



Statistische Trends

# Oversterfte tijdens de corona-epidemie: toepassing van een dynamisch regressiemodel

Trond Husby (PBL)  
Lenny Stoeldraijer (CBS)  
Hans Visser (PBL)

2020

# Inhoud

<b>1. Inleiding</b>	<b>3</b>
<b>2. Belangrijkste uitkomsten</b>	<b>4</b>
<b>3. Data en methode</b>	<b>6</b>
3.1 Gegevens	6
3.2 Model	6
3.3 Gevoeligheidsanalyse	7
<b>4. Resultaten</b>	<b>8</b>
4.1 Uitkomsten model	8
4.2 Vergelijking met andere modellen	10
<b>5. Discussie en conclusie</b>	<b>12</b>
Referenties	13
Bijlage	14

Corona heeft een duidelijke invloed gehad op de sterfte sinds begin maart 2020. Hoe groot de impact precies is, is echter nog niet zeker. In dit artikel wordt een nieuwe methode gepresenteerd om de wekelijkse oversterfte door COVID-19 te schatten. De methode, een zogeheten dynamisch regressiemodel, biedt een consistente inschatting van de onzekerheid rondom de oversterfte en kan de relatie duiden tussen de geregistreerde COVID-19 sterfte en de totale sterfte gerelateerd aan corona. Door de geregistreerde sterfte aan corona als factor mee te nemen in het model, wordt voorkomen dat de schattingen voor oversterfte door COVID-19 'vervuild' worden door sterfte aan andere oorzaken. Een vergelijking met andere modellen laat zien dat de orde van grootte van de oversterfte in alle gevallen vergelijkbaar is, maar dat het dynamische regressiemodel preciezer is wanneer het gaat om kortere tijdsperiodes en wanneer de geregistreerde COVID-19 sterfte sterk toe- of afneemt. De geschatte oversterfte in de eerste elf weken van de corona-epidemie (week 11 tot en met week 21) komt uit op 10 164. Met 95 procent zekerheid lag de oversterfte in die weken tussen 8 593 en 11 691. Dit betekent dat naast elke 10 geregistreerde doden als gevolg van corona er tussen de 5 en 10 mensen overleden aan corona die niet geregistreerd werden.

## 1. Inleiding

Sinds de eerste positief geteste patiënt in Nederland met COVID-19 op 27 februari 2020 en het eerste sterfgeval in Nederland op 6 maart nam de totale sterfte in Nederland wekelijks toe (CBS, 29 mei 2020). Tot en met week 10 (tot en met 8 maart) overleden gemiddeld 3 136 mensen per week. Daarna steeg de sterfte naar een maximum van 5 080 in week 14 (30 maart tot en met 5 april). In week 20 (11 tot en met 17 mei) was de wekelijkse sterfte op een vergelijkbaar niveau als dezelfde periode in eerdere jaren.

Het RIVM rapporteerde dagelijks het aantal gemelde overleden COVID-19-patiënten (RIVM, 2020a). Echter, het werkelijke aantal overleden aan het coronavirus zal hoger zijn dan het aantal meldingen aan het RIVM omdat niet alle overledenen getest zijn op COVID-19. Als alleen naar de overleden COVID-19-patiënten van het RIVM wordt gekeken (5 900 overleden tot en met week 21), zal de impact van corona op de sterfte sterk worden onderschat.

Om een beter beeld te krijgen van de impact van de corona-epidemie, kan een schatting gemaakt worden van de oversterfte, dat is het verschil tussen het waargenomen aantal overledenen en het aantal dat kan worden verwacht als er geen corona-epidemie zou hebben plaatsgevonden. Ook voor andere epidemieën of incidenten wordt een dergelijke aanpak gebruikt (RIVM, 2020b).

Het CBS heeft in samenwerking met Amsterdam Medisch Centrum (AMC) bij de rapportage van de oversterfte tijdens de corona-epidemie een methode gehanteerd waarbij cijfers van het aantal overledenen in de weken voorafgaand aan de corona-epidemie zijn gebruikt en gecorrigeerd is voor seizoengebonden factoren. Er bestaan echter ook andere methoden om de oversterfte te schatten. Vanwege de onzekerheid rondom het schatten van het verwachte aantal overledenen en de wekelijkse fluctuaties in de sterfte, is een vergelijking met andere methoden gewenst.

In dit artikel wordt met een dynamisch regressiemodel de oversterfte tijdens de corona-epidemie geschat. Met deze alternatieve methode is het ook mogelijk om een consistente inschatting te maken van de onzekerheid rondom de oversterfte – voor zowel elke week als voor perioden over meerdere weken. Bovendien geeft de methode duiding van de relatie tussen geregistreerde sterfte aan corona en de totale sterfte.

