

Eindrapportage

# *Audit prestatie-indicatoren vervoer- en beheerconcessie*



**Andersson Elffers Felix**

Maliebaan 16  
Postbus 85198  
3508 AD Utrecht

+31 30 236 30 30  
mail@aef.nl  
www.aef.nl

Kamer van Koophandel  
30096560

 **BAKER TILLY  
BERK**

**Datum**

30 november 2016

**Opdrachtgever**

Ministerie van Infrastructuur en Milieu

**Contact**

Robert-Jaap Voorn

**Referentie**

GI72/Eindrapportage

## *Inhoud*

<b>Managementsamenvatting</b>	<b>4</b>
<b>1 Inleiding</b>	<b>8</b>
1.1 Achtergrond .....	8
1.2 Scope van de audit .....	9
1.3 Aanpak.....	11
1.4 Leeswijzer.....	11
<b>2 Prestatie-indicatoren NS</b>	<b>12</b>
2.1 Zitplaatskans reizigers in spits HRN.....	12
2.2 Top 10 reizen met de slechtste vervoerplaatskans in de spits.....	18
2.3 Zitplaatskans reizigers in spits HSL.....	20
<b>3 Prestatie-indicatoren ProRail</b>	<b>29</b>
3.1 Klantoordeel reizigersvervoerders.....	29
3.2 Klantoordeel goederenvervoerders.....	33
3.3 Klantinder als gevolg van storingen infra.....	36
<b>4 Gezamenlijke prestatie-indicatoren</b>	<b>42</b>
4.1 Reizigerspunctualiteit < 5 minuten HRN .....	42
4.2 Reizigerspunctualiteit < 15 minuten HRN .....	48
4.3 Reizigerspunctualiteit < 5 minuten HSL.....	49
4.4 Reizigerspunctualiteit < 15 minuten HSL.....	52
<b>5 Aanvullende aandachtspunten</b>	<b>54</b>
5.1 Opvolging eerdere generieke aanbevelingen .....	54
5.2 Generieke aanbevelingen volgend uit huidige audit.....	56
<b>Bijlage 1. Overzicht aanbevelingen</b>	<b>57</b>
<b>Bijlage 2. Gesprekspartners</b>	<b>63</b>
<b>Bijlage 3. Brondocumenten</b>	<b>65</b>
<b>Bijlage 4. Beoordelingskader</b>	<b>68</b>

## Managementsamenvatting

### Inleiding

Het ministerie van Infrastructuur en Milieu (IenM) heeft het voornemen om per 2017 een aantal prestatie-indicatoren bij te stellen en toe te voegen. Het doel hiervan is beter aan te sluiten bij de klantbeleving en metingen betrouwbaarder te maken. Om de aanpassingen in de prestatie-indicatoren te toetsen, heeft het ministerie van IenM aan Baker Tilly Berk (BTB) verzocht een onafhankelijke audit op de doorontwikkelde prestatie-indicatoren uit te voeren.

Deze audit richt zich op de volgende prestatie-indicatoren uit de vervoer- en beheerconcessie:

- Zitplaatskans voor reizigers in de spits op het HRN (NS)
- Top 10 reizen met slechtste vervoerplaatskans in spits (NS)
- Zitplaatskans voor reizigers in de spits op de HSL (NS)
- Klantoordeel van reizigersvervoerders (ProRail)
- Klantoordeel van goederenvervoerders (ProRail)
- Klanthinder als gevolg van storingen van de infrastructuur (ProRail)
- Reizigerspunctualiteit < 5 minuten voor het HRN (Gezamenlijk<sup>1</sup>)
- Reizigerspunctualiteit < 15 minuten voor het HRN (Gezamenlijk)
- Reizigerspunctualiteit < 5 minuten voor de HSL (Gezamenlijk)
- Reizigerspunctualiteit < 15 minuten voor de HSL (Gezamenlijk)

Prestatie-indicatoren die in de vervoer- en beheerconcessie gebruikt worden, dienen te voldoen aan de volgende voorwaarden:

- De publieksdefinities en meetmethoden dienen helder, concreet en eenduidig interpreteerbaar beschreven te zijn
- De meetpraktijk dient overeen te komen met de beschreven meetmethoden en de metingen worden juist gerapporteerd
- De metingen dienen zodanig betrouwbaar te zijn dat een vergelijking over meerdere jaren mogelijk is.
- Er dient sprake te zijn van betekenisvolle bodem- en streefwaarden, waartegen de gemeten waarden kunnen worden afgezet.

