

Coagulatie-flocculatie-sedimentatie: modelleringsmogelijkheden en toekomstperspectief

Jasper Immink, Tessa Vrijhoeven (KWR), Bas Jacobs, Jennifer Lee, Joost de Munk (Waternet)

Coagulatie-flocculatie-bezinking is een complex proces. Dit maakt het lastig om uitkomsten te voorspellen of het proces in een model te vatten. KWR en Waternet hebben in recente projecten verschillende modellen opgezet om beter grip op de uitkomsten te krijgen. Het eerste is een mechanistisch model, gebaseerd op fysisch-chemisch inzicht. Het tweede model is een volledig datagedreven machine-learningmodel. Het derde is een hybride mechanistisch-AI-model, waarin de kracht van beide modeltypes naar voren komt. Alle drie de modellen geven nuttige inzichten en kunnen een nuttige bijdrage leveren aan optimalisatie van de zuiveringstechniekbesturing.

Coagulatie-flocculatie-bezinking (de vorming van vlokken die vervolgens kunnen bezinken) is een eeuwenoude waterzuiveringstechniek die in de watersector breed toegepast wordt. Toch zou deze zuiveringsmethode in de praktijk verbeterd kunnen worden met een model-gebaseerde sturing die leidt tot een lager chemicaliënverbruik. Daarom is er behoefte aan een overkoepelend model dat adviseert hoeveel vlokmiddel toegevoegd moet worden, of een model dat goed inzicht kan geven in de huidige of verwachte condities in een coagulatie-flocculatie-sedimentatieproces. [1] In de laatste jaren zijn verschillende modellen opgezet die aan deze vraag voldoen of nuttige inzichten kunnen bieden. Ook zijn nieuwe datagedreven methodes interessant om in deze modellen in te bouwen. In meerdere recente projecten bij KWR en Waternet zijn beide type modellen ontwikkeld en is hun toegevoegde waarde getoetst [2], [3].

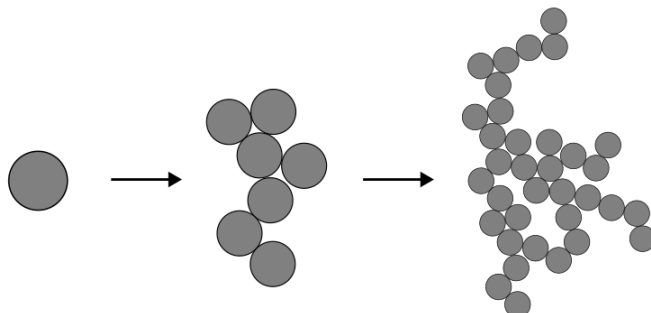
Mechanistische aanpak

Een van de methodes om coagulatie-flocculatie te modelleren, is een mechanistische aanpak. Deze methode bestaat het langst en houdt in dat het zuiveringsproces wordt voorgesteld als een aaneenschakeling van chemische, fysische, procestechnische en mechanistische processen. Deze processen worden in wiskundige formules gevat en de uitkomsten door computers uitgerekend [4]. Veelgebruikte mechanistische modellen in de waterzuivering variëren van enkele vergelijkingen tot uitvoerige modellen. Een voorbeeld van een relatief eenvoudig model is de Carman-Kozeny-vergelijking, die de drukval berekent over een zandfilter. Een uitvoeriger model is bijvoorbeeld AquaPriori, dat uitkomsten kan berekenen voor verschillende zuiveringstechnieken, en berust op verschillende opeenvolgende formules en parameters.

In twee recente projecten hebben onderzoekers van KWR mechanistische modellen voor coagulatie-flocculatie-bezinking opgezet, die belangrijke uitkomsten van deze zuiveringsmethode kunnen voorspellen: één voor de coagulatie-flocculatiestappen, en één voor de bezinkingsstap.