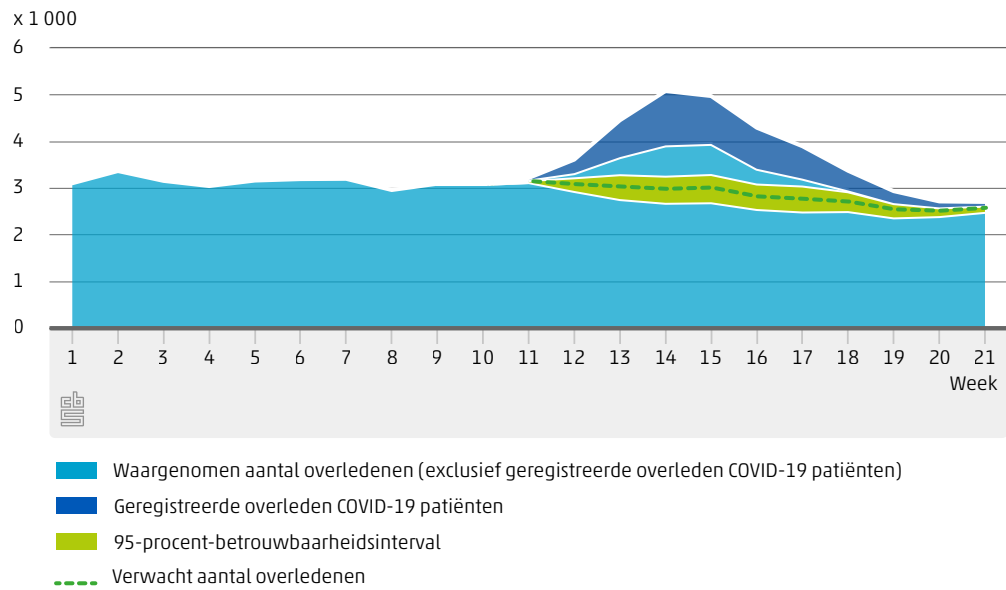
In hoofdstuk 2 worden de belangrijkste resultaten getoond van dit regressiemodel. Hoofdstuk 3 behandelt de data en de methode. Hoofdstuk 4 geeft meer gedetailleerde uitkomsten van de methode en biedt een vergelijking met andere modelberekeningen. In Hoofdstuk 5 volgen de discussie en conclusie.

## 2. Belangrijkste uitkomsten

In de eerste tien weken van 2020 overleden gemiddeld 3 136 mensen per week (Figuur 2.1). Op basis van het dynamische regressiemodel zou de wekelijkse sterfte daarna langzaam gaan dalen als er geen corona-epidemie was geweest (weergegeven door de gestippelde lijn in Figuur 2.1 met het 95-procent-betrouwbaarheidsinterval in lichtgroen). In werkelijkheid steeg de wekelijkse sterfte naar 5 080 in week 14 om daarna te dalen.

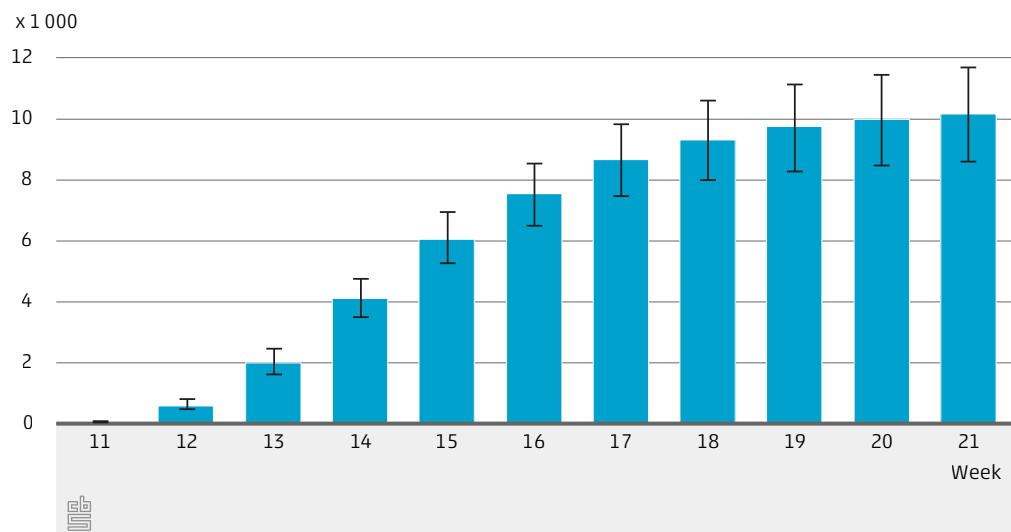
Vanaf week 11 was er sprake van significante oversterfte: het waargenomen aantal overledenen lag hoger dan het 95-procent-betrouwbaarheidsinterval van het verwachte aantal overledenen. Een deel van die oversterfte bestaat uit geregistreerde overleden COVID-19 patiënten (weergegeven met het donkerblauwe gebied in Figuur 2.1). Niet alle overledenen worden getest op corona en de onderregistratie van mensen die aan corona overleden is dan ook duidelijk te zien in Figuur 2.1. Daarnaast is het waarschijnlijk dat de relatie tussen het aantal geregistreerde coronadoden en de oversterfte niet noodzakelijk één-op-één is vanwege het feit dat het aantal overledenen dat is getest op COVID-19 per week kan verschillen.

## 2.1 Waargenomen en verwachte sterfte, week 1 tot en met week 21 van 2020



De schatting van totale oversterfte tijdens de eerste elf weken van de corona-epidemie (week 11 tot en met week 21) komt uit op 10 164 overledenen. Met 95 procent zekerheid ligt de totale oversterfte tijdens de corona-epidemie tussen 8 593 en 11 691 overledenen. De geschatte oversterfte in de weken 11 tot en met 21 bedroeg rond een kwart van het totaal aantal overledenen in die weken.

