
Over de adoptie van automatisering in hydrologisch en hydraulisch modelleren

HYDROLIB WP5: Modellerkeuzes en automatisering
Lieke Melsen, Janneke Remmers, Carine Wesselius

Datum: 1 November 2022
Contactpersoon: Lieke Melsen, lieke.melsen@wur.nl

1. Introductie

Aanleiding voor deze korte memo is het TKI-project 'HYDROLIB', getrokken door Deltares, in een consortium met zes waterschappen, zes adviesbureaus, en Wageningen University. Doel van dit project is om een software-bibliotheek te ontwikkelen die het mogelijk maakt om een deel van het modeller-proces te automatiseren door gebruik te maken van scripts die al eerder ontwikkeld zijn en die tot nettere centrale versies samen te brengen.

Automatiseren is gedefinieerd als “*de vervanging van menselijke arbeid door computers*”. Dit is een gradueel proces: iets kan volledig geautomatiseerd zijn (een druk op de knop leidt tot een volledig resultaat) of slechts kleine onderdelen kunnen geautomatiseerd zijn. Modelleren is al voor een groot deel een geautomatiseerd proces. HYDROLIB focust met name op pre- en post-processing van de input en output data en het opzetten van het model op basis van de input data, om zo het modeller-proces verder te automatiseren. Het idee is dat dit leidt tot meer efficiëntie, verbeterde reproduceerbaarheid, betere onderlinge vergelijkbaarheid, en het maakt de kans op menselijke fouten kleiner.

Pre- en post-processing vergt op dit moment vaak nog een aantal menselijke handelingen zoals het downloaden van de juiste input-data, het omzetten van data in het juiste format of het manipuleren van data naar de juiste resolutie. Deze stappen zijn op het eerste oog relatief makkelijk te automatiseren, waar met automatiseren in deze context bedoeld wordt: in de HYDROLIB software-bibliotheek staat een 'kant-en-klaar' script die deze stappen uit kan voeren. Kanttekening is dat deze stappen vaak gepaard gaan met bepaalde keuzes die soms “expert judgement” vereisen. Het automatiseren van deze stappen middels het gebruik van kant-en-klare scripts uit de software-bibliotheek vormt dus een balans tussen de keuzevrijheid en expert judgement van de modelleur enerzijds, en het versimpelen en versnellen van het modeller-proces anderzijds. Het automatiseringsproces verplaats de besluitvorming van deze keuzes van de modelleur naar diegene die de automatiseringsscripts schrijft. In sommige gevallen zal dat dezelfde persoon zijn, maar met het groeien van de bibliotheek en de gebruikers-community van HYDROLIB zullen dit steeds vaker verschillende personen zijn.

Daarnaast maken veel modelleurs al gebruik van enige vorm van automatisatie voor pre- en post-processing, door middel van zelfgeschreven scripts of scripts geschreven door collega's of adviesbureaus. Overstappen op de HYDROLIB software-bibliotheek moet dus ten eerste voldoende keuzevrijheid aan de modelleur laten daar waar nodig, en ten tweede een direct voordeel hebben boven het gebruik van de al bestaande, maar slechts lokaal gebruikte, automatiserings-scripts.

Middels interviews met modelleurs van waterschappen en adviesbureaus zijn we nagegaan

- 1) Voor welke stappen in het modeller-proces de modelleurs automatiseren wenselijk zouden vinden**
- 2) of en wanneer modelleurs over zouden stappen van lokale automatiseringsscripts naar een centrale bibliotheek zoals HYDROLIB.**

Deze informatie kan gebruikt worden om de automatisatie die HYDROLIB faciliteert af te stemmen op de wensen van de gebruiker.

