



Realtidsfortolkning af Vandkvalitetsmålinger

WP2: Fremtidens Vandforsyning

Marts 2016

Titel:

Realtidsfortolkning af Vandkvalitetsmålinger

Forfattere:

Rasmus Boe-Hansen, Erling Fischer

Udgiver:

Miljøstyrelsen
Strandgade 29
1401 København K
www.mst.dk

Foto:

-

Illustration:

-

År:

2016

Kort:

-

ISBN nr.

[xxxxxx]

Ansvarsfraskrivelse:

Miljøministeriet offentliggør rapporter og indlæg vedrørende forsknings- og udviklingsprojekter inden for miljøsektoren, som er finansieret af Miljøministeriet. Det skal bemærkes, at en sådan offentliggørelse ikke nødvendigvis betyder, at det pågældende indlæg giver udtryk for Miljøministeriets synspunkter. Offentliggørelsen betyder imidlertid, at Miljøministeriet finder, at indholdet udgør et væsentligt indlæg i debatten omkring den danske miljøpolitik. Må citeres med kildeangivelse.

Indhold

Indhold	4
Forord.....	5
Konklusion og sammenfatning	6
UK summary	7
1. Indledning	8
1.1 Baggrund	8
1.2 Formål	8
1.3 Projektforløb	8
2. Realtidsovervågning.....	10
2.1 Koncept	10
2.2 Væsentlige driftshændelser	11
2.3 Detektionsalgoritmer.....	11
2.3.1 Målerfejl	12
2.3.2 Ekstremværdi	13
2.3.3 Multikriterie	13
2.4 Frontend (Brugerflade).....	13
2.5 Backend	14
2.6 Opsætning og implementering.....	15
2.7 Test og Demonstration	15
3. Resultater og diskussion.....	17
4. Formidling	20
5. Konklusion.....	21

Forord

FutureWater

Dette projekt blev udført som en selvstændig Workpackage under Fyrtårnsprojektet ”Fremtidens Vand”. Projektet blev støttet med midler fra Vandsektorens Teknologiudviklingsfond (VTUF) og Miljøteknologisk Udviklings- og Demonstrationprogram (MUDP).

Følgende partnere deltog i projektet:

- Krüger A/S
- Aarhus Vand A/S
- VandCenter Syd as
- DTU Miljø, Danmarks Tekniske Universitet

Konklusion og sammenfatning

Forsyningernes indsamling og generering af data er i hastig vækst. Det er ressourcekrævende og ekstremt komplekst at fortolke og anvendelige disse store mængder data.

En af de største udfordringer i at fortolke store mængder data fra mange kilder er at opstille og beskrive relationerne mellem datakilderne. I en traditionel deterministisk model beskrives relationerne som en matematisk sammenhæng mellem data. Dette er en effektiv metode når systemet er velbeskrevet og når antallet af datakilder er begrænsede og uændrede. Metoden kommer imidlertid hurtigt til kort overfor store datastrømme og dynamiske systemer, fordi ressourceforbruget til konfiguration og vedligeholdelse af modellerne vil være kraftigt stigende og i sidste ende begrænsende for udviklingen.

I projektet "Realtidsfortolkning af vandkvalitetsdata" har Krüger arbejdet med VCS, Århus Vand og DTU på at udvikle smarte algoritmer til at detektere driftsafvigelser baseret sensorer i vandbehandlingen.

UK summary

The amount of data generated by water utilities has increased dramatically during the recent years and the rate of data being collected seems to accelerate even further at present. This is to a large extent caused by an increased use of online sensors, which consequently calls for advanced data management and interpretation. Today, the amount of data being generated by a modern utility surpass the capabilities of manual interpretation, thus automated interpretation is no longer an option more than a need.

Our project aimed at developing real time interpretation software for the purpose of detecting water quality abnormalities using online sensors and event detection algorithms. The focus of the project was to:

- Secure and improve raw data quality
- Detect abnormal water quality events
- Develop a suitable GUI for visualizing and interacting with events

Following the development, the software was tested in the field at two different water works at the utilities of VCS Denmark and Aarhus Vand A/S.

