den numeriske models prognose. EnviDan har benyttet denne metodik i forbindelse med forudsigelse af biogas-produktion, se Figur 27 og [Hansen et al. 2020](#).



Figur 27 - Sammenligning mellem målte værdier for Methan produktionen (blå), en fysisk model og fysikmodel (orange), der er korrigeret med deltaværdier prædikeret af Machine Learning model (grøn).

Som det fremgår af Figur 27, så blev der opnået væsentligt bedre resultater end den fysiske model. ML-modellen var altså i stand til at finde en sammenhæng mellem input data (biomasse indfødning) og den fejl, som den fysiske model havde i sin prognose af biogasproduktionen. Dette er ligeledes en formodning i konteksten af Gudenå-modellen.

Den tredje metodik (C) er state-of-the art og er forsat meget forskningsorienteret. Her benyttes ML-metodikker til at optimere beregningstiden af en numerisk model. I den videnskabelig litteratur, er dette opnået i mange forskellige applikation bl.a. i løsningen af Navier-Stokes Equation og turbulente flows. Dette har langt rækkende konsekvenser, hvis det ligeledes bliver kommercielt implementeret, da f.eks. Computational Fluids Dynamics (CFD) vil kunne beregnes med en hastighed, hvor det kan benyttes til styring. Med andre ord bliver det realistisk at lave numeriske modeller med meget høj spatial og temporal opløsning.

Af mulighederne for at opstille en hybrid model, så er det i denne kontekst primært metodik (B), der er af relevans. Anbefalingen er derfor, at hvis f.eks. DMI kommer med en prognosemodel eller det ønskes, at nuværende model fra DHI skal optimeres, så er metodik (B) en direkte implementerbar tilgang. Dette kan optimere prognose-performance i så fald ML-modellen kan finde sammenhænge mellem input og den numeriske modelprognosefejl, hvilket i andre tilfælde er eftervist. En lidt uvidenskabelig måde at beskrive dette på, så kan den kommende DMI-prognose tilpasses lokalt til Gudenå uden at være specifikt kalibreret hertil. Yderligere giver det en række fordele fra begge modelleringsverdener og kan være med til at forbedre performance bl.a. ved målestation Silkeborg Langsø, Ringvejsbroen.

5.6 Driftsomkostning for produktionsmoden ML-prognosemodel

Nærværende Proof-of-concept (POC) for prognostisering af vandstand vha. ML-model er netop en POC. Det betyder, at den ikke kører i realtid, men er baseret på en historisk analyse af performance. For at operationalisere ML-modellen, så skal der tages højde for en række andre parametre såsom

- IT-infrastruktur
 - Data pipelines & monitorering
 - Data validering
 - Data imputering (håndtering af udfald af input data)
 - Gentræningscyklus for ML-modellen & herunder triggere for gentræning (primært ændringer i forudsætninger for input-data og model-performance)
- Implementering af Non-Functional-Requirements (NFR)
- Udvikling af front-end til visualisering af prognoser
- Opbygning af varslingsystem og advisering

Der er mange uafklarede parametre i forhold til ønsker fra Silkeborg Kommune i forhold til ovenstående, hvorfor en nærmere prissætning for udviklingen kan være svær for nuværende at fastsætte. Udgifter kan opdeles i henholdsvis en del, der er medgået tid i forhold til udvikling og implementering af ovenstående punkter samt en "abonnement" pris, der dækker over de udgifter der løbende er i forbindelse med at drive systemet.

Et indikativt overslag vil være, at udviklingen koster i omegnen af 500-750 TDKK med en månedlig abonnementspris på 8,5 - 10 TDKK (100-120 TDKK/år).

6. Afklaring af spørgsmål

Silkeborg Kommune har d. 5/4 sendt nogle afklarende spørgsmål, som her besvares:

Varsler modellen kun på punkter eller kan den også sige noget mellem punkterne - hvis ikke hvor tæt vil det så være fordelagtigt at loggerpunkterne står i systemet?