2. Methode

Via email en via verschillende project-overleggen hebben we gevraagd naar modelleers die geïnterviewd zouden willen worden (“convenience sampling”). In enkele gevallen werden er via de geïnterviewde modelleers weer nieuwe namen genoemd die vervolgens weer gevraagd zijn voor een interview (“snowball sampling”). Daarnaast hebben we specifiek gezocht naar respondenten van verschillende adviesbureaus (“stratified sampling”). De interviews zijn gehouden tussen september en november 2021. Op dat moment was bij de meeste geïnterviewden nog niet bekend welke scripts er concreet binnen HYDROLIB ontwikkeld zouden worden.

De interviews waren semi-gestructureerd: vragen waren voorbereid in een interview-protocol, maar op basis van het antwoord konden opvolg-vragen gesteld worden (“waarom?”, “hoe?”). Elke geïnterviewde heeft geïnformeerd toestemming gegeven voor deelname en voor het opnemen van het interview. Op basis van de opnames werden de interviews getranscribeerd en teruggestuurd naar de geïnterviewde voor definitieve goedkeuring.

Interviews duurden gemiddeld ongeveer anderhalf uur. Hierbij werden verschillende onderwerpen aangestipt over de verschillende stappen van het modelleerproces en wat de rol van automatisering hierin zou kunnen zijn.

3. Resultaten

3.1. Automatisering en welke model-stappen hiervoor in aanmerking zouden komen

Op basis van de modelleer-stappen zoals die in het Good Modelling Practice Handbook (van Waveren e.a., 1999) genoemd worden, zijn de modelleers gevraagd waar ze mogelijkheden zien voor automatiseren in het modelleer-proces.

Modelopbouw

Een aantal geïnterviewden zag mogelijkheden om de modelopbouw verder te automatiseren. Modelopbouw direct vanuit de legger werd expliciet genoemd, zodat berging, profielen, duikers en bruggen automatisch meegenomen en geactualiseerd kunnen worden. Nu is er nog vaak handmatige controle of alles up-to-date is. Tegelijkertijd gaf een andere geïnterviewde aan dat bij het adviesbureau waar deze persoon werkzaam is, een groot deel hiervan al geautomatiseerd is in wat zij noemen “de hydrologische gereedschapskist”.

Forcering

Voor de meteorologische input data wordt al veel gebruik gemaakt van geautomatiseerde tools, vaak ontwikkeld specifiek voor een waterschap in opdracht van dat waterschap, en de geïnterviewden leken tevreden met de tools die ze op dit moment tot hun beschikking hadden.

Onderliggende data

Enkele geïnterviewden gaven aan mogelijkheden te zien voor het automatiseren van de controle van data. Duikerdata wordt nu bijvoorbeeld vaak handmatig gecontroleerd, terwijl er wellicht bepaalde stelregels te bepalen zijn die aan kunnen geven of bepaalde data verdacht of onverwacht is.

Daarnaast zou het inlezen van data van centrale bronnen geautomatiseerd kunnen worden. Er werd bijvoorbeeld genoemd dat Tygron dit al heeft: automatisch data uitlezen uit de Basisregistratie Ondergrond en basisregistratie Gewaspercelen.

Simulatieperiode

Wordt nu vaak bepaald met trial-and-error, bijvoorbeeld op basis van wanneer de afvoer weer stabiel is na een hoogwatergolf. Een geïnterviewde gaf aan dat je soms wat te ruim zit met je inschatting, wat rekentijd kost. Een aantal geïnterviewden gaven aan dat dit te automatiseren zou zijn, maar sommigen vonden ook dat je hier graag zelf controle op wilt houden.

Ruimtelijke resolutie

Wat betreft ruimtelijke resolutie zien een aantal geïnterviewden mogelijkheden voor het automatisch verfijnen van de resolutie rondom kunstwerken. Dit wordt nu vaak nog handmatig aangepast. Ook zou een verfijnd grid op basis van opgegeven polygonen prettig zijn. Daarnaast gaf iemand aan dat resolutie

nu vaak een trial-and-error proces is, waarbij stap voor stap bepaald wordt in welke gridcel er veel variatie ontstaat en dus verfijning nodig is. Hiervoor worden achter elkaar een aantal keer dezelfde stappen herhaald, en zijn er dus mogelijkheden om te automatiseren.