1. Indledning

1.1 Baggrund

Vandforsyningerne forventes i de kommende år at blive mødt med stigende krav til optimering af den daglige drift. Kravene udspringer af forskellige interessenters ønsker og ambitioner for den danske vandsektor, herunder:

- Ønske om økonomisk effektivisering (udmøntet ved Vandsektorloven)
- Ambition om energireduktion (DANVA og El-sparefonde)
- Forbedret sikkerhed i driften, DDS (Drikkevandsdirektivet og forbrugerkrav om færre forureningshændelser)

Optimering af den daglige drift nødvendiggør, at der i langt højere grad end tidligere opstilles målere, der overvåger vandkvaliteten. Udviklingen ses i særdeleshed hos de største forsyninger, men vinder efterhånden også indpas på mindre vandværker.

Antallet af vandkvalitetsmålinger på vandværkerne de senere år har således været stærkt stigende. I flere danske vandværker måles således turbiditet, opløst ilt, ledningsevne, pH, UV-transmissivitet, temperatur, flow, tryk og filteropstuvning. Hertil kommer en række mere eksotiske målinger, der kun lejlighedsvist ses benyttet. Adskillige af parametrene måles flere steder i vandbehandlingssystemet. Herudover foretages rutinemæssige stikprøver af hensyn til myndighederne, der måles med akkrediterede vandkvalitetsanalyser.

Den øgede anvendelse af online målere har i sagens natur betydet, at mængden af data der indsamles er steget dramatisk. Det har vist sig vanskeligt i praksis at anvende disse store mængder rådata meningsfuldt. Bedre forståelse og udnyttelse af data åbner op for en bred vifte af spændende nye muligheder indenfor early warning systemer, DDS dokumentation, fejlfinding, optimering og forbedret design af anlæg/systemer.

Ovennævnte skitserede problemstilling er ikke kun et dansk fænomen. Trenden mod stigende antal målere og dramatisk forøgelse i datamængden ses i vandforsyningssektoren i hele den vestlige verden. Der vurderes således at være et stort behov for at finde løsninger der muliggør realtidsfortolkning af vandkvalitetsmålinger.

Det er i den forbindelse afgørende at data behandles og fortolkes i realtid. Dette stiller krav til udvikling af avancerede algoritmer og modeller med særlig fokus på at:

- Overvåge og forbedre kvaliteten de indsamlede data
- Udpege specifikke uønskede hændelser eller afvigelser fra normal drift
- Visualisere den aktuelle og historiske vandkvalitet
- Input til SCADA systemet med henblik på at optimere vandkvaliteten

1.2 Formål

Det overordnede mål med projektet er at give vandforsyninger nye muligheder for at behandle, visualisere og fortolke data fra forskellige online vandkvalitetsmålere i real tid og koble disse med øvrige SCADA inputs, samt relevante off-line målinger med henblik på optimeret drift og øget vandkvalitet.

I projektet vil der blive udviklet software, der kan understøtte operatørerne i den daglige drift samtidig med at det vil være et væsentligt værktøj ved optimering og fejlfinding. Værktøjet vil blive udviklet med en brugerflade, der gør det muligt at håndtere for driftspersonale med et minimum af træning.

1.3 Projektforløb

Projektet bestod af 6 delelementer:

- Konzeptudvikling
- Algoritmeudvikling

- Brugerfladeudvikling
- Opsætning og implementering
- Test og Demonstration
- Formidling

I de følgende afsnit beskrives hovedaktiviteterne for hvert delement.

2. Realtidsovervågning

2.1 Koncept

Konceptet for dette projekt omkring realtidsfortolkning af vandkvalitetsdata blev udviklet som resultatet af en række arbejds møder i projektgruppen, med driftsfolk fra vandforsyninger og med IT udviklere. Formålet var at afdække behov og tekniske muligheder samt tilpasse projektets ambition med de forhåndværende ressourcer.

Krüger udviklede på den baggrund et koncept til realtidsovervågning, hvor matematiske algoritmer på baggrund af store datamængder kan udpege specifikke hændelser af driftsmæssig interesse. Softwareudviklingen tog udgangspunkt i følgende fire kerneværdier:

Driftsorienteret. Softwaren forventes primært anvendt som beslutningsstøtteredskab af operatørerne i den daglige drift som en hjælp til at håndtere store mængder data fra mange kilder.

Brugervenlig. Softwaren skal give et hurtigt overblik og kunne anvendes af mange med kun lidt træning.

Hændelsesdetektion. Softwaren søger at detektere driftssituationer der kræver operatørens opmærksomhed eller indgriben.