Het mechanistische coagulatie-flocculatiemodel gebruikt een aantal meetbare parameters, zoals de zoutconcentratie, de temperatuur en de zetapotential (een maat van elektrische lading op de deeltjes), om de vloggrootte te voorspellen op verschillende punten in de coagulatie-flocculatie-sedimentatiereactor. Deze vloggrootte is van belang voor bijvoorbeeld de zuiveringscapaciteit van

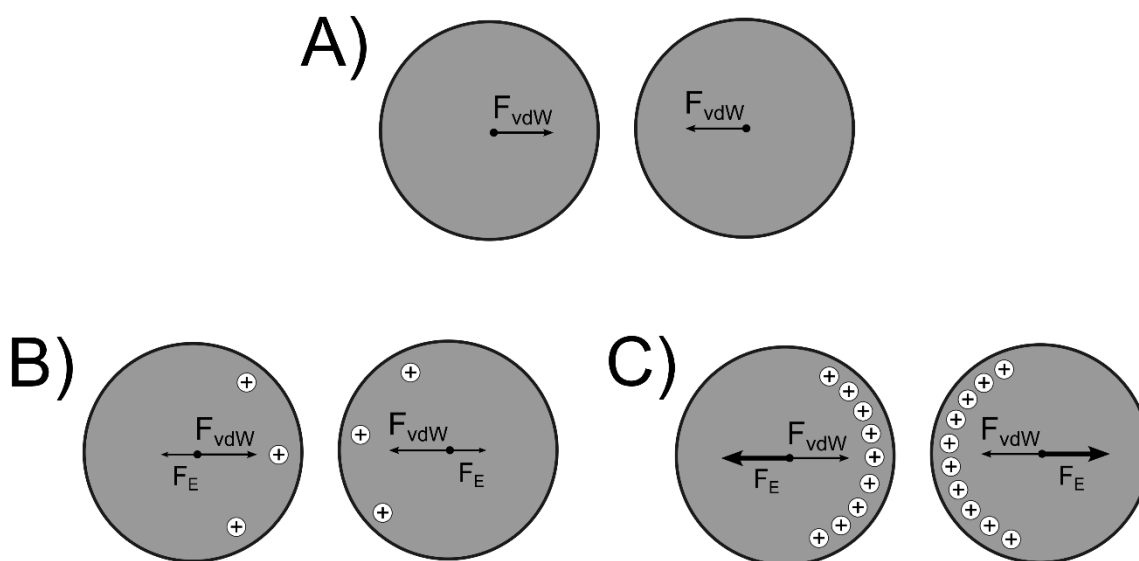
coagulatie-flocculatie: het is een van de belangrijkste parameters om te bepalen hoeveel verontreinigingen zich aan de vlokken kunnen hechten [5].



Afbeelding 1. Schematische weergave van het groeiproces van een vlokje. Kleine deeltjes vinden elkaar, aggregeren, en groeien uit tot een grote vlok die uiteindelijk zwaar genoeg wordt om uit te zakken

Dit model berust op twee verschillende sub-modellen die allebei een eigen fysisch-chemisch of mechanistisch fenomeen beschrijven. Ten grondslag van het model ligt het aggregatieproces: bij coagulatie-flocculatie wordt een vlokmiddel aan water toegevoegd, dat kleine vaste deeltjes (vlokjes) vormt in het water. Doordat deze vlokjes elkaar tegenkomen, botsen ze en blijven ze aan elkaar plakken, oftewel ze aggregeren. In het coagulatie-flocculatiemodel worden twee belangrijke parameters bepaald: de kans dat de vlokjes elkaar vinden ('botskans'), en de kans dat als de vlokjes elkaar gevonden hebben, ze ook blijven plakken ('plakkans').

Vlokjes vinden elkaar voornamelijk door drie processen: roeren, zwaartekracht en diffusie. [6] Door het water te roeren, komen vlokjes in feite in een (turbulente) stroom terecht waarin ze elkaar tegenkomen. Zwaartekracht zorgt ervoor dat de grootste vlokjes snel naar beneden zakken, terwijl de kleinere wat langer hoog blijven. Tijdens het zakken komen vlokjes elkaar tegen. Tenslotte is er nog de diffusie: de allerkleinste vlokjes bewegen ook, door een microscopische diffusiekracht, die ook meegenomen moet worden.



