

Betere financiële prognoses in business cases door de toepassing van big data

Bas van Raaij, Arco van de Ven

Received 5 June 2018 | Accepted 21 October 2018 | Published 2 November 2018

Samenvatting

Investeringsbeslissingen worden vaak gebaseerd op een business case waarin investeringsalternatieven met elkaar worden vergeleken. Prognoses van de financiële impact van de investeringsalternatieven spelen hierin een belangrijke rol. De gebruikte historische data op basis waarvan financiële prognoses tot stand komen en de wijze waarop deze data worden gebruikt bij het bepalen van verwachte inkomsten en uitgaven bepalen in hoge mate de betrouwbaarheid van de prognoses. Big data en moderne analysetechnieken hebben de potentie om de betrouwbaarheid van financiële prognoses te vergroten. Het empirische bewijs van de werkelijke toegevoegde waarde van big data op dit vlak is echter schaars. In dit artikel wordt beschreven welke big data-toepassingen kunnen bijdragen aan betere financiële prognoses in business cases. Daarnaast wordt een casestudie gepresenteerd waarin big data-toepassingen daadwerkelijk tot betere financiële prognoses in business cases zouden hebben geleid.

Relevantie voor de praktijk

De explosieve toename van de hoeveelheid en variëteit aan data en de geavanceerde analysemogelijkheden die bedrijven tot hun beschikking hebben, bieden kansen voor het bepalen van financiële prognoses van investeringsalternatieven.

1. Inleiding

Een business case is een instrument ter ondersteuning van besluitvorming en planning om een probleem op te lossen of een kans te verzilveren. Om besluitvorming te ondersteunen bevat een business case voorspellingen van toekomstige effecten van het te nemen besluit: financiële resultaten en andere consequenties (Braaksma et al. 2006; Schmidt 2005; Stratton 2004). In dit artikel staan de mogelijkheden van het verbeteren van deze voorspellingen, de financiële prognoses, centraal.

In de literatuur worden diverse risico's met betrekking tot de kwaliteit van financiële prognoses in business cases onderkend. Allereerst worden prognoses vaak gebaseerd op input van personen in de organisatie die een belang hebben bij de investering. Hierdoor ontstaat het risico van *confirmation bias* (Klayman 1995): een psychologisch effect waarbij men zoekt naar informatie die hun opvatting – in dit geval de noodzaak om al dan niet te investeren – bevestigt. Door dit effect kan de juistheid

en volledigheid van de input voor de financiële prognose in het geding komen.

Een andere valkuil is dat prognoses vaak worden gebaseerd op het meest waarschijnlijke toekomstscenario. Op basis van dit scenario wordt dan een verwacht rendement op de investering berekend. Een goede investeringsbeslissing kan echter niet alleen worden genomen op basis van het meest waarschijnlijke scenario, de kansverdeling van mogelijke uitgaven en inkomsten speelt ook een rol. Zo moeten de consequenties van de investeringsvoorstellen binnen de risicobereidheid van de organisatie vallen (Committee of Sponsoring Organizations 2017). De financiële consequenties van negatieve scenario's spelen hierbij een belangrijke rol. Om meer gevoel te krijgen bij mogelijke afwijkingen van het rendement worden naast het meest waarschijnlijke scenario dan ook vaak andere scenario's doorgerekend zoals een *best case*- en een *worst case*-scenario. De kans dat alle variabelen in werkelijkheid de slechtst of best denkbare waarde realiseren is echter klein. In werkelijkheid zal een mix optreden: het

werkelijke rendement ligt waarschijnlijk tussen het worst case- en het best case-scenario en er wordt aangenomen dat het in de buurt van het meest waarschijnlijke scenario ligt. Een betere onderbouwing van de kansdistributie van het rendement ondersteunt de besluitvorming en het zicht op de consequenties van het optreden van afwijkingen van het meest waarschijnlijke scenario.

Technieken in het veld van big data en moderne analyse-technieken bieden mogelijkheden om de genoemde risico's te beperken en tot betere prognoses te komen. De hoeveelheid data en de variëteit aan bronnen die deze data herbergen neemt explosief toe en moderne analysetechnieken maken het mogelijk om die data te transformeren tot nauwkeuriger voorspellingen. Dit is mogelijk door de objectiviteit van de feitelijke informatie (input) te verhogen en daarmee tegenwicht te geven aan de *confirmation bias*. Daarnaast bieden analysetechnieken mogelijkheden om de input om te zetten in nauwkeuriger prognoses door niet alleen rekening te houden met het meest waarschijnlijke scenario.

Van drie business cases van Port of Amsterdam die in het verleden hebben geleid tot positieve investeringsbeslissingen wordt bepaald of de toepassing van big data en moderne analysetechnieken tot nauwkeuriger en daarmee betere financiële prognoses zou hebben geleid.

Het artikel is als volgt opgebouwd. In paragraaf 2 wordt de theoretische achtergrond geschetst van aspecten die van belang zijn voor de kwaliteit van financiële prognoses in business cases. In paragraaf 3 worden big data-technieken besproken en in paragraaf 4 wordt de relatie gelegd tussen deze big data-technieken en de aspecten die van belang zijn voor de kwaliteit van financiële prognoses. In de volgende paragrafen worden de onderzoeksmethode en de toepassing van de big data-technieken op drie business cases van Port of Amsterdam besproken en geanalyseerd. Het artikel wordt afgesloten met conclusies en beperkingen van het onderzoek.

2. Kwaliteit van financiële prognoses in business cases

Het kloppend hart van een business case wordt gevormd door een financieel model (Messner 2013; Schmidt 2005;

Stratton 2004). De toekomstige kasstromen die in verband kunnen worden gebracht met het investeringsalternatief leiden hierin tot een verwacht rendement op de investering. De betrouwbaarheid van de financiële prognose hangt af van de kwaliteit van de gebruikte data en de methodiek die voor de analyse wordt gehanteerd.