Skalérbarhed. Mængden af data der skal håndteres i realtid forventes at være stigende og det er derfor nødvendigt at udvikle løsninger, der let kan skaleres med datamængderne.



Figur 2.1. Realtidsfortolkning af sensordata

Softwaren blev udviklet af Krüger som en del af STAR Utility Solutions, der er en programpakke af intelligent software målrettet vandforsynings- og spildevandsanlæg.

STAR Utility Solutions™
WATER MANAGER

Figur 2.2. Logo for Water Manager som en del af STAR Utility Solutions.

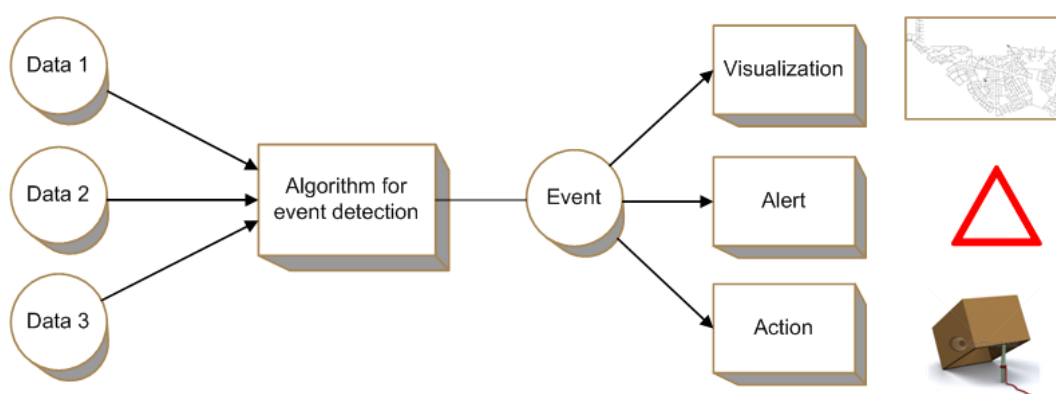
Softwaren indgår i dag som en del af Water Manager produktet (se figur 2.2).

2.2 Væsentlige driftshændelser

Tidligt i projektføreløbet udarbejdede projektgruppen en bruttoliste over væsentlige hændelser på vandværker, der ville kunne have relevans for driften og som potentielt burde kunne detekteres med sensorer (eksisterende eller fremtidige). Krüger udvidede listen med forslag til relevante sensorsignaler og skitser til detektionsalgoritmer, hvorefter projektgruppen prioriterede listen med henblik på at fokusere algoritmeudviklingen på netop hændelser/detektionsalgoritmer, der havde mest relevans for den daglige drift.

2.3 Detektionsalgoritmer

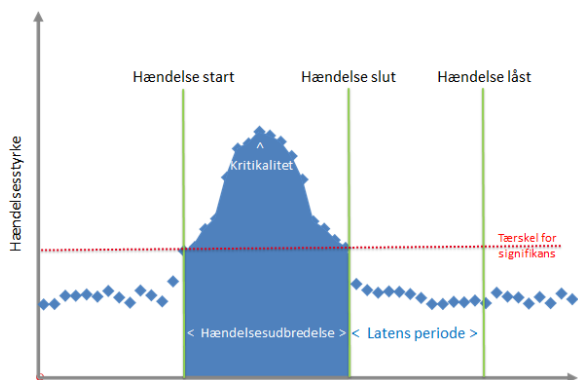
Som nævnt søger projektet at opstille af matematiske algoritmer, der leder efter bestemte mønstre i data med henblik på at detektere hændelser der afviger fra det normale (se figur 2.3).



Figur 2.3. Koncept for hændelsesdetektion.

Hver algoritme er programmeret til at lede efter et specifikt datamønster. Den enkelte algoritme kan være målerettet meget en specifik hændelse med et karakteristisk datamønster (f.eks. lækage) eller mere uspecifikke ekstreme værdier (f.eks. unormal rentvandskvalitet). Antallet af algoritmer der analyserer data er i princippet ubegrænset (men i praksis begrænset af computerens regnekraft). Nye algoritmer kan tilføjes løbende ligesom eksisterende algoritmer selvfølgelig kan tilpasses og forbedres.