Afbeelding 2. Schematische weergave van de krachten die samen de plakkans bepalen. Twee ongeladen deeltjes in A) ervaren alleen vanderwaalskracht (F_{vdW}) en hebben een grote plakkans. Laag geladen deeltjes in B) ervaren zowel vanderwaalskracht als elektrostatische kracht (F_E), resulterend in een kleinere plakkans. Een grotere lading in C) betekent een kleinere plakkans

Deze drie processen worden in formules gevat. Voor het uitrekenen van deze formules zijn inputparameters nodig. Een belangrijke inputparameter is de roersnelheid (gegeven in G-factor), waarmee de roercontributie bepaald kan worden. De dichtheid van de vlokjes, of hun gewicht, is ook van belang voor bijvoorbeeld de sedimentatiesnelheid. Zo zijn er nog een aantal parameters nodig, die allemaal uiteindelijk op te zoeken of bekend zijn.

Op een vergelijkbare manier kan de plakkans berekend worden. Vlokken die bij elkaar komen, ervaren over het algemeen twee krachten: een afstotende elektrostatische kracht en een aantrekkende vanderwaalskracht. De vanderwaalskracht is zonder meetparameters te berekenen. De elektrostatische afstoting kan worden bepaald door de zetapotential vast te stellen; een relatief eenvoudige meting die een indicatie geeft van de lading op een deeltje.

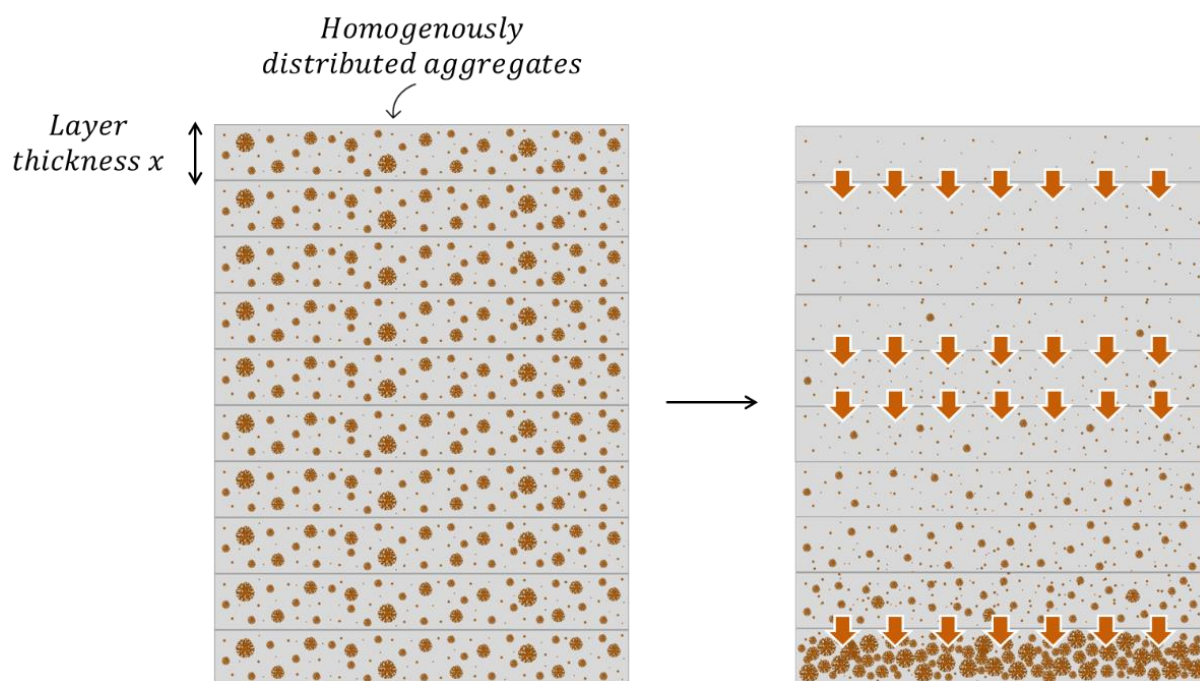
Met de berekende bots- en plakkans kan een computermodel de situatie in een coagulatie-flocculatiereactor voorspellen. Dit geeft (i) inzicht voor operators en procestechnologen, (ii) de mogelijkheid tot (bij)sturen, een lager chemicaliënverbruik en kleinere milieu-impact, en (iii) eerder inzicht in calamiteiten, waardoor het deze kan helpen voorkomen.

