



Voorspellen van drinkwatertemperaturen aan de tap met sensoren en data science

Joost van Summeren, Dirk Vries (KWR Watercycle Research Institute), Johan van Erp (Brabant Water)

Volgens de Drinkwaterwet mag drinkwater niet warmer zijn dan 25°C. Vanwege klimaatverandering en verstedelijking zal de kans op hoge temperaturen in het leidingnet naar verwachting toenemen. Om overschrijdingen te voorkomen is het belangrijk te weten waar, wanneer en waarom in het leidingnet (te) hoge temperaturen optreden. Met dataonderzoek is het KWR samen met Brabant Water en Nelen & Schuurmans gelukt om de temperatuur ‘aan de tap’ op 0,5°C nauwkeurig te voorspellen door een algoritme te trainen met actuele weergegevens en temperatuurmetingen op de taplocatie. Een volgende uitdaging is om in een groter en diverser gebied temperaturen te bepalen.

De zomer van 2018 was extreem warm, zeer zonnig en droog. Klimatologen voorspellen dat dergelijke zomers de komende decennia in Nederland vaker voor zullen komen. Vanwege de klimaatverandering en verstedelijking neemt de kans op langdurig aanhoudende hoge temperaturen toe, waardoor ook de temperatuur van drinkwaterleidingen in de bodem toeneemt.

In de Drinkwaterwet (2011) is een maximumwaarde voor drinkwater aan de tap vastgesteld van 25 graden. Deze grens garandeert het comfort van fris drinkwater voor de klant en draagt bij aan levering zonder microbiologische risico's. Om overschrijdingen in de toekomst te voorkomen is het belangrijk te weten waar, wanneer en waarom er in het leidingnet (te) hoge temperaturen optreden.

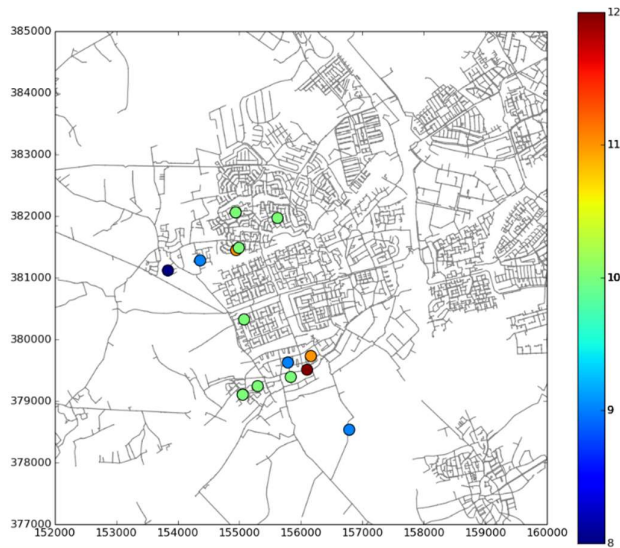
Tegen deze achtergrond heeft KWR Watercycle Research Institute samen met Brabant Water en ingenieursbureau Nelen & Schuurmans dataonderzoek uitgevoerd om met omgevingsdata en sensorgegevens de kennis over de temperatuur in het distributienet te vergroten. Dataminingstechnieken zijn bij uitstek geschikt voor het analyseren van grote datasets. Brabant Water heeft recent het databeheer gestroomlijnd van gegevens van met temperatuurmeters uitgeruste watermeters. Dit maakt de data toegankelijk en beperkt de inspanning wanneer nieuwe gegevens of sensoren beschikbaar komen. Met een visualisatieplatform, gekoppeld aan het databeheersysteem, zou het mogelijk zijn om in de toekomst grote hoeveelheden data samen te brengen, inzichtelijk te maken en te gebruiken voor operationele doeleinden. Bij dit alles is het belangrijk om te bedenken hoe de verkregen informatie verder benut kan worden in de bedrijfsvoering van het waterbedrijf.

Gevarieerde databronnen voor datagedreven onderzoek

Brabant Water experimenteert met een nieuw type watermeter waarmee ook temperatuurmetingen mogelijk zijn. Ten tijde van het onderzoek waren 40 van deze watermeters geïnstalleerd bij klanten. Om de batterij te sparen, registreren de watermeters alleen dagelijks de minimumtemperatuur van het water en van de binneninstallatie. De hoeveelheid temperatuurgegevens is dus beperkt. Ondanks deze kleine hoeveelheid gegevens is het interessant om te zien of de sensorgegevens de modelberekeningen kunnen aanvullen.

