

Aktivitet	Forskning og udvikling		
Aktivitetsplan (titel):	Brugerværdi ved udnyttelse af store datamængder	Aktivitetsplan nr.:	9
Resumé	Aktiviteten tager afsæt i Vandvisionen 2015 (http://mim.dk/arbejdsmraader/jord-luft-og-vand/vandvision-2015/) med fokus på at gøre Danmark verdensførende i at levere intelligente, bæredygtige og effektive vandløsninger. Vandløsninger er her varslingsystemer, og aktivitetens hovedformål er at gøre den danske vand- og miljøkonsulentsektor (vand, grundvand, afløb og hav) i stand til at opsætte og vedligeholde varslingsystemer. Herved opnår danske konsulenter samt SMV'er en styrket position på det nationale og internationale vandmarked. Forudsætningerne for at kunne udstyre målgruppen med denne styrkede markedsposition er: • Let og smidig adgang til vandsektorrelaterede datamængder • Hurtig afvikling af modelkørsler • Adgang til IT-infrastruktur, der muliggør afvikling af modeller med den nødvendige fine opløsning • Adgang til IT-komponenter til at præsentere resultater via internettet Der vil derfor blive udviklet: • En stærk MIKE integration med dataportalen, der sikrer automatisk modelopsætning og opdatering • Hurtigere modelafviklingsmetoder til at sikre, at resultater bliver leveret i tide • Metoder til at håndtere de store datamængder, der produceres gennem varslingsmodellerne • En række præsentationsværktøjer til at formidle resultater gennem internettet. De hovedaktiviteter, der skal gennemføres for at realisere aktivitetsplanen, er: Udvikling af algoritmer til prækvalificering af data Identifikation af nye numeriske koblingsmetoder. Målet er en øget modelafviklingshastighed ved brug af nye innovative, matematiske løsningsmetoder samt kraftfulde beregnings- og lagringsenheder. Fokus vil være på integrerede modeller (f.eks. by-land-hav) til afvikling af oversvømmelsesberegninger. Implementering af metoder til smart håndtering af store data. Smart håndtering af store data er nødvendig grundet størrelsen på de modeller, der anvendes. Der er behov for optimal håndtering af store distribuerede data, tilegnet gennem den seneste forskning inden for området Udvikling af web-softwareværktøjer, hvormed konsulent selv kan skabe en 'web' eller en 'smart device' løsning til at visualisere sine resultater.		
1) Målgruppe og behov	Med den hastige udvikling i hardware og beregningsalgoritmer vil varslings af oversvømmelser, badevandskvalitet, bølgehøjder etc. kunne gøres lettere tilgængelig for danske konsulenter samt kommuner. Visionen er at kunne tilbyde den nødvendige teknologi til det danske erhvervsliv. Det at opstille og gennemføre vandforudsigelser (fx oversvømmelsesprædiktioner i realtid) har hidtil været en specialistopgave for de få, men ved brug af den skitserede serviceydelse vil omkostningerne reduceres, og ydelsen vil derved kunne bredes ud. Denne omkostningsreduktion vil ske gennem øget modelautomatisering, optimeret modelafvikling, intelligent anvendelse af data samt smarte visualiseringskomponenter til formidling af resultater på internettet.		

	Målgrupperne for ydelserne i nærværende aktivitetsplan er: • <i>Danske vandkonsulenter:</i> Denne sektor (hvoraf størstedelen er MIKE software brugere) vil med den nye teknologi blive mere omkostningseffektiv og dermed kunne byde på flere opgaver inden for vandforudsigelser nationalt såvel som internationalt. Yderligere vil de enkelte projekter kunne afvikles hurtigere. Den sparede arbejdstid er vurderet til 10-20 %, hvilket kan omsættes til en tilsvarende reduktion i projektomkostninger. De beskrevne ydelser vil være attraktive hos samtlige af de danske konsulenter, der i dag benytter MIKE software. Denne gruppe dækker alle de største vandkonsulentfirmaer. • <i>Små og mellemstore virksomheder,</i> som arbejder med vandmodellerings-konsulenttydelser (i størrelsesordenen 10-20) og installation af vandteknologier til monitoring af vand (i størrelsesordenen 5-10). Fordelen for SMV'er er en reduktion af omkostninger til eget køb af hardware samt databearbejdning. Herved kan SMV'er byde på større opgaver uden at skulle foretage store investeringer i digital infrastruktur. • <i>Alle danske kommuner:</i> Med den valgte teknologi vil det være lettere at opbygge og vedligeholde systemer til eksempelvis oversvømmelsesforudsigelser. Det skønnes, at 10 af de mest oversvømmelsesudsatte kommuner vil adoptere teknologien inden for 2 år. • Andre vandrelaterede segmenter inden for det danske erhvervsliv som havne, offshore vindindustri og shipping industri. Behovet for de beskrevne ydelser er udtrykt som ønsker på DHI's MIKE software brugerseminarer samt gennem den løbende kontakt med kunder via support. Hertil kommer, at dialogen på BedreInnovation.dk viser et ønske om en effektiv håndtering og visualisering af de store datamængder, der er knyttet til varsling af oversvømmelser i forbindelse med eksempelvis skybrud. Derudover er der udtrykt et behov for intelligent tilgang til simuleringresultater, som vil gøre selv de store modelleringsopgaver tilgængelige for SMV'er i Danmark. Fra shippingindustrien vil ydelsen imødekomme ønsket om præcise data om strøm, bølger, dønning og vind til anvendelse ved ruteplanlægning samt overvågning af et skibs performance.