In het kader van *confirmation bias* hebben diverse auteurs gewezen op het belang van objectieve data (Klayman 1995). De kwaliteit van de input van de financiële prognose is voldoende wanneer er geen relevante zaken ontbreken (volledigheid) en de input overeenkomt met de werkelijkheid (juistheid) (Van Leeuwen and Bergsma 2014). Natuurlijk betreft het prognoses en is een mate van onzekerheid over de volledigheid en juistheid inherent aan prognosticeren, maar gegeven de *confirmation bias* kan het objectiveren van data de betrouwbaarheid van de prognoses verhogen.

Naast objectieve input is van belang op welke wijze de input wordt vertaald in prognoses. Messner (2013) en Schmidt (2015) stellen dat er een beoordeling moet worden gemaakt van het risico dat kleeft aan de besluitvorming: de gevoeligheid van de voorspellingen in het financiële model. Deze analyse helpt om de uitkomsten van de berekening op de juiste manier te interpreteren en maakt de berekening waardevoller doordat er betere besluitvorming kan plaatsvinden. Dit is mogelijk door middel van *sensitivity*- en *probability*-analyses (Alkaraan and Northcott 2006). Hierbij wordt respectievelijk inzichtelijk gemaakt in hoeverre het resultaat afhankelijk is van de afzonderlijke variabelen en wordt onzekerheid in de berekening meegenomen, wat leidt tot een kansverdeling van de mogelijke uitkomsten van de business case.

Tot slot wordt in de literatuur gesproken over de wijze waarop de analyseresultaten worden gepresenteerd. “*Data summaries and analysis should be presented objectively and directly, keeping interpretations and explanatory text to a minimum*” (Schmidt 2005, p. 17).

In deze paragraaf zijn aspecten in kaart gebracht die van belang zijn voor de kwaliteit van financiële prognoses in business cases. Deze aspecten hebben betrekking op de kwaliteit van de input en de methodiek die voor de analyse wordt gehanteerd. In tabel 1 worden de aspecten per onderdeel samengevat.

Tabel 1. Aspecten met betrekking tot de kwaliteit van financiële prognoses in business cases.

Aspecten met betrekking tot de kwaliteit van de input van de financiële prognose:

- de inputvariabelen zijn volledig;
- de inputvariabelen kennen een zekere mate van juistheid;
- de inputvariabelen komen objectief tot stand.

Aspecten met betrekking tot de methodiek die voor de analyse van het financiële model wordt gehanteerd:

- de analyse wordt voorzien van een sensitivity analyse;
 - de analyse wordt voorzien van een probability analyse;
 - de analyse wordt objectief gepresenteerd.
-

3. Big data-toepassingen

Big data is een verzamelnaam voor de methoden en technieken waarmee organisaties proberen om concurrentievoordeel te behalen uit de data die ze tot hun beschikking hebben (Provost and Fawcett 2013). In de definitie van big data worden vaak drie termen genoemd die staan voor de uitdagingen waar traditionele systemen en methoden niet mee om kunnen gaan, en waar big data een antwoord op biedt: het grote *volume* aan data dat organisaties tot hun beschikking hebben, de *velocity* (snelheid) waarmee deze data wordt gecreëerd en moet worden omgezet in bruikbare informatie en de grote *variety* aan bronnen en dataformaten waar de data uit afkomstig is (Heusinkveld 2014; Katal et al. 2015).

Moderne analysetechnieken maken het voor organisaties mogelijk om op basis van deze data toekomstige ontwikkelingen te voorspellen en besluitvorming te ondersteunen (Bughin et al. 2010; George et al. 2014). In de literatuur worden drie vormen van analysetechnieken onderscheiden (Evans and Lindner 2012; Haas et al. 2011):

1. *Descriptive analytics*: het vinden van patronen en relaties in historische data. Technieken in deze categorie zijn visualisatie en *data mining*. Visualisatie is volgens Evans and Lindner (2012, p. 3) de meest bruikbare component van big data waarmee verrassende patronen en relaties inzichtelijk kunnen worden gemaakt en data effectief kan worden gecommuniceerd binnen organisaties. *Data mining* wordt ingezet om op basis van historische data nieuwe en belangrijke patronen te ontdekken. Hierbij worden statistische methoden gebruikt (Chen et al. 2012; Kohavi et al. 2002).
2. *Predictive analytics*: het voorspellen van toekomstige waarschijnlijkheden en trends, met name door het extrapoleren van trends en patronen die worden waargenomen in historische data. Dit werkt zo lang de toekomst lijkt op het verleden. Een bekende techniek in deze categorie is *what-if*-simulatie waarbij inzichtelijk wordt gemaakt wat het effect is indien de variabelen in een model zich anders gedragen dan vooraf verwacht (Haas et al. 2011, p. 1). Een geavanceerdere techniek is *simulation & risk* waarbij niet alleen de impact wordt onderzocht, maar de (on-)zekerheid van het gehele model wordt bepaald. Hierbij worden wiskundige, statistische modellen ingezet om grote hoeveelheden scenario's door te rekenen en zo te komen tot een kansverdeling – en daarmee waarschijnlijkheid – van mogelijke uitkomsten van het totale model (Davenport and Harris 2007; Katal et al. 2013; Stratton 2004). Ook gevoeligheidsanalyses, waarbij modellen meermaals worden doorgerekend om de gevoeligheid van het model voor de afzonderlijke variabelen te bepalen, vallen onder deze toepassing (Ibrahim and Koch 2008).
3. *Prescriptive analytics*: in onzekere omgevingen ontstaan de patronen die in het verleden zijn waargenomen niet meer. Bij deze vorm van *analytics* worden die inzichten gecombineerd met modellen die rekening houden met de onzekerheid in de data en zelf in staat zijn om voorspellingen aan te passen op basis van nieuwe inzichten. *Machine learning* is een vorm van *artificial intelligence* waarbij

algoritmen worden ontwikkeld waarmee systemen gedrag ontwikkelen dat volledig wordt gebaseerd op data. Uiteindelijk kunnen, doordat het systeem zichzelf verbetert en zo leert, geautomatiseerd intelligente besluiten worden genomen door het systeem zonder dat bij dat bewuste besluit menselijk handelen komt kijken (Chen et al. 2012).