Algoritmerne beregner løbende en "hændelsesstyrke" der anvendes til at afgøre om hændelsen er signifikant. Hvis tærsklen for signifikans overskrides startes en hændelsesforløbet, der først afsluttes når hændelsesstyrken igen er under signifikanstærsklen og en latensperiode er gennemløbet. I hændelsens udbredelsesperiode beregnes en kritikalitet der er proportional med den højeste hændelsesstyrke på det pågældende tidspunkt (se figur 2.4).



Figur 2.4. Hændelsesforløb.

Det betyder altså at hændelsens kritikalitet normalt vil øges efterhånden som hændelsen udvikler sig.

Projektets realtidsfortolkning benytter sig af fire forskellige typer algoritmer, der ofte anvendes i kombination, nemlig:

- Målerfejl
- Ekstremværdi
- Multikriterie

I de følgende afsnit er den grundlæggende funktion for de forskellige typer beskrevet i korte træk.

2.3.1 Målerfejl

Et centralt element i behandling af indkomne data er kvalitetskontrollen. Fejlbehæftede data kan let forstyrre fortolkningen og præcision, men kan heldigvis ofte let adskilles fra de øvrige målinger. Fejlbehæftede data opstår typisk ved deciderede fejl på en sensor eller kommunikationen, men kan også optræde periodisk i den almindelige drift i forbindelse med målerkalibrering, luftbobler osv. Visse typer af målerfejl kan selvsagt være meget vanskelige at detektere f.eks. et langsvarigt skred til måleværdier som følge af fouling eller ved fejlkalibrering. Datakvalitetskontrollen kan således ikke være en 100% garanti mod dårlig datakvalitet, men tjener til at øge troværdigheden af de indkomne data samt at advisere operatøren om potentielle måletekniske problemer.

Alle indkomne rådata gennemgik følgende fire tests:

Rate. Nye rådata sammenlignes med tidligere rådata med henblik på at vurdere om forskellen er indenfor acceptable rammer.

Range. Nye rådata sammenlignes med interval for normale målinger med henblik på at vurdere om værdien ligger indenfor acceptable rammer.

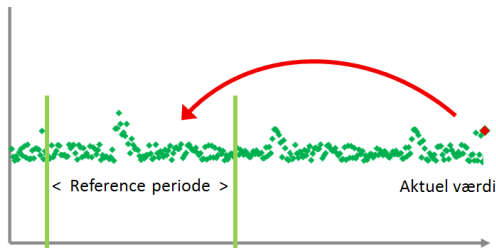
St.dev. Der beregnes en standardafvigelse for de seneste måleværdier med henblik på at vurdere om standardafvigelsen ligger indenfor acceptable rammer.

Constant value. Nye rådata sammenlignes med tidligere rådata med henblik på at vurdere om måleværdien er fastfrosset.

Hvis et datapunkt fejler én eller flere af de fire tests, stemples datapunktet med "dårlig datakvalitet" og dette udgår herefter af eventuelle efterfølgende beregninger. "Dårlig datakvalitet" vil herudover udløse en "Målerfejl" hændelse for den pågældende måling.

2.3.2 Ekstremværdi

Abnormale hændelser detekteres med en algoritme, der sammenligner nye dataværdier med en referenceperiode med henblik på at fastlægge, hvad sandsynligheden for at netop denne værdi optræder på et givent tidspunkt (se figur 2.5). Algoritmens hændelsesstyrke er proportional med sandsynligheden.



Figur 2.5. Princip for ekstremværdialgoritme.

Ekstremværdialgoritmen tilpasser sig således dynamisk de aktuelle datastrøm og kræver således et minimum af konfiguration inden brug. Dog gælder det at opsætning specifikke detektionsalgoritmer normalt vil kræve, at der opstilles en række forudsætninger for hvorvidt data skal indgå i beregningen. Typisk vil disse forudsætninger være af typen:

- Hvis ventil A er åben, så ...
- Hvis flow B er større end X, så ...
- Hvis timer C er aktiv, så ...