## 2.2 Cumulatieve oversterfte, week 11 tot en met week 21 van 2020



# 3. Data en methode

## 3.1 Gegevens

Voor de schatting van de oversterfte is gebruikgemaakt van het aantal overledenen onder de bevolking van Nederland, ongeacht de doodsoorzaak, leeftijd en geslacht (CBS Statline, 5 juni 2020). Deze tijdreeks loopt van 1995 tot en met 2020. De cijfers over 2020 hebben een voorlopig karakter. Deze cijfers kunnen bij elke vernieuwing van de publicatie wijzigen doordat later ontvangen overlijdensberichten worden meegeteld. De gegevens die zijn gebruikt voor de analyse bevatten wijzigingen tot en met 5 juni 2020, en de tijdreeks loopt dan tot en met week 21.

Daarnaast worden gegevens over geregistreerde sterfte per overlijdensdatum van het RIVM gebruikt (RIVM 2020a, 8 juni 2020). De tabel wordt dagelijks gewijzigd.

## 3.2 Model

Als uitgangspunt voor de schatting van de oversterfte is gebruikgemaakt van tijdreeksanalyse waarbij de wekelijks waargenomen sterfte onderverdeeld wordt in vier componenten.

De eerste component geeft een trendmatige ontwikkeling in de reeks. Deze beschrijft de langzame veranderingen die zichtbaar zijn in de sterfte gedurende de jaren. Als tweede bevat het model een seizoencomponent. Deze beschrijft de variaties die zichtbaar zijn in de sterfte gedurende het jaar, aangezien in de winter meer mensen overlijden dan in de zomer. Als derde is de geregistreerde coronasterfte toegevoegd als verklarende variabele. Toevoeging van deze component voorkomt dat de schatting voor corona-oversterfte vervuld wordt door sterfte aan andere oorzaken, zoals een griepvirus. De vierde en laatste component bestaat uit statistische ruis, de onregelmatige bewegingen die niet kunnen worden verklaard uit de andere componenten.

Het model dat hier wordt toegepast, is een zogeheten dynamisch lineair model (DLM) en kan als volgt worden opgeschreven:

$$\text{Sterfte}_t = \text{Trend}_t + \text{Seizoen}_t + \beta_t \text{Coronasterfte}_t + \varepsilon_t$$

De coëfficiënt  $\beta_t$  beschrijft de relatie tussen de geregistreerde overledenen aan corona en de totale oversterfte,  $\varepsilon_t$  is de ruis.

De exacte formulering van het dynamische regressiemodel is opgenomen in de bijlage. Voor meer details over dergelijke modellen verwijzen we naar Petris et al. (2009, 2018), Durbin en Koopman (2012) en Laine (2019). In deze literatuur wordt ook uiteengezet hoe de onbekenden in het model, zoals de tijdafhankelijke coëfficiënt  $\beta_t$ , geschat worden met het Kalmanfilter.

Zoals eerder opgemerkt, is het waarschijnlijk dat de relatie tussen de geregistreerde coronasterfte en de oversterfte niet een-op-een is. Om die reden kan de coëfficiënt in het model ook variëren met de tijd. Op deze manier ontstaat er een beeld van de dynamiek van de relatie. De coëfficiënt kan geïnterpreteerd worden als een ophogingsfactor, ofwel het aantal overledenen (oversterfte) per geregistreerde corona-overledene.

Met het model kan de oversterfte gerelateerd aan corona berekend worden door de geschatte ophogingsfactor te vermenigvuldigen met de geregistreerde coronasterfte. Deze berekening is consistent met de standaardmethode waarin oversterfte wordt berekend als het verschil tussen de verwachte en waargenomen sterfte.

Het voordeel van het gebruiken van de ophogingsfactor is dat de betrouwbaarheidsintervallen die worden geschat met het model gebruikt kunnen worden als maat van betrouwbaarheid van de geschatte oversterfte. Als bijvoorbeeld de ophogingsfactor in een week niet significant verschillend is van nul, is er ook geen significante oversterfte in die week. Als de ophogingsfactor in een week significant negatief is, is er zelfs sprake van ondersterfte (door het zogenoemde oogsteffect). De verwachte sterfte kan in het model eenvoudig worden berekend als de som van de trend- en seizoencomponenten.

Een ander voordeel van het gebruikte model is dat alle parameters stochastisch zijn, waardoor betrouwbaarheidsintervallen voor afgeleiden kunnen worden berekend met Monte-Carlo-simulatie. Zo is het mogelijk om consistente betrouwbaarheidsintervallen te berekenen voor zowel de oversterfte per week als de totale oversterfte in alle weken sinds week 11.

### 3.3 Gevoeligheidsanalyse

Het dynamische regressiemodel bevat een aantal aannames. Als een gevoeligheidsanalyse wordt daarom de oversterfte die is berekend met dit model vergeleken met drie andere modellen om de oversterfte te schatten, te weten twee varianten van het dynamische regressiemodel en de methode die het CBS en het AMC in eerdere publicaties hebben gebruikt.

In de eerste variant van het dynamische regressiemodel, het statische regressiemodel, varieert de ophogingsfactor niet met de tijd. Dit is hetzelfde als een standaard (statische) regressie via OLS. In dit statische model is de waarde van de geschatte ophogingsfactor dan een gemiddeld effect over de hele tijdsperiode.

In de tweede variant, het interventiemodel, wordt de geregistreerde coronasterfte vervangen door een dummyvariabele. Deze dummyvariabele is een indicator van een (tijdelijke) interventie en bij een significante coëfficiënt signaleert het een significante afwijking van een trend. Deze dummyvariabele is 1 in weken met geregistreerde coronasterfte en 0 in weken zonder geregistreerde coronasterfte. Omdat perioden van oversterfte jaarlijks voorkomen in verband met griep epidemieën, krijgt de dummyvariabele ook in de weken van recente perioden met oversterfte de waarde 1. In het model zijn 5 griep epidemieën opgenomen, te weten die van 2017/2018, van 2016/2017 en 2015/2016 (RIVM 2020b), en die van 2014/2015 en 2012/2013 (RIVM 2015).

