

Verlag van de Pastas-gebruikersdag 9 februari 2023, TU Delft

VINCENT POST

Op 9 februari 2023 vond in Delft een Pastas-gebruikersdag plaats. Pastas is een opensourcerekeningtool voor tijdreeksanalyse geschreven in de programmeertaal Python. Circa vijftig hydrologen verzamelden zich in een collegezaal van de TU om kennis uit te wisselen over de toepassing en toekomst van Pastas. De dag werd geopend door Frans Schaars van Artesia. Hij gaf een kort overzicht van de mogelijke toepassingen van Pastas, waaronder meetreeksvalidatie, voorspelling van grondwaterstanden tijdens droogte en als hulpmiddel bij de ijking van numerieke modellen.

Verlag

Presentaties van de ontwikkelaars

Daarna was Mark Bakker (TU Delft) aan de beurt met een presentatie over de geschiedenis van tijdreeksanalyse-software die begon met een korte uitleg van tijdreeksanalyse met responsfuncties. Dit soort modellen zijn een voorbeeld van relatief eenvoudige modellen waarin slechts enkele onbekende parameters gefit hoeven te worden. Een van de eerste software *tools* voor tijdreeksanalyse was Menyanthes, maar er bestond een wens om een opensourceprogramma te ontwikkelen dat buiten een grafische gebruikersomgeving om kon worden gebruikt. Daarom is Pastas ontwikkeld en uitgebracht onder de *MIT license*, wat betekent dat iedereen de code kan gebruiken en aanpassen. De code staat op GitHub, en via dit platform kunnen ontwikkelaars en gebruikers verbeteringen en uitbreidingen aanleveren, maar het Pastas-team besluit uiteindelijk of deze in de releaseversie van de code terechtkomen. Pastas is dus continu in ontwikkeling. Zo konden vroegere versies alleen lineaire systemen simuleren, maar inmiddels zijn er ook mogelijkheden toegevoegd om niet-lineaire systemen te beschrijven.

Het huidige versienummer van Pastas is 1.0.1, en deze versie heeft als bijnaam Arrabiata. Een overzicht van de huidige mogelijkheden van de code werd gegeven door Raoul Collenteur, één van de hoofdontwikkelaars en inmiddels werkzaam bij EAWAG in Zwitserland. Hij begon met een overzicht van de ontwikkeling, die ongeveer zeven jaar geleden begon. Het hoofddoel bij de ontwikkeling van versie 1.0 was een stabiele code die dient als basis voor verdere uitbreidingen in de toekomst. Het kortst mogelijke model bestaat uit slechts vijf regels Pythoncode. Een nieuwe functionaliteit is de mogelijkheid om verschillende modellen grafisch te vergelijken. Ook is het mogelijk om de verandering van de parameters tijdens de iteraties grafisch te volgen en is er een module waarmee statistische waarden die de *goodness of fit* uitdrukken kunnen worden berekend. Voor de bruikbaarheid is er heel veel aandacht besteed aan

de documentatie, die via ReadtheDocs toegankelijk is. Gebruikers die nog versie 0.23 gebruiken ondervinden een aantal gevolgen door de veranderingen in de code, die zijn gedocumenteerd op het discussieplatform op GitHub. Het inlezen van meetreeksen (zoals neerslag- en verdampingsdata van het KNMI) is nu gescheiden van Pastas zelf en ondergebracht in een nieuwe Pythonbibliotheek, HydroPandas. Dergelijke functionaliteit is vooral gericht op de Nederlandse context en door deze te scheiden van het rekengedeelte wordt Pastas geschikter voor gebruik buiten Nederland.