Het tweede, recent ontwikkelde model kan aan de hand van de berekende vloggroottes bepalen hoe lang de sedimentatiestap zou duren – dat wil zeggen hoelang het zou duren voordat de vlokken tot op de bodem van het sedimentatiebassin gezonken zijn. Hoe snel een vlok naar de bodem zakt, is niet alleen afhankelijk van de grootte van de vlok, maar ook van de porositeit van de vlok. Wanneer een vlok heel poreus is, bevat die meerdere kleine gaatjes en kanalen waar water doorheen kan stromen. Dit zorgt ervoor dat de vlok tijdens het uitzakken op een andere manier weerstand ervaart in vergelijking met een dicht vlokje, en dus met een andere snelheid naar de bodem uitzakt.

Met wiskundige formules kan voor vlokken van verschillende grootte en porositeit de sedimentatiesnelheid ('zaksnelheid') berekend worden. In het sedimentatiemodel bestaat het bassin uit bijvoorbeeld 10 lagen. In elke laag bevinden zich in de beginsituatie evenveel vlokken van elke

grootte (zie afbeelding 3). Met de berekende zaksnelheid wordt per laag bepaald welke vlokken in een bepaalde tijd uitzakken naar de volgende laag. Deze berekening wordt herhaald totdat het meeste materiaal in de onderste laag zit. Als 90% van het materiaal in de onderste laag zit, stopt de berekening en wordt genoteerd hoe lang het duurde voordat dit moment bereikt was. Dan zijn de vlokken in het model 'bezonken'.

In deze berekeningen duurt de bezinking net iets langer dan een uur per meter bassindiepte, wat relatief goed overeenkomt met de werkelijkheid. Met het inzicht in de sedimentatieduur op basis van dit bezinkingsmodel kan het coagulatie-flocculatie-bezinkingsproces efficiënter worden toegepast.



Afbeelding 3 Schematische weergave van het bezinkingsmodel waarin de sedimentatiesnelheid berekend wordt aan de hand van de grootte en porositeit van de vlokken. Op $t=0$ bevatten alle lagen evenveel vlokken van een bepaalde grootte. Wanneer sedimentatie bereikt is, $t=t_{\text{sedimentatie}}$, is een groot deel van de vlokken naar de bodem gezakt. De bovenste lagen bevatten alleen nog kleine vlokjes

Datagedreven aanpak

Een andere modelleeraanpak maakt gebruik van de data die waterbedrijven vergaren. Deze data kun je zien als een weerspiegeling van de ervaring van medewerkers van een waterbedrijf: de data beschrijven ook specifieke omstandigheden of effecten die niet in formules te vatten zijn. Met deze data kunnen de ervaringen van waterbedrijven en deze specifieke effecten meegenomen worden in de modellen. Dit zijn dan vaak statistische of AI-modellen.

Aan de hand van vergaarde data heeft Waternet heeft een datagedreven model gemaakt voor haar coagulatie-flocculatieproces in Nieuwegein. Dit model heeft als doel om de troebelheid te voorspellen aan de effluentkant van de coagulatie-flocculatiereactor. Dit is een waardevolle parameter voor operators: als zij de troebelheid van het water kunnen voorspellen, kunnen ze goed inschatten hoeveel vlokmiddel moet worden toegevoegd aan de reactor. Om deze troebelheid te voorspellen, wordt een aantal parameters gemeten aan de influentkant, zoals temperatuur, toegevoegde hoeveelheid vlokmiddel en pH.