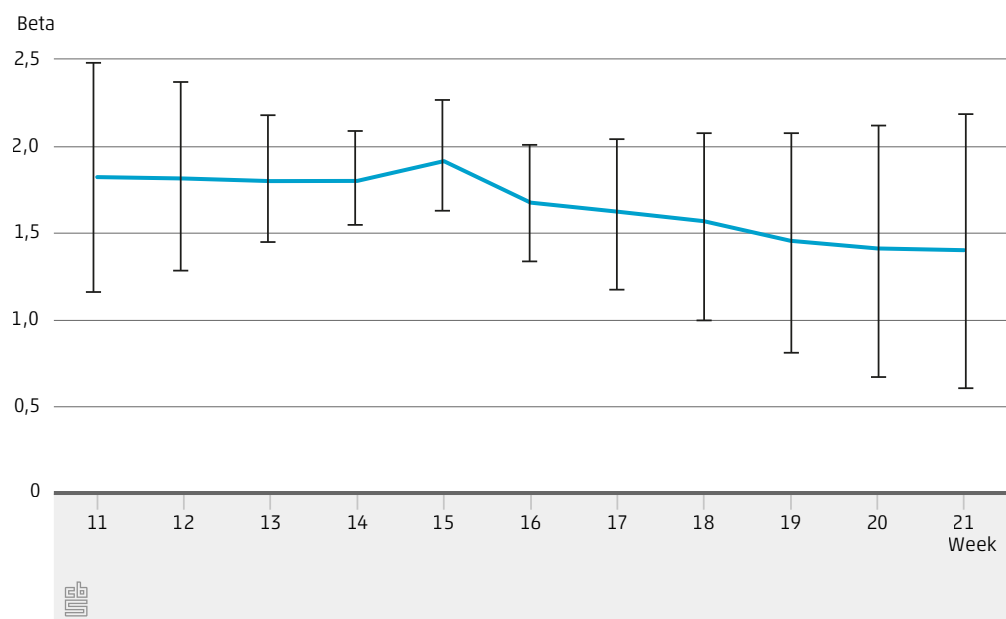
In de methode van het CBS/AMC wordt het verwachte aantal overledenen geschat alsof er geen corona-epidemie zou zijn geweest, op basis van het gemiddeld aantal overledenen in de voorafgaande weken (weken 3 tot en met 10), waarbij wordt gecorrigeerd voor de gemiddelde weektemperatuur (beschikbaar via KNMI.nl). De aanname is dat elke graad Celsius hoger dan de gemiddelde temperatuur in de weken 3 tot en met 10 leidt tot 1 procent minder sterfte. Voor week 11 tot en met week 14 was de correctie zeer gering vanwege een klein verschil in temperatuur ten opzichte van de weken 3 tot en met 10. Vanaf week 15 was het warmer en bedroegen de correcties enkele procentpunten.

## 4. Resultaten

### 4.1 Uitkomsten model

Figuur 4.1.1 laat de geschatte coëfficiënten (inclusief 95-procent-betrouwbaarheidsintervallen) per week zien die de relatie tussen de geregistreerde overledenen aan corona en de totale oversterfte beschrijven. De coëfficiënt, welke geïnterpreteerd kan worden als een ophogingsfactor, is significant hoger dan 1,0 in de weken 11 tot en met 18. Dat betekent dat als in deze weken een geregistreerde corona-overledene wordt waargenomen, er in werkelijkheid meer dan een extra overleden persoon gerelateerd aan corona was. Na week 16 daalt de ophogingsfactor geleidelijk tot onder 1,5 en nemen de betrouwbaarheidsintervallen toe.