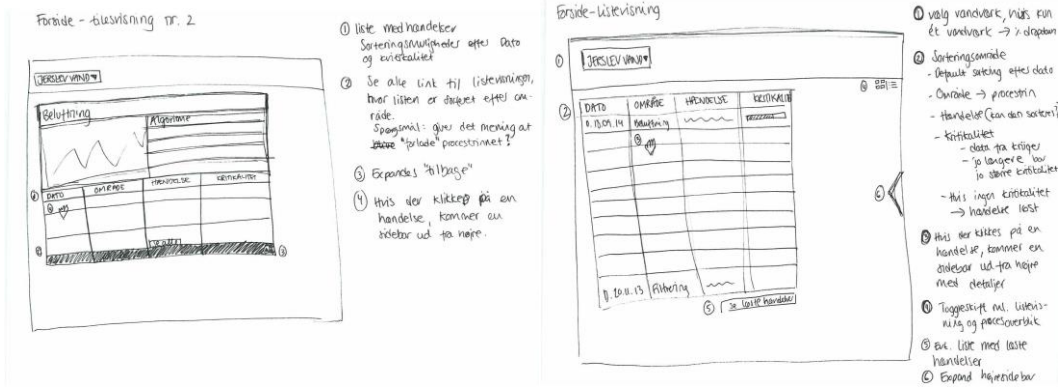
2.3.3 Multikriterie

Multikriteriealgoritmen fungerer som en overbygning til ekstremværdialgoritmen. Algoritmen beregner for hver datastrøm en sandsynlighed for den pågældende værdi. Den samlede sandsynlighed for den aktuelle kombination af måleværdier kan derefter beregnes som produktet af sandsynligheder for de enkelte datastrømme. Den samlede sandsynlighed omregnes herefter til en hændelsesstyrke.

2.4 Frontend (Brugerflade)

Da den udviklede software har til formål skal kunne anvendes af driftsoperatøren i det daglige arbejde med et minimum af træning var det af afgørende betydning af brugerfladen var nem og intuitiv at bruge samtidig med at den skal kunne give overblik over en potential meget stor datamængde. Når en hændelse detekteres vises den på softwarens brugerflade, hvorefter operatøren skal foretage en aktiv vurdering hændelsens driftsmæssige relevans.

Brugerfladeudviklingen startede med et konceptuelt design som oplæg til en UX workshop (User Experience) med potentielle brugere. På workshoppen blev der arbejdet med forskellige måder at visualisere forsyningens hændelser. Efter workshoppen blev der udarbejdet en række skitser af brugerfladen (se figur 2.6), der efterfølgende dannede grundlag for den endelige programmering.



Figur 2.6. Eksempler på wireframes for brugerflade.

Den færdige brugerflade blev udviklet som et webinterface, der kan fungere på tværs af forskellige platforme f.eks. computer, tablet og telefon. Brugerfladen blev særligt udviklet, så den var egnet til touch interaktion i en flad struktur. Det betyder at operatøren på en tablet med få taps ville kunne få adgang til alt væsentlig indhold (se figur 2.7).



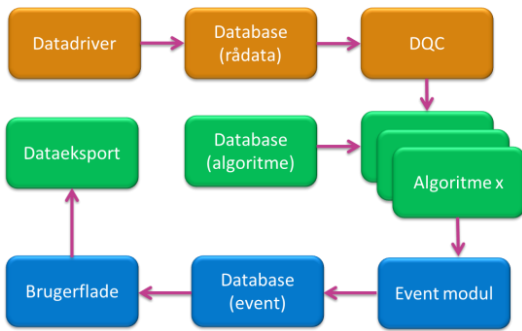
Figur 2.7. Eksempler på skærbilleder fra Water Manager.

Inden det endelige design var fastlagt blev der gennemført en brugertest, hvor brugere skulle løse en række mindre opgaver af driftsmæssig karakter i en prototype af brugerfladen.

Brugerfladens design og UX blev efterfølgende evalueret med udgangspunkt i tilbagemeldingen fra brugertesten, hvilket førte til en række mindre ændringer.

2.5 Backend

Backend blev udviklet sideløbende med brugerfladen. Selve backend blev udviklet på basis Krügers STAR Utility Solution platform.



Figur 2.8. Arkitektur for backend.

Backend arkitekturen er opbygget med en datadrivers på SCADA maskine, der indsamler og skubber data til en database på STAR serveren opstillet indenfor forsyningens firewall.

Realtidsfortolkningen af data foregår i en række moduler der læser og skriver i STAR databasen som vist i figur 2.8.

Algoritmerne, der analyserer data er udviklet i et særligt programmeringsmiljø, der muliggør en hurtig implementation af nye algoritmer ligesom gamle algoritmer på enkel vis kan tilpasses og forbedres.