2) Den nye teknologiske serviceydelse	Ydelsen vil bestå af følgende delydelser: • Automatiseret modelopsætning • Optimerede modelafviklinger for integrerede modeller • Rammeværk for smart håndtering af store distribuerede datamængder • Mulighed for præsentation af resultater gennem konsulentens egen 'web' eller app Hver af disse ydelser er uddybet nedenfor. <u>Automatiseret modelopsætning</u> Konsulentprojekter inden for vandsektoren er kendetegnet ved en del dataindsamling og -behandling. Dataindsamling foregår manuelt gennem afsøgning af diverse datakilder - typisk via internettet - og kan være tidskrævende. Databehandling fokuserer på kvalitetssikring, interpolation i tid og sted, re-sampling af opløsning etc. Ved brug af den nye serviceydelse vil det være muligt automatisk at få konstrueret de matematiske MIKE modeller, der benyttes i konsulentarbejdet, ved brug af DHI's dataportal. Dataportalen, som blev udviklet med støtte fra forrige resultatkontrakt, vil i denne sammenhæng blive udbygget med en automatiseringsmodelopsætnings-service. Denne service vil udstyre brugeren med en MIKE model for et givet område på basis af data, der er til rådighed. Typisk vil der være tale om topografi samt randbetingelser. Brugerens input til denne service er angivelse af det geografiske område af interesse samt en ønsket opløsning. Servicen kan herved reducere tiden, der benyttes på det initiale modelopsætningsarbejde. De automatisk skabte modeller kan herefter danne basis for mere detaljerede

	undersøgelser ved inklusion af hydrauliske strukturer samt rørledninger, men kan også benyttes direkte til at udføre screeningskørsler tidligt i projektforsløbet. <u>Optimerede modelafviklinger for integrerede modeller</u> En forudsætning for at kunne levere realtidsforudsigelser er, at modelsimuleringerne kan afvikles effektivt. Til dette vil de tilgængelige <i>hardware optimeringsteknologier</i> blive benyttet, dels High Performance Computere og dels Grafiske Processor enheder (units) (GPU). Støtte fra forrige resultatkontrakt har været medvirkende til, at adgang til HPC nu kan tilbydes som en service. Denne adgang sikrer MIKE brugere en hurtig modelafvikling af eksempelvis strømning hen over terræn. Disse terrænbaserede modeller er uhyre nyttige for situationer med kraftig regnskyl, hvor afløbssystemets kapacitet kun udgør en lille del af den samlede vandmængde. Hændelser, hvor afløbssystemets-kapaciteten ikke kan ignoreres, kræver inddragelse af det detaljerede rør/afløbssystem. Afløbsmodel og terrænmodel kan let kobles og afvikles i MIKE modellerne, men koblingen har ofte den konsekvens, at den samlede afviklingstid for modelkørslen stiger. Dette skyldes forskel i tidsskala i de endimensionale rørmodeller og de todimensionale terrænmodeller kombineret med den numeriske metode, der anvendes i forbindelse med koblingen. Vandforudsigelser skal benyttes for hændelser på mange skalaer - også dem, der kræver en detaljeret rørmodel. Derfor er optimering af afviklingshastigheden af de koblede modeller fundamental for at kunne tilbyde forudsigelser inden for den nødvendige tidshorisont. <u>Rammeværk for smart håndtering af store distribuerede datamængder</u> De datamængder, der dels indgår i modellerne og dels produceres gennem modelafviklingerne, er store. Eksempelvis vil en oversvømmelsesmodel for hele København i en opløsning på 1 meter - med resultater hvert minut og med en forudsigelseshorisont på 24 timer - producere mere end 20 GB data. Dertil kommer, at data er gemt på en eller flere centrale lagringsenheder. En effektiv håndtering af disse data er en forudsætning for at kunne tilbyde vandforudsigelsesservicen bredt. <u>Mulighed for præsentation af resultater gennem konsulentens egen 'web' eller app</u> Resultaterne af modelafviklingerne skal formidles til interessenterne, som dels udgøres af beslutningstagere, men også i stor grad af befolkningen. Med en så bred interessentgruppe er internettet det rigtige medie til formidling. For at styrke denne formidling vil der blive stillet en række visualiseringskomponenter til rådighed gennem MIKE's Software Development Kit. Disse komponenter kan benyttes som en del af brugerens egen webside til at præsentere resultater. Komponenterne vil blive baseret på teknologi, som er egnet til PC, tablets samt mobiltelefoner. Herved sikres den størst mulige forankring samt udbredelse af servicen. Vandforudsigelser