

Voorspelmodellen voor toerisme

Is het mogelijk om toeristische drukte te voorspellen?

Onderdeel van het Data & Development Lab



Centraal Bureau
voor de Statistiek



NBTC
Nederlands Bureau
voor Toerisme &
Congressen



Centre of Expertise
leisure, tourism & hospitality

Auteurs

Jasper Heslinga

Programmamanager NHL Stenden Hogeschool CELTH

Harm IJben

Programmamanager HZ University of Applied Sciences CELTH

31 mei 2021



Is het mogelijk om toeristische druk te voorspellen?

Voorspelmodellen zijn een grote belofte als het gaat om het inschatten van toeristische druk op bestemming. Zij kunnen helpen bij het (beter) inschatten waar bezoekersstromen (als indicatoren van toerismedrukke) van verblijfs- en dagbezoekers (op nationaal-, regionaal- en lokaalniveau) mogelijk druk(te) gaan opleveren.

Een inzicht in (verwachte) bezoekersstromen (en daarmee toerismedrukke) is op die manier (als beleids- en sturingsinstrument) voor een bewuste bestemming (toeristische regio) zeer wenselijk in haar streven naar een meer duurzame balans tussen inwoner, bezoeker, ondernemer en (natuurlijke) omgeving met oog voor de draagkracht van het gebied. Een goed werkend voorspelmodel is er echter helaas nog niet.



Doel

Het doel van dit DDL-project is antwoord geven op de vraag: in hoeverre het mogelijk is om voorspelmodellen te ontwikkelen (als middel) om inzicht te geven in toerismedrukke en dus in verwachte bezoekersstromen (doel)?

Wat is er gedaan?

Het project bestaat uit drie met elkaar samenhangende onderdelen:

1. Literatuurstudie

Voor een goede basis is een literatuurreview ('quickscan') uitgevoerd in de vorm van een brede verkenning om te bepalen wat er al bekend is over voorspelmodellen voor toerisme(drukke) en hoe hierin bezoekersstromen als indicator worden gehanteerd, met speciale aandacht voor de (on)mogelijkheden, de onderzoekstechnische en methodologische uitdagingen (en de benodigde en beschikbare databronnen);

2. Experimenten

Werken de gevonden modellen ook in de praktijk? Om die vraag te beantwoorden zijn er een aantal experimenten opgezet om dit te testen, waarbij verschillende databronnen zijn gebruikt. In het eerste experiment is verkend in hoeverre 'Google Trends' (geaggregeerde data verkregen via zoekopdrachten) zinvol zijn. Het tweede experiment zoomt in op het toepassen van online data op basis van 'webscraping' om daarmee vanuit een inzicht in de online beschikbare verblijfsaccommodaties duiding te kunnen geven aan de bezetting van diezelfde verblijfsaccommodaties en daarmee de te verwachten toerismedrukke. In het derde experiment tenslotte is stil gestaan bij de waarde die 'GPS-data' (verkregen via tracking software) mogelijk heeft voor een inzicht in bezoekersstromen;