Na de koffie gingen Martin Vonk, Ruben Caljé en Raoul Collenteur in op de keuze van een modelstructuur. Pastas bevat talrijke opties voor het kiezen van een *stress*-model, responsfunctie, foutmodel en *solver*. Ruben gaf een voorbeeld van een niet-lineair systeem waarin de stijging van de grondwaterstand wordt begrensd door drainerende waterlopen. Naast drainage door waterlopen is de niet-lineariteit het gevolg van een afnemende verdamping met toenemende diepte van de grondwaterstand. De stijghoogte voor dit voorbeeld was gegenereerd met een numeriek model gemaakt met NLMOD (nog een Pythonbibliotheek uit de Artesiastal). De reeks bestaat uit de stijghoogte onder de deklaag. Met een lineair model in Pastas (exponentiële responsfunctie) is het percentage verklaarde variantie al 86%, wat normaal gesproken een goede score is, maar voor dit synthetische voorbeeld verwacht je nog een beter resultaat. Het gebruik van de Kraijenhoff- of Gamma-responsfunctie leverde slechts een marginale verbetering op. De demping als gevolg van de drainage kan worden meegenomen met de ThresholdTransformmethode die in Pastas is ingebouwd, waardoor het percentage verklaarde variantie stijgt met 10% punten. De niet-lineaire TarsoModel-respons optie (Tarso staat voor *Threshold AutoRegressive Self-exciting Open loop*), waarbij de stijghoogte bepaalt welke van twee responsfuncties actief is, resulteert in een bijna perfecte fit tussen de gesimuleerde reeks en de synthetische reeks uit het numerieke model. Na dit voorbeeld lichtte Martin toe hoe een niet-lineair model gebruikt kan worden om de fit te verbeteren (ten opzichte van een lineair model) in systemen waarin processen in de onverzadigde zone belangrijk zijn. Dergelijke modellen kunnen ook worden toegepast om de grondwateraanvulling te schatten. Ook het effect van sneeuw, die pas infiltreert nadat het begint te smelten, kan worden gesimuleerd. Deze module is ontwikkeld door Raoul en gevalideerd met lysimeterdata uit Oostenrijk. Martin liet zien dat ook systemen waarin een stapsgewijze verandering optreedt (bijvoorbeeld het sluiten van een stuw of een plotselinge landgebruiksverandering), gesimuleerd kunnen worden. Hij sloot de presentatie af met een aantal visualisatiemogelijkheden in Pastas.

David Brakenhoff presenteerde daarna een voorbeeld van het inschatten van de verlaging wanneer er sprake is van meerdere onttrekkingen op vier winlocaties van Vitens in Overbetuwe. De winningen zitten allemaal in het tweede watervoevende pakket waar neerslag, verdamping, rivierpeilfluctuaties en onttrekkingen invloed hebben op de stijghoogte. Het doel was het schatten van de verlaging van de freatische grondwaterstand ten gevolge van de onttrekkingen, en hoe deze zich verhouden tot de verlaging als gevolg van droogte. Voor het simuleren van het effect van meerdere onttrekkingen is gekozen voor een benade-

ring met één responsfunctie, gebaseerd op Hantush, voor alle onttrekkingen die geschaald worden met de afstand van de winput tot de waarnemingsbuis. Hierdoor wordt het aantal modelparameters gereduceerd ten opzichte van een model waarin elke winning een eigen responsfunctie heeft. Bij het maken van het tijdreeksmodel zijn de processen die de stijghoogte beïnvloeden stapsgewijs toegevoegd, wat acht modelstructuren opleverde die werden toegepast op 250 stijghoogtereeksen. De beoordeling van de resulterende 2000 modellen is geautomatiseerd, waarbij is gekeken naar de kwadratensom, de autocorrelatie van de ruis, realistische responstijden en de onzekerheid over de *gain*. Van alle modellen voldeden 247 modellen (12%) aan alle criteria, waarbij voor sommige tijdreeksen meerdere betrouwbare modellen werden geselecteerd. In die gevallen werd het beste model gekozen op basis van het Aikake information criterium, wat modellen bestraft die relatief veel parameters hebben. Hierdoor bleven voor 126 van de 250 meetlocaties betrouwbare tijdreeksmodellen over. Met behulp van de responsfunctie van de onttrekkingen kon de stijghoogteverandering als gevolg van de onttrekking worden bepaald, die op drie van de winlocaties duidelijk afnam met de afstand van het waarnemingspunt tot aan de winning. Bij de winning waar dit niet het geval was, kon dit worden verklaard uit de aanwezigheid van kleilagen. Met het model kunnen ook de responstijden worden afgeleid, die, zoals verwacht, voor de freatische grondwaterstanden duidelijk langer waren dan voor de grondwaterstanden in het diepere watervorende pakket. Uiteindelijk zijn de modellen gebruikt om de bijdrage van de winningen aan de verlaging van de freatische grondwaterstand te bepalen, inclusief mogelijke maatregelen om deze tijdens droge zomers te verminderen. De optie om meerdere onttrekkingen te simuleren is ook ingebouwd in Pastas. Een Jupyter-notebook met een voorbeeld is gegeven op de documentatie-*website* van Pastas.