Als onderzoeksgebied is gekozen voor Veldhoven, omdat daar de meeste meters in gebruik zijn. De data van de watermeters over de periode van 1 mei 2016 tot en met oktober 2017 zijn verzameld. Ter

illustratie zijn de meetlocaties inclusief de temperatuur op 1 mei 2016 getoond in afbeelding 1. Te zien is dat de gemeten temperatuur op sommige locaties hoger is dan elders.

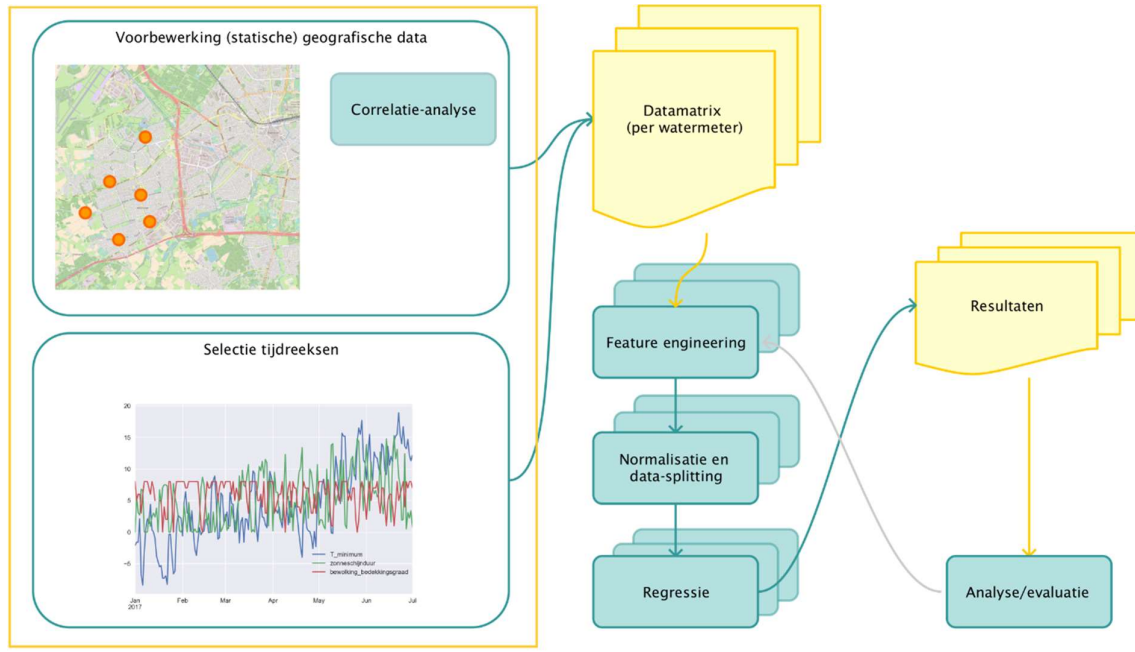


Afbeelding 1. Temperatuurmetingen op 9 locaties in Veldhoven op 1 mei 2016

Om met dataminingstechnieken relaties tussen de gemeten temperatuur en invloedfactoren te onderzoeken is een aantal datasets geselecteerd. De selectie van datasets komt voort uit onderzoek dat KWR de afgelopen jaren heeft uitgevoerd (zie kader). Het gaat om:

- leidingneteigenschappen (diameter, materiaal)
- sensorgegevens van temperatuur en volumestroom
- weergegevens van het KNMI (luchttemperatuur, zonschijnduur, zonnestraling, bewolgingsgraad)
- omgevingskenmerken uit de de Klimaateffectatlas [1]
- omgevingskenmerken uit het register Publieke Dienstverlening Op de Kaart (PDOK - Bodemsoort, gebouwhoogtedata) [2]

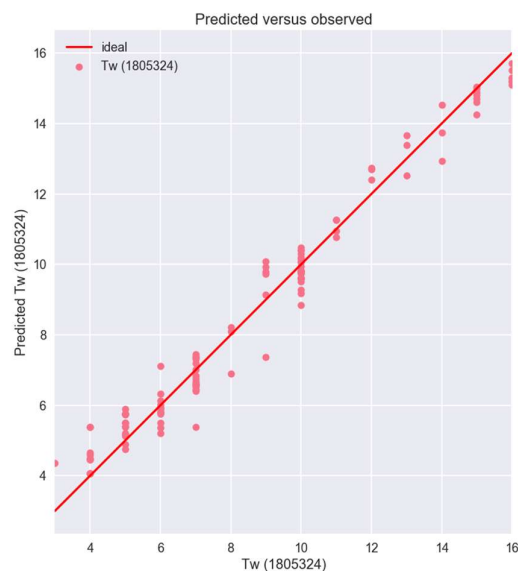
Het data-analyseproces is schematisch weergegeven in afbeelding 2. Met trainingsgegevens leert het algoritme door iteratie over de relaties tussen omgevingsdata en de gemeten temperatuur. Na deze trainingsfase kan het algoritme uit nog niet eerder geziene omgevingsdata de temperatuur aan de tap voorspellen.