4. Toepassing van big data in financiële prognoses

In deze paragraaf worden verbanden gelegd tussen de aspecten die van belang zijn voor kwalitatief goede financiële prognoses in business cases en hoe de hiervoor behandelde big data-toepassingen daaraan kunnen bijdragen.

4.1 Toepassing van big data bij het bepalen van input voor de financiële prognose

Davenport and Harris (2007, p. 59) noemen *future value analysis* als een van de toepassingsgebieden van big data bij investeringsbeslissingen. Big data-toepassingen kunnen volgens hen worden ingezet om de input voor de financiële prognoses te bepalen.

De kwaliteit van input is allereerst afhankelijk van de volledigheid. In traditionele financiële modellen worden alleen de variabelen opgenomen waar betrokkenen aan denken. Mogelijk kennen zij niet alle variabelen die invloed uitoefenen op de investering. Door middel van *data mining* kunnen nieuwe en belangrijke patronen worden ontdekt in de historische data. Op basis van deze patronen kan het financiële model worden uitgebreid met extra inputvariabelen waardoor de volledigheid van de input van de financiële prognose toeneemt.

Een tweede aspect is de juistheid van de input. Big data kan hier op twee manieren aan bijdragen. Allereerst kan door middel van big data-toepassingen de voorspelling worden gemaakt. Mogelijke toepassingen hierbij zijn visualisatie, waarbij historische patronen inzichtelijk worden gemaakt en kunnen worden omgezet in een voorspelling, en *data mining* waarbij nieuwe patronen aan het licht komen die, al dan niet geautomatiseerd, worden vertaald in voorspellingen. Daarnaast kunnen big data-toepassingen worden gebruikt om de voorspelling te valideren.

Het derde en laatste aspect met betrekking tot de input is dat ze objectief tot stand komt. De voorspellingen mogen niet worden gekleurd door het belang van de betrokkenen die ze vaststellen. Visualisatie kan hierbij een rol spelen. Door historische patronen van data te visualiseren en de voorspelde waarden daarbij te presenteren kan door de besluitvormers relatief eenvoudig worden bepaald of de voorspelling sterk afwijkt van het verleden waarna een discussie kan worden gevoerd of dit terecht is. Ook *data mining* kan bijdragen aan objectievere inputvariabelen door onbekende patronen aan het licht

te brengen waar men – bewust of onbewust – nog niet aan had gedacht.

4.2 Toepassing van big data als methodiek voor het analyseren van het financiële model

In het voorafgaande werd geconcludeerd dat de eisen aan het analyseren van het financiële model voornamelijk betrekking hebben op de risico's die voorspellingen in business cases met zich meebrengen. Als antwoord op deze risico's werd in paragraaf 3 de big data-toepassing *simulation and risk* genoemd, bestaande uit *sensitivity*- en *probability*-analyses. Davenport and Harris (2007, p. 59) noemen voor het uitvoeren van *probability*-analyses de term Monte Carlo-simulatie: “*computerized technique used to assess the probability of certain outcomes or risks by mathematically modeling a hypothetical event over multiple trials and comparing the outcomes with predefined probability distributions*”. Dergelijke simulaties kunnen worden ingezet bij de analyse van financiële modellen in business cases. Naast *probability*-analyses zijn Monte Carlo-simulaties ook bruikbaar voor *sensitivity*-analyses (Ibrahim and Koch 2008).

Op het gebied van scenarioanalyses werden ook *what-if*-analyses genoemd. Aangezien zowel *sensitivity*- als *probability*-analyses worden uitgevoerd met behulp van scenarioanalyses zou *what-if*-analyse als big data-toepassing ook geschikt zijn. *What-if*-analyses wijken af van Monte Carlo-simulatie doordat bij *what-if* scenario's specifieke scenario's vooraf gespecificeerd worden om vervolgens te worden doorgerekend. Bij Monte Carlo-simulaties worden de inputvariabelen gemodelleerd waarna met een druk op de knop vele duizenden scenario's worden doorgerekend. Hierdoor kost het substantieel minder tijd om een statistisch verantwoorde analyse uit te voeren met Monte Carlo-simulaties dan met traditionele *what-if*-analyses.

De kwaliteit Monte Carlo-simulaties hangt sterk af van de kwaliteit van de vooraf gemodelleerde inputvariabelen. Voor iedere variabele wordt een kansverdeling oftewel distributie bepaald (Ibrahim and Koch 2008, p. 39). Dit kan handmatig worden ingesteld, wat dezelfde kwaliteitsrisico's met zich meebrengt als de aspecten die in paragraaf 3 zijn beschreven voor het bepalen van inputvariabelen. Zo kan men onvoldoende inzicht hebben in de bandbreedte of verdeling van variabelen, of kan de objectiviteit onder druk komen te staan wanneer men belangen heeft bij een bepaalde uitkomst. Het inzetten van visualisatie- en *data mining*-technieken kan hierbij helpen.