er det næste skridt inden for eksempelvis oversvømmelsesstudier. Hidtil har en stor del af studierne kun omhandlet, hvad der kan ske ved fx en 20 års hændelse. Disse studier vil med den skitserede teknologi kunne udvides med, hvor, hvornår og hvordan det sker i realtid. Teknologien samt kompetencen, der vil blive udviklet og tilegnet, vil være ny inden for vand- og miljødomænet. Den beskrevne kombination af intelligent data-håndtering, automatisering samt optimeret afvikling af modeller vil være unik. Konkurrenterne har ikke den samme stærke treenighed, der udgør kernen af ydelsen. Der findes således ikke lignende services nationalt eller internationalt. Tiltaget, der vil være markedsmodent inden for tre år, giver MIKE produkterne en yderligere førerposition inden for vandmodelleringssoftware.
--	---

3) Aktiviteter	For at kunne realisere ”Vandforudsigelser for alle”, skal der gennemføres følgende hovedaktiviteter: A. Udvikling af algoritmer til prækvalificering af data B. Identifikation af nye numeriske koblingsmetoder C. Implementering af metoder til smart håndtering af store data D. Udvikling af web-softwareværktøjer, hvormed konsulenten selv kan skabe en ’web’ eller en ’smart device’ løsning til at visualisere sine resultater. De fire listede aktiviteter er beskrevet nedenfor. A. Udvikling af algoritmer til prækvalificering af data Modelopsætning kan være tidskrævende grundet dataindsamling, validering af data og konvertering mellem forskellige dataformater. For at reducere tiden brugt på dataindsamling tilbyder DHI i dag en service med adgang til vandrelaterede data. Der udvikles en smart overbygning til denne dataservice, som kan automatisere modelopsætningen. B. Identifikation af nye numeriske koblingsmetoder Et kontinuerligt fokus for brugere af DHI’s modeller er hurtig modeleksekvering. Dette har afstedkommet store landvindinger med hensyn til optimering af de enkelte modelkomponenter til brug ved 2D- og 1D-modellering. Dog har de integrerede modeller, hvor både 1D- og 2D/3D-komponenter benyttes, ikke altid kunnet udnytte alle implementerede optimeringer grundet forskelle i tidsskala for dynamikken i 1D- i forhold til 2D/3D-modeller samt begrænsninger i de valgte koblingsmetodikker. Der udvikles en ny ikke-begrænsende koblingsmetodik, der kan underbygge variationen i de dynamiske tidsskalaer. C. Implementering af metoder til smart håndtering af store data MIKE modeller dækker hele byer og kystområder i meget høj opløsning. Beregningsmodellerne afvikles på store High Performance Computer (HPC) systemer, hvor de opdeles i en række delmodeller tilpasset den hardware, der benyttes. Hver af disse delmodeller producerer en resultatfil svarende til det delområde, som de repræsenterer. Resultater fra en simulering består således af mange filer, der er distribueret over lagringsenhederne på den anvendte hardware. Traditionelt flettes disse filer sammen til en stor fil, der herefter overføres til den lokale PC til videre behandling. Modellerne, der arbejdes med her, er af en sådan størrelse, at konventionelle metoder til at tilgå resultatfiler kommer til kort. Den optimale løsning er at lade resultatfiler/data forblive i deres distribuerede form på den lokalitet, hvor de er skabt, og tilgå data fra distancen. Denne metode fordrer stor fleksibilitet mht. data-/resultatselektion og reducerer derved den samlede modelafviklingstid. Metoden gavner desuden standard MIKE installationer på PC’er med flere beregningskerner og vil således også optimere løsninger af i dag. Der udvikles en ny optimeret metodik til håndtering og fjerntilgang af distribueret data samt resultatfiler. D. Udvikling af web-softwareværktøjer Kommunikation vedrørende vandforudsigelser skal foregå på de mest gængse medier, dvs. ’web’ og ’smart devices’. Til dette formål udvikles en web-service, der kan tilgå de beregnede forudsigelser og præsentere dem på en nem og overskuelig måde til borgeren. Med denne web-service vil det være muligt for rådgivere og kommuner at skræddersy deres egen ’smart app’ til fx varsling af oversvømmelser, bølgehøjder, vandkvalitet etc. Den beskrevne web-service indlejres i MIKE’s Software Development Kit (SDK). Herved sikres det, at funktionaliteten bliver generelt tilgængelig for tredje part, samt at komponenterne ”åbnes op” ved hjælp af fleksible, veldefinerede og veldokumenterede softwaregrænseflader (API’er).