2.6 Opsætning og implementering

Software blev indledningsvis sat i drift med data fra Lindvedværket, VandCenter Syd (VCS). Der blev til lejligheden indkøbt en STAR server, der blev fysisk opstillet hos VCS indenfor deres firewall. STAR serveren har til formål at afvikle Water Manager softwaren samt at modtage udvalgte data fra SCADA serveren. En tilsvarende opsætning er blevet udført med data fra Truelsbjergværket, Aarhus Vand.

Tabel 2.1 viser et overblik over signaler og algoritmer på de to testlokationer.

Tabel 2.1. Algoritmeopsætning på Lindvedværket og Truelsbjergværket.

	Lindved	Truelsbjerg
Antal sensorer	37	34
Antal algoritmer i alt	36	52
Algoritmer for målerfejl	6	6
Algoritmer for ekstremværdi	22	38
Algoritmer for multikriterie	8	8
Sensor typer	Flow Tryk Niveau Turbiditet Ilt Ammonium pH Ledningsevne	Flow Tryk Niveau Turbiditet Ilt pH Ledningsevne Temperatur

2.7 Test og Demonstration

Software og algoritmer blev testet med brugere gennem fire testperioder. Testperioderne havde til formål at validere systemet med henblik på at sikre, at de hændelserne, der blev genereret var relevante for for driften.

I testperioderne havde driftsoperatørerne fået til opgave at vurdere den driftsmæssige relevans af alle indkomne hændelser og afgive denne vurdering i systemet. Testperioderne blev gennemført på Truelsbjergværket og Lindvedværket i perioder af 3 til 4 ugers varighed.

De to første testperioder blev gennemført i juni 2015 og de sidste to testperioder i juli 2015.

Herudover har projektets partnere aftalt at samarbejde om at gennemføre yderligere en demonstrationsperiode på begge værker, der løber frem til årsskiftet 2016/2017.

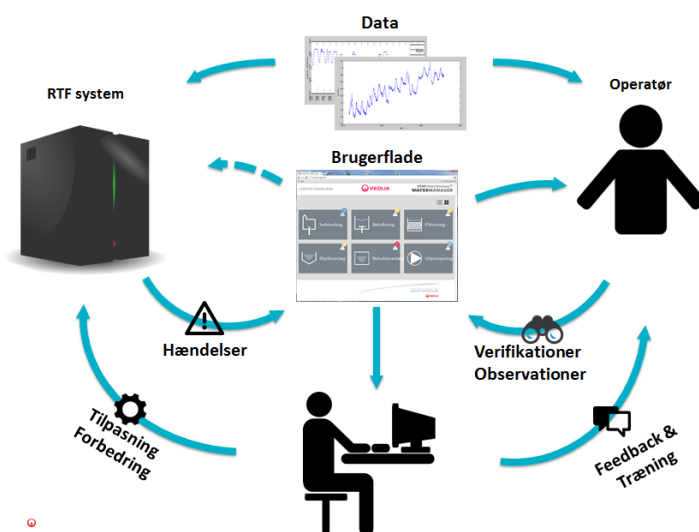
3. Resultater og diskussion

Da den overvejende del af de algoritmer, der blev udviklet i dette projekt var målrettede hændelser der må anses som sjældne i normal drift (få hændelser pr. år eller mindre) blev detektionsfølsomheden øget i testperioden for at sikre at der blev generet et "rimeligt" antal hændelser. Det betød altså, at testens algoritmer var hyperfølsomme og derfor generede op i mod 10 gange så mange hændelser som normalt forventet. Af praktiske hensyn vurderede projektgruppen, at det i første omgang af ville være mere hensigtsmæssig at gennemføre korte og intensive testperioder med hyperfølsomme algoritmer fremfor længere testperioder med normal følsomhed. Det betød imidlertid, at de hyperfølsomme algoritmer genererede et stort antal hændelser under normal drift uden væsentlig driftsmæssig relevans. Det var herefter driftsoperatørerne opgave adskille relevante og ikke-relevante hændelser.

På baggrund af driftsoperatørerne vurdering af de enkelte hændelser kunne Krüger løbende tilpasse og forbedre algoritmerne, så de i højere grad var overensstemmelse mellem operatørernes vurdering.