- Modellen bruger data fra loggere til at forudsige sammenhænge mellem nedbør og vandstand, så det vil primært være der, hvor der allerede er en målestation, at ML-modellen kan forudsige værdier.
- Hvis der opstilles en hybrid mellem ML- og hydraulisk model, kan denne bruges til at beregne vandstande mellem målestationer, der er korrigeret ud fra ML-modellen, se tidligere beskrivelse.
- Loggere bør være placeret der, hvor der sker forandringer og hydraulikken ikke er så forudsigelig samt hvor det er af interesse at kende vandstanden i en varslingskontekst. Det er med andre ord svært at give en optimal afstand for disse. Uanset, så vil flere målestationer være brugbare både i forhold til numeriske modeller og ML-modellerne. Vi vil gerne kigge yderligere ind i dette emne.

Hvor lange tidsserie anbefales før at der ML er validt? Tænker ift. opsætning af nye loggere hvor længe skal der gå før der er "brugbare" i ML regi?

- Tidsserierne i nærværende projekt er ca. 3-4 år lange, men fx ved Silkeborg Langsø er perioden kun 2 år, hvor det ene år er brugt som 'lærings-data' mens det andet år er brugt som 'validerings-data'. Modellen hér viser dårligere performance, så forventeligt skal der være min. 3 års data for at modellen kan lære sammenhængene godt. Dette kan dog variere fra lokation til lokation. Yderligere, så er det også en spørgsmål om, hvor godt en ML-model kan generalisere - det har bl.a. betydning for, hvor godt en model repræsenterer "edge-cases" f.eks. ekstreme tilfælde.
- Nye opstillede loggere kan fra dag 1 indgå som validering/kalibrering af modellen, så de vil ikke være forgæves at stille op, da de vil indgå i en allerede opstillet ML-model, hvor der også er længere tidsserier.

DHI-modellen kan i dag varsle selvom der er udfald i loggerdata. Den viser også en oversvømmelses-polygon på varslingsiden. Vil en fuldstændig ML-model også kunne det?

- ML-modellen bruger både observationer og nedbør til at forudsige vandstanden, så selvom der sker nedbrud i en af loggerne, vil det stadig være muligt at få resultater, om end mindre præcise. Der er mange muligheder for at håndtere dette både med modeller, som er i stand til at lave en prædiktions med udfald i input data og foretage imputering af data. Imputering af data er "udfyldning af data", som igen kan gøres med en middelværdi, statistiske metodikker eller ligefrem prædiktes af støttende ML-modeller.
- Hvis modellen opstilles med flere input (p.t. er der én nedbørsmåler og én opstrøms vandløbslogger) er den også mere robust for udfald.

Silkeborg Kommune

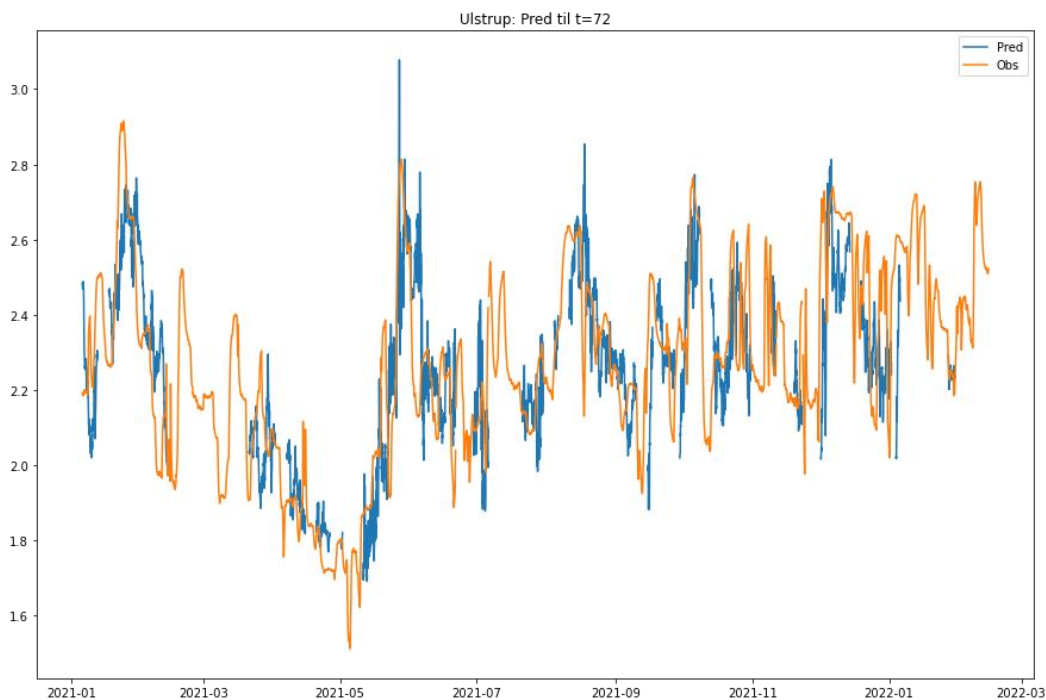
Proof-of-Concept: Gudenå Machine Learning