Op basis van voornoemde voorwaarden zijn door het ministerie van IenM vragen per prestatie-indicator geformuleerd. In bijlage 4 bij deze rapportage is beschreven welke criteria BTB/AEF hanteert om de vragen te beantwoorden.

---

<sup>1</sup> De proces-verantwoordelijkheid voor de gezamenlijke prestatie-indicatoren is ondergebracht bij NS. Dit betekent onder meer dat NS de data verzameld en de score op de prestatie-indicator berekend.

In deze managementsamenvatting beschrijven we op generiek niveau de belangrijkste bevindingen en aanbevelingen. In bijlage 1 van deze rapportage is een overzicht van de aanbevelingen opgenomen.

### **Belangrijkste bevindingen en aanbevelingen**

De definities zijn op enkele aandachtspunten na voldoende helder, concreet en eenduidig beschreven. Belangrijk is dat de definities zijn afgebakend in zowel tijd als scope. Voor alle prestatie-indicatoren geldt dat de beknopte beschrijving van het meetsysteem bijdraagt aan de meetbaarheid van de definitie. Enkele aanbevelingen in dit verband zijn:

- Top 10 reizen met slechtste vervoerplaatskans in spits: Neem in het eerste deel van de publieksdefinitie op hoe de top-10 wordt samengesteld.
- Klantoordeel reizigersvervoerders: Maak helder wat de ondergrens is voor de vervoerders die meegenomen worden in het onderzoek en verwerk deze in de publieksdefinitie.

De meetmethodiek is over het algemeen nauwkeurig en compleet beschreven en voldoende gevalideerd. De onderverdeling in (proces)stappen van de meetmethoden voor de gezamenlijke prestatie-indicatoren en die van NS is verhelderd. De steekproeven die zijn gebruikt voor deze prestatie-indicatoren zijn een voldoende afspiegeling van de totale populatie reizigers in de spits. Tevens heeft NS verschillende tests met het meetsysteem gedraaid, waardoor fouten zijn opgespoord en hersteld. ProRail moet met name de wijze waarop de scores op de prestatie-indicatoren klantoordeel tot stand komen gedetailleerder inzichtelijk maken. Ook kan de representativiteit van deze indicatoren worden verbeterd. Voor zowel NS als ProRail zijn verschillende aanbevelingen gedaan om (de vastlegging van) de meetmethoden te verbeteren, bijvoorbeeld:

- Prestatie-indicatoren NS en Gezamenlijk: Leg de aannames bij de meetmethoden in hetzelfde document vast als de beschrijving van de meetmethoden of zorg als alternatief voor centrale vastlegging van beide.
- Klantoordeel goederenvervoerders en Klantoordeel reizigersvervoerders: Leg beter vast hoe de ingevulde vragenlijsten verwerkt worden tot een score op de prestatie-indicator.
- Klantoordeel goederenvervoerders en Klantoordeel reizigersvervoerders: Borg de representativiteit van het klantoordeel door in de instructie voor het invullen van de vragenlijst meer richtlijnen op te nemen en de wegging van vervoerders te herijken.

BTB/AEF constateert dat de meetpraktijk bij NS - inzake de prestatie-indicatoren Vervoercapaciteit en Reizigerspunctualiteit - op grote lijnen overeenkomt met de beschreven meetmethodiek. NS voert voldoende controles uit op de meetpraktijk in het algemeen, en de kwaliteit van de (bron)data in het bijzonder. Medewerkers hebben oog voor de risico's en sturen actief bij om deze te beheersen. Onderhavige audit heeft echter wel kwetsbare elementen aan het licht gebracht waarvoor aanbevelingen zijn gedaan. Deze overlappen deels met de uitkomsten van een risicoanalyse die NS medio 2016 zelf heeft uitgevoerd. Het betreft onder meer de volgende aanbevelingen:

- Prestatie-indicatoren NS en Gezamenlijk: Leg de interne en externe aanlevering van, de (kwaliteit)controle op en de verwerking van de brondata vast in een rapportagecyclus.
- Prestatie-indicatoren NS en Gezamenlijk: Evalueer de meetpraktijk halfjaarlijks met alle betrokkenen om het lerend vermogen te vergroten en formaliseer deze evaluatie.
- Prestatie-indicatoren NS en Gezamenlijk: Breng de onzekerheden in het meetsysteem ook kwantitatief in beeld en leg deze in een document vast. Leg daarnaast ook de belangrijkste risico's en de bijbehorende beheersmaatregelen vast.