Gevoeligheidsanalyse

Veel geïnterviewden gaven aan nu niet of nauwelijks een gevoeligheidsanalyse uit te voeren (of uit te besteden aan de adviesbureaus), waarbij enkelen aangaven dit wel te gaan doen als dit makkelijker gemaakt zou worden, bijvoorbeeld door middel van een tool. Anderen gaven aan weleens een gevoeligheidsanalyse te doen, veelal handmatig, en wellicht vaker een gevoeligheidsanalyse te gaan doen als het gemakkelijker zou worden. De modelleur moet wel altijd kunnen bepalen welke parameters er meegenomen worden, en voor heel specifieke gevallen zal dit niet geautomatiseerd kunnen worden.

Kalibratie

Kalibratie is sterk afhankelijk van de beschikbare data en de kwaliteit van de beschikbare data. Op veel plekken gebeurt dit daarom nu nog handmatig en visueel, of helemaal niet. Een geïnterviewde geeft aan dat een hybride vorm, deels geautomatiseerd en deels handmatig, wel zou kunnen helpen, terwijl een andere geïnterviewde aangaf dat het misschien nuttig zou kunnen zijn geautomatiseerde kalibratieresultaten naast de handmatige kalibratie te kunnen leggen. Weer een ander vond dat kalibratie echt geautomatiseerd zou moeten worden. Tegelijkertijd gaf een aantal geïnterviewden aan dat er een kalibratie-tool in Sobek zat, maar dat die zelden werd gebruikt.

Validatie

Gebeurt nu nog vaak visueel en handmatig. Zou deels geautomatiseerd kunnen worden, maar meerdere geïnterviewden gaven aan dat de interpretatie wanneer iets goed genoeg is bij de modelleur moet blijven liggen. Een geïnterviewde gaf aan dat validatie altijd gepaard moet gaan met een basiscontrole van de kwaliteit van data waarmee gevalideerd wordt.

Onzekerheidsanalyse

De geïnterviewden gaven vrijwel unaniem aan dat een onzekerheidsanalyse heel zelden tot nooit gedaan wordt. Daarbij gaven velen ook aan het wellicht wel te gaan doen als het geautomatiseerd zou kunnen worden, met als kanttekening van een modelleur van een adviesbureau dat dit ook afhangt van de interesse van de opdrachtgever.

3.2 De overstap naar HYDROLIB

Er zijn een aantal punten die in de interviews naar voren kwamen, die relevant zouden kunnen zijn voor het creëren van een gebruikers-community voor HYDROLIB.

Ten eerste kwam uit de interviews naar voren dat kritische massa relevant is. Veel waterschappen gebruiken een bepaald model 'omdat dit al veel gebruikt wordt bij de waterschappen'. Daarnaast hebben Deltares en de adviesbureaus een goede naam bij de waterschappen, en hebben de waterschappen over het algemeen veel vertrouwen in hun producten. Hetzelfde zou dus kunnen gelden voor HYDROLIB: zodra een bepaalde kritische massa is bereikt, wordt het gebruik hiervan vanzelfsprekender. Maar de uitdaging zit natuurlijk met name in het bereiken van die kritische massa.