-----------------------	--

	SDK'et gør konsulenter samt kommuner i stand til at integrere DHI-viden og teknologi i egne produkter og løsninger.
4) Viden-samarbejde og -hjemtagning	DHI har gennem en årrække haft et frugtbart samarbejde med Det Tekniske Universitet i Ostrava - Tjekkiet. Fokus er at udnytte kraftfuldt HPC udstyr til at eksekvere DHI's MIKE modeller for derved at øge beregningshastigheden signifikant. Det tekniske Universitet i Ostrava besidder en viden og hardware, der komplementerer DHI og sikrer, at vi kan nå det skitserede mål. Det tekniske Universitet i Ostrava leder en EU Horizon 2020 (H2020-DRS-2015) ansøgning, hvor DHI er en af 14 deltagende partnere. Der ansøges om støtte til at udvikle metoder til brug i forbindelse med katastrofe-risikoledeelse (Disaster Risk Management). DHI's bidrag er på modelsiden, hvor realtidssystemer og hurtige simuleringer er hovedfokus. På DHI vil viden fra samarbejdet med Ostrava bliver indlejret som en HPC as a Service (HPCaaS), som vil blive udbudt til DHI's MIKE kunder. Af andre samarbejdspartnere inden for HPC kan nævnes DTU Compute samt HPC Wales. Inden for anvendelse af de nye teknologier nævnes følgende samarbejdspartnere: staten, kommuner, rådgivere m.fl. Staten samt kommunerne har en interesse i at kunne forudsige mulige oversvømmelser i realtid, således at de kan tage de nødvendige forholdsregler med hensyn til sikring af værdier, men også i form af håndtering af den trafikale infrastruktur.
5) Inddragelse og videnspredning	De udviklede serviceydelser vil blive stillet til rådighed gennem MIKE produktporteføljen. MIKE software har 1300+ aktive produktinstallationer i Danmark alene og 10.000+ internationalt. Installationer findes hos de store danske rådgivere inden for vand og miljø samt hos 40+ små og mellemstore rådgivere. Dertil kommer, at softwaren er installeret hos 42 danske kommuner og samtlige 7 miljøcentre. MIKE softwaren stilles gratis til rådighed for alle danske universiteter og anvendes bredt i undervisningen på bl.a. DTU, KU, AU, AUC og Syddansk Universitet. Der afholdes MIKE brugergruppemøder i projektperioden, hvor resultater samt funktionalitet bliver præsenteret. Derudover er der brugerforum for de forskellige produkter, hvor information også bliver spredt. Funktionalitet forankret i MIKE teknologien vil som en del af den serviceaftale, de enkelte MIKE brugere indgår med DHI, blive udbudt også fremover. Herved sikres det, at ydelsen bliver vedligeholdt, samt at investeringen bliver sikret fremover. Det planlægges at publicere to videnskabelige artikler i projektperioden: den ene i samarbejde med Det Tekniske Universitet i Ostrava med fokus på HPCaaS og den anden med fokus på den optimerede modelafviklingshastighed. Derudover vil der blive skrevet to konference papers om emnet. Der vil blive udarbejdet kurser og seminarer inden for aktivitetsplanens forskellige delaktiviteter under THE ACADEMY by DHI. Der forventes afholdt i alt 3 kurser/seminarer i resultatkontraktperioden med deltagelse af minimum 30 aftagere fra de primære målgrupper. Der vil videre blive afholdt minimum ét åbent "F&U Highlights" Webinar pr. år, hvor en DHI-ekspert vil fortælle nyt om igangværende F&U projekter inden for aktivitetsplanen. Disse "F&U Highlights" arrangeres gennem THE ACADEMY by DHI. Derudover vil aktiviteterne bliver præsenteret ved MIKE Powered by DHI's danske brugergruppemøde (ca. 50 deltagere), der afholdes med jævne mellemrum. Væsentlige dele af de nye udviklinger vil desuden blive stillet til rådighed gennem MIKE software, der har mere end 1.300 brugere i Danmark og således er en meget effektiv metode til spredning af aktivitetsplanens resultater.