Het datagedreven model is gebaseerd op beslisbomen. Een beslisboom leert regels afleiden op basis van patronen in de data. Een voorbeeld van een regel kan zijn: ‘als de temperatuur lager is dan 10 °C, én de troebelheid in het influent is 15 FTU (formazine turbidity unit), én er is 2 mg vlokmiddel per liter toegevoegd, dan is de troebelheid van het effluent 5 FTU’. In het datagedreven model wordt een groot aantal van deze beslisbomen gemaakt en op een slimme manier gecombineerd. Hierdoor kan het model complexe relaties leggen en patronen leren die in de data te vinden zijn.

Het model wordt de goede richting op gestuurd door bepaalde relaties af te dwingen op basis van domeinkennis van procestechnologen. Zo wordt bijvoorbeeld afgedwongen dat een hogere troebelheid van het influent - bij gelijke omstandigheden - niet kan leiden tot een lagere troebelheid van het effluent. Binnen deze begrenzingsen kan het model de kracht van de relaties zelf afleiden. Dit zorgt ervoor dat het model minder gevoelig is voor ongewenste patronen, die enkel zijn ontstaan door de manier dataverzameling, maar niet het coagulatie-flocculatieproces vertegenwoordigen.

De door het model gevonden patronen zijn door de procestechnologen van Waternet gebruikt om beslissingen te nemen over de dosering van het vlokmiddel. Dit heeft bijgedragen aan een besparing op chemicaliënverbruik. Daarnaast kan het model gebruikt worden om een automatische regeling te maken voor de dosering van het vlokmiddel. Voor een gegeven temperatuur, pH en troebelheid van het influent, kan de dosering dusdanig worden gekozen dat de troebelheid van het effluent op een bepaalde streefwaarde uitkomt. Dit kan leiden tot een optimaal gebruik van vlokmiddel en daarmee een minimale dosering.

Hybride aanpak

Het meest veelbelovend voor de watersector lijken modellen die gebruik maken van zowel datagedreven als mechanistisch modelleren. Op die manier kunnen de twee methodes elkaar aanvullen: een mechanistisch model is gelimiteerd door de kennis over het zuiveringsproces, maar dit kan opgevangen worden door ervaring en datamodellen. Aan de andere kant zijn datamodellen minder goed in het interpreteren van procesuitkomsten. De methode van modeltraining kan soms ook leiden tot modellen in lokale minima: modeltraining zoekt het model met de minimale fout, maar het trainingsalgoritme kan soms vast komen te zitten in een lokaal minimum in plaats van de best mogelijk minimum. Het is lastig om dit te verhelpen. Bij hybride modellering is dit een minder groot probleem. Daarnaast zijn de voorspellingen met hybride modellen veel beter te interpreteren. Een combinatie van deze twee methodes zou dus de kracht van beide kunnen combineren. [7]