Bilag 1 - Resultater fra ML-modellerne

EnviDan

1. Bredstenbro

Her kigger vi på den sidste observation i hver prognose fra januar 2021 til marts 2022. CNN modellen og den rigtige vandstand er vist i figuren herunder.



Den gennemsnitlige performance for den sidste observation i hver prognose i perioden januar 2021 til marts 2022 fremgår af følgende tabel.

Station	Model	R^2	MAE	RMSE
Bredstenbro	CNN	0.802	0.061	0.083
	CNN+offset	0.745	0.072	0.094
	Persistence	0.545	0.087	0.126

I tabellen nedenfor er den gennemsnitlige performance for 3-døgnsprognerne fra januar 2021 til marts 2022

Station	Model	MAE	RMSE
Bredstenbro	CNN	0.036	0.042
	CNN+offset	0.04	0.047
	Persistence	0.048	0.056

2. Langsø, Ringvejsbroen

Af tabellen nedenfor fremgår den gennemsnitlige performance for den sidste observation i hver prognose i perioden januar 2021 til marts 2022.

Station	Model	R^2	MAE	RMSE
Langsø, Ringvejsbroen	CNN	0.897	0.058	0.078
	CNN+offset	0.914	0.052	0.072
	Persistence	0.942	0.044	0.059

Af figuren nedenfor kan CNN+offset modellen forudsigelse med en 72 timers horisont ses sammen med den faktiske vandstand.



I tabellen nedenfor er den gennemsnitlige performance for 3-døgnsprognerne fra januar 2021 til marts 2022

Station	Model	MAE	RMSE
Langsø, Ringvejsbroen	CNN	0.058	0.078
	CNN+offset	0.052	0.072
	Persistence	0.044	0.059

3. Kongensbro

Den gennemsnitlige performance for den sidste observation i hver prognose i perioden januar 2021 til marts 2022 kan aflæses i nedenstående tabel.

Station	Model	R^2	MAE	RMSE
Kongensbro	CNN	0.923	0.05	0.061
	CNN+offset	0.935	0.044	0.056
	Persistence	0.927	0.044	0.592

Nedenfor kan CNN+offset modellen forudsigelse med en 72 timers horisont ses sammen med den faktiske vandstand.



I tabellen nedenfor er den gennemsnitlige performance for 3-døgnspogoserne fra januar 2021 til marts 2022