#### 4.1.1 Geschatte coëfficiënten



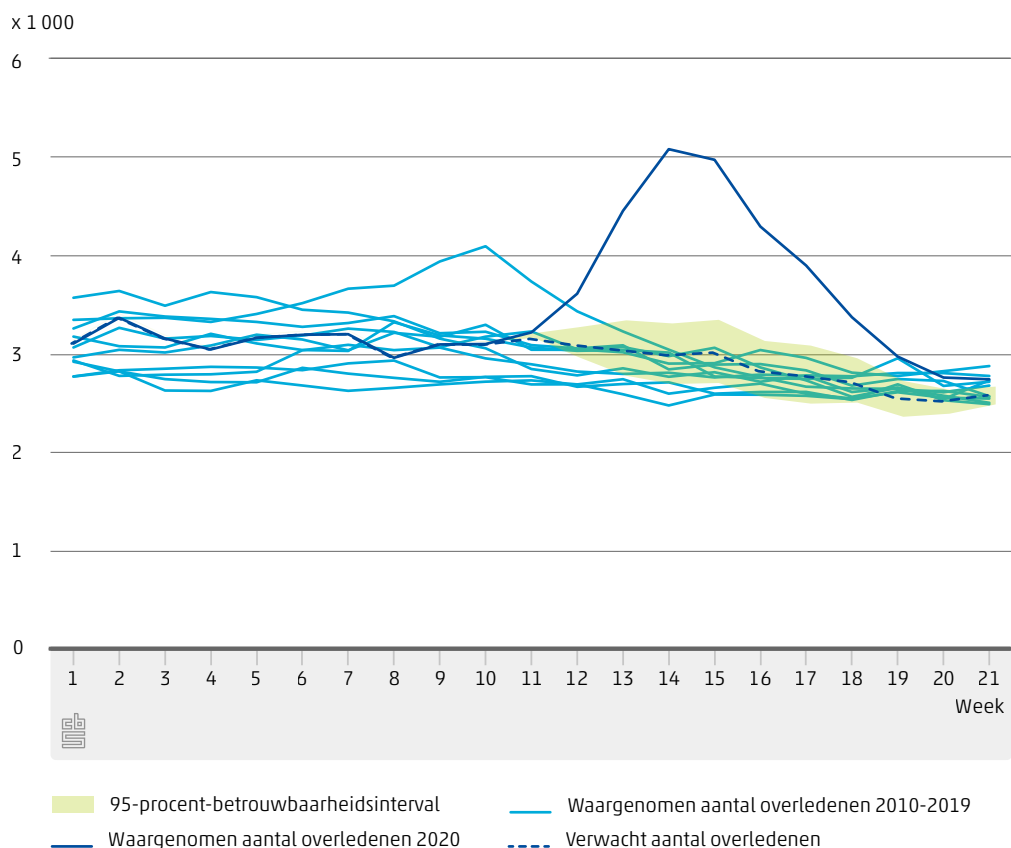


De betrouwbaarheidsintervallen tot en met week 15 worden steeds kleiner gedurende de weken, en daarna weer groter. In weken met kleine betrouwbaarheidsintervallen is coronasterfte een belangrijke verklaring voor oversterfte. Vanaf week 20 zijn de coëfficiënten niet meer significant groter dan 1,0, terwijl er geregistreerde coronasterfte is. In deze weken voegt de coronasterfte weinig verklaringskracht toe aan het model.

In Figuur 4.1.2 worden de waargenomen en de verwachte sterfte weergegeven (met 95-procent-betrouwbaarheidsintervallen) als er geen corona-epidemie was geweest. De oversterfte in deze figuur is het verschil tussen de lijn met de waargenomen sterfte en de lijn met de verwachte sterfte. Er is sprake van significante oversterfte als de lijn van de waargenomen sterfte buiten het interval van de verwachte sterfte ligt. Dit is in alle weken het geval.

Vergeleken met voorgaande jaren is de verwachte sterfte in de weken na de uitbraak van het coronavirus aan de hoge kant. Echter, de sterfte van voorgaande jaren valt over het algemeen binnen de betrouwbaarheidsintervallen van de verwachte sterfte. In week 19 ligt de verwachte sterfte onder het niveau van de laatste 10 jaar. Dat heeft te maken met een plotselinge sterftetoename tussen week 18 en week 19 in eerdere jaren (de lichtblauwe lijnen in Figuur 4.1.2). Maar ook in week 19 valt de sterfte in voorgaande jaren binnen de betrouwbaarheidsintervallen. In de weken 20 en 21 beweegt de verwachte sterfte zich richting de geobserveerde sterfte, wat suggereert dat de oversterfte sterk gereduceerd is.

#### 4.1.2 Waargenomen en verwachte sterfte, 2010-2020



## 4.2 Vergelijking met andere modellen

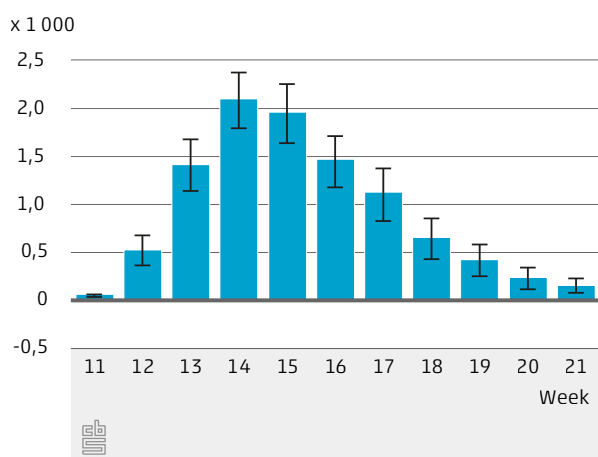
Om de uitkomsten van het dynamische regressiemodel in context te kunnen plaatsen, wordt de geschatte oversterfte afgezet tegen de berekende oversterfte in de drie andere modellen.

De wekelijkse oversterfte die met deze vier modellen kan worden berekend, is weergegeven in Figuur 4.2.1. In de figuur is te zien dat vooral in de eerste vijf weken van de coronapandemie de geschatte oversterfte in de vier modellen ongeveer even hoog is. De schattingen tot en met week 18 zijn nog redelijk vergelijkbaar bij het dynamische regressiemodel, het interventiemodel en het CBS/AMC-model.