Het laatste aspect met betrekking tot de methodiek die voor de analyse van het financiële model wordt gehanteerd is dat de analyse objectief plaatsvindt en dat interpretatieverschillen tijdens en naar aanleiding van de analyse zo veel mogelijk worden voorkomen. Aangezien Monte Carlo-simulaties de mogelijkheid bieden om alle mogelijke uitkomsten door te rekenen, ook de uitkomsten die intuïtief niet als mogelijkheid werden gezien, is er sprake van een rationele en objectieve analyse (Ibrahim and Koch 2008). De objectiviteit valt of staat echter met de objectiviteit waarmee de inputvariabelen in het simulatiemodel worden gemodelleerd. Visualisatie kan hierbij een rol spelen. Door historische patronen van data te visualiseren, en de voorspelde waarden daarbij te presenteren, kan door de besluitvormers gemakkelijker worden bepaald of de voorspelling sterk afwijkt van het verleden waarna een discussie kan worden gevoerd of dit terecht is. Ook *data mining* kan bijdragen aan objectieve modellering van variabelen door onbekende patronen aan het licht te brengen waar men daarvoor – bewust of onbewust – nog niet aan had gedacht.

In tabel 2 worden de mogelijke toepassingsgebieden van big data om de kwaliteit van financiële prognoses te verbeteren ten behoeve van business cases schematisch weergegeven.

Tabel 2. Mogelijke toepassing van big data bij de totstandkoming van financiële prognoses in business cases.

		Visualisatie	Data mining	What-if	Simulation and risk	Machine learning
Input	Volledigheid		✓			
	Juistheid	✓	✓			
	Objectiviteit	✓	✓			
Analyse	Sensitivity analysis	✓	✓	✓	✓	
	Probability analysis	✓	✓	✓	✓	
	Objectiviteit	✓	✓		✓	

5. Onderzoeksmethode

Voor de casestudie zijn drie business cases geselecteerd die in het verleden bij Port of Amsterdam tot een positieve investeringsbeslissing hebben geleid. In alle drie de gevallen betreft het investeringen in haveninfrastructuur zoals steigers, kades en afmeerpalen met investeringsvolumes van €0,5 mln., €1 mln. en €26 mln. De investeringen hebben betrekking op verschillende toepassingsgebieden, variërend van de overslag van olie tot afmeerfaciliteiten voor cruiseschepen. De looptijd van de berekeningen is tussen de twintig en 25 jaar.

In de casestudie zijn de financiële prognoses van de business cases opnieuw berekend, maar nu door middel van de toepassing van big data-technieken. Bij de nieuwe prognoses is enkel gebruik gemaakt van data die op het moment van de oorspronkelijke berekening beschikbaar was.

Op basis van het literatuuronderzoek werd gesteld dat big data-technieken op twee onderdelen kunnen bijdragen aan betere financiële prognoses:

1. bij het bepalen van kwalitatief goede input en zo bijdragen aan de juistheid, volledigheid en objectiviteit;
2. als methodiek voor het analyseren van financiële prognoses door middel van simulatie (*sensitivity- en probability-analyses*) en door bij te dragen aan objectieve presentatie van de resultaten.

Beide onderdelen zijn in de casestudie afzonderlijk getoetst door voor ieder onderdeel een afzonderlijke financiële prognose te ontwikkelen. Dit betekent dat voor alle drie de geselecteerde business cases uiteindelijke de volgende financiële prognoses zijn gepresenteerd:

- de initiële financiële prognose;
- twee nieuwe financiële prognoses die zijn opgesteld door middel van toepassing van big data-technieken:
 - een prognose waarbij dezelfde statistische methode wordt gebruikt als bij de initiële prognose (deterministisch), maar waarbij de inputvariabelen door middel van big data-technieken zijn bepaald;
 - een prognose waarbij een andere statistische methode wordt gebruikt (Monte Carlo-simulatie; probabilistisch). Om deze methode te kunnen toepassen wordt voor iedere business case een simulatiemodel gebouwd. Bij het definiëren van de variabelen in deze modellen worden ook big data-technieken toegepast.

De analyses van de kwaliteit van de prognoses heeft plaatsgevonden op basis van de *Forecast Accuracy* (Van der Laan et al. 2016). De accuraatheid van de prognose wordt hierbij aangegeven met een percentage. Hoe dichter het percentage bij 100% ligt, hoe dichter de prognose bij de werkelijke realisatie ligt en hoe accurater de prognose is.

Na beëindiging van de projecten zou het mogelijk zijn geweest om de prognoses af te zetten tegen de daadwerkelijke realisatiegegevens over de gehele looptijd van het

project. Voor alle geselecteerde business cases geldt echter dat de looptijd waar in de initiële financiële prognose vanuit is gegaan nog niet is verstreken. Daarom zijn er twee analyses uitgevoerd. Ten eerste is de *forecast accuracy* bepaald door de prognoses van de beëindigde jaren af te zetten tegen de daadwerkelijke resultaten in die jaren. Daarnaast is ter controle een tweede analyse verricht waarbij de realisatiegegevens zijn aangevuld met prognoses waarbij rekening is gehouden met alle nu aanwezige kennis.

6. Resultaten van de casestudie

In deze paragraaf wordt eerst stilgestaan bij de totstandkoming van de twee nieuwe financiële prognoses met behulp van big data-toepassingen: allereerst bij het bepalen van input en vervolgens bij het analyseren van het financiële model. Vervolgens worden de analyseresultaten behandeld.