6) Sammenhæng med institut-strategi	DHI's MIKE software er den teknologiske hovedakse i DHI's kompasstrategi. Aktivitetsplanen "Brugerværdi ved udnyttelse af store datamængder" er koordineret med instituttets generelle strategi og med DHI's øvrige aktivitetsplaner.
7) Milepæle år 1	Vidensamarbejde, - hjemtagning og kompetenceopbygning A. <i>Udvikling af algoritmer til prækvalificering af data</i> A.1 Konceptbeskrivelse for intelligent afsøgning og udvælgelse af data på baggrund af et brugerdefineret område B. <i>Identifikation af nye numeriske koblingsmetoder</i> B.1 Konceptbeskrivelse for ikke begrænsende koblingsteknik for 1D/2D/3D modelafvikling. Rapport af samme C. <i>Implementering af metoder til smart håndtering af store data mængder</i> C.1 Udvikling af koncept og specifikation for tilgang til distribueret fjerndata. Rapport med brugerinddragelse gennem MIKE brugerforum D. <i>Udvikling af web-softwareværktøjer</i> D.1 Specifikation af web-komponenter til brug i varslingsystemer Udvikling af teknologisk service B. <i>Identifikation af nye numeriske koblingsmetoder</i> B.2 Koncept af ikke-begrænsende koblingsteknik i DHI's koblede model MIKE FLOOD Inddragelse og videnspredning F. <i>Formidling</i> F.1 Kursus i remote execution, afholdt. F.2 En konferenceartikel F.3 Webinar om koblingsmetoder, afholdt.
Milepæle år 2	Vidensamarbejde, - hjemtagning og kompetenceopbygning A. <i>Udvikling af algoritmer til prækvalificering af data</i> A.1 Koncept for intelligent afsøgning og udvælgelse af data på baggrund af et brugerdefineret område Udvikling af teknologisk service B. <i>Identifikation af nye numeriske koblingsmetoder</i> B.1 Implementering af ikke-begrænsende koblingsteknik i DHI's koblede model MIKE FLOOD C. <i>Implementering af metoder til smart håndtering af store data</i> C.1 Prototype af distribueret datatilgang D. <i>Udvikling af web-softwareværktøjer</i> D.1 Prototype af web-komponenter udviklet Inddragelse og videnspredning B.1 Afholdelse af kursus i forbindelse med dansk brugerseminar med fokus på kobling af modeller F. <i>Formidling</i> F.1 Kursus i koblingsteknik, afholdt. F.2 Peer-reviewed artikel indsendt F.3 Afholdelse af MIKE Brugergruppemøde F.3 Webinar om smart håndtering af store data, afholdt.
Milepæle år 3	Udvikling af teknologisk service A. <i>Udvikling af algoritmer til prækvalificering af data</i> A.1 Release af MIKE produkter til danske kunder med intelligent overbygning af dataportal B. <i>Identifikation af nye numeriske koblingsmetoder</i> B.2 Release af MIKE produkter med ikke-begrænsende koblingsteknik C. <i>Implementering af metoder til smart håndtering af store data</i> C.1 Distribueret datatilgang tilgængelig gennem MIKE produkterne D. <i>Udvikling af web-softwareværktøjer</i> D.1 Release af web-komponenter i MIKE SDK Inddragelse og videnspredning D.2 Afholdelse af kursus med fokus på anvendelse af web komponenter i SDK

	<i>F. Formidling</i> F.1 Kursus i web-komponenter i MIKE SDK, afholdt. F.2 Peer-reviewed artikel indsendt F.3 1 konferenceartikel F.4 Webinar om automatisk opdatering af modeller, afholdt.
Titel ved præsentation på BedreInnovation.dk	Brugerværdi ved udnyttelse af store datamængder