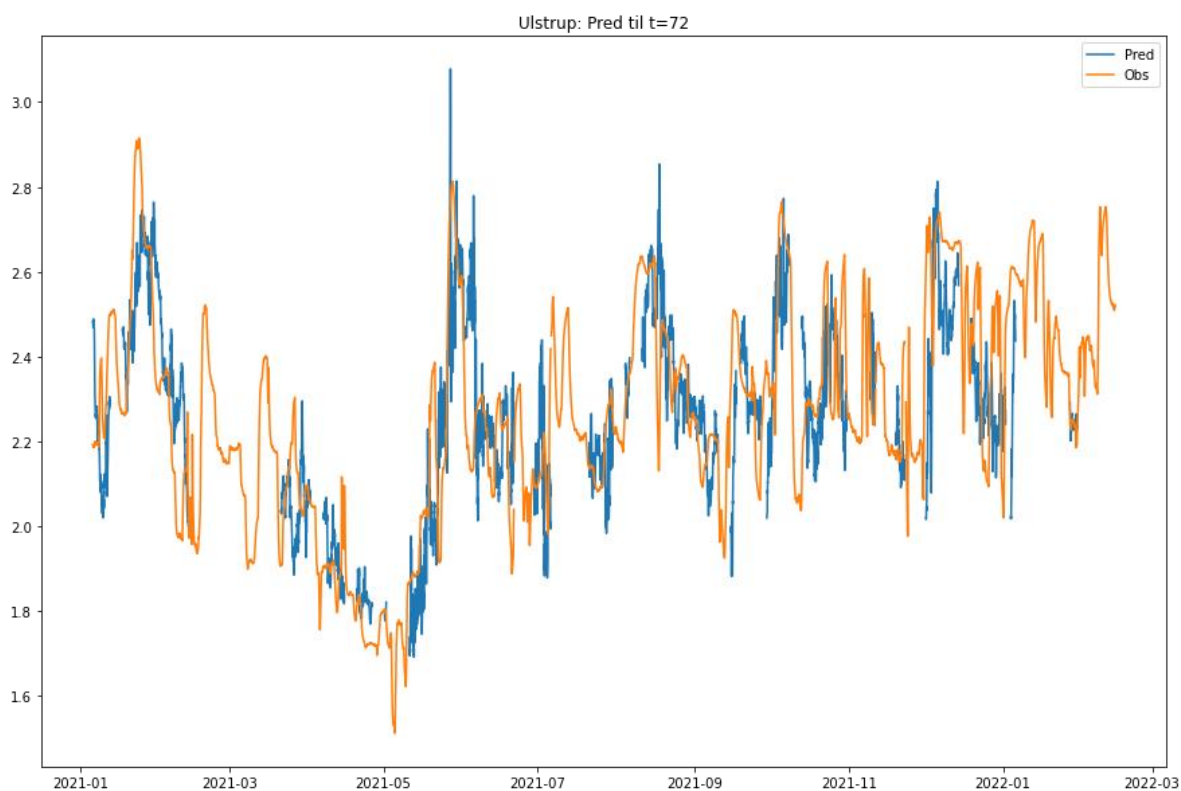
Station	Model	MAE	RMSE
Kongensbro	CNN	0.033	0.037
	CNN+offset	0.026	0.03
	Persistence	0.025	0.03

4. Ulstrup

Den gennemsnitlige performance for den sidste observation i hver prognose i perioden januar 2021 til marts 2022 fremgår af følgende tabel.

Station	Model	R^2	MAE	RMSE
Ulstrup	CNN	0.493	0.137	0.178
	CNN+offset	0.401	0.149	0.193
	Persistence	0.313	0.154	0.207

Nedenfor kan CNN modellen forudsigtelse med en 72 timers horisont ses sammen med den faktiske vandstand.



I tabellen nedenfor er den gennemsnitlige performance for 3-døgnspgnooserne fra januar 2021 til marts 2022

Station	Model	MAE	RMSE
Ulstrup	CNN	0.097	0.114
	CNN+offset	0.104	0.122
	Persistence	0.094	0.112

Silkeborg Kommune

Proof-of-Concept: Gudenå Machine Learning

Bilag 2 - Værdier for peaks

EnviDan

5. Peaks

Tabeller med performance for modellerne når vandstanden ved den givende station danner et "peak".