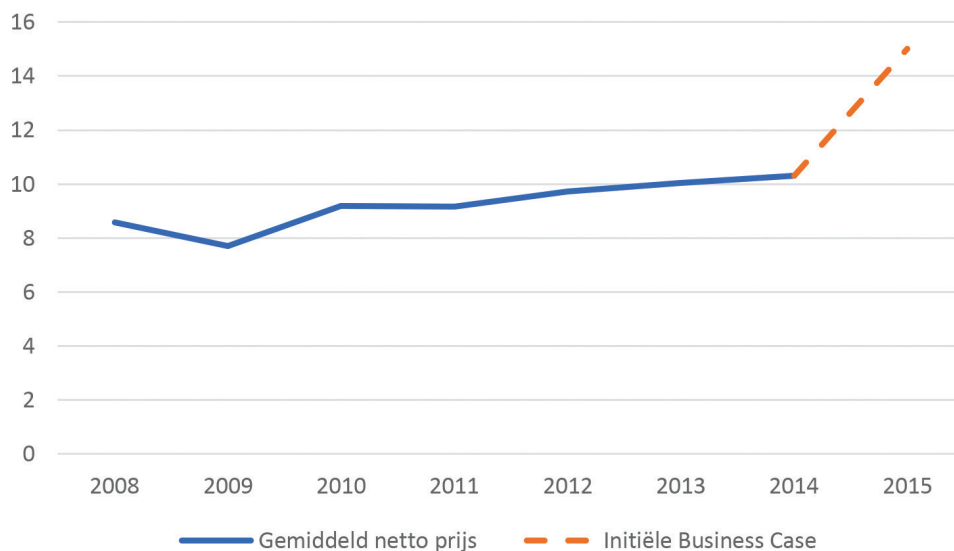
6.1 Toepassing van big data bij het bepalen van input

Op basis van de literatuur zijn twee big data-toepassingen benoemd die een bijdrage kunnen leveren aan het vaststellen van de input voor de financiële prognoses: visualisatie en *data mining*. In de casestudie zijn beide technieken toegepast.

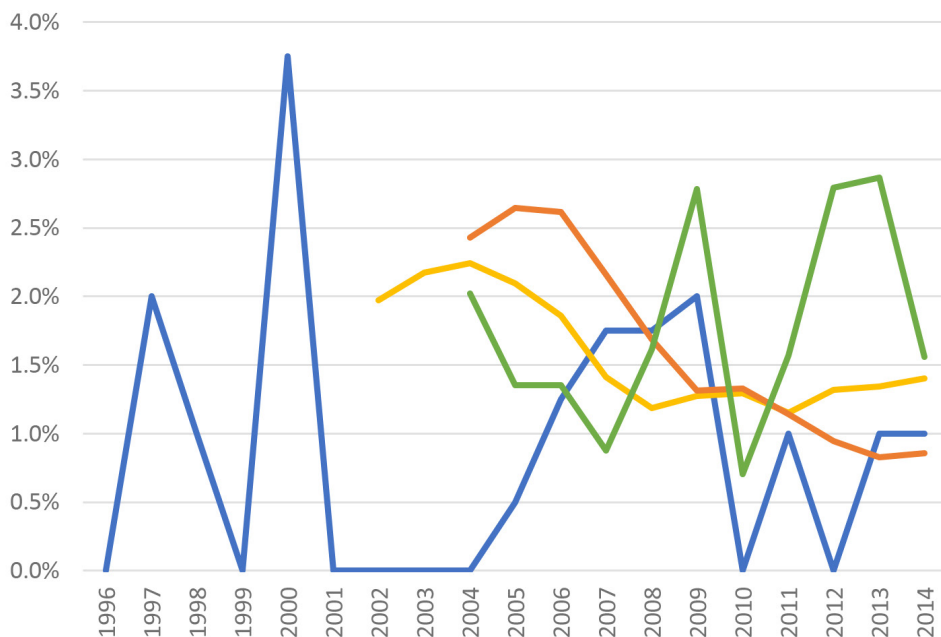
Visualisatie bleek een goed hulpmiddel om historische prijsontwikkelingen van inkomende kasstromen inzichtelijk te maken en zo te valideren of de prijs die men in de initiële business cases gebruikte – en dus door de betrokkenen als meest waarschijnlijk werd gezien – plausibel was. Figuur 1 toont een dergelijke (vereenvoudigde) visualisatie waaruit blijkt dat de historische prijsontwikkeling (de blauwe doorgetrokken lijn) flink lager ligt dan het gehanteerde prijsniveau waar in de initiële financiële berekening vanaf 2015 mee werd gerekend: het eerste jaar van de business case-berekening. De oranje stippellijn toont aan dat de sprong die moet worden gemaakt van de realisatie in 2014 naar het geschatte prijsniveau in 2015 niet realistisch lijkt. Door middel van visualisatie werd inzichtelijk dat de gehanteerde prijs in dit voorbeeld te hoog is en naar beneden moest worden bijgesteld. Visualisatie droeg in dit voorbeeld bij aan validatie van de input waarop de prijs in neerwaartse richting is bijgesteld.

In de initiële financiële prognoses werd voor alle inkomstenstromen een en hetzelfde indexatiepercentage gehanteerd. Door middel van visualisatie (figuur 2) – waarin iedere lijn het indexatiepercentage van een afzonderlijke inkomstenstroom representeert – werd inzichtelijk dat de historische ontwikkeling van de verschillende percentages sterk fluctueerde en van elkaar afweek. Om die reden is het niet terecht om voor alle inkomstenstromen van één indexatiepercentage uit te gaan.

In de nieuwe berekeningen is daarom allereerst niet meer uitgegaan van een en hetzelfde indexatiepercentage



Figuur 1. Vereenvoudigde visualisatie van de historische prijsontwikkeling en de sprong naar het prijsniveau waar in de initiële business case mee werd gerekend vanaf 2015 (de start van de business case).



Figuur 2. Visualisatie van de historische ontwikkeling van indexatiepercentages van diverse inkomstenstromen.

voor alle inkomstenstromen, maar zijn de indexaties per inkomstenstroom gesplitst. Daarnaast zijn door middel van eenvoudige statistische technieken op basis van historische indexatiepercentages gemiddelden en standaarddeviaties bepaald die in de nieuwe berekeningen zijn opgenomen.