Uit de interviews bleek dat het modelleerproces bij zowel de waterschappen als de adviesbureaus vaak al een aaneenschakeling is van automatiserings-scripts en tools, ontwikkeld door verschillende partijen en vaak alleen voor intern gebruik. Sommige tools werden al meer dan tien jaar gebruikt, en "*je moet bij deze scripts wel echt alles netjes stapje voor stapje uitvoeren, anders crashen ze*". In het gebruik van al deze tools en scripts kwam de rol van vertrouwen heel duidelijk naar voren. Veel modelleurs waren niet altijd exact op de hoogte zijn van wat er precies in deze scripts en tools gebeurde, maar vertrouwden erop dat het goed was omdat het al zolang gebruikt werd. Dit geeft aan dat de keuzes die in deze scripts vastgelegd liggen niet meer bevraagd worden nadat er een bepaalde mate van vertrouwen in de scripts bereikt is. Dit werd ook aangegeven door een geïnterviewde: "*Soms word je blind voor de inhoud bij automatisering*". Wij kregen de indruk dat dit vertrouwen intergenerationeel is: niet elke net beginnende modelleur hoeft opnieuw vertrouwen in de scripts te krijgen, maar vertrouwt hierin op het oordeel van collega's.

Tegelijkertijd geven de geïnterviewden, in de context van HYDROLIB, wel aan dat het overstappen op automatiseringsscripts weloverwogen en in stappen moet gebeuren. Een geïnterviewde stelt voor de automatisatie in etappes in te voeren om zo vertrouwd te worden met de mogelijkheden. Een ander

geeft aan het belangrijk te vinden inzicht te blijven houden in wat er gebeurt, en altijd een check uit te voeren achteraf, zeker in het begin. Een derde geïnterviewde waarschuwt voor het verlies van inzicht door automatisatie, en het risico van gebruik van automatisaties door mensen zonder verstand van de processen erachter, "*Het is voor het gemak van de modelleur, niet voor het vervangen van de keuzes*". Dit lijkt enigszins in contrast te zijn met de manier waarop nu al met automatiseringsscripts omgegaan wordt binnen de organisaties, maar geeft aan dat zeker in het begin het vertrouwen in de scripts nog moet groeien.

In het algemeen zien velen de voordelen van automatisering. Dit geeft meer tijd voor interpretatie, en dat blijft de meest cruciale stap. Uitdagingen ziet men met name in hoe generiek de scripts kunnen en moeten zijn, en daarmee hoe toepasbaar ze zijn voor specifieke vragen. Daarnaast werd aangegeven dat het aan elkaar kunnen breien van verschillende scripts belangrijk is: dat de output van de ene tool direct bruikbaar is als input voor de volgende. Ook is er behoefte aan heldere, korte handleidingen bij de scripts (hierbij werd bijvoorbeeld verwezen naar de handleiding van Sobek als een slecht voorbeeld: te dik en te lastig om te vinden wat je zoekt), of werd een '*personalized living document*' gesuggereerd: centrale documentatie waarin de waterschappen aanpassingen kunnen doen voor hun eigen specifieke wensen. Hiermee wordt voor de ontwikkelaars ook meteen inzichtelijk waar de tools voor gebruikt worden en waar de knelpunten zitten. Een andere suggestie was om een automatisch logboek bij te houden van alle stappen en keuzes in het modelleer-proces.

4. Discussie

4.1 Modelkeuzes

Uit de interviews blijkt dat 1) veel keuzes in automatiseringsscripts na een bepaalde tijd niet meer bevraagd worden, terwijl 2) men daar wel kritisch over is als men over zou stappen op HYDROLIB. Hierin herkennen wij de rol van vertrouwen en dat dat moet groeien over tijd, maar ook de rol van gewoontes in modelleren – na verloop van tijd worden bepaalde stappen niet meer bevraagd omdat ze een gewoonte zijn geworden. Dit brengt risico's met zich mee.

De discussie over welke keuzes waar moeten liggen blijft continu relevant. Er zullen bepaalde modelkeuzes zijn die altijd veel invloed hebben op het resultaat en daardoor altijd bij de modelleur moeten liggen (bijvoorbeeld met betrekking tot schematisatie), maar er zullen ook modelkeuzes zijn die bij bepaalde toepassingen weinig invloed op het resultaat hebben, terwijl ze bij andere toepassingen wel heel relevant zouden kunnen zijn. Als dit soort keuzes eenmaal vastgelegd zijn in een automatiserings-script en dit script vertrouwd wordt en onderdeel is van de modelleer-gewoontes, dan zal deze keuze niet snel meer ter discussie worden gesteld. Het blijft een continue afweging welke keuzes vastgelegd kunnen worden, welke keuzes de modelleur kan en moet maken, en welke keuzes probabilistisch benaderd moeten worden. Transparantie in welke keuzes in de scripts zitten, bijvoorbeeld met duidelijke markers en goede documentatie, faciliteert deze afweging.