Afbeelding 2. Stappen in de gebruikte datamining: (a) inventariseren van sensorgegevens en omgevingsdata (correlatie-analyse). (b) samenvoegen van relevante data in een datamatrix. (c) voorbereken van de data: feature engineering, normalisatie en datasplitting en (d) analyseren van de relaties tussen temperatuur en omgevingsfactoren (regressie). Om relaties uit de data te onderzoeken, worden stappen (c) en (d) herhaald

Resultaat onderzoek

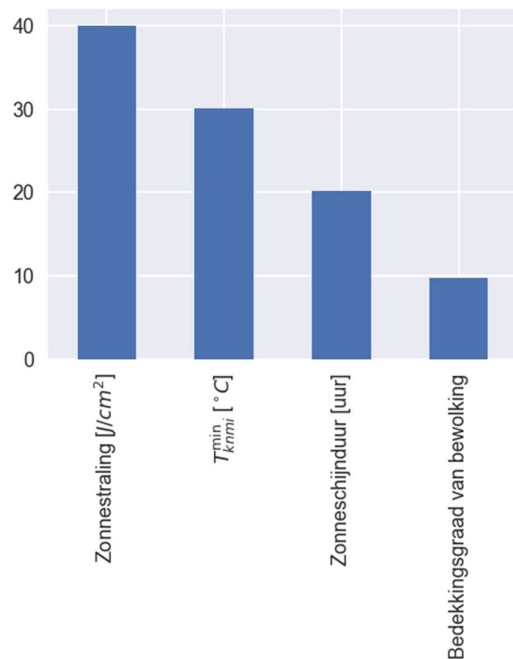
Met de gevolgde dataminingstechniek is aangetoond dat de drinkwatertemperatuur op een locatie nauwkeurig (tot op 0,5°C) is te voorspellen als op die locatie eerdere metingen en actuele weergegevens beschikbaar zijn. Als voorbeeld zijn voor één meetlocatie de verschillen tussen de voorspelde en gemeten temperatuur tegen elkaar uitgezet, in afbeelding 3.



Afbeelding 3. De voorspelde drinkwatertemperatuur ("Predicted Tw") uitgezet tegen de gemeten temperatuur ("Tw") voor meetlocatie 1805324. Hoe kleiner de afstand tot de diagonale lijn, hoe kleiner de onnauwkeurigheid van de voorspelling

De voorspellingen worden nauwkeuriger als perioden van laag verbruik worden weggelaten uit de analyse. Als klanten bijvoorbeeld langdurig afwezig zijn, zal het water de temperatuur van de omgeving (bijv. de meterkast) aannemen in plaats van de temperatuur in het leidingnet. De temperatuurvoorspellingen blijken nauwkeuriger als op de luchttemperatuur een maandelijkse correctie wordt toegepast voor hitte-eilandeffecten (het fenomeen dat de temperatuur in stedelijke gebieden kan toenemen ten opzichte van omliggende landelijke gebieden). Om de temperatuur voor een gebied in detail te monitoren en een beeld over *hotspots* te verkrijgen, is het aan te bevelen om een substantieel aantal watermeters in een divers gebied te installeren en over meerdere jaren te meten.

Naast de watertemperatuurvoorspellingen, geeft het dataminingsalgoritme weer in welke mate omgevingsfactoren procentueel bijdragen aan de voorspellingen. Een voorbeeld voor één sensorlocatie is weergegeven in afbeelding 4: de zonnestraling, buitentemperatuur, zonneschijnduur en bewolgingsgraad blijken het belangrijkste in het maken van een betrouwbare voorspelling.



Afbeelding 4. Factoren die het meest van invloed zijn op de modelvoorspellingen, procentueel uitgedrukt op de y-as

Conclusie onderzoek

Het is gelukt om per meetlocatie de watertemperatuur aan de tap tot op 0,5°C nauwkeurig te voorspellen door een algoritme te trainen met actuele weergegevens en eerdere temperatuurmetingen op de taplocatie. Ook zijn de bepalende factoren voor een goede voorspelling aan te geven; in dit geval zijn dat weersfactoren.