Ook bij het bepalen van toekomstige onderhoudsuitgaven bleek *data mining* toepasbaar. Omdat de beschikbaarheid en bruikbaarheid van onderhoudsdata een beperkende factor bleek is slechts gebruik gemaakt van eenvoudige

statistische technieken. Toch hebben deze analyses geleid tot aanpassingen van de verwachte onderhoudsuitgaven.

Samenvattend zijn zowel visualisatie als *data mining*-technieken toegepast om de juistheid van de inputvariabelen te vergroten. De visualisaties hebben bijgedragen aan de validatie van de onderliggende aannames van de oorspronkelijke financiële berekening. Eenvoudige statistische technieken zijn toegepast om patronen in data te analyseren en zo de input te bepalen.

6.2 Toepassing van big data als methodiek voor het analyseren van financiële modellen

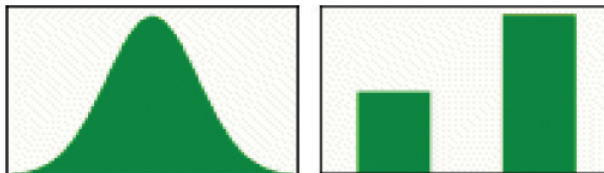
Het tweede deel van de casestudie is gericht op de analyse van de financiële berekeningen. Op basis van de literatuur is gesteld dat de toepassingsmogelijkheid van big data hierbij tweeledig is. Allereerst wordt onder de toepassing van big data verstaan het uitvoeren van *sensitivity*- en *probability*-analyses door middel van *simulation and risk*. Door voor iedere onzekere variabele een distributie te definiëren ontstaat een simulatiemodel dat softwarematig kan worden doorgekend. In de casestudie is gebruik gemaakt van simulatiesoftware van Oracle, genaamd Crystal Ball, waarmee zogenaamde Monte Carlo-simulaties kunnen worden uitgevoerd. De simulatiemodellen zijn uitgevoerd met $n = 10.000$ wat betekent dat de drie simulatiemodellen ieder 10.000 keer zijn gesimuleerd. Bij iedere simulatie wijst de software voor alle variabelen waar een distributie voor is gedefinieerd *at random* een waarde toe. Op die manier passeerden voor iedere business case 10.000 mogelijke uitkomsten van de financiële berekening de revue wat leidt tot een kansdistributie van het mogelijke rendement op de investering.

Ten tweede kunnen big data-technieken bijdragen aan de totstandkoming van de simulatiemodellen. In deze modellen wordt voor iedere variabele met een onzekere toekomstverwachting zogenaamde *probability distributions* of assumpties gedefinieerd: een reeks van mogelijke waarden, inclusief een kansverdeling van optreden van die mogelijke waarden. Afhankelijk van de variabele kan de distributie verschillen. Bij het definiëren van deze distributies wordt verondersteld dat big data-toepassingen visualisatie en *data mining* een rol kunnen spelen.

In figuur 3 zijn twee distributies die in de casestudies zijn gebruikt schematisch weergegeven. Wanneer er sprake is van een bekende, meest waarschijnlijke waarde en het waarschijnlijker is dat de onbekende variabele dicht bij dat punt ligt dan verder ervan af (circa 68% ligt binnen één standaarddeviatie van het gemiddelde), dan wordt de normaalverdeling toegepast (Oracle, z.d.). In de casestudies is voor een van de prijscomponenten gekozen voor deze distributie. Met behulp van eenvoudige *data mining*-technieken is op basis van historische data een gemiddelde en standaarddeviatie vastgesteld. Deze waarden hebben als input gediend bij het definiëren van de normaalverdelingen. De ja/nee-distributie wordt toegepast bij variabelen waarbij slechts twee waarden mogelijk zijn (ja of nee; kop of munt, etc.; Oracle, z.d.). In de casestudie is deze distributie toegepast bij de vraag of er in een specifiek jaar wel of geen onderhoud plaats zou vinden. Door middel van *data mining*-technieken is op basis van historische data bepaald hoe groot de kans is dat voor specifieke assets onderhoud noodzakelijk is. Op basis van die analyse zijn de ja/nee-distributies geprogrammeerd.

6.3 Analyseresultaten

Op basis van de uitkomsten van de berekeningen die in de casestudie zijn gemaakt, de initiële berekeningen en de werkelijke realisatie is een nauwkeurigheidsscore berekend: de *forecast accuracy* (FA). Bij een score van 100% is de prognose optimaal nauwkeurig: de prognose komt in dat geval volledig overeen met de werkelijke realisatie. Een FA-percentage van 95% betekent dat de prognose 5% lager is dan de werkelijke realisatie. Een



Figuur 3. Normaalverdeling en ja-nee distributie (Oracle, z.d.).

FA-percentage van 105% betekent dat de prognose 5% hoger is dan de werkelijke realisatie. In beide gevallen is de prognose 5% minder accuraat dan wanneer de realisatie perfect was voorspeld. Hoe dichter de score bij de 100% ligt, hoe nauwkeuriger – en daarmee beter – de financiële prognose is.

In de tabellen 3 en 4 wordt per business case de *forecast accuracy* vergeleken tussen de initiële financiële berekening, de berekeningen waarbij de input door middel van big data-technieken opnieuw is bepaald en de berekeningen die door middel van simulatie zijn uitgevoerd. Tabel 3 toont de berekeningen waarbij alleen de periode in de berekening is betrokken die achter ons ligt; tabel 4 toont de berekeningen waarbij ook de jaren in de business cases die nog niet zijn gepasseerd, worden betrokken in de berekening.

De berekeningen waarvan het rendement het dichtst bij de werkelijke realisatie ligt zijn groen gekleurd, de tweede oranje en de waarde die het minst nauwkeurig is ten opzichte van de werkelijke realisatie rood.