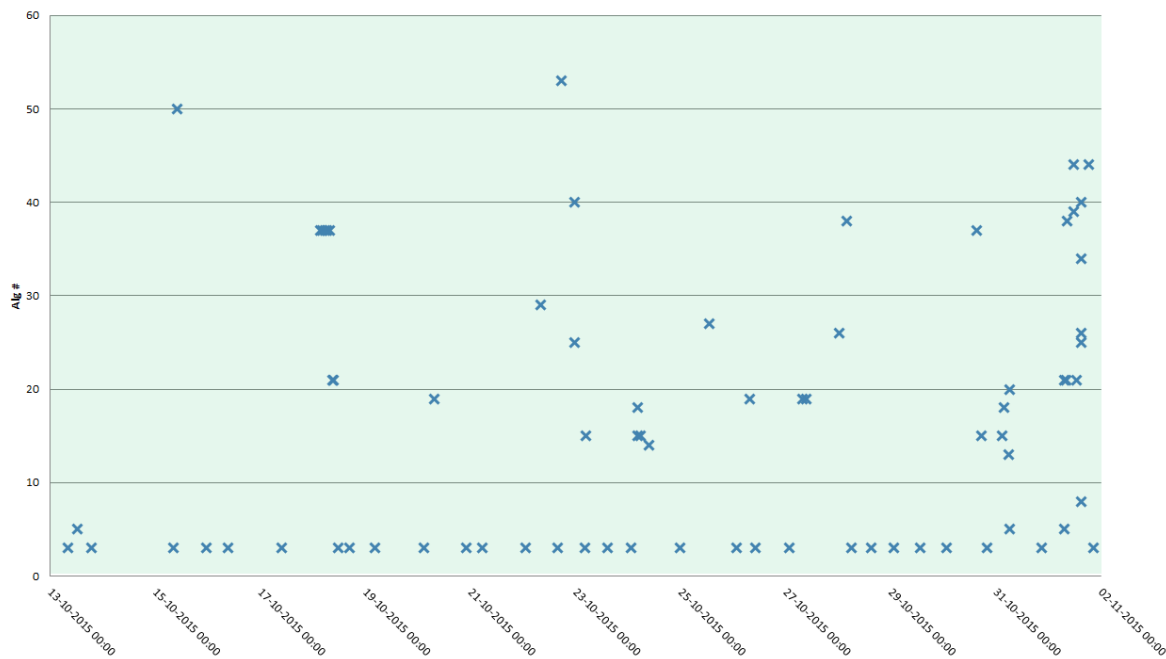
Figur 3.1 viser workflow mellem software, operatører og udviklere. Realtidsfortolkningssystemet (RTF system) gennemgår data med henblik på at detektere ekstreme eller usædvanlige hændelser. Operatøren kan tilgå rådata via det almindelige SRO system. Når RTF systemet detekterer en hændelse vises den på Water Manager brugerfladen, hvorfra operatøren kan gennemgå og verificere de seneste hændelser.

Periodisk kan udviklerne vurdere de enkelte algoritmers effektivitet med udgangspunkt i det samlede antal detekterede hændelser og hvor mange af disse der er blevet verificerede af driftsoperatørerne. På den baggrund kan udviklerne forbedre algoritmernes præcision ved at tilpasse/forbedre beregningen og eller give feedback/træning til driften.