5.1 Bredstenbro

Periode	Model	MAE	RMSE
2021-01-19	CNN	0.09	0.115
	CNN+offset	0.09	0.112
2021-01-28	Persistence	0.27	0.307
	DHI	-	-
2021-05-25	CNN	0.098	0.13
	CNN+offset	0.109	0.145
2021-06-02	Persistence	0.189	0.228
	DHI	-	-
2021-07-28	CNN	0.07	0.085
	CNN+offset	0.087	0.102
2021-08-25	Persistence	0.084	0.099
	DHI	-	-
2021-09-10	CNN	0.049	0.061
	CNN+offset	0.071	0.082
2021-09-25	Persistence	0.08	0.09
	DHI	0.092	0.096
2021-09-30	CNN	0.063	0.075

2021-10-11	CNN+offset	0.064	0.076
	Persistence	0.108	0.124
	DHI	0.084	0.097
2021-10-19	CNN	0.067	0.089
	CNN+offset	0.077	0.101
2021-10-30	Persistence	0.155	0.175
	DHI	0.12	0.135
2021-12-01	CNN	0.062	0.087
	CNN+offset	0.079	0.105
2021-12-12	Persistence	0.184	0.199
	DHI	0.29	0.486
Gennemsnit	CNN	0.071	0.092
	CNN+offset	0.083	0.103
	Persistence	0.153	0.175
	DHI	0.148	0.204

5.2 Klostermølle

Periode	Model	MAE	RMSE
2021-01-20	CNN	0.019	0.024
	CNN+offset	0.029	0.032
2021-02-05	Persistence	0.034	0.037
	CNN	0.042	0.046
2021-05-28	CNN+offset	0.025	0.029
	Persistence	0.034	0.041
2021-06-10	CNN	0.015	0.019
	CNN+offset	0.016	0.016
2021-08-08	Persistence	0.013	0.015
	CNN	0.015	0.019
2021-08-28	CNN+offset	0.016	0.016
	Persistence	0.013	0.015

2021-09-10	CNN	0.016	0.021
	CNN+offset	0.014	0.018
2021-09-30	Persistence	0.018	0.021
2021-10-20	CNN	0.018	0.020
	CNN+offset	0.012	0.015
2021-11-02	Persistence	0.19	0.021
2021-12-01	CNN	0.029	0.035
	CNN+offset	0.02	0.024
2021-12-24	Persistence	0.025	0.033
Gennemsnit	CNN	0.023	0.027
	CNN+offset	0.019	0.023
	Persistence	0.24	0.028

5.3 Langsø

Periode	Model	MAE	RMSE
2021-01-20	CNN	0.056	0.078
	CNN+offset	0.054	0.075
2021-01-30	Persistence	0.055	0.071
	DHI	-	-
2021-05-25	CNN	0.047	0.063
	CNN+offset	0.052	0.070
2021-06-15	Persistence	0.055	0.078
	DHI	-	-
2021-08-06	CNN	0.117	0.132
	CNN+offset	0.105	0.121
2021-08-28	Persistence	0.034	0.037
	DHI	-	-

2021-10-19	CNN	0.077	0.089
	CNN+offset	0.089	0.102
2021-10-26	Persistence	0.034	0.04
	DHI	0.035	0.041
Gennemsnit	CNN	0.074	0.09
	CNN+offset	0.075	0.092
	Persistence	0.045	0.057
	DHI	0.035	0.041

5.4 Kongensbro

Periode	Model	MAE	RMSE
2021-01-20	CNN	0.05	0.058
	CNN+offset	0.042	0.054
2021-01-28	Persistence	0.098	0.113
	DHI	-	-
2021-05-24	CNN	0.069	0.078
	CNN+offset	0.09	0.101
2021-06-02	Persistence	0.132	0.147
	DHI	-	-
2021-08-13	CNN	0.09	0.094
	CNN+offset	0.07	0.074
2021-08-27	Persistence	0.025	0.035
	DHI	-	-
2021-10-01	CNN	0.037	0.042
	CNN+offset	0.034	0.039
2021-10-08	Persistence	0.055	0.064
	DHI	0.035	0.045

2021-10-18	CNN	0.042	0.05
	CNN+offset	0.031	0.038
2021-10-28	Persistence	0.033	0.04
	DHI	0.05	0.066
2021-12-01	CNN	0.053	0.059
	CNN+offset	0.039	0.047
2021-12-12	Persistence	0.057	0.069
	DHI	0.08	0.103
Gennemsnit	CNN	0.056	0.064
	CNN+offset	0.051	0.059
	Persistence	0.067	0.078
	DHI	0.056	0.071