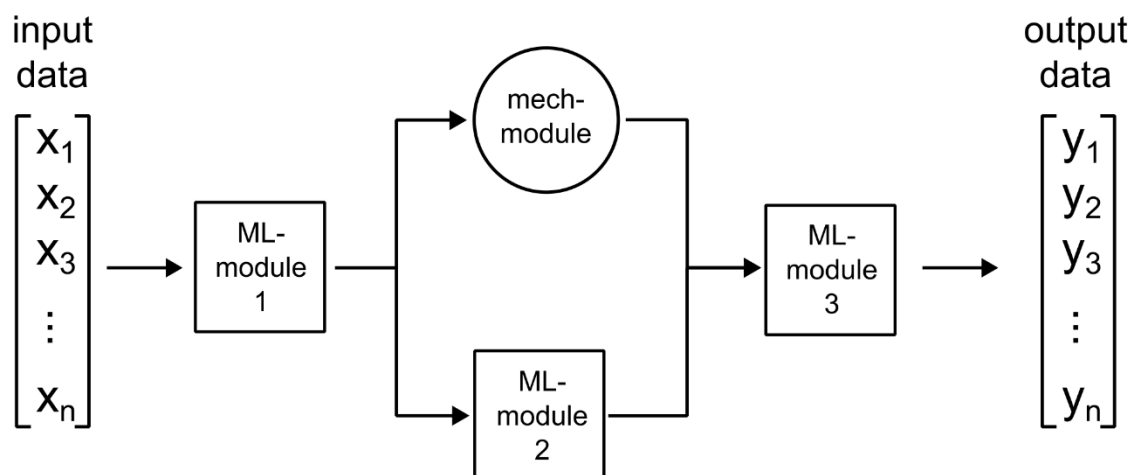
Dit is getest in een ander recent project van KWR. Daarin is een hybride model ontwikkeld voor de coagulatie-flocculatie van Waternet in Nieuwegein, waarin de verschillende modeltypen bij elkaar komen. Het mechanistische gedeelte van het model is gebaseerd op de modellen die in de vorige paragraaf beschreven staan. In de huidige versie van het model is de plakkans geïmplementeerd, en in latere versies zal ook de botskans worden ingebouwd. De plakkans en de botskans kunnen echter niet alle aspecten van het proces beschrijven. Zo zit er bijvoorbeeld een heel scala aan chemische stoffen in het water, die niet worden meegenomen in de plak- en botskans, maar wel invloed hebben op de vlokvorming. Dit soort effecten zit niet in de mechanistische modellen, maar wel in de meetdata: de troebelheid van het in- en effluent worden in de praktijk gemeten, en hierin wordt de eventuele invloed van aanwezige verontreinigingen dus wel degelijk meegenomen.

Het hybride model heeft, net als het datagedreven model, als doel om de troebelheid te voorspellen aan de effluent-kant van de coagulatie-flocculatiereactor. In het hybride model worden op twee

manieren datamodel-modules gebruikt om de gemeten parameters om te zetten naar een troebelheid aan de effluent-kant.

Allereerst wordt gekeken of een machine-learning (ML)-module deze parameters kan gebruiken om een troebelheid te voorspellen. Dit betekent dat de computer een model traint om zo goed mogelijk de uitkomst (troebelheid aan effluentkant) te voorspellen aan de hand van de inputparameters.

Zo'n ML-module kan ook ingebouwd worden in een groter geheel. Zo wordt in het KWR-model een twee-takmodel gebruikt: de data gaan aan één kant het model in en kopieën van de data worden in een van de twee takken geleid (zie **Fout! Verwijzingsbron niet gevonden.**).



Afbeelding 4. Grafische weergave van het hybride model. Het is ontworpen om de inputdata zo goed mogelijk te interpreteren, zowel met behulp van bekende en onbekende relaties tussen in- en output. Het model kan zo verbanden leggen op basis van kennis en data

Om het model goed werkend te krijgen, wordt ook een ML-module vóór de vertakking gezet. De mechanistische module heeft parameters als zoutconcentratie nodig, die in de praktijk niet gemeten worden. Daarentegen worden wel de pH en de toegevoegde hoeveelheid vlokmiddel gemeten, die iets zeggen over de zoutconcentratie. Een ML-module (afbeelding 4, module 1) kan iets zeggen over parameters als de zoutconcentratie, gegeven de beschikbare informatie. Deze informatie wordt vervolgens aan het mechanistische model doorgegeven. Iets vergelijkbaars gebeurt zodra de vertakkingen bij elkaar komen: het mechanische model voorspelt vloggrootte. Deze vloggrootte is gerelateerd aan troebelheid, maar er is meer nodig om de troebelheid te voorspellen. Een ML-module (**Fout! Verwijzingsbron niet gevonden.**, module 3) kan deze interpretatiestap uitvoeren. Tenslotte is de ML-module in de vertakking (module 2) bedoeld om verbanden tussen in- en output te leggen, die niet in het mechanistische model gevat zijn, maar wel in de data te vinden zijn. Daarmee zijn fenomenen te vatten die complex zijn en niet door de beschikbare kennis beschreven kunnen worden. Toepassing van dit modelontwerp heeft geleid tot modellen die beter de effluent-troebelheid kunnen voorspellen dan modellen die alleen gebruik maken van mechanistische of alleen van machine-learning-componenten. Om dit te kwantificeren wordt vaak de determinatiecoëfficiënt of R^2 gebruikt: een waarde tussen 0 en 1 waarbij 0 betekent dat het model niet beter is dan simpelweg het gemiddelde voorspellen, en 1 dat het model perfecte voorspellingen doet. Het door Waternet ontwikkelde model, puur gebaseerd op machine learning-componenten, had een R^2 tot 0,54; het hybride model bereikte R^2 -waarden tot 0,64. Aangezien een coagulatie-flocculatieproces sterk onderhevig is aan