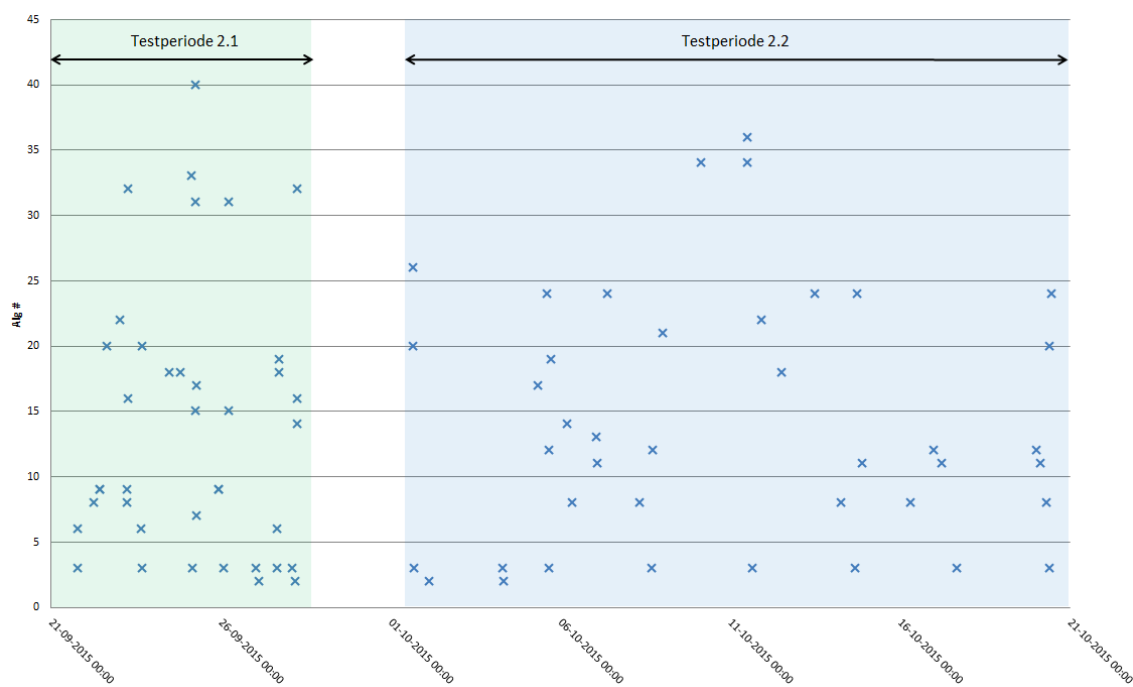


Figur 3.1. Workflow realtidsfortolkningsalgoritmer på Truelsbjergværket og Lindvedværket.

Figur 3.2 og 3.3 viser resultatet fra en testperiode på hhv. Truelsbjergværket og Lindvedværket.



Figur 3.2. Eksempler på hændelser på Truelsbjergværket. Y-aksen angiver den specifikke algoritmes løbenummer.



Figur 3.3. Eksempler på hændelser på Lindvedværket i testperiode. Y-aksen angiver den specifikke algoritmes løbenummer. Efter den første del af testperioden blev algoritmerne s følsomhed generelt reduceret.

Som det ses af figuren har forskellige hændelser til tider en tendens til at ”klumbe” sammen som det f.eks er tilfældet omkring d. 1/11. På andre tidspunkter er optræder meget få hændelser i anlægget f.eks. i weekender.

Tilbage meldingen fra systemet brugere har været at softwaren er enkel at anvende, men at det kan være svært i den givne situation at vurdere relevansen af de enkelte hændelser.

Kruger arbejder løbende på at forbedre præcisionen af algortimerne og i den forbindelse analyseres de enkelte hændelser og sammenfald mellem hændelser. Kruger arbejder ligeledes på at forbedre

driftsoperatørernes mulighed for at vurdere hændelsen samt at gøre algoritmernes beregninger mere gennemskuelige for brugerne.

4. Formidling

Dette workpackage har givet anledning til følgende formidlingsaktiviteter:

- Præsentation ved Dansk Vandkonference 2014
- Præsentation ved Dansk Vandkonference 2015
- Abstract indsendt til Nordisk Vandkonference 2016

Herudover har projektet ved flere lejligheder være præsenteret i forbindelse med fyrtårnsprojektet Fremtidens Vand bl.a. på Aquatech 2016 i Amsterdam og på projektets hjemmeside <http://futurewatercity.com/>.

5. Konklusion

Projektet havde succes med at udvikle software, der kunne analysere store datamængder i realtid med henblik på at udpege ekstreme hændelser. Softwaren blev udviklet som et beslutningsstøtte værktøj for driftoperatører i vandforsyninger.

Softwaren blev demonstreret i fuldskala med realtidsdata fra to vandværker:

- Truelsbjergværket ved Aarhus Vand
- Lindvedværket ved VandCenter Syd

Brugerne af softwaren meldte om at softwaren var let at bruge og forstå, men at de hændelser, der blev generet af algoritmerne kunne være svære at verificere.

Strukturen i softwaren er i høj grad skalérbar, hvorfor det vil kunne anvendes på datamængder der er væsentlig større end det var tilfældet i dette projekt. Softwaren er som udgangspunkt enkelt at sætte op, da algoritmerne i høj grad konfigurerer sig selv baseret på den aktuelle data.

Krüger forventer at arbejde videre med at forbedre og udvikle algoritmer til realtidsfortolkning baseret på Water Manager. Projektgruppen har foreløbigt aftalt at videreføre projektet med fuldskala demonstration af softwaren på begge projektets værker i 2016.



Miljøstyrelsen
Strandgade 29
1401 København K

www.mst.dk