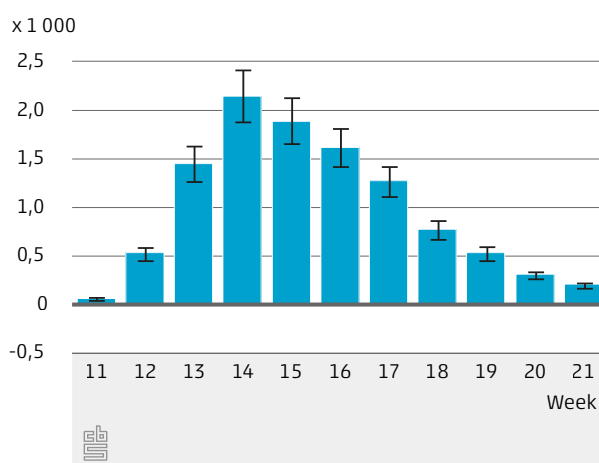
Het verschil tussen deze drie modellen zit met name in weken 20 en 21. In deze weken schat het dynamische regressiemodel significante oversterfte, het CBS/AMC-model ondersterfte en het interventiemodel niet-significante oversterfte. Het statische regressiemodel geeft vanaf week 16 een duidelijk hogere inschatting van de oversterfte

### 4.2.1 Oversterfte per week, vier modellen

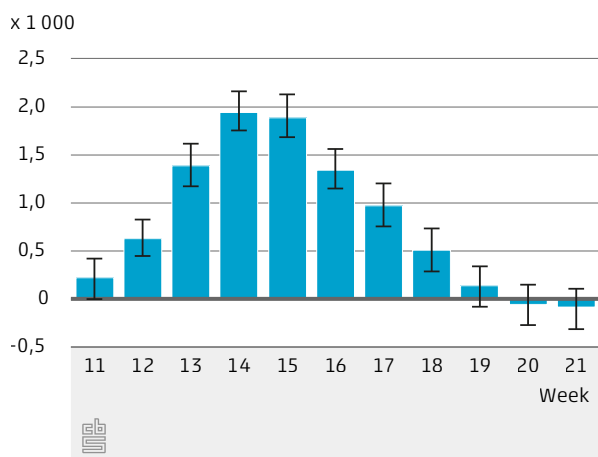
#### Dynamische regressiemodel



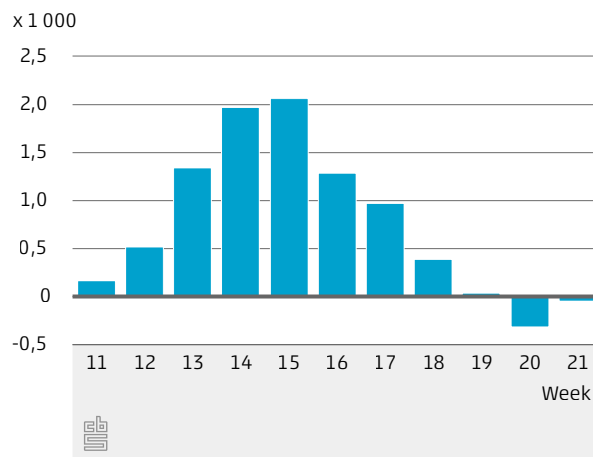
#### Statische regressiemodel



#### Interventie regressiemodel



#### CBS/AMC-model

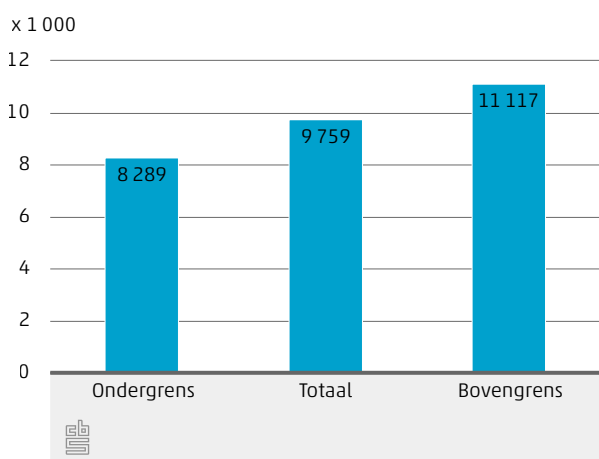


dan de andere modellen. Dit suggereert dat de flexibiliteit van de dynamische regressie belangrijk is voor de schatting, met name in de fase waarin de sterftcijfers van de epidemie kleiner worden.

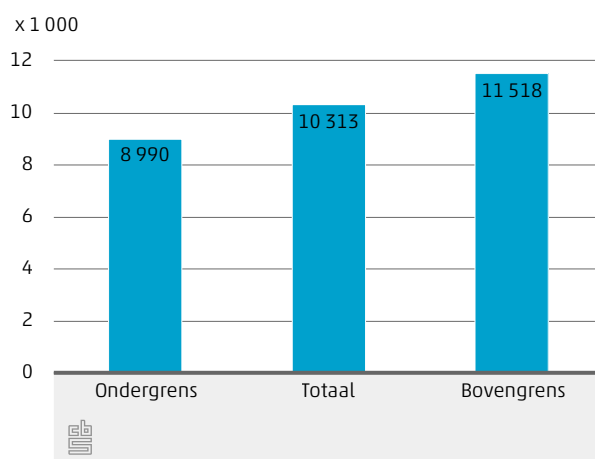
De totale oversterfte van week 11 tot en met week 19, met de onder- en bovengrens van de 95-procent-betrouwbaarheidsintervallen, is weergegeven in Figuur 4.2.2. In de figuur is te zien dat de gemiddelde oversterfte van het dynamische en het statische regressiemodel, het interventiemodel en het CBS/AMC-model dezelfde orde van grootte heeft. Toch zijn er verschillen tussen de modellen: terwijl de oversterfte van het CBS/AMC-model en het interventiemodel zeer vergelijkbaar zijn, valt de gemiddelde schatting van het CBS/AMC-model niet binnen het betrouwbaarheidsinterval van het statische regressiemodel. Het betrouwbaarheidsinterval van het dynamische regressiemodel plaatst het aantal overledenen door corona tussen 8 289 en 11 117 personen. Wederom is de oversterfte in het model met vaste regressiecoëfficiënt een flink stuk hoger dan in de andere modellen.