Presentaties door gebruikers

Na de lunch (geen pasta) was het woord aan de gebruikers. Pytrik Graafstra gaf een overzicht van de toepassingen van Pastas bij Aveco de Bondt. Deze behelzen het bepalen van de plausibiliteit van oude reeksen, het relateren van systeemingenrepen of droogte aan de grondwaterstand, het bepalen van dominante invloeden in een systeem en als ondersteuning bij de rapportageplicht van vergunningen. Bij plausibiliteitsbepalingen komen soms dingen naar voren zoals meetfouten door het foutief inmeten van de bovenkant van de peilbuis. Bij droogtestudies kan de invloed van verschillende factoren worden onderscheiden en gekwantificeerd, zoals een lage netto-neerslag, een laag rivierpeil of een tijdelijk hogere onttrekking. Door de resultaten in een ruimtelijke context te plaatsen (bijvoorbeeld de ligging van stroomgeulen in het rivierengebied of de afstand vanaf een stuwwal), kunnen tijdreeksmodellen veel kennis over de werking van het grondwatersysteem opleveren.

Anouk Sprong (Vitens) sprak over tijdreeksanalyse rondom winning Van Heek in het Montferland, waar nitraatproblemen ertoe hebben geleid dat het debiet in de jaren tachtig moest worden verminderd. Een aantal jaar geleden is gekeken of het mogelijk zou zijn om met een diepere onttrekking de drinkwaterwinning weer te verhogen, maar in de omgeving is veel grondwaterafhankelijke natuur.

Voor elk peilfilter in de omgeving van de winning zijn twee tijdreeksmodellen met neerslag en verdamping gemaakt: één zonder de winning en één met de winning. Het meenemen van de winning aan het model geeft wel een verbetering van de fit maar de onzekerheid over de gefitte parameters is erg hoog. Uit de analyse bleek geen duidelijk causaal verband tussen de winning en de grondwaterstand. Dit valt toe te schrijven aan kleischotten in de ondergrond die zijn ontstaan door stuwning van de pakketten tijdens de ijstijd. Hierdoor ontstaat compartimentalisering van de ondergrond, waardoor elk filter zijn eigen karakteristieke verloop heeft. Hierdoor kan tijdreeksanalyse veelal niet uitwijzen wat het effect is van onttrekkingen op de grondwaterstand.

Jorn van de Velde (Sumaqua) gaf een aantal voorbeelden van tijdreeksanalyse in België, te beginnen met de grondwaterstandsindicator die informatie geeft over de huidige toestand van het freatische grondwater. Tijdreeksanalyse wordt hierin voor 200 peilfilters toegepast om te interpoleren tussen handmatige metingen en om scenario's (nat weer, droog weer) door te rekenen. De modellen worden maandelijks automatisch opnieuw gekalibreerd op basis van de meest recente data. Een tweede voorbeeld was het bepalen van de invloed van klimaatverandering op het freatische grondwater op 217 locaties in Vlaanderen. Hiervoor werden, naast Pastas, SWAP en een *machine learning* (ML) algoritme gebruikt. SWAP en Pastas gaven gelijkwaardige resultaten. Het ML-algoritme had een iets betere gemiddelde kwadratensom maar overschatte de laagste grondwaterstanden. Uiteindelijk is voor doorrekening met SWAP gekozen omdat het fysisch gebaseerd is en de opdrachtgever meer vertrouwd was met het modelconcept. Het derde voorbeeld ten slotte was het bepalen van de invloed van terreinmaatregelen op droogte. Door op 186 verschillende locaties in Vlaanderen tijdreeksmodellen te kalibreren tot 2016 en de modelvoorspelling van de daaropvolgende jaren te vergelijken met de metingen, kon bepaald worden of er sprake was van verdroging of vernatting. Verdroging werd waargenomen op 57% van de locaties, terwijl slechts 9% vernatting vertoonde. Het effect van de terreinmaatregelen zal pas later duidelijk worden, omdat ze op dit moment worden uitgevoerd. De waarde van Pastas is bij de gepresenteerde grootschalige grondwateronderzoeken gebleken, maar de kwaliteit van de beschikbare data is vaak een beperkende factor.