4.2 Diffusion of innovations

De "*Diffusion of Innovations*"-theorie, uitgewerkt door Everett Rogers (Rogers, 1962), beschrijft hoe en waarom bepaalde innovaties wel of niet opgepikt worden. Hoewel er ook kritiek is op deze theorie (onder anderen door Rogers zelf), biedt het wel een nuttig raamwerk om de resultaten van de interviews mee te classificeren. Hiervoor putten we uit twee onderdelen van de theorie: de vijf karakteristieken die helpen bij de spreiding van innovaties, en de vijf stappen van het innovatie-besluitvormingsproces.

De vijf karakteristieken die helpen bij de spreiding van innovaties

Rogers heeft de vijf onderdelen geïdentificeerd als bepalend in de spreiding van innovaties (Rogers, 1962, Hfdstk. 6). Wij plaatsen deze onderdelen in de context van HYDROLIB, met behulp van de interviews.

1) Voordeel ten opzichte van huidige oplossing

Omdat veel waterschappen en adviesbureaus al interne tools hebben die delen van het modelleerproces automatiseren, en men al een hoge mate van vertrouwdheid heeft met deze tools, is het voordeel van overstappen op HYDROLIB ten opzichte van de huidige situatie niet direct evident. Om een gebruikers-community te creëren zou dit voordeel verder uitgewerkt en benadrukt kunnen worden. Voordelen zouden kunnen zijn; tijd- en geldbesparing ten opzichte van zelf (laten) ontwikkelen, een bepaalde kwaliteitsgarantie, aanzien, goede documentatie, en stabiliteit van de scripts (zoals uit de quote eerder bleek zijn sommige van deze tools erg gevoelig voor crashen).

2) Compatibiliteit met bestaande systeem/context

Het is gunstig dat HYDROLIB zich (nu nog) met name focust op D-HYDRO, en veel waterschappen op dit moment de overstap naar D-HYDRO gemaakt hebben, maken, of overwegen. Veel van de tools die nu intern gebruikt worden zijn wellicht minder compatibel met dit pakket, wat mogelijkheden biedt voor HYDROLIB.

3) Complexiteit, gebruiksvriendelijkheid

Zoals al naar voren kwam uit de interviews, moet het belang van een heldere handleiding niet onderschat worden. Dit zei één van de geïnterviewden bijvoorbeeld over de interne tools die al beschikbaar zijn: *“Veel tooltjes die we gebruiken hebben geen documentatie, dat mis je soms wel.”* (Dit is ook relevant voor punt 1, voordeel ten opzichte van huidige oplossing). Onder gebruiksvriendelijkheid valt ook de vindbaarheid van de benodigde scripts binnen de bibliotheek (hoe kan een gebruiker het script vinden voor de taak die deze persoon heeft?) en vindbaarheid van antwoorden in de documentatie.

4) Uitprobeerbaarheid; goede/duidelijke/makkelijke test-cases beschikbaar?

Om de adoptie nog verder te bevorderen, zou een goede test-case beschikbaar moeten zijn voor elk script in de bibliotheek. Zoals uit de interviews naar voren kwam, is vertrouwen belangrijk voor de adoptie van scripts. Zowel gebruiksvriendelijkheid (punt 3) als uitprobeerbaarheid (punt 4) kunnen hieraan bijdragen.