Een volgende uitdaging is om in een groter, meer divers gebied en op locaties waar geen metingen voorhanden zijn *hotspots* te bepalen: locaties in het distributienet met de grootste kans op te hoge temperaturen. Hoe zullen geografische variaties van parameters doorwerken in de drinkwatertemperatuur, zoals gebouwhoogte, de aanwezigheid van lokale warmtebronnen en het gebruik van het landoppervlak of de grondsoort? Dit zal bijdragen aan het doorgronden van hotspots

en inschatten van overschrijdingen, zelfs op locaties waar geen metingen voorhanden zijn. Met negen sensorlocaties in een voornamelijk uniform gebied is dit nog niet mogelijk. Een uitgebreidere set watermeters over een gebied met grotere diversiteit (rurale en stedelijke kenmerken, verschillende soorten bebouwing), eventueel aangevuld met metingen van de temperatuur op verschillende plekken, biedt een startpunt om deze vragen verder te onderzoeken met dataminingtechnieken.

Overwegingen en inzichten voor de toekomst

De datamining-aanpak in combinatie met temperatuursensoren op de watermeter biedt een perspectief om hotspots te identificeren en te voorspellen, indien veel meetgegevens voorhanden zijn. Er zijn echter meer benaderingen mogelijk op de lijn van modelstudie naar datagedreven aanpak. Een keuze in de te volgen aanpak hangt samen met de bedrijfsspecifieke vragen: welke nauwkeurigheid en detailniveau van een temperatuurkaart gewenst zijn. In tabel 1 worden de voor- en nadelen van verschillende benaderingen besproken.

Door bij elke klant een temperatuurmeter te plaatsen ontstaat een goede dekking van het meetgebied. De invloed van lokale warmtebronnen, zoals hoogspanningskabels, wordt automatisch meegenomen. Het is zeer kostbaar om dit op grote schaal te doen. Als een waterbedrijf echter sowieso geavanceerde meters bij klanten wil gaan plaatsen, bijvoorbeeld om het verbruik nauwkeuriger te meten, dan is het interessant om temperatuurmetingen hierop te laten 'meeliften'.

Tabel 1. Overzicht van benaderingen om de hotspots in kaart te brengen. Van links naar rechts verschuift de nadruk van kennis uit sensormetingen naar kennis uit fysische modellen

	Volledige dekking sensoren	Datagedreven aanpak	Hybride aanpak	Modelstudie (o.b.v. fysische wetten) gevalideerd met metingen
Inzicht hotspotlocaties (nauwkeurigheid en resolutie)	Optimaal	Afhankelijk van sensordichtheid en interpolatie	Gebaseerd op combinatie modellen & sensoren	Afhankelijk van modelaannames + validatie
Inzicht in bepalende factoren	Nee (tenzij aangevuld met omgevingsdata)	Ja (voor zover aanwezig in data)	Ja (voor zover aanwezig in data en model)	Ja, voor zover in het model aanwezig
Inzicht in causale verbanden	Nee	Nee	Ja	Ja
Aantal benodigde sensoren	Bij elke aansluiting	Afhankelijk van gewenst inzicht	Afhankelijk van gewenst inzicht	Ter validatie
Omgevingsdata nodig?	Nee	Ja	Ja	Ja
Scenario's modelleren mogelijk?	Nee	Ja	Ja	Ja

Dankwoord

Dit onderzoek is uitgevoerd binnen het kader van het bedrijfstakonderzoek (BTO) van de Nederlandse waterbedrijven en de Toeslag voor Topconsortia voor Kennis en Innovatie (TKI's) van het ministerie van Economische Zaken en Klimaat. De auteurs danken Jan-Maarten Verbree (Nelen & Schuurmans) voor zijn bijdrage aan dit onderzoek.

Achtergrondinformatie: huidige onderzoeken omtrent temperatuur in drinkwaterleidingen

Het hier beschreven onderzoek geeft de mogelijkheid vast te stellen wat de bepalende factoren zijn voor warme locaties. Datamining is een verzamelnaam voor technieken om nuttige informatie te winnen uit grote sets (verschillende typen) data. Deze technieken worden succesvol ingezet voor heel uiteenlopende toepassingen: van consumentenspecifieke advertenties op basis van klik- en 'post'-gedrag tot het ontdekken van planeten buiten het zonnestelsel uit satellietgegevens. De dataminingstechniek die voor deze studie is toegepast, is in een eerder project succesvol gebleken in het relateren van storingsfrequenties van drinkwaterleidingen van verschillend materiaaltype en het drukregime van pompstations [3].