#### 4.2.2 Totale oversterfte, week 11 tot en met week 19 van 2020

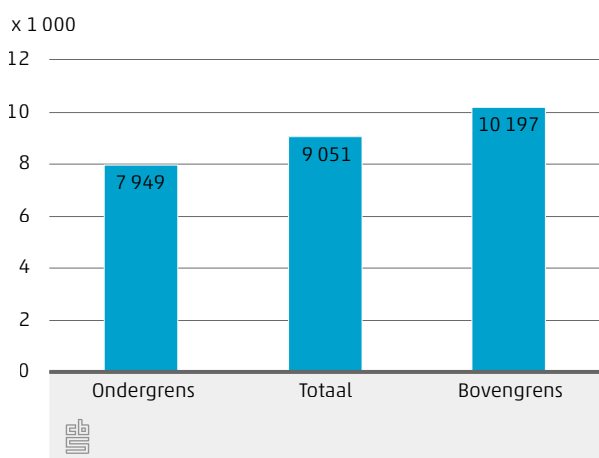
##### Dynamische regressiemodel



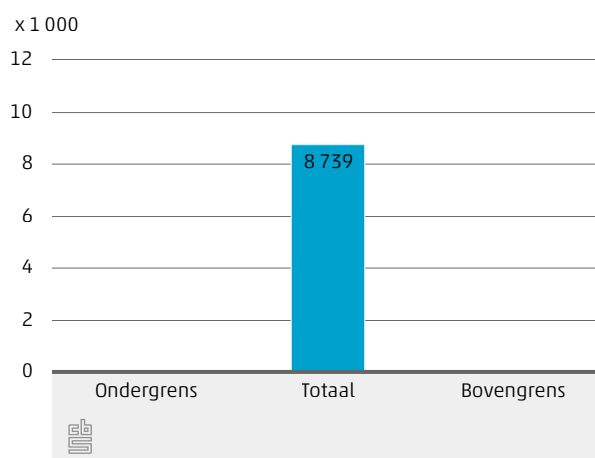
##### Statische regressiemodel



##### Interventie model



##### CBS/AMC-model



## 5. Discussie en conclusie

Op basis van het dynamische regressiemodel is er sprake van significante oversterfte van week 11 tot en met week 21 van 2020, die gerelateerd is aan de corona-epidemie. De oversterfte in deze periode wordt geschat op 10 164. Echter, als rekening wordt gehouden met onzekerheid, kan het aantal overledenen zowel hoger als lager zijn. Het betrouwbaarheidsinterval plaatst de bovengrens van de oversterfte op 11 691 personen en de ondergrens bij 8 593 personen. Dit zijn dus het maximale en minimale aantal personen overledenen door het coronavirus. Naast elke 10 geregistreerde corona-sterfgevallen zijn er dus tussen de 5 en 10 mensen overleden aan corona die niet geregistreerd werden.

De drie andere methoden om de oversterfte tijdens de corona-epidemie te schatten, waaronder de methode die het CBS (in samenwerking met AMC) in eerdere publicaties heeft gebruikt, geven soortgelijke uitkomsten. Een oversterfte van rond 9 duizend in de eerste negen weken van de corona-epidemie (CBS, 29 mei 2020) is dus plausibel.

De resultaten van het dynamische regressiemodel geven aan dat de oversterfte ook in de weken 20 en 21 significant positief is, al gaat het om een klein aantal. Dit staat enigszins haaks op de resultaten van het CBS/AMC-model en het interventiemodel. Dat na een periode van oversterfte ondersterfte volgt, lijkt aannemelijk. Dit wordt in de epidemiologie het oogsteffect genoemd, en is een compensatie na uitzonderlijke oversterfte. Op basis van de resultaten van alle modellen moet er worden geconcludeerd dat er geen sterke bewijzen zijn voor zowel over- als ondersterfte in de weken 20 en 21. Ook is de corona-epidemie nog niet ten einde en kan de periode van lagere sterfte ook in de komende weken nog optreden.

In Nederland, waar alleen positief geteste overleden personen worden geregistreerd als coronadoden, is de geschatte oversterfte een goede en belangrijke manier om enige duiding te kunnen geven aan het aantal overledenen door COVID-19. Op een later moment komt wel informatie over doodsoorzaken (met ook vermoedelijke COVID-19) via het doodsoorzakenformulier ter beschikking. Het is dan interessant om de uitkomsten van de hier gepresenteerde modellen te vergelijken met de verdeling van de sterfte over de doodsoorzaken. Echter, het is niet zeker dat er met de statistiek van doodsoorzaken een compleet overzicht van alle COVID-19 overledenen te verwachten is, omdat het een nieuwe en derhalve niet uitgebreid bestudeerde ziekte betreft (Office for National Statistics, 5 juni 2020). Daarnaast is het mogelijk dat de *lockdown* een effect heeft op de sterfte aan andere oorzaken. De geschatte oversterfte geeft een beeld van alle extra overledenen tijdens corona, ook indirecte.