ProRail heeft de meetpraktijk inzake de prestatie-indicator Klanthinder op een aandachtspunt na goed op orde. Metingen vinden plaats conform de beschreven

meetmethode. Er worden periodiek steekproeven gedaan om de kwaliteit van de registratie van storingen te bepalen. De kwaliteit van de meetgegevens wordt goed gemonitord.

BTB/AEF ziet mogelijkheden voor verbetering bij de prestatie-indicatoren Klantoordeel. De risico's en beheersmaatregelen van het onderzoek zijn in beeld maar zijn niet centraal vastgelegd en de kwaliteitsborging is deels gedocumenteerd.

- Klantoordeel goederenvervoerders en Klantoordeel reizigersvervoerders: Stel een risicomatrix op waarin de risico's van het onderzoek beschreven zijn, evenals de maatregelen om deze te beheersen.
- Klantoordeel goederenvervoerders en Klantoordeel reizigersvervoerders: Maak in de rapportage van het onderzoek inzichtelijk door hoeveel respondenten de vragen (generiek en specifiek) beantwoord zijn.
- Klantoordeel goederenvervoerders: Monitor veranderingen in marktaandeel van de goederenvervoerders en pas hierop indien nodig jaarlijks de weging aan.

NS en ProRail hebben over het geheel genomen de berekening en de opbouw van de bodem- en streefwaarden voldoende inzichtelijk gemaakt. De gebruikte datasets zijn eveneens getoetst. BTB/AEF onderschrijft de wijzen waarop de nieuwe bodem- en streefwaarden zijn bepaald en bevestigt dat de uitkomsten van de meeste prestatie-indicatoren - waar dat gewent was - een ambitieneutrale doorvertaling zijn. De prestatie-indicatoren Zitplaatskans HSL, Reizigerspunctualiteit < 5 min HSL en Klantoordeel reizigersvervoerders verdienen in dit verband bijzondere aandacht. De vastlegging is voorts nog een algemeen aandachtspunt. De belangrijkste aanbevelingen zijn de volgende:

- Zitplaatskans reizigers in de spits HSL: Bepaal op basis van de conclusies van de review op het Verbeterplan IC direct met het ministerie van IenM in welke mate de door NS en ProRail voorgestelde verbeter-maatregelen haalbaar zijn en aanleiding geven tot aanscherping van de door NS voorgestelde streefwaarde van 94,4%. Is de conclusie van de review bijvoorbeeld dat 6% uitval haalbaar is, verhoog de streefwaarde dan met 0,5%.
- Reizigerspunctualiteit < 5 min. HSL: Monitor de geïmporteerde punctualiteit uit Brussel en stel vast wanneer deze een te negatieve invloed heeft op de realisatiewaarde. Bepaal in samenspraak met het ministerie van IenM welke consequenties dit heeft gehad op de gerealiseerde prestaties en of dit aanleiding geeft om de bodem- en/of streefwaarde aan te passen.
- Klantoordeel van reizigersvervoerders: Werk voor de streefwaarde met een 'afronding' op halve punten en stel de streefwaarde vast op 7,5 om daarmee voldoende ambitie in te bouwen.

### **Vergelijking met vorig jaar**

Ten opzichte van de generieke aanbevelingen van vorig jaar hebben NS en ProRail de inhoud en het gebruik van de prestatie-indicatoren verder verbeterd:

- Het beeld uit de audit is dat de meetpraktijk binnen NS en ProRail goed wordt gemonitord. De uitkomsten van interne monitoring kunnen nog beter inzichtelijk worden gemaakt.
- Er bestaat nog onduidelijkheid in de onderlinge relatie tussen ProRail en NS over wie waarvoor verantwoordelijk is in het geval sprake is van de onderlinge doorlevering van gegevens om de score van een prestatie-indicator te bepalen. Binnen NS is een traject gestart om de betrouwbaarheid van de meetgegevens te verbeteren waarin dit punt wordt meegenomen.
- De interne vastlegging van de prestatie-indicatoren ten opzichte van de voorgaande audit is verbeterd. De vastlegging van de bodem- en streefwaarde is een aandachtspunt.

- Over het detailniveau van de afspraken over de prestatie-indicatoren tussen het ministerie van IenM enerzijds en NS en ProRail anderzijds zijn inmiddels heldere afspraken gemaakt.

# 1 Inleiding

## 1.1 Achtergrond

In de organisatie van het spoorvervoer en -beheer in Nederland wordt gebruikgemaakt van concessies: de Nederlandse Staat is concessieverlener en de Nederlandse Spoorwegen (NS) en ProRail zijn concessiehouder voor respectievelijk het spoorvervoer op het Hoofdrailnet (HRN) inclusief de Hoge Snelheidslijn (HSL) en het beheer van het spoorwegnet voor zover niet in particulier bezit. Om te sturen op de dienstverlening gedurende de looptijd van de vervoer- en de beheerconcessie werken de concessieverlener en -houders met prestatie-indicatoren. Monitoring aan de hand van de afgesproken prestatie-indicatoren stelt de Nederlandse Staat in de gelegenheid om toe te zien op deugdelijk en kwalitatief hoogstaand spoorvervoer en -beheer in Nederland.

Prestatie-indicatoren die in de vervoer- en beheerconcessie gebruikt worden, dienen te voldoen aan de volgende voorwaarden:

- de publieksdefinities en meetmethoden dienen helder, concreet en eenduidig interpreteerbaar beschreven te zijn
- de meetpraktijk dient overeen te komen met de beschreven meetmethoden en de metingen worden juist gerapporteerd
- de metingen dienen zodanig betrouwbaar te zijn dat een vergelijking over meerdere jaren mogelijk is.

Daarnaast dient er sprake te zijn van betekenisvolle bodem- en streefwaarde, waartegen de gemeten waarden kunnen worden afgezet.

Het ministerie van Infrastructuur en Milieu (IenM) heeft het voornemen om per 2017 een aantal prestatie-indicatoren bij te stellen en toe te voegen. Het doel hiervan is beter aan te sluiten bij de klantbeleving en metingen betrouwbaarder te maken. Er is gekomen tot nieuwe publieksdefinities, aangepaste meetmethoden en daar waar mogelijk een ambitieneutrale omzetting van bodem- en streefwaarde. Om de aanpassingen in de prestatie-indicatoren te toetsen, heeft het ministerie van IenM aan Baker Tilly Berk (BTB) verzocht een onafhankelijke audit op de doorontwikkelde prestatie-indicatoren uit te voeren. Deze audit heeft betrekking op drie prestatie-indicatoren uit de vervoerconcessie van NS, drie prestatie-indicatoren uit de beheerconcessie van ProRail en vier gezamenlijke prestatie-indicatoren.

Voor het uitvoeren van de audit heeft BTB samenwerking gezocht met Andersson Elffers Felix (AEF). De uitkomsten van de audit zijn opgenomen in de voorliggende rapportage.



## 1.2 Scope van de audit

### 1.2.1 Vraagstelling

In de audit dient per prestatie-indicator een aantal vragen te worden beantwoord. Deze vragen verschillen per indicator. Opgeteld gaat het om de volgende vragen:

- a. Zijn de publieksdefinities helder, concreet en eenduidig interpreteerbaar beschreven?
- b. Zijn de meetmethoden helder, concreet en eenduidig interpreteerbaar beschreven?
- c. Zijn de aangepaste bodem- en streefwaarde een ambitieneutrale doorvertaling van de bodem- en streefwaarde in de concessie bijlage 2? Is de gekozen methode van ambitieneutrale omzetting voldoende navolgbaar en statistisch voldoende betrouwbaar? Specifiek ten aanzien van de prestatie-indicatoren Zitplaatskans HRN en negatieve uitschieters wordt naast een beoordeling van de toegepaste methode gevraagd of er een of meerdere betere methoden zijn dan de toegepaste methode en zo ja, welke waarden de toepassing daarvan oplevert.
- d. Voor ProRail is het voornemen de informatie indicator klanthinder als gevolg van storingen aan de infra om te zetten naar prestatie-indicator. Bevatten de bodem- en streefwaarde voldoende ambitie t.o.v. de tot nu toe gerealiseerde data? Volgt de totstandkoming van de bodem- en streefwaarde logisch uit de sturingsmogelijkheden van ProRail? Klantoordeel goederenvervoerders: bevat de per 2016 aangepaste streefwaarde voldoende ambitie t.o.v. de tot nu toe gerealiseerde data? Is de gekozen methode m.b.t. de totstandkoming van de streefwaarde voldoende navolgbaar en statistisch betrouwbaar?
- e. Komt de meetpraktijk overeen met de beschrijving van de meetsystemen (conformiteit) en wordt er gerapporteerd wat er gemeten wordt (juistheid)?
- f. Is de informatie afkomstig uit de meetsystemen voldoende representatief en statistisch betrouwbaar om een vergelijking tussen verschillende kalenderjaren mogelijk te maken, en is de kwaliteitscontrole van de brongegevens voldoende geborgd?
- g. Zijn de - in het kader van deze opdracht relevante - aanbevelingen uit de rapportage van BTB/AEF 'Audit prestatie-indicatoren Vervoer- en Beheerconcessie' van 16 oktober 2015 en uit de rapportage van KPMG 'Audit prestatie-indicatoren voortvloeiende uit vervoer- en beheerconcessie' van 17 september 2014 door NS en ProRail overgenomen of zijn er in elk geval stappen in gezet?
- h. Belangrijk onderdeel van deze auditopdracht is het geven van een expert judgement over de door NS in beeld gebrachte sturingsmogelijkheden en aannames en het beoordelen of NS mogelijkheden over het hoofd heeft gezien om te drukke treinen op de HSL te voorkomen. M.b.t. de prestatie-indicator zitplaatskans HSL dient te worden beoordeeld of de voorgestelde bodem- en streefwaarde zowel vanuit historisch perspectief alsook gebaseerd op recente realisaties, sturingsmogelijkheden van NS en te verwachten toekomstige ontwikkelingen haalbaar en ambitieus zijn. Zo nee, wat is dan een realistische en ambitieuze bodem- en streefwaarde?
- i. Is de voorgestelde bodem- en streefwaarde zowel vanuit historisch perspectief alsook gebaseerd op recente realisaties, sturingsmogelijkheden van NS en ProRail en te verwachten toekomstige ontwikkelingen haalbaar en ambitieus?

In bijlage 4 bij deze rapportage is beschreven welke criteria BTB/AEF hanteert om de vragen te beantwoorden.

## 1.2.2 Prestatie-indicatoren

### NS

Voor NS richt de audit zich op de volgende drie prestatie-indicatoren:

- zitplaatskans voor reizigers in de spits op het HRN
- top 10 reizen met slechtste vervoerplaatskans in spits
- zitplaatskans voor reizigers in de spits op de HSL.

De per prestatie-indicator te beantwoorden vragen zijn opgenomen in de onderstaande tabel.

Prestatie-indicatoren NS	a	b	c	d	e	f	g	h	i
Zitplaatskans reizigers in spits HRN	■	■	■		■	■	■		
Top 10 reizen met slechtste vervoerplaatskans in spits	■	■	■		■	■	■		
Zitplaatskans reizigers in spits HSL	■	■			■	■	■	■	

### ProRail

Voor ProRail richt de audit zich op de volgende drie prestatie-indicatoren:

- klantoordeel van reizigersvervoerders
- klantoordeel van goederenvervoerders
- klanthinder als gevolg van storingen van de infrastructuur.

De per prestatie-indicator te beantwoorden vragen zijn opgenomen in de onderstaande tabel.

Prestatie-indicatoren ProRail	a	b	c	d	e	f	g	h	i
Klantoordeel reizigersvervoerders	■	■	■		■	■	■		
Klantoordeel goederenvervoerders	■	■		■	■	■	■		
Klanthinder o.g.v. storingen infra	■	■		■	■	■	■		

### Gezamenlijk

Tot slot is er een viertal gezamenlijke prestatie-indicatoren waar de audit zich op richt:

- reizigerspunctualiteit < 5 minuten voor het HRN
- reizigerspunctualiteit < 15 minuten voor het HRN
- reizigerspunctualiteit < 5 minuten voor de HSL
- reizigerspunctualiteit < 15 minuten voor de HSL.

De per prestatie-indicator te beantwoorden vragen zijn opgenomen in de onderstaande tabel.

Prestatie-indicatoren ProRail	a	b	c	d	e	f	g	h	i
Reizigerspunctualiteit < 5 min. HRN	■	■	■		■	■	■		
Reizigerspunctualiteit < 15 min. HRN	■	■	■		■	■	■		
Reizigerspunctualiteit < 5 min. HSL	■	■	■		■	■	■		
Reizigerspunctualiteit < 15 min. HSL	■	■			■	■	■		■

### **1.3 Aanpak**

De audit is in samenspraak met het ministerie van IenM in vier stappen uitgevoerd:

- In de eerste stap is een beoordelingskader uitgewerkt voor de beantwoording van de bovenstaande vragen. Een beschrijving van dit beoordelingskader is opgenomen in een bijlage bij deze rapportage.
- Vervolgens is in de tweede stap een analyse gemaakt van de brondocumenten die door NS en ProRail beschikbaar zijn gesteld. Een overzicht van deze documenten is eveneens in een bijlage bij deze rapportage opgenomen.
- In de derde stap hebben verdiepende gesprekken plaatsgevonden met medewerkers van NS en ProRail. Een overzicht van de gesprekspartners is ook als bijlage bij deze rapportage opgenomen.
- In de vierde stap is een synthese gemaakt van de bevindingen en zijn op basis hiervan conclusies en aanbevelingen geformuleerd.

De stappen zijn doorlopen met een nauwe betrokkenheid van medewerkers van zowel NS als ProRail. In het kader van hoor en wederhoor hebben beide organisaties de gelegenheid gehad om op een conceptversie van de voorliggende rapportage te reageren, zonder daarbij afbreuk te doen aan de onafhankelijkheid van de onderzoekers.

### **1.4 Leeswijzer**

De opbouw van de rapportage is als volgt. In hoofdstuk twee gaan we in op de bevindingen uit de audit voor de prestatie-indicatoren voor NS. Vervolgens beschrijven we in hoofdstuk drie de bevindingen voor de prestatie-indicatoren van ProRail. In hoofdstuk vier gaan we in op de bevindingen voor de prestatie-indicatoren waarvoor NS en ProRail gezamenlijk verantwoordelijk zijn. In hoofdstuk vijf komen we terug op de generieke bevindingen uit de voorgaande audit van BTB/AEF en beschrijven we hoe hier door NS en ProRail mee is omgegaan.

## 2 Prestatie-indicatoren NS

NS meet de prestatie-indicatoren 'Zitplaatskans reizigers in spits HRN', 'Top 10 reizen met de slechtste vervoerplaatskans in de spits' en 'Zitplaatskans reizigers in spits HSL' grotendeels volgens dezelfde methodiek. Ook de opzet van de publieksdefinities van voornoemde prestatie-indicatoren verschilt slechts beperkt. De (algemene) bevindingen en aanbevelingen die gelden voor 'Zitplaatskans reizigers in de spits HRN' zijn daarom ook van toepassing op de overige twee prestatie-indicatoren. Bevindingen en aanbevelingen die specifiek zijn voor deze prestatie-indicatoren zijn opgenomen in de betreffende paragrafen.

### 2.1 Zitplaatskans reizigers in spits HRN

#### 2.1.1 Publieksdefinitie

##### Bevindingen

De publieksdefinitie is, op een beperkt aantal aandachtspunten na, helder, compleet en consistent. De beknopte beschrijving van het meetsysteem in de publieksdefinitie geeft een adequate samenvatting van de stappen die NS zet om tot bepaling van de indicator te komen. Hierbij merken we op, dat de beschreven stappen voor een relatieve buitenstaander naar verwachting moeilijk te volgen zijn. Wij geven om die reden in overweging om de beschrijving van het meetsysteem in de publieksdefinitie te versimpelen. Dit is bijvoorbeeld mogelijk door het gebruik van een stroomdiagram, waarin een schematische voorstelling van de beschreven processtappen is opgenomen.

Voor een eenduidige interpretatie van de publieksdefinitie verdient de formulering met betrekking tot de afbakening van de scope van de prestatie-indicator aandacht. De voorlaatste zinnen in de publieksdefinitie zijn multi-interpretabel. De zinsnede '*exclusief Thalys en Eurostar*' kan door een lezer zowel worden geïnterpreteerd als uitzondering op de zinssnede '*behalve in een internationale trein met zitplaatsreservering of toeslag*' als een beperking ten opzichte van de gehele zin '*In scope zijn alle reizigers die inchecken én uitchecken op een Hoofdrailnet station en met NS reizen behalve in een internationale trein met zitplaatsreservering of toeslag*'. De publieksdefinitie moet dusdanig worden geformuleerd dat duidelijk is dat de Thalys en de Eurostar buiten scope zijn, omdat reizigers die in- en uitchecken (via de poortjes en paaltjes) niet met deze treinen kunnen reizen. In de ICE, IC Brussel en IC Berlin is in- en uitchecken wel mogelijk, wat maakt dat deze treinen wel onderdeel zijn van de scope.

Aanvullend beperkt de indicator zich tot de spits. In de beschrijving van het meetsysteem in de publieksdefinitie wordt de spits als volgt afgebakend: een reiziger reist in de spits als de check-in op een werkdag tussen 7:00 en 9:00 uur of tussen 16:00 uur en 18:00 uur plaatsvindt.

Gelet op de gevoerde gesprekken en overige documentatie is deze tijdsafbakening incorrect. Dit is namelijk tussen 06:35 en 08:55 uur en tussen 16:05 en 18:25 uur.

#### **Aanbevelingen**

- Scherp de formulering van de scope zodanig aan dat helder is welke treinen hier wel/niet in vallen.
- Pas het tijdframe voor de spits in het niet-publieke deel van de publieksdefinitie aan naar 06:35-08:55 en 16:05-18:25.

### **2.1.2 Meetmethode**

#### **Bevindingen**

De processtappen in de meetmethode zijn, eveneens op een paar aandachtspunten na, nauwkeurig, compleet en consistent beschreven.

In de beschrijving van de meetmethode wordt gerefereerd aan de comfortnorm (onderdeel van de Comfortabel/Acceptabel/Vol-normen (CAV)). De comfortnorm is echter niet gelijk aan de afbakening die in de publieksdefinitie wordt gegeven: '*Het aantal zitplaatsen in de spits omvat zowel eerste- als tweede klaszitplaatsen en klapzittingen*'. Om verwarring te voorkomen geven wij in overweging om de voornoemde beschrijving uit de publieksdefinitie in de meetmethode over te nemen. Daarnaast wordt ten behoeve van het bepalen van de reismogelijkheden een tijds marge van 5 minuten genomen. We stellen voor om in de documentatie nader te specificeren waarom voor deze tijds marge wordt gekozen (ruimere populatie ten opzichte van een kleinere tijds marge om meer reizigers aan treinen te kunnen toewijzen). Inhoudelijk kan BTB/AEF zich in deze tijds marge vinden.

De aannames die ten grondslag liggen aan de meetmethode zijn vastgelegd in een hiervoor opgesteld document. In een separaat rapport is de toetsing van enkele kritieke aannames terug te vinden. Voor een goed begrip van de meetmethode is het van belang dat deze gegevens centraal toegankelijk worden gemaakt voor medewerkers van NS (meer in het bijzonder medewerkers van Business Intelligence & Analytics en het NSR Kenniscentrum, omdat zij respectievelijk data verzamelen en verwerken en de indicator berekenen) dan wel in eenzelfde document worden opgenomen.

De functionaliteit voor het berekenen van prestatie-indicatoren in de applicatie SOFA is op hoofdlijnen verdeeld in vier afzonderlijke deelapplicaties:

1. Deelapplicatie Rockt.
2. Fractiemodel levering.
3. Deelapplicatie bepaling treinbezetting en instappers.
4. Deelapplicatie KPI berekening (sluitstuk architectuur SOFA).

Het fractiemodel wordt gebruikt om de met check-in-check-out (CICO) toegedeelde reizigers op te hogen naar het totaal aantal reizigers. Per 1/10/2015 is dit fractiemodel herijkt.

Op basis van de gesprekken en de ontvangen documentatie is BTB/AEF van mening, dat de eerste twee deelapplicaties voldoende zijn gevalideerd (zie in dit verband ook het volgende aandachtspunt). Voor wat betreft de derde en vierde deelapplicatie is niet inzichtelijk welke validaties NS heeft uitgevoerd. NS geeft wel aan dat eventuele afwijkingen in de praktijk desondanks wel aan het licht komen, aangezien de resultaten van deze deelapplicaties ook (al enige tijd) door andere bedrijfsonderdelen van NS worden gebruikt. Dit gebruik draagt bij aan de validiteit van de meetmethode. Hierbij merken we echter wel op dat deze

bedrijfsonderdelen de deelapplicaties voor andere doeleinden inzetten. Mogelijk worden daarom niet alle voor de prestatie-indicatoren relevante afwijkingen geconstateerd. BTB/AEF acht validatie van deelapplicaties 3 en 4 voor de berekening prestatie-indicatoren vervoercapaciteit daarom wel van belang.

Tot slot geldt dat de indicator geen onderscheid maakt tussen 1e- en 2e klasse reizigers. Eerste klas zitplaatsen worden meegenomen in de berekening van de zitplaatskans voor reizigers met een tweede klas abonnement. Andersom is dat ook het geval. Structurele reizigers (met abonnementen) hebben overigens in geval van een volle tweede klas ook de mogelijkheid om te upgraden naar eerste klas.

#### **Aanbevelingen**

- Leg de aannames bij de meetmethode in hetzelfde document vast als de beschrijving van de meetmethode of zorg als alternatief voor centrale vastlegging van beide.
- Maak inzichtelijk welke validaties zijn uitgevoerd op de deelapplicaties 3 en 4 van SOFA. Indien er geen validaties zijn uitgevoerd, voer deze dan alsnog uit.

### **2.1.3 Meetpraktijk**

#### **Bevindingen**

Door middel van de gesprekken en een korte demonstratie heeft BTB/AEF zich een beeld gevormd van de meetpraktijk. Er is sprake van een volledig draaiend productiesysteem, waarin de juiste basisgegevens worden gebruikt in termen van omvang en tijdvak. We constateren dat de meetpraktijk overeenkomt met de beschreven meetmethodiek.

Voor de sturing en de organisatie van de meetpraktijk zijn de taken, verantwoordelijkheden en bevoegdheden van de verschillende actoren in het proces vastgelegd en is de rapportagecyclus beschreven:

- De berekening van de zitplaatskans en de aantallen reizigers per dag zijn belegd bij het Kenniscentrum van NS. Binnen het Kenniscentrum zijn geen specifieke verantwoordelijken aangewezen voor de berekening.
- De afdeling Business Intelligence & Analytics is verantwoordelijk voor het verzamelen en verwerken van de (primaire) data en ontbreekt in de beschrijving van het organisatiemodel. Gelet op de belangrijke rol, vindt BTB/AEF het van belang om deze afdeling hierin wel op te nemen.
- De aanlevering en verwerking van (rapportages over de kwaliteit van de) interne en externe brondata is niet vastgelegd in de rapportagecyclus. Dit terwijl het kenniscentrum wel afhankelijk is van interne en externe aanlevering van, de (kwaliteit)controle op en de verwerking van de brondata.

Op de meetpraktijk voert NS voldoende controles uit, waaronder controles op de kwaliteit van reismogelijkheden en de verdeling van reizigers over treinen door middel van gebruik van controles door conducteurs in de trein en matching met de reismogelijkheid. Ook is het beeld dat er sprake is van voldoende controle op de kwaliteit van de brondata. De taken en verantwoordelijkheden op dit punt zijn echter niet vastgelegd. Ook ontbreekt het aan rapportages. Wij geven in overweging om een en ander beter vast te leggen.

Voor wat betreft het niveau van de eindverantwoordelijken voor de prestatie-indicator heeft NS niet inzichtelijk gemaakt, in hoeverre zij actief de gehele meetpraktijk van betreffende prestatie-indicatoren (op basis van de beschikbare (kwaliteits)rapportages) monitoren.

Daarnaast ontbreekt het aan vaste momenten, waarop de gehele meetpraktijk onder leiding van de eindverantwoordelijken wordt geëvalueerd.

NS heeft de marge van onzekerheden in het meetsysteem kwalitatief inzichtelijk gemaakt door de beperkingen en afwegingen van het meetsysteem en de meetpraktijk in kaart te brengen. Een voorbeeld van een onzekerheid is dat er geen weging naar reisafstand plaatsvindt. De onderliggende afweging is dat elke reiziger voor NS even belangrijk is.

Uit de audit is naar voren gekomen dat de marge van onzekerheden in het meetsysteem ook kwantitatief in beeld kan worden gebracht. Denk bijvoorbeeld aan het aantal reizigers dat uit de (reguliere) dataset valt door bijvoorbeeld een verkeerde waarneming (zoals reizigers die vergeten om uit te checken). Vaak blijken deze foutmarges zeer beperkt te zijn. Dat betekent echter niet dat deze niet (structureel) in beeld kunnen worden gebracht en worden vastgelegd.

Tot slot valt op dat er niet wordt gewerkt met een (algemene) matrix waarin de risico's en beheersmaatregelen zijn vastgelegd. Een voorbeeld betreft het risico dat data niet tijdig wordt aangeleverd. Ondanks dat een dergelijke matrix ontbreekt, hebben medewerkers van NS wel oog voor de risico's en wordt er actief gestuurd om deze te beheersen.

### **Aanbevelingen**

- Maak medewerkers binnen het Kenniscentrum verantwoordelijk voor het verzamelen van de benodigde data en de berekening van de prestatie-indicator om daarmee de kwaliteit en continuïteit van de berekening en de resultaten beter te borgen.
- Neem de (interne en externe) leveranciers van de brondata op basis waarvan de indicator wordt bepaald en de rol van Business Intelligence & Analytics op in het organisatiemodel. Hierdoor is makkelijker te sturen op tijdigheid en kwaliteit van de brondata.
- Leg de interne en externe aanlevering van, de (kwaliteit)controle op en de verwerking van de brondata vast in een rapportagecyclus.
- Maak inzichtelijk hoe de