Hoewel computerberekeningen veel inzicht geven, hebben ze ook beperkingen. In de berekeningen wordt bijvoorbeeld geen rekening gehouden met bodemkoeling of -verwarming door de aanwezigheid van grote drinkwaterleidingen, of met de invloed van naburige leidingen. KWR werkt op dit moment aan de aanscherping van de modellen op deze punten. Ook zijn de mens aangebrachte warmtebronnen vaak niet goed in kaart gebracht, terwijl deze plaatselijk wel de temperatuur verhogen.

De temperatuur bij pompstations is goed bekend uit online temperatuurmetingen, maar in het leidingnet ontbreken zulke metingen meestal. Onderweg verandert de temperatuur van het water door warmte-uitwisseling met de bodem. Hierdoor is de watertemperatuur bij de klant niet exact bekend. Eerder ontwikkelde KWR computermodellen voor het berekenen van de watertemperatuur in ondergrondse leidingen [4], [5]. Die berekeningen gaan uit van warmteoverdracht en houden rekening met de verschillende invloedsfactoren:

- *weersomstandigheden*, zoals luchttemperatuur en hoeveelheid zonne-instraling,
- *oppervlaktegebruik*, zoals verhard of onverhard, schaduwrijk of schaduwarm, mate van groenvoorziening, wateroppervlak en mate van verdamping,
- *bodemeigenschappen*: zoals grondsoort en door de mens gecreëerde warmtebronnen, zoals elektrische verdeelstations, waternetwerken en hoogspanningskabels.

De modelberekeningen laten zien dat het water in distributieleidingen meestal binnen enkele uren de bodemtemperatuur aanneemt [4]. De temperatuur bij het pompstation wordt daarbij overschreven, meestal voordat het water de klant bereikt.

Het effect van een veranderende omgeving is van belang voor de drinkwatertemperatuur [6], [7]. Door verstedelijking kan de oppervlaktetemperatuur een aantal graden toenemen ten opzichte van landelijke gebieden. Dit effect staat bekend als het stedelijk hitte-eilandeffect ('urban heat island effect').

Metingen voor Nederland zijn te vinden in de *Klimaat-effectatlas* [1]. In de zomer kan hierdoor het drinkwater in stedelijke gebieden enkele graden opwarmen ten opzichte van het platteland [8]. De Nederlandse situatie is geen uitzondering: temperatuurverschillen tussen metingen (van de bodem rond waterleidingen op 45 locaties) in Rotterdam in dezelfde periode kunnen oplopen tot 10°C.

Omstandigheden die bijdragen zijn: (i) locaties in de zon, (ii) locaties onder tegeloppervlak, (iii) de nabijheid van minstens één ondergrondse warmtebron [5]. Naast verstedelijking draagt ook klimaatverandering bij aan risico's op te hoge drinkwatertemperaturen. Zo volgt uit berekeningen voor het meest extreme 'W+'-toekomstscenario van het KNMI dat het aantal overschrijdingen op hotspotlocaties per jaar stijgt van 9 dagen voor de huidige situatie naar 55 dagen in 2030 en 83 dagen in 2050 [5].

Referenties

1. Klimateffectatlas. <http://www.klimateffectatlas.nl>.
2. Publieke Dienstverlening Op de Kaart. <https://www.pdok.nl/>.
3. Vries, D., Vonk, E., Erp, J. van, Diemel, R., Jong, W. de (2016). Waterdatamining – de eerste ervaringen uit de drinkwaterpraktijk. *H2O-Online*, 1 april 2016. <https://www.h2owaternetwerk.nl/vakartikelen/512-waterdatamining-de-eerste-ervaringen-uit-de-drinkwaterpraktijk>.
- 4 .Blokker, E. J. M. & Pieterse-Quirijns, E. J. (2013). *Modeling temperature in the drinking water distribution system*. Journal – American Water Works Association, 105(1), E19-E29.
5. Agudelo-Vera, C. M. (2015). *Drinking water temperature in future urban areas*. KWR Watercycle Research Institute. BTO 2015.012.
6. Ondergrondse hittestress uitdaging voor watersector (2017). *H2O Nieuws*, 22 mei 2017. <https://www.h2owaternetwerk.nl/h2o-nieuws/1127-ondergrondse-hittestress-uitdaging-voor-watersector>.
7. Richtlijnen nodig om te warm drinkwater te voorkomen (2018). *H2O Nieuws*, 19 april 2018. <https://www.h2owaternetwerk.nl/h2o-nieuws/1882-richtlijnen-nodig-om-te-warm-drinkwater-te-voorkomen>.
8. Stewart, I. D. & Oke, T. R. (2012). Local climate zones for urban temperature studies. *Bulletin of the American Meteorological society*, 93(12), 1879-1900.