Als laatste van de vier gebruikers sprak Marjolein van Huijgevoort (KWR) over het analyseren van de recente droogtes in Nederland met Pastas. Het doel was het doen van een regionale grondwaterdroogte-analyse met een groot aantal stijghoogte-meetreeksen en het in kaart brengen van de ontwikkelingen in Zuid- en Oost-Nederland tijdens de droogte van 2018. De geselecteerde reeksen moesten voldoen aan een aantal criteria, zoals de reekslengte in verhouding tot de responstijd, het aantal ontbrekende metingen en het laatste meetpunt moest in 2018 vallen. Ook statistische waardes van de modelfit moesten aan bepaalde voorwaarden voldoen. Na toepassing van de criteria bleven 1931 van de 5818 beschikbare reeksen over. Vervolgens zijn de modellen doorgerekend vanaf oktober 1995 tot en met 'vandaag'. Op basis van de frequentieverdeling van de meetreeksen wordt de *Standardised Groundwater Index* (SGI) op dagbasis berekend. Deze kunnen dan ruimtelijk worden weergegeven zodat de ontwikkeling

van de grondwaterdroogte op verschillende tijdstippen in beeld kan worden gebracht. De gegevens zijn ontsloten via www.droogteportaal.nl.

Andere tools en de toekomst van Pastas

Na de gebruikerspresentaties vertelde Onno Erbens van Artesia iets over Hydropandas en Pastastore. Hydropandas is gemaakt om het inlezen van grondwaterstanden, neerslag en verdamping te vergemakkelijken. Onno liet zien hoe grondwaterstanden uit de BRO-database kunnen worden opgehaald in Python. De opgehaalde informatie bestaat zowel uit de meetreeks als de bijbehorende metadata. Op dezelfde manier kunnen neerslag en verdampingsreeksen van het KNMI worden aangevraagd, waarbij de gebruiker enkel de coördinaten van het meetpunt hoeft op te geven en de software het dichtstbijzijnde KNMI-station opzoekt. Vervolgens kan met een handvol aanvullende coderegels een Pastas-model worden doorgerekend. Met de Pastastore kunnen Pastas-modellen worden gebundeld, zodat het veel gemakkelijker wordt om voor meerdere locaties in één keer tijdreeksmodellen op te stellen. De functionaliteit werd live gedemonstreerd, waarbij in slechts enkele minuten voor zes locaties de benodigde gegevens werden opgehaald en in Pastas-modellen werden omgezet. Naast de modelresultaten kunnen ook de locaties van de KNMI-meetstations en de gebruikte peilbuizen ruimtelijk in beeld worden gebracht. De data en modellen kunnen ook opgeslagen worden in een zip-bestand, waardoor het mogelijk is om op een later tijdstip weer met dezelfde data en de doorgerekende modellen verder te werken.

Het laatste programmaonderdeel bestond uit een discussie over de toekomst van Pastas waarbij de aanwezigen werden opgedeeld in groepen. Bij de discussie werd stilgestaan bij vier vragen, die tezamen met de antwoorden te vinden zijn op het Pastas-discussieforum op GitHub. Raoul sloot de discussie af met een aantal dingen die het Pastas-team zelf graag zou willen, waaronder een voorspellingsmodule, het inbouwen van betere hydrologische modellen, tools voor kalibratie en onzekerheidsanalyse en het *benchmarken* van Pastas met andere (fysisch gebaseerde) modellen.

Auteur

VINCENT POST
EDINSI Groundwater
vincent@edinsi.nl