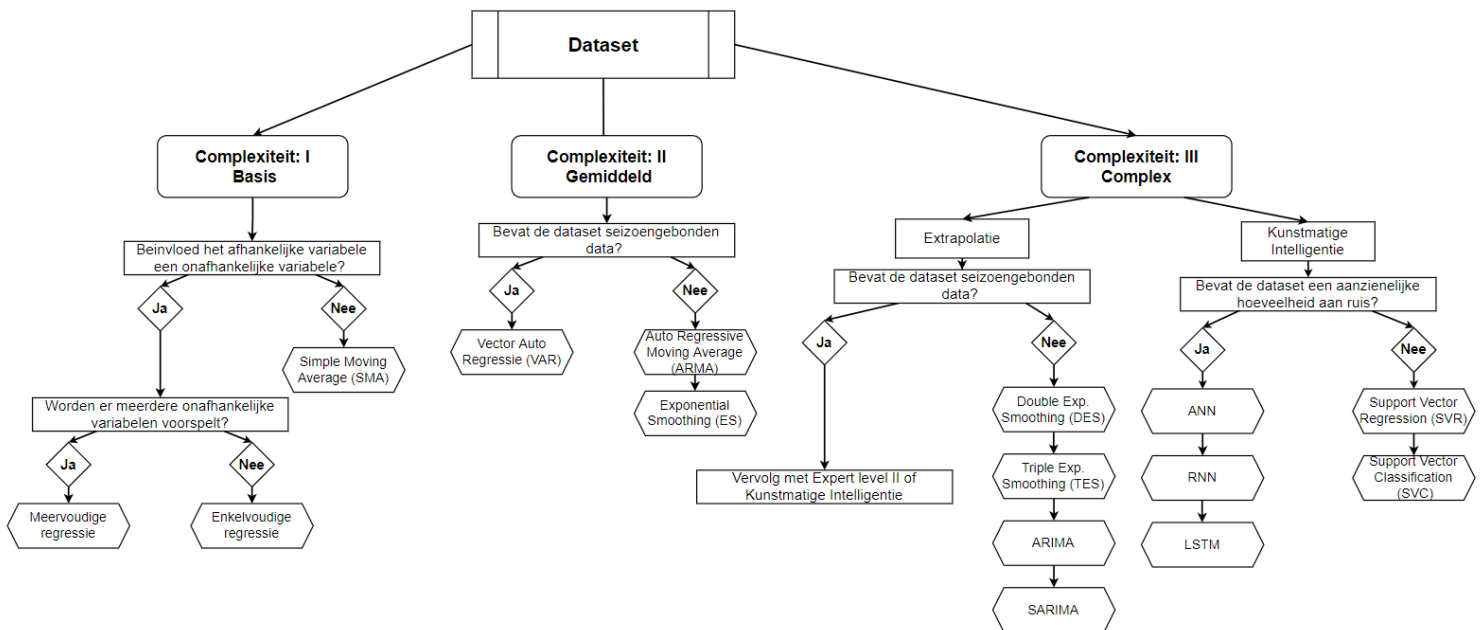
3. Expertsessie

Voor de afsluiting is er een expertsessie georganiseerd om tot een uitwisseling te komen van de opgedane inzichten per experiment (in aansluiting op de literatuurreview), voor het vaststellen van de leeropbrengsten, en kijken naar gewenste vervolgstappen.

Wat heeft het literatuuronderzoek opgeleverd?

De literatuurstudie biedt een goed overzicht van wat een voorspelmodel is, welke methoden er zijn en wanneer deze onder welke voorwaarden te gebruiken zijn. Dit overzicht is vooral bedoeld voor onderzoekers die nog geen experts zijn in modelleren, maar wel daarmee wil beginnen of meer van wil weten. Deze 'quickscan' biedt een handig overzicht om een goede start te maken.

De studie is vrij uitvoering gedaan en geeft zoals eerder gezegd een goed overzicht, maar riskeert daardoor een wat 'platte' opsomming te worden die veel detail bevat. Om daar wegwijz in de worden is er een keuze matrix ontwikkelt die helpt bij het maken van keuzes voor het model en methoden (zie Figuur 1). De volledige achtergrond informatie is uiteraard beschikbaar in een losse publicatie (Middelhoek & Heslinga).



Figuur 1: Flowchart & overzicht van kwantitatieve voorspelmodellen

Wat hebben de experimenten opgeleverd?