onberekenbare omgevingsfactoren is een waarde van 0,5 is al interessant, maar een R^2 -verbetering is een mooie stap vooruit.

Samenvattend kan hybride modellering een interessante bijdrage leveren aan het voorspellen van zuiveringsprocesuitkomsten. Een ‘grey-boxmethode’ – een methode waar de berekening tot op zekere hoogte kan worden ingezien - kan meer inzicht bieden in waaróm het model voorspellingen doet. Ook draagt de proceskennis bij aan het model, en komt het voor dat hybride modellen met een kleinere dataset dezelfde voorspellingsaccuraatheid bereiken, vergeleken als een puur datagedreven model. Ten slotte kunnen hybride modellen in sommige gevallen tot accuratere voorspellingen komen dan pure machine-learning- of pure procesmodellen.

Stand van zaken en toekomststappen

Mechanistische, datagedreven en hybride modellen kunnen alle drie een nuttige bijdrage leveren aan de waterzuivering. Er zijn veel mogelijkheden om de modellen te combineren, elk met hun eigen voor- en nadelen. Voor elk proces kan een model ontworpen worden dat inspeelt op de beschikbare kennis en bestaande modellen, de beschikbare data en de sensoren die procesparameters meten. Met deze maatwerkmodellen kan een beter inzicht verkregen worden voor belangrijke processen. Daarnaast kunnen modellen ontworpen worden die direct advies geven over de operatie of die meer informatie teruggeven aan de operators. Dit geldt voor coagulatie-flocculatie-bezinking, maar andere zuiveringstechnieken kunnen ook profijt hebben van deze manier van modellering. De mogelijkheden zijn talrijk en afhankelijk van de wensen van een bedrijf, kunnen er modellen van hoge kwaliteit worden ontworpen en ingezet.

Referenties

- Jiang, J.-Q. (2015). ‘The role of coagulation in water treatment’. *Curr. Opin. Chem. Eng.* Volume 8, pp. 36-44 (2015).
2. Korevaar, M. W. et al. (2022). *BO Softsensor coagulatie*. KWR, Nieuwegein.
3. Immink, J. N. et al. (2023). *BO Softsensor flocculatie*. KWR, Nieuwegein.
4. Thomas, D. N. et al. (1999). ‘Flocculation modelling: a review’. *Water Research*, Volume 33, Issue 7, pp. 1579-1592 (1999)
5. Guérin, L. et al. (2019). ‘Fractal dimensions and morphological characteristics of aggregates formed in different physico-chemical and mechanical flocculation environments’. *Colloids Surf. A*, Volume 560, pp. 213-222.
6. Jeldres, R. I. et al. (2018). ‘Population balance modelling to describe the particle aggregation process: A review’. *Powder Technol.*, Volume 326, pp. 190-207 (2018).
7. Cuomo, S. et al. (2002). ‘Scientific Machine Learning Through Physics-Informed Neural Networks: Where we are and What’s Next’. *J. Sci. Comput.*, Volume 92, p. 82