De resultaten in dit artikel laten zien dat de flexibiliteit van het dynamische regressiemodel belangrijk is om een goed beeld te krijgen van de ontwikkeling van de oversterfte. Tijdens de corona-epidemie veranderde het testbeleid met betrekking tot corona en dit wordt opgepakt door de ophoogfactor. In de weken 14 en 15 was de ophoogfactor significant hoger dan 1,5, wat een suggestie geeft van onderregistratie van het aantal doden. Vanaf week 20 voegt de coronasterfte weinig verklaringskracht toe aan het model. De resultaten suggereren dat de flexibiliteit van de dynamische regressie belangrijk is met name in de fase waarin de sterftcijfers van de epidemie kleiner worden.

De modellen die in dit artikel gepresenteerd zijn, zijn uiteraard niet volledig. Een vergelijking met de uitkomsten van andere modellen, zoals de methode die het RIVM gebruikt voor de sterfmonitoring (RIVM, 2020b), geeft mogelijk een nog vollediger beeld van de impact van corona op de sterfte.

De modellen die in dit artikel zijn gebruikt, zijn geschikt om de oversterfte aan andere oorzaken dan corona te schatten. Het interventiemodel is bijvoorbeeld ook geschikt om de oversterfte te berekenen tijdens hittegolven. Daarnaast zijn de modellen geschikt voor een regionale analyse (CBS 2020b). Bovendien kunnen de modellen toegepast worden op data van andere landen om op die manier de impact van de corona-epidemie in verschillende landen te vergelijken.

## Referenties

CBS (29 mei 2020). [Sterfte in coronatijd](#). Nieuwsbericht, 29 mei 2020.

CBS Statline (5 juni 2020). [Overledenen; geslacht en leeftijd, per week](#).

Durbin, J. and Koopman, S.J. (2012). Time series analysis by state space models. Oxford statistical sciences series 38.

Laine, M. (2020). Introduction to Dynamic Linear Models for time series analysis. Chapter 4 in: Geodetic time series analysis in earth sciences, p. 139–156.

Nielsen, J., Krause, T. G., & Mølbak, K. (2018). Influenza-associated mortality determined from all-cause mortality, Denmark 2010/11–2016/17: The FluMOMO model. Influenza and other respiratory viruses, 12(5), 591–604.

Office for National Statistics (5 juni 2020). [Analysis of death registrations not involving coronavirus \(COVID-19\)](#), England and Wales, 5 juni 2020.

Petris, G., Petrone, S. and Campagnoli, P. (2009). Dynamic Linear Models with R. Springer Verlag.

Petris, G. and Gilks, W. (2018). Package DLM. Cran package for R.

RIVM (2020a). [Actuele informatie over het nieuwe coronavirus \(COVID-19\)](#).

RIVM (2020b). [Monitoring sterftcijfers Nederland](#).

RIVM (2015). [Annual report Surveillance of influenza and other respiratory infections in the Netherlands: winter 2014/2015](#).

# Bijlage

Het Dynamic Linear Model zoals geformuleerd in 3.2, kan als volgt wiskundig weergegeven worden:

$$\begin{aligned}y_t &= \alpha_t + \beta_t x_t + g_t + y_t^{ar} + v_t & v_t &\sim N(0, \sigma_v^2) \\ \alpha_t &= \alpha_{t-1} + w_{\alpha,t} & w_{\alpha,t} &\sim N(0, \sigma_\alpha^2) \\ \beta_t &= \beta_{t-1} + w_{\beta,t} & w_{\beta,t} &\sim N(0, \sigma_\beta^2) \\ g_t &= \sum_{j=1}^2 \left( a_j \cos\left(t \frac{2\pi j}{52.18}\right) + b_j \sin\left(t \frac{2\pi j}{52.18}\right) \right) + w_{g,t} & w_{g,t} &\sim N(0,0) \\ y_t^{ar} &= \phi y_{t-1} + w_{ar,t} & w_{ar,t} &\sim N(0, \sigma_{ar}^2)\end{aligned}$$

De modellering van seizoenseffecten, met een combinatie van harmonische en autoregressieve elementen, volgt de internationale literatuur van modellering van griep-epidemieën (Nielsen et al., 2018). Onbekende varianties en de parameter voor de autoregressieve component zijn geschat met maximum likelihood estimation.

De andere onbekenden in dit model worden geschat met het discrete Kalmanfilter waarbij de software-implementatie in R gegeven is door Petris et al. (2009, 2018). Het R-script kan op aanvraag toegestuurd worden.

## Verklaring van tekens

Niets (blanco)	Een cijfer kan op logische gronden niet voorkomen
.	Het cijfer is onbekend, onvoldoende betrouwbaar of geheim
*	Voorlopige cijfers
**	Nader voorlopige cijfers
2019-2020	2019 tot en met 2020
2019/2020	Het gemiddelde over de jaren 2019 tot en met 2020
2019/'20	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2019 en eindigend in 2020
2017/'18-2019/'20	Oogstjaar, boekjaar, enz., 2017/'18 tot en met 2019/'20

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.

## Eindredactie Statistische Trends

Math Akkermans  
Ronald van der Bie  
Marion van den Brakel  
Moniek Coumans  
Kees Groenenboom  
Annelie Hakkenes-Tuinman  
Brigitte Hermans  
Suzanne Loozen

## Colofon

*Uitgever*  
Centraal Bureau voor de Statistiek  
Henri Faasdreef 312, 2492 JP Den Haag  
[www.cbs.nl](http://www.cbs.nl)

*Prepress*  
Centraal Bureau voor de Statistiek

*Ontwerp*  
Edenspiekermann

*Inlichtingen*  
Tel. 088 570 70 70  
Via contactformulier: [www.cbs.nl/infoservice](http://www.cbs.nl/infoservice)

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2020.  
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.