

Van tellen naar voorspellen - Sturen op risico's met een voorspellend wiskundig model op basis van historische brandweerdeata

Martine School, Maurits de Graaf, Marie-Colette van Lieshout, Emiel Sanders & Ron de Wit

Dit artikel presenteert een wiskundig model voor het voorspellen van schoorsteenbranden. Het model is opgesteld met behulp van gegevens over de jaren 2004 tot en met 2015, beschikbaar gesteld door de veiligheidsregio Twente.

Als verklarende variabelen gebruiken wij het aantal inwoners per standardeenheid van 500 bij 500 meter, de dagelijkse gemiddelde temperatuur en een indicatorvariabele voor de maand oktober. Hieraan voegen wij ruis toe om ook latente variabelen zoals menselijk gedrag te modelleren. Het model is gevalideerd aan de hand van de waarnemingen in 2016 en 2017. Op basis van dit model is een dashboard geïmplementeerd waarmee via een publiek toegankelijke webpagina een kansenrapport voor schoorsteenbranden in Twente gegenereerd wordt, op basis van de weersvoorspelling voor de komende zes dagen. Daarmee kan het dashboard gebruikt worden voor preventiedoeleinden en planning.

Inleiding

Elke dag staan de brandweerkorpsen in Nederland klaar om branden te blussen en hulp te verlenen. Op ieder moment van de dag kunnen ze door de meldkamer gealarmeerd worden, waarna de brandweermannen en -vrouwen naar de auto's snellen om binnen een paar minuten op locatie te zijn. Hier verhelpen ze veel branden en ongemakken. Het tijdstip van een melding, de aard van de melding en de hoeveelheid meldingen per dag is voor elk korps weer afwachten. Het vooraf inzicht krijgen in de te verwachten aantallen en typen meldingen biedt mogelijkheden om efficiënter voor te lichten, te plannen en te sturen op risico's.

Sinds 2010 ontwikkelt de Nederlandse brandweer een nieuw bedrijfsmodel met drie speerpunten: een veiligere maatschappij, een financieel beheersbare organisatie en meer sturen op risico's. Dit model wordt beschreven in 'De Brandweer over morgen' (Brandweer Nederland, 2010). Hierin komt de wens om datagestueerd op te treden sterk naar voren. Data kunnen zowel voor repressie als preventie worden gebruikt. Repressie behelst bijvoorbeeld het voorspellen van het aantal meldingen om door capaciteitsplanning beter te kunnen inspelen op branden. Preventie gaat juist om het creëren van veiligheidsbewustzijn, bijvoorbeeld het gericht geven van voorlichting, in de hoop daardoor meldingen te voorkomen. Sinds 2004 verzamelt Brandweer Twente gegevens (bijv. locatie, tijd en type) van iedere brandmelding. Inmiddels is dit uitgegroeid tot een grote hoeveelheid data. Brandweer Twente en de Toegepaste Wiskunde-opleiding van de Universiteit Twente werken al geruime tijd samen vanuit de ambitie om brandmeldingen te gaan voorspellen.

Voorafgaand werk

Het gebruik van wiskundige modellen voor het analyseren en voorspellen van patronen van bosbranden in ruimte en tijd heeft een lange geschiedenis. Wij beperken ons hier tot enkele recente illustratieve voorbeelden. Zo combineren Peng et al. (2005) een door de brandweer van de regio Los Angeles ontwikkelde brandindex met historische data om het risico op een bosbrand te voorspellen. In Møller en Dias-Ávalos (2010) worden verklarende variabelen met betrekking tot topografie, begroeiing en weer in combinatie met een stochastisch effect gebruikt om de kans op een brand te modelleren. Het stochastisch effect dient ertoe om ruis en niet beschikbare, zogenoemde latente, variabelen mee te kunnen nemen. Ten slotte stellen Serra et al. (2014) voor om aan soortgelijke verklarende variabelen gecorreleerde Gaussische ruis als stochastisch effect toe te voegen. Moderne numerieke instrumenten (Lindgren et al., 2011) worden ingezet om grote hoeveelheden data te kunnen analyseren.

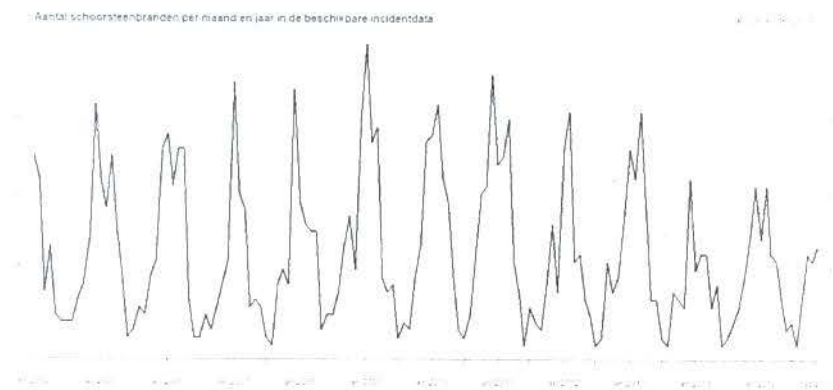
Er is veel minder literatuur beschikbaar over het analyseren van branden in stedelijk gebied. Bij Brandweer Nieuw Zeeland is vooral naar oorzaken van branden gezocht. De slag naar een voorspellend model is nog niet gemaakt (Holmes et al., 2009).

In 2013 is door de Universiteit Twente en Brandweer Twente op basis van een beperkte hoeveelheid data onderzoek gedaan naar het voorspellen van autobranden (Borchers et al., 2013). Dit werd gevolgd door een uitgebreider onderzoek voor brandincidenten in het algemeen (Wendels, 2017). Een belangrijke conclusie van dit onderzoek was dat een voorspelmodel het meest betrouwbaar lijkt als het gericht wordt op één specifiek type branden gerelateerd aan een beperkt aantal verklarende variabelen. In dit kader bespreken we hieronder voor een breed publiek de hoofdlijnen van het onderzoek van School (2018) gericht op schoorsteenbranden. Dit onderzoek heeft een sterk verkennend karakter. Wij geven daarom ook aanbevelingen voor vervolgonderzoek.

Schoorsteenbranden

In het vervolg van dit artikel richten wij ons op schoorsteenbranden. De ervaring van de afdeling brandveiligheid is dat er duidelijke omstandigheden zijn waaronder de open haard of kachel aangestoken wordt. Daarnaast zijn mensen zich bewust van deze actie, dat maakt dat er een aantal aanknopingspunten zijn voor risicocommunicatie: laat in het najaar als je je haard weer gaat gebruiken de schoorsteen vegen (Brandweer Nederland, 2020). Ook is een schoorsteenbrand het meest voorkomende type woningbrand. Een goed aangepast model biedt de mogelijkheid tot gerichte voorlichting op het juiste moment, kan helpen om de effecten van voorlichting te kwantificeren en kan gebruikt worden voor bijvoorbeeld personeelsplanning.

Figuur 1 *Het aantal schoorsteenbranden per maand over de jaren 2004-2015 op basis van data beschikbaar gesteld door Brandweer Twente*



Schoorsteenbranden zijn seizoensafhankelijk. Figuur 1 toont het aantal schoorsteenbranden per maand over de jaren 2004-2015 op basis van data beschikbaar gesteld door Brandweer Twente. Er is een piek te zien in de wintermaanden. Figuur 2 toont waar in Twente schoorsteenbranden zijn voorgekomen in de periode 2004-2015. De steden Enschede, Hengelo en Almelo zijn duidelijk herkenbaar.

Figuur 2 De locaties waar een schoorsteenbrand is ontstaan in de jaren 2004-2015 op basis van data beschikbaar gesteld door Brandweer Twente



Voorspellingsmethode

Het ontwikkelen van een statistisch model is vaak een iteratief proces zoals weer-gegeven in het stappenplan van figuur 3.

Figuur 3 Het blokschema ter verduidelijking van dit onderzoek



In de eerste stap worden data verzameld. In een tweede stap wordt de statistische samenhang van een groot aantal potentiële verklarende variabelen met de waargenomen schoorsteenbranden onderzocht. Na selectie van de belangrijkste verklarende variabelen wordt een statistisch model gemaakt. Vervolgens wordt dit model gevalideerd en zo nodig aangepast, bijvoorbeeld door meer verklarende variabelen toe te voegen of door de parameters of vorm van het statistisch model te wijzigen.

Na een aantal iteraties is de definitieve verzameling verklarende variabelen duidelijk en bepalen we het uiteindelijke model. Ten slotte wordt dit model geïmplementeerd in een softwareprogramma dat geschikt is voor operationeel gebruik. We lichten deze stappen een voor een toe.

Data

Zoals gezegd is Brandweer Twente sinds 2004 bezig met het verzamelen van informatie over meldingen. Dit gegevensbestand bevat informatie over alle incidenten vanaf 2004, zoals het type incident, de datum en tijd en de plaats van het incident in het RD-coördinatenstelsel. Daarnaast voegt de meldkamer aan elk incident verschillende omschrijvingen en karakteristieken toe. Ons onderzoek is gebaseerd op de gegevens van schoorsteenbranden van 2004 tot en met 2015. Dit zijn branden waarbij het woord 'schoorsteen' voorkomt in de omschrijving. De gegevens uit 2016 en 2017 gebruiken we om het model te valideren. De data uit de periode 2004 tot en met 2015 noemen we dan ook 'testdata', de gegevens uit 2016 en 2017 'validatiedata'.

De gegevens over de branden worden gecombineerd met ruimtelijke informatie van het Kadaster, het Centraal Bureau voor de Statistiek en het BAG-register, zoals het aantal inwoners opgesplitst naar leeftijdscategorie en het aantal woningen, scholen, bedrijfspanden, enzovoort per standaardvak van 500 bij 500 meter. Twente bestaat uit meer dan 6000 van deze vakken. Het zou voor de hand liggen ook het aantal schoorstenen per vak mee te nemen, maar helaas is deze informatie niet beschikbaar. Naast ruimtelijke informatie maken we ook gebruik van diverse waarnemingen van het weerstation 'Twenthe' van het KNMI, zoals de gemiddelde temperatuur gedurende een etmaal, de hoeveelheid neerslag en de aanwezigheid van mist.

Verklarende variabelen

De volgende stap in het blokschema van figuur 3 is het identificeren van die ruimtelijke informatie en weergegevens die de sterkste relatie vertonen met de waargenomen aantallen schoorsteenbranden. Als maat hiervoor gebruiken we de Pearson correlatiecoëfficiënt. Een waarde van nul betekent dat de betreffende variabele en de aantallen schoorsteenbranden onafhankelijk zijn, positieve (negatieve) waarden betekenen dat grote aantallen schoorsteenbranden gepaard gaan met grote (kleine) waarden van de variabele. Op grond van praktijkervaring van Brandweer Twente is een voorselectie gemaakt.

Ten aanzien van het type woning wordt verwacht dat schoorsteenbranden vaker ontstaan in vrijstaande en twee-onder-een-kapwoningen dan in appartementsgebouwen waar zelden een schoorsteen aanwezig is. Naast het type van een woning zou ook het bouwjaar een factor kunnen zijn in het ontstaan van schoorsteenbranden. Door de aanwezigheid van een verouderde versie van de schoorsteenpijp, bijvoorbeeld een gemetselde variant, in oudere woningen lijkt een schoorsteenbrand daar waarschijnlijker dan in moderne woningen.

Tabel 1 De tien variabelen met de hoogste absolute Pearson correlatie-coëfficiënt

Variable	Ruimtelijk/Tijd	Correlatiecoëfficiënt
Dagelijkse gemiddelde temperatuur	Tijd	-0,3769
Dagelijkse gemiddelde gevoelstemperatuur	Tijd	-0,3758
Het aantal inwoners per vak	Ruimtelijk	0,2727
Het aantal 65+ inwoners per vak	Ruimtelijk	0,2706
Het aantal vrouwelijke inwoners per vak	Ruimtelijk	0,2685
Het aantal woningen per vak	Ruimtelijk	0,2669
Het aantal vrijstaande en twee onder een kap huizen per vak	Ruimtelijk	0,2664
Het aantal gebouwen per vak	Ruimtelijk	0,2637
Het aantal inwoners tussen de 45 en 64 jaar per vak	Ruimtelijk	0,2539
Booleaanse variabele die aangeeft of het winter is of niet	Tijd	0,2509

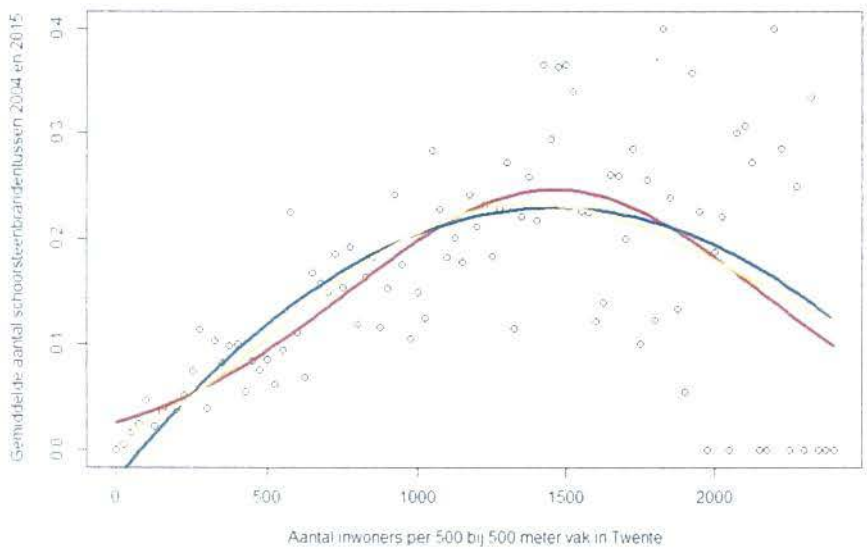
Daarnaast zouden sommige weerscondities aanleiding kunnen geven tot een hogere kans op schoorsteenbrand, bijvoorbeeld een lage (gevoels)temperatuur, eventueel aangevuld met de aanwezigheid van mist. Bij kou worden immers thuis de open haarden aangestoken en mist – zo is althans de ervaring van de brandweer – zorgt voor minder trek in de schoorsteenpijp.

De tien variabelen (ruimtelijk en tijdsgerelateerd) met de hoogste absolute waarde van de Pearson correlatiecoëfficiënt zijn weergegeven in tabel 1. Bovenaan staan twee variabelen met een sterk negatieve correlatie, de gemiddelde temperatuur en gemiddelde gevoelstemperatuur. De coëfficiënten zijn negatief, omdat over het algemeen het aantal schoorsteenbranden afneemt als de temperatuur stijgt. Na de temperatuurgerelateerde variabelen volgen andere variabelen waarvan de correlatiecoëfficiënten dichtbij elkaar liggen. Ter wille van hanteerbaarheid en interpretatie hebben we gekozen voor één ruimtelijke en één tijdsgerelateerde variabele met de hoogste correlatiecoëfficiënt. Dit zijn de *dagelijkse gemiddelde temperatuur* en het *aantal inwoners per vak*.

Aanzet tot een model

Op basis van de geselecteerde variabelen construeren we allereerst een naïef model waarmee we testvoorspellingen gaan doen. We bepalen per variabele hoe de invloed daarvan er uit zou kunnen zien door families van functies aan te passen aan de testdata. Daartoe zet figuur 4 het gemiddelde aantal schoorsteenbranden in alle 500 bij 500 meter vakken met een gegeven aantal inwoners af tegen het aantal inwoners van zo'n vak. De verschillende kleuren representeren verschillende wiskundige families van functies: de parabool (blauw), de exponentiële functie (rood) en de dubbel-exponentiële functie (geel).

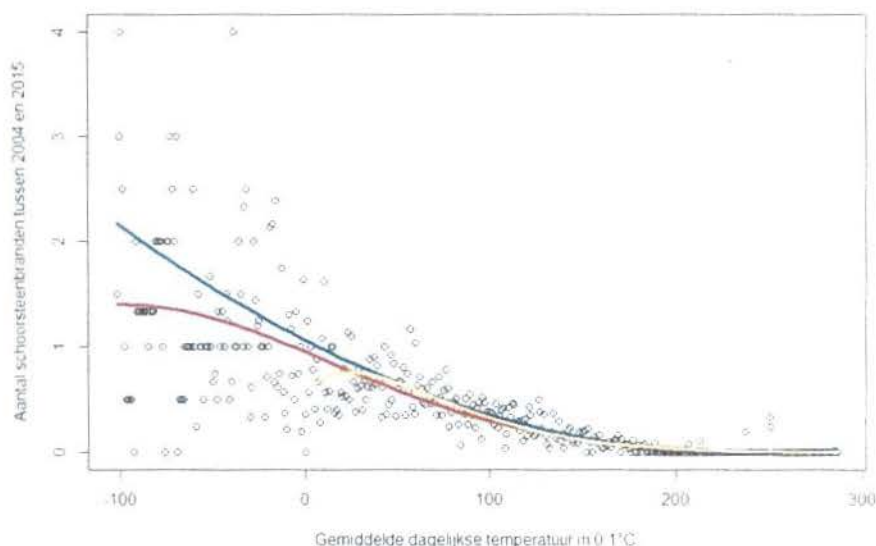
Figuur 4 *Het gemiddeld aantal schoorsteenbranden gedurende de periode 2004-2015 per vak tegenover het aantal inwoners per 500 bij 500 meter vak. De best passende functies uit drie families worden getoond: de parabool (blauw), de exponentiële (rood) en de dubbel-exponentiële (geel) functie*



In figuur 4 is een piek in de functie te zien rond 1400 inwoners. Bij meer of minder inwoners is de kans op schoorsteenbranden lager. Dit is mogelijk te verklaren doordat bij een hogere adressendichtheid meer appartementen in een vak opgenomen zijn. Zoals al eerder genoemd, hebben appartementen zelden een schoorsteen en daardoor is een schoorsteenbrand ook uitgesloten. Bij een lagere adressendichtheid zijn er eenvoudigweg minder huizen.

In figuur 5 is het gemiddelde aantal schoorsteenbranden in Twente afgezet tegen de temperatuur met de best passende parabool (blauw), exponentiële functie (rood) en dubbel-exponentiële functie (geel).

Figuur 5 *Het gemiddeld aantal schoorsteenbranden in Twente per dag tegenover de dagelijkse gemiddelde temperatuur. De best passende functies uit drie families worden getoond: de parabool (blauw), de exponentiële (rood) en de dubbele exponentiële (geel) functie*



In figuur 5 is te zien dat het aantal schoorsteenbranden afneemt naarmate de temperatuur toeneemt, waarbij vanaf ongeveer 20 graden zelden een schoorsteenbrand verwacht wordt.

Voor beide variabelen hebben we de best passende functie gekozen. Voor het aantal inwoners is dat de exponentiële functie, terwijl voor de gemiddelde temperatuur de parabool de beste functie is. Ons eerste model maakt nu de aanname dat het aantal schoorsteenbranden op zekere dag in een gegeven vak Poisson verdeeld is met het product van deze twee functies als gemiddelde. Bovendien nemen we aan dat verschillende vakken en dagen elkaar niet beïnvloeden. Zo'n model heet een *inhomogeneous Poissonproces* (Diggle, 1983; Lieshout, 2019; Wendels, 2017).

Validatie

Het zojuist beschreven model bepaalt op basis van de gemiddelde temperatuur op een dag en het aantal inwoners in een vak een kansdichtheidsfunctie. Deze voorspelt voor ieder vak en voor iedere dag, met welke kans een gegeven aantal branden kan optreden. We valideren het model met behulp van de validatiedata, door voor heel Twente het gemiddelde aantal schoorsteenbranden per maand te bepalen. De resultaten hiervan zijn te zien in tabel 2 voor 2016. Het jaar 2017 laat een soortgelijk beeld zien.

Tabel 2 *De voorspelling op basis van het aantal inwoners en de gemiddelde temperatuur met daaronder het aantal waargenomen schoorsteenbranden in 2016*

Maand	J	F	M	A	M	J	J	A	S	O	N	D	Totaal
Voorspelling	18	15	15	10	4	2	2	2	2	9	15	16	110
Realiteit	17	13	16	14	6	1	4	1	5	18	14	16	125

Om het aangepaste model te beoordelen, kijken we naar de som van de kwadraten van de residuen. Residuen zijn de verschillen tussen de modelvoorspelling en de waarneming. Deze zijn over het algemeen klein. Het valt echter op dat in april en oktober (dik gedrukt) het model veel minder branden verwacht dan er in werkelijkheid voorkomen. Het blijkt dat door de functie die de verwachting in ons Poisson-model weergeeft in oktober met een constante factor te vermenigvuldigen, de modelvoorspellingen verbeteren. (Het toevoegen van een soortgelijke variabele voor april zorgt voor een verslechtering van de voorspelling.) Aldus krijgen we een nieuw model gebaseerd op drie variabelen: het aantal inwoners per vak, de dagelijkse gemiddelde temperatuur en een indicatorvariabele die aangeeft of de huidige dag in oktober valt. Validatie van dit model leidt tot tabel 3.

Tabel 3 *De voorspelling op basis van het aantal inwoners, de gemiddelde temperatuur en een indicatorvariabele die aangeeft of het oktober is met daaronder het aantal waargenomen schoorsteenbranden in 2016*

Maand	J	F	M	A	M	J	J	A	S	O	N	D	Totaal
Voorspelling	18	15	15	10	4	2	1	2	2	16	15	16	116
Realiteit	17	13	16	14	6	1	4	1	5	18	14	16	125

De voorspellingen zijn voor een eenvoudig model verrassend goed te noemen. Toch zijn er wel enkele kanttekeningen te plaatsen. Ten eerste zijn de functies die de invloed van de verschillende verklarende variabelen modelleren op basis van gemiddelden bepaald en eenvoudigweg met elkaar vermenigvuldigd. Bovendien hebben we per variabele, onafhankelijk van de anderen, bekeken wat het effect op het aantal schoorsteenbranden is. Het is echter denkbaar dat de verklarende variabelen zowel een afzonderlijk als een gezamenlijk effect hebben, wat kan worden beschreven door een zogenoemd gegeneraliseerd regressiemodel in ruimte en tijd. De hiervoor benodigde theorie wordt uitputtend beschreven in Kutoyants (1998), de numerieke implementatie in Baddeley et al. (2015). Ten slotte zijn veel verklarende variabelen die volgens tabel 1 correleren met het optreden van schoorsteenbranden buiten beschouwing gelaten en is in het Poissonmodel ook geen rekening gehouden met mogelijke afhankelijkheden tussen naburige vakken en opeenvolgende dagen.

Een verbeterd model

Ons eerste model kan worden verbeterd door een stochastisch effect toe te voegen. Omdat naburige vakken en opeenvolgende dagen doorgaans meer op elkaar lijken dan vakken of dagen die ver van elkaar verwijderd zijn, voegen wij normaal verdeelde gecorreleerde ruis toe (een 'Gaussisch stochastisch veld'). Op deze manier kunnen dus afhankelijkheden tussen naburige vakken en dagen worden meegenomen. Een stochastisch effect compenseert ook voor toevallige ('latente') factoren, bijvoorbeeld variaties in menselijk gedrag. In tegenstelling tot de verklarende variabelen heeft de ruis geen invloed op het verwachte aantal schoorsteenbranden, wel op hun onderlinge samenhang. We komen aldus tot een uitbreiding van het naïeve inhomogene Poissonproces, een zogenoemd 'log-Gaussisch Coxproces' (Coles & Jones, 1991). Voor het schatten van de modelparameters maken we gebruik van de aanpak in (Møller et al., 1998; Brix & Diggle, 2001; Gabriel & Diggle, 2009). Een meer geavanceerd alternatief wordt beschreven in Serra et al. (2014) en Lindgren et al. (2011). Voor de details verwijzen we naar School (2018).

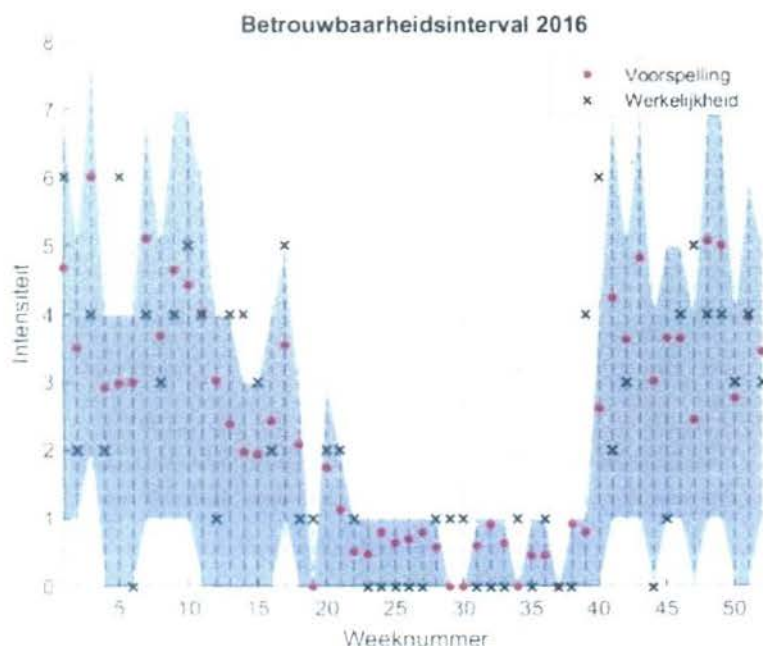
Met deze toevoeging is het model compleet en kunnen opnieuw voorspellingen gedaan worden voor 2016, zoals te zien in tabel 4. De voorspellingen voor 2017 geven een vergelijkbaar beeld.

Tabel 4 *Voorspelling van het gemiddeld aantal schoorsteenbranden in Twente op basis van het aantal inwoners, de gemiddelde temperatuur, een indicatorvariabele die aangeeft of het oktober is en de ruis met daaronder het aantal waargenomen schoorsteenbranden in 2016*

Maand	J	F	M	A	M	J	J	A	S	O	N	D	Totaal
Voorspelling	18	15	16	11	5	2	2	3	2	17	15	17	123
Realiteit	17	13	16	14	6	1	4	1	5	18	14	16	125

Wat direct opvalt, is dat het totaal aantal voorspelde branden dichterbij de realiteit ligt en dat oktober nog beter voorspeld wordt. Om informatie te kunnen geven over de nauwkeurigheid van de voorspelling tonen we in figuur 6 benaderde 80%-betrouwbaarheidsintervallen.

Figuur 6 De grijze 80%-betrouwbaarheidsintervallen behorend bij de voorspellingen van 2016 met daarbij de absolute voorspelling en realiteit in een respectievelijk rode stip en zwart kruis



De betekenis hiervan is dat bij het toepassen van de voorspellingsmethode op verschillende jaren (met exact dezelfde gemiddelde temperatuur en bevolkingsaantallen) voor 80% van de jaren het werkelijke aantal branden binnen het berekende betrouwbaarheidsinterval zal liggen. Voor details verwijzen we weer naar School (2018).

Uit figuur 6 blijkt dat de realiteit meestal in het betrouwbaarheidsinterval valt en daarbij vaak heel dichtbij de rode stip van de voorspelde waarde komt. Voor de jaren 2016 en 2017 viel 86% van de wekelijkse voorspellingen in het betrouwbaarheidsinterval.

Implementatie

Het uiteindelijke model is geïmplementeerd in een 'dashboard' gebaseerd op Microsoft Power BI met onderliggende scripts in R. Dit programma wordt inmiddels al enige tijd gebruikt bij Brandweer Twente en biedt een eenvoudig gebruikersinterface. In het dashboard kan de weersvoorspelling voor de komende zes dagen ingevoerd worden op basis waarvan dan een kansenrapport voor Twente wordt gegenereerd (zie figuur 7). Als visualisaties worden drie plaatjes weergegeven: (1) het verwachte aantal schoorsteenbranden; (2) het kansenverloop per dag; (3) de wijken met de grootste kans (zie figuur 8).

Figuur 7 De invoerpagina van het voorspellende dashboard voor schoorsteenbranden. De dag van het jaar en de temperatuur van de komende zes dagen kunnen hier ingevoerd worden

BRANDWEER TWENTE **DE WERKLIJKHEID** **SCHOORSTEENBRANDEN VERWACHTING** **UNIVERSITY OF TWENTE**

De dag van vandaag.
Deze staat in principe altijd op de huidige dag.

Vandaag: 15-8-2018

De temperatuur de komende zes dagen
Vul hier de verwachting in minimum en maximum temperatuur voor de komende zes dagen in.

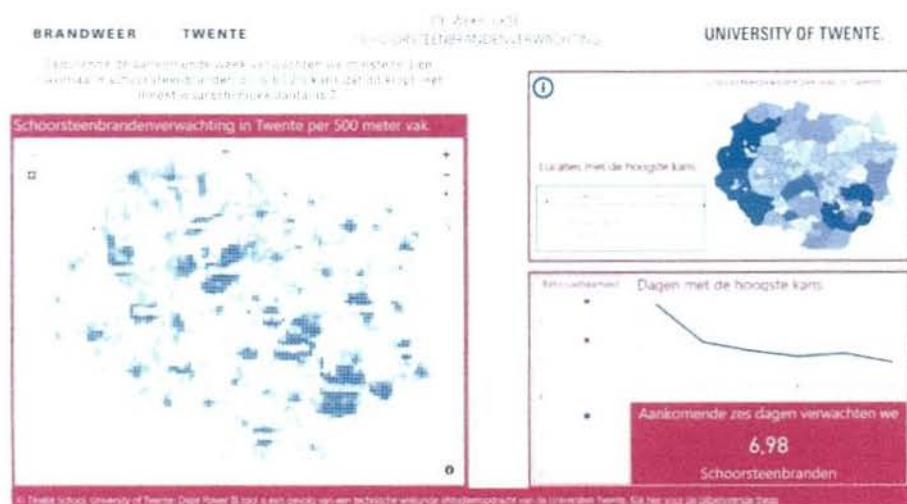
15-8-18	16-8-18	17-8-18	18-8-18	19-8-18	20-8-18
Minimum: 24.00	Minimum: 18.00	Minimum: 12.00	Minimum: 12.00	Minimum: 12.00	Minimum: 12.00
Maximum: 17.00	Maximum: 12.00	Maximum: 12.00	Maximum: 12.00	Maximum: 12.00	Maximum: 12.00

© Twente School, University of Twente. Deze Pagina is een product van een technische wiskunde afvalverwerking van de Universiteit Twente. Het is een product van de Universiteit Twente.

Dit dashboard is bedoeld om de mogelijkheden van de wiskunde in de organisatie zichtbaar te maken en daarmee anderen op de verborgen potentie van de aanwezige data te attenderen. Het levert een frisse kijk op de huidige manier van werken voor de brandweer zelf, maar ook voor de burgers die nu het verwachte aantal schoorsteenbranden in de eigen wijk kunnen zien. Daarmee kan het dashboard gebruikt worden voor preventiedoeleinden en planning: wijkgerichte 'just in time' voorlichting en het beschikbaar houden van voldoende personeel in regio's en weken waar een grote kans is op schoorsteenbranden.

In het najaar van 2018 is het dashboard in de praktijk getest. De resultaten waren dermate bemoedigend, dat NWO een beurs beschikbaar heeft gesteld om het door te ontwikkelen in het project 'Data driven risk management for fire services'.

Figuur 8 Het kansenrapport van het voorspellende dashboard voor schoorsteenbranden



Conclusie

Dit artikel presenteert een wiskundig onderbouwde methode voor het analyseren en voorspellen van schoorsteenbranden. Gebruikmakend van gegevens die beschikbaar zijn gesteld door Brandweer Twente over de jaren 2004 tot en met 2015 is een model opgesteld om schoorsteenbranden te voorspellen. Voor de jaren 2016 en 2017, die als validatiedata zijn gebruikt, viel 86% van de wekelijkse voorspellingen in het 80%-betrouwbaarheidsinterval. Dit betekent dat het model goed presteert. Verklarende variabelen zijn het aantal inwoners per vak in Twente, de dagelijkse gemiddelde temperatuur en een indicatorvariabele voor de maand oktober. Hieraan is gecorreleerde normaal verdeelde ruis toegevoegd om ook latente variabelen en afhankelijkheden in ruimte en tijd te modelleren. Op basis van dit model is een dashboard geïmplementeerd waarmee gedurende 2018-2019 via een publiek toegankelijke webpagina de weersvoorspelling voor de komende zes dagen ingevoerd kan worden. Het dashboard genereert dan een kansenrapport voor schoorsteenbranden per wijk in Twente. Daarmee kan het dashboard gebruikt worden voor preventiedoeleinden en planning.

De gekozen procedure biedt een leidraad voor de ontwikkeling van modellen voor andere typen incidenten waar de brandweer bij betrokken is. In de inleiding verwezen we al naar eerder werk met betrekking tot bosbranden. Men kan ook denken aan verkeersongevallen. Deze lijken sterk weersafhankelijk. Daarnaast is de ruimtelijke indeling van Twente mogelijk van belang. Zo kan bijvoorbeeld de aanwezigheid van een snelweg het risico vergroten, maar is het ook een ervaringsfeit dat in het buitengebied vaak te snel wordt gereden met ongevallen tot gevolg.

De in dit artikel beschreven aanpak heeft ten slotte ook potentie voor andere onderdelen binnen het veiligheidsthema, zoals het analyseren van woninginbraakda-

ta of het voorspellen van schade aan woningen na storm. Ondanks het verkennen- de karakter van dit onderzoek, zijn we ervan overtuigd dat de beschreven methode breed inzetbaar is binnen het veiligheidsdomein.

Verder onderzoek

Tijdens dit verkennende onderzoek is een aantal aspecten geïdentificeerd die verbeterd en uitgebreid zouden kunnen worden.

Allereerst valt op dat slechts één van de oorspronkelijk belangrijk geachte verklarende variabelen (type en bouwjaar woning, mist) in tabel 1 terug te zien is, namelijk het aantal vrijstaande en twee-onder-een-kapwoningen. Verklaringen hiervoor zouden kunnen zijn dat mist relatief weinig en zeer lokaal voorkomt of dat de opdeling in 500 bij 500 meter vakken te grof is om de invloed van type woning en bouwjaar zichtbaar te maken.

Het zou interessant zijn om deze indeling te verfijnen naar bijvoorbeeld buurten met dezelfde gebiedskenmerken, zodat gericht kan worden getoetst op specifieke kenmerken, maar deze gegevens zijn momenteel niet beschikbaar.

Wij hebben ervoor gekozen om slechts de twee verklarende variabelen met de hoogste correlatie met de data mee te nemen in onze modellen. De verschillen tussen de verklarende variabelen in de top tien zijn echter maar klein. Men zou dus moeten toetsen of toevoegen van meer verklarende variabelen leidt tot een statistische verbetering van het model. Merk op dat vanwege de nesting tussen verschillende van deze verklarende variabelen, de modellen dan wel complexer worden en er aanmerkelijk meer parameters moeten worden geschat.

Ten derde hebben we voor figuur 6 benaderde betrouwbaarheidsintervallen gebruikt op basis van simulaties. Een betere onderbouwing voor deze intervallen is zeer gewenst, maar lijkt niet beschikbaar in de literatuur. Wel zijn er enkele simulatiestudies uitgevoerd, bijvoorbeeld door Davies en Hazelton (2013).

Ten vierde wordt een brand gecategoriseerd als schoorsteenbrand als in het meldingsoverzicht van de meldkamer het woord 'schoorsteen' voorkomt. Dit kan betekenen dat sommige geclassificeerde schoorsteenbranden geen daadwerkelijke schoorsteenbrand zijn, bijvoorbeeld als genoemd is dat de open haard nog aanstond, maar dat de brand in de gordijnen is begonnen. Een verbeterde classificering van de schoorsteenbranden zou daarom ook helpen om het model preciezer te maken.

Ten slotte zou het ook interessant zijn te kijken naar multivariate modellen, dat wil zeggen, naar modellen voor verscheidene typen branden tegelijkertijd.

Literatuur

- Baddeley, A., E. Rubak & R. Turner (2015) *Spatial point patterns: Methodology and applications with R*. Boca Raton: Chapman & Hall/CRC.
- Borchers, M., T. Gellenbeck, G. Maduro, M. Visser & S. Visser (2013) *Op zoek naar het operationeel centrum van autobrandstichters in Twente* (B.Sc thesis). Enschede: University of Twente.

- Brandweer Nederland (2010) *De brandweer over morgen*. Enschede: Projectteam Strategische Reis Brandweer.
- Brandweer Nederland (2020) *Voorkom Schoorsteenbrand*. www.brandweer.nl/brandveiligheid/kachel-en-schoorsteen.
- Brix, A. & P. Diggle (2001) Spatio-temporal prediction for log-Gaussian Cox processes. *Royal Statistics Society: Series B (Statistical Methodology)*, 63, 823-841.
- Coles, P. & B. Jones (1991) A lognormal model for the cosmological mass distribution. *Monthly Notices of the Royal Astronomical Society*, 248(1), 1-13.
- Davies, T.M. & M.L. Hazelton (2013) Assessing minimum contrast parameter estimation for spatial and spatiotemporal log-Gaussian Cox processes. *Statistica Neerlandica*, 67(4), 355-389.
- Diggle, P. (2014) *Statistical analysis of spatial and spatio-temporal point processes* (3rd ed.). Boca Raton: Chapman and Hall / CRC Press.
- Diggle, P., P. Moraga, B. Rowlingson & B. Taylor (2013) Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. *Statistical Science*, 28, 542-563.
- Gabriel, E. & P. Diggle (2009) Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica*, 63, 43-51.
- Holmes, M., Y. Wang & I. Ziedins (2009) *The application of data mining and statistical techniques to identify patterns and changes in fire events*. Auckland: University of Auckland.
- Kutovants, Y.A. (1998) *Statistical inference for spatial Poisson processes*. New York: Springer.
- Lieshout, M.N.M. van (2019) *Theory of spatial statistics: A concise introduction*. Boca Raton: Chapman and Hall/CRC Press.
- Lindgren, F., H. Rue & J. Lindstrom (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach (with discussion). *Journal of the Royal Statistical Society, Series B*, 73(4), 423-498.
- Møller, J. & C. Dias-Ávalos (2010) Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires. *Scandinavian Journal of Statistics*, 37(1), 2-25.
- Møller, J., A. Syversveen & R. Waagepetersen (1998) Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25, 451-482.
- Peng, R.D., F.P. Schoenberg & J.A. Woods (2005) A Space-Time conditional intensity model for evaluating a wildfire hazard index. *Journal of the American Statistical Association*, 100(469), 26-35.
- School, M.L. (2018) *A log Gaussian Cox process for predicting chimney fires at fire department Twente* (M.Sc thesis). Enschede: University of Twente.
- Serra, L., M. Saez, J. Mateu, D. Varga, P. Juan, C. Dias-Ávalos & H. Rue (2014) Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: the case of Catalonia, 1994-2008. *Environmental and Ecological Statistics*, 21, 531-563.
- Wendels, M. (2017) *A spatio-temporal point process model for firemen demand in Twente* (B.Sc thesis). Enschede: University of Twente.