5) Zichtbaarheid van de innovatie en/of het effect ervan

Allereerst moeten de potentiële gebruikers op de hoogte zijn van het product. In het geval van HYDROLIB is er met het TKI-project al naamsbekendheid in de Nederlandse waterwereld gegenereerd. Belangrijk is om alle potentiële gebruikers (waterschappen, adviesbureaus) uiteindelijk op de hoogte te brengen van het bestaan.

De vijf stappen van het innovatie-besluit proces.

In de besluitvorming om over te stappen op een innovatie, in dit geval op de HYDROLIB software-bibliotheek, onderscheidt Rogers vijf stappen.

Allereerst moet de potentiële gebruiker op de hoogte zijn van het bestaan. Dit borduurt voort op punt 5 hierboven. Deze kennis-overdracht kan georganiseerd worden door middel van workshops. Vervolgens moet de potentiële gebruiker overtuigd raken van het voordeel van de nieuwe aanpak (HYDROLIB) ten opzichte van de oude (interne tools en scripts). Dit wordt bepaald op basis van ervaringen, zoals gebruikersvriendelijkheid en testbaarheid, en discussies met anderen. Vervolgens moet de beslissing worden gemaakt om over te stappen. In de context van HYDROLIB is het relevant dat deze beslissing waarschijnlijk niet puur bij de modelleur zelf ligt, maar ook bepaald wordt door institutionele context (de visie van het waterschap, het adviesbureau, de leidinggevenden) en dat er dus ook geïnvesteerd moet worden in het overtuigen op dit niveau. In de implementatie-fase raakt men bekend en vertrouwd met de innovatie en wordt deze uiteindelijk gebruikt. De laatste stap is bevestiging: de overstapper ondervindt de positieve effecten van het besluit.

5. Conclusie

Uit de interviews bleek dat modelleers voor vrijwel elke stap in het modelleerproces mogelijkheden zien tot automatiseren. Tegelijkertijd kwam naar voren dat veel waterschappen en adviesbureaus al automatiseringsscripts en tools hebben en gebruiken, en daarmee dus al regelmatig 'keuzes uit handen geven'. Vooral gevoeligheidsanalyse, kalibratie, validatie, en onzekerheidsanalyse gebeuren op dit moment niet of nauwelijks, terwijl de geïnterviewden wel aangaven deze stappen vaker uit te voeren als dit makkelijker gemaakt zou worden. Hier liggen dus specifiek kansen voor HYDROLIB om gebruikers kennis te laten maken met de software-bibliotheek, zonder dat dit direct de vertrouwde interne scripts of tools hoeft te vervangen.

Uit de interviews blijkt dus dat veel waterschappen en bureaus al eigen scripts hebben om delen van het modelleerproces te automatiseren. Het vergt een investering om over te stappen van deze vertrouwde omgeving naar de HYDROLIB bibliotheek. De eerdergenoemde modelstappen voor automatisering (kalibratie, validatie, gevoeligheidsanalyse) zouden kunnen dienen als eerste kennismaking met HYDROLIB. Vervolgens is het belangrijk dat men op basis van deze ervaring

vertrouwen krijgt in de HYDROLIB scripts en de overstap positief ervaart. Dit kan bewerkstelligd worden door middel van goede en heldere documentatie, goede vindbaarheid van de benodigde scripts, en goede voorbeelden en testcases.

Referenties

Rogers, E.M., (1962, reprint 1983) Diffusion of Innovations, Third edition, The Free Press, London

Rogers, E.M., (1962, reprint 1983) Diffusion of Innovations, Third edition, Hfdstk 6: Attributes of Innovations and their rate of adoption (p.210-238).

Van Waveren, R., S. Groot, H. Scholten, F. van Geer, J. Wösten, R. Kroeze en J. Noort (1999) Vloeiend modelleren in het waterbeheer – handbook Good Modelling Practice; Technical Report 99-05/99.036.STOWA/RIZA, <https://www.stowa.nl/publicaties/handboek-good-modelling-practice-vloeiend-modelleren-het-waterbeheer>