De volgende resultaten vallen op:

- In vijf van de zes gevallen is het resultaat van de simulatieberekeningen het meest nauwkeurig met maximaal 25% afwijking van de werkelijke realisatie. Het verschil bij business case 2 waarin de simulatieberekening niet het meest nauwkeurig is, bedraagt slechts 1%.
- De rendementen van de simulaties waarbij alleen de jaren zijn gebruikt die achter ons liggen (tabel 3), zijn nauwkeuriger dan die waarbij de volledige looptijd van de business case is meegenomen door middel van geactualiseerde voorspellingen (tabel 4).
- De verschillen in nauwkeurigheid van de initiële berekeningen en de berekeningen waarbij input met behulp van big data-technieken tot stand is gekomen, zijn beperkt.

Het toepassen van big data-technieken in business cases-berekeningen blijkt bij twee van de drie business cases te hebben geleid tot een nauwkeuriger berekening.

Tabel 3. Forecast accuracy vergelijkingen van het rendement op de drie business cases.

Rendement: Alleen realisatie			
	Initieel	Input d.m.v. big data	Input en analyse d.m.v. big data (simulatie)
Business Case 1	-5%	-24%	110%
Business Case 2	107%	106%	107%
Business Case 3	97%	98%	99%

Business case 2 toont dusdanig kleine variaties in de *forecast accuracy*-percentages dat kan worden gesteld dat de toepassing van big data niet heeft geleid tot betere business case-berekeningen.

De berekeningen waarbij gebruik is gemaakt van big data-technieken zijn overwegend nauwkeuriger dan de initiële berekening. Hierbij zijn de simulatiemodellen het meest nauwkeurig, gevolgd door de berekeningen waarbij de input opnieuw is bepaald door middel van visualisatie en *data mining*. Deze bevindingen komen voort uit business cases 1 en 3.

Om de validiteit van het onderzoek te borgen zijn naast de vergelijking van totaalrendementen tussen de verschillende berekeningen ook analyses gemaakt op de verschillende onderliggende kasstromen van de berekeningen. Onderliggende *forecast accuracy*-berekeningen van investeringsuitgaven, inkomsten en exploitatie-uitgaven lieten hetzelfde beeld zien als de *forecast accuracy* van de rendementen. De componenten in de derde berekening zijn overwegend het meest accuraat, gevolgd door de componenten waarvan de input door middel van big data-toepassingen zijn bepaald.

7. Conclusie en beperkingen van het onderzoek

7.1 Conclusie

In dit artikel is onderzocht of big data kan bijdragen aan betere financiële prognoses ten behoeve van business cases.

Door middel van literatuurstudie is bepaald welke eisen aan business case-berekeningen worden gesteld om bruikbaar te zijn in de besluitvorming. Daarnaast zijn big data-technieken in kaart gebracht en is bepaald welke van die technieken kunnen bijdragen aan de kwaliteit van financiële prognoses in business cases. Vervolgens is door middel van een casestudie in de praktijk getoetst of die big data-technieken daadwerkelijk leiden tot nauwkeuriger – en daarmee betere – financiële prognoses.

De volgende aspecten zijn van belang voor de kwaliteit van financiële prognoses in business cases:

- de input moet juist en volledig zijn en objectief tot stand komen;

Tabel 4. Forecast accuracy vergelijkingen van het rendement op de drie business cases.

Rendement: Alleen realisatie			
	Initieel	Input d.m.v. big data	Input en analyse d.m.v. big data (simulatie)
Business Case 1	77%	73%	90%
Business Case 2	80%	80%	81%
Business Case 3	518%	313%	125%

- bij de analyse van het financiële model moet rekening worden gehouden met de onzekerheid van de uitkomsten en moeten resultaten objectief worden gepresenteerd.

De casestudie laat zien hoe bij Port of Amsterdam big data een bijdrage zou kunnen hebben geleverd aan de kwaliteit van financiële prognoses. Door middel van visualisatie en *data mining* zijn onder andere prijsniveaus geobjectiveerd, toekomstige onderhoudsuitgaven bepaald en is er onderscheid gemaakt in indexatie van verschillende soorten uitgaven. Bij de analyse van de resultaten zijn Monte Carlo-simulaties toegepast die inzichtelijk hebben gemaakt hoe waarschijnlijk de resultaten van de berekeningen zijn en welke variabelen het grootste effect hebben op de uitkomsten. Om deze simulaties te kunnen uitvoeren zijn simulatiemodellen gebouwd. Visualisatie en *data mining* hebben een belangrijke rol gespeeld bij het definiëren van de variabelen in deze modellen.

De resultaten laten zien dat bij twee van de drie business cases de genoemde big data-technieken inderdaad een positief effect hebben op de nauwkeurigheid van de prognoses. De simulatieberekeningen waarbij de inputvariabelen zijn bepaald door middel van en waarbij de analyse van het financiële model heeft plaatsgevonden met behulp van big data-technieken zijn het meest nauwkeurig. De berekeningen waarbij de inputvariabelen door middel van big data zijn bepaald en waarbij de wijze waarop de analyse plaatsvindt onveranderd is, volgen op de tweede plaats.

7.2 Beperkingen van het onderzoek

Dit onderzoek kent een aantal beperkingen. Allereerst is het aantal van drie business cases te klein om significante conclusies te trekken over de bruikbaarheid van big data-toepassingen bij het verbeteren van business case-berekeningen. Het verdient daarom aanbeveling om meer casestudies uit te voeren. Het zou interessant zijn om hierbij andere organisaties en investeringen te betrekken zodat de resultaten breder toepasbaar zijn.