Er is gekeken naar een drietal (big)-data bronnen om experimenten mee uit te voeren. Daar vielen een aantal zaken op:

- ▶ Google Trends is niet nuttig bij het doen van voorspellingen. Hoewel de bron openbaar beschikbaar is, is het niet optimaal verkrijgbaar. Voor het doen van een goede voorspelling is het hebben van absolute cijfers wenselijk. Google Trends laat slechts mutatie/trends zien over tijd (tussen 0-100). Bovendien kijkt deze bron naar trends uit het verleden, terwijl voor voorspelmodellen er ook toekomst variabelen nodig zijn.
- ▶ Zowel bij het Zeeuwse als Friese experiment werden voorspellingen op basis van tijdsreeksen ("time series") als minder succesvol ervaren. Dit heeft vooral te maken met het tekort aan data van verschillende jaargangen. Hierdoor komen de seizoensinvloeden niet goed uit de verf. Dit model op basis van tijdreeksen is doorgaans een model met een hoge mate van complexiteit. De experimenten laten zien dat een complex model geen garantie is voor een beter model.
- ▶ In de experimenten doen regressies het een stuk beter. In de modellen op basis van Airbnb en GPS zijn er correlaties te zien. Waar in het begin het coëfficiënt (sterkte van het verband tussen variabelen) veelal laag is, werd dit steeds beter naar mate verschillende andere variabelen werden toegevoegd (vakanties, weer, evenementen, etc.).
- ▶ Het technisch modelleren wil op zich wel, maar wordt uiteraard gedaan op basis van veel aannames. Het is belangrijk om hier goed bewust van te zijn. Een model is en blijft een abstractie van de werkelijkheid.
- ▶ Het coronavirus en de maatregelen zorgen voor ruis in de data en op die manier voor een bepaalde mate van onzekerheid in de analyses.

Welke lessen kunnen getrokken worden voor toekomstig onderzoek?

- ▶ Het 'beste' model bestaat helaas niet. Er is een breed scala aan modellen in kaart gebracht, maar welke het beste is, is afhankelijk van de context en de exacte vraag. Voor analyses in de toekomst wordt geadviseerd om verschillende modellen en methoden te combineren. Een dergelijke mix zorgt voor een zo optimaal mogelijk resultaat.
- ▶ Het is van belang om van tevoren scherp te hebben wat je voorspelhorizon is. Dit helpt om helder te krijgen wat je precies wil voorspellen en waarvoor het model wordt gebruikt.
- ▶ Om te corrigeren voor de effecten van corona in het model biedt de zogenaamde 'Oxford Stringency index' een oplossing. Deze index geeft het niveau van maatregelen in een land weer, waardoor er voor corona gecorrigeerd zouden kunnen worden.
- ▶ Vervolgstappen voor een regressiemodel zijn het toevoegen van meer databronnen en te kijken in hoeverre deze correleren. Hierbij is het wenselijk om overzicht te krijgen in welke variabelen wel/niet werken, en welke wel/niet beschikbaar zijn.
- ▶ 'Ingewikkelde modellen zijn niet altijd beter dan simpele' Hou het daarom simpel ook in de communicatie naar stakeholders voor draagvlak.
- ▶ Tenslotte kan er worden geconcludeerd dat dit project een eerste goede eerste stap is op het gebied van voorspelmodellen voor toerismedrukke. De gedane analyses geven perspectief, maar vragen tegelijk ook om vervolgonderzoek.

Disclaimer

Dit project is onderdeel van het Data & Development Lab van het NBTC, CELTH en CBS. De projectresultaten zijn tot stand gekomen door intensieve samenwerking en kennisuitwisseling tussen verschillende instellingen. Zo werd er door het European Tourism Futures Institute (onderdeel van NHL Stenden) een handig literatuuroverzicht rondom voorspelmodellen ontwikkeld. Er zijn experimenten uitgevoerd met DataFryslân (over Airbnb data), met HZ University of Applied Sciences (over GPS data) en de gemeente Amsterdam heeft analyses gedeeld (hotelcijfers). In de expertsessie hebben betrokkenen van de bovengenoemde organisaties met elkaar van gedachten gewisseld over de uitkomsten van deze studie, hoe deze beter te kunnen duiden en te komen tot suggesties voor vervolgonderzoek.



Vragen over deze publicatie? Neem contact op met:

Centre of Expertise Leisure, Tourism & Hospitality

Mgr. Hopmansstraat 2
4817 JS Breda

Email: info@celth.nl
Website: www.celth.nl