Bij de totstandkoming van de nieuwe business case-berekeningen met behulp van big data-toepassingen is enkel gebruik gemaakt van interne data: data die beschikbaar was binnen Port of Amsterdam. Brancherelevante macro- en micro-economische data heeft enorme potentie om de business case-berekeningen verder te verrijken en nauwkeuriger te maken. Aanbevolen wordt om vervolgonder-

zoek te richten op de toegevoegde waarde van externe data in business case-berekeningen.

Tot slot is het beperkte aantal jaren met realisatiedata in de casestudie een beperking voor het onderzoek. Idealiter zou de volledige looptijd van de oorspronkelijke business case zijn verstreken waardoor er voor de gehele periode

realisatiedata beschikbaar is en nauwkeurigheid van de berekeningen daarmee kan worden vergeleken. Dit was echter bij de drie geselecteerde business case-berekeningen niet het geval. Het verdient aanbeveling om deze analyse nogmaals uit te voeren zodra er business cases beschikbaar zijn waarvan de gehele looptijd is verstreken.

-
- **Drs. B.T.G. van Raaij RC** is coördinator Business Control bij Havenbedrijf Amsterdam nv en studeerde in 2017 af als Executive Master of Finance and Control aan TIAS School for Business and Society, Tilburg University. Dit artikel is gebaseerd op zijn afstudeerscriptie.
 - **Prof. dr. A.C.N. van de Ven RA** is hoogleraar Bestuurlijke Informatieverzorging aan TIAS, School for Business and Society, Tilburg University en begeleidde Bas bij zijn afstudeeronderzoek.
-

Literatuur

- Alkaraan F, Northcott D (2006) Strategic capital investment decision-making: A role for emergent analysis tools? A study of practice in large UK manufacturing companies. *British Accounting Review* 38(2): 149–173. <https://doi.org/10.1016/j.bar.2005.10.003>
- Braaksma J, Commandeur A, Berghout EW (2006) The business case for (ICT) investment evaluation in nonprofit organisations. The 13th European Conference on Information Technology Evaluation (pp. 1–13). https://aisel.aisnet.org/sprouts_all/178/
- Bughin J, Chui M, Manyika J (2010) Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly*, August: 75–86. <https://www.mckinsey.com/industries/high-tech/our-insights/clouds-big-data-and-smart-assets-ten-tech-enabled-business-trends-to-watch>
- Chen H, Chiang RHL, Storey VC (2012) Business intelligence and analytics: From big data to big impact. *MIS Quarterly* 36(4): 1165–1188. <https://dl.acm.org/citation.cfm?id=2481683>
- Committee of Sponsoring Organizations (2017) Enterprise Risk Management—Integrating with strategy and performance. www.coso.org
- Davenport TH, Harris JG (2007) *Competing on analytics: The new science of winning*. Harvard Business Press.
- Evans JR, Lindner CH (2012) Business analytics : The next frontier for decision sciences. *Decision Line* 43(2): 4–6.
- George G, Haas MR, Pentland A (2014) From the editors: Big data and management. *Academy of Management Journal* 57(2): 321–326. <https://doi.org/10.5465/amj.2014.4002>
- Haas PJ, Maglio PP, Selinger PG, Tan W-C (2011) Data is dead... Without what-if models. *Proceedings of the VLDB Endowment* 4(12): 1486–1489. <http://www.vldb.org/pvldb/vol4/p1486-haas.pdf>
- Heusinkveld B (2014) Analyse van relevante aandachtspunten bij de toepassing van Big Data vanuit het IT-audit perspectief. Vrije Universiteit Amsterdam.
- Ibrahim S, Koch A (2008) Risicobeheer met Monte-Carlo technieken. *Tijdschrift Controlling* 2008(12): 38–41. <https://cmweb.nl/2008/12/risicobeheer-met-monte-carlo-technieken/>
- Katal A, Wazid M, Goudar RH (2013) Big data: Issues, challenges, tools and good practices. 2013 6th International Conference on Contemporary Computing, IC3 2013 (pp. 404–409). <https://ieeexplore.ieee.org/document/6612229/>
- Klayman J (1995) Varieties of confirmation bias. *Psychology of Learning and Motivation* 32: 385–418. [https://doi.org/10.1016/S0079-7421\(08\)60315-1](https://doi.org/10.1016/S0079-7421(08)60315-1)
- Kohavi R, Rothleder NJ, Simoudis E (2002) Emerging trends in business analytics. *Communications of the ACM* 45(8): 45–48. <https://doi.org/10.1145/545151.545177>
- Laan E van der, Dalen J van, Rohrmoser M, Simpson R (2016) Demand forecasting and order planning for humanitarian logistics: An empirical assessment. *Journal of Operations Management*, 45, 114–122. <https://doi.org/10.1016/j.jom.2016.05.004>
- Leeuwen OC van, Bergsma JBT (2014) *Algemene grondslagen Starreveld*. Noordhoff Uitgevers bv (Groningen / Houten).
- Marr B (2015) Big data: using smart big data analytics and metrics to make better decisions and improve performance. John Wiley & Sons.
- Messner W (2013) *Making the compelling business case*. Palgrave Macmillan (New York).
- Oracle (z.d.). Normal Distribution. Geraadpleegd op https://docs.oracle.com/cd/E17236_01/epm.1112/cb_user/frameset.htm?normal_distribution.htm
- Oracle (z.d.). Yes-No Distribution. Geraadpleegd op https://docs.oracle.com/cd/E17236_01/epm.1112/cb_user/yes_no_distribution.htm
- Provost F, Fawcett T (2013) Data science and its relationship to big data and data-driven decision making. *Big Data* 1(1), 51–59. <https://doi.org/10.1089/big.2013.1508>
- Schmidt MJ (2005) *Business case essentials: a guide to structure and content*.
- Stratton MJ (2004) *Business case development guideline*. Crystal Ball User Conference (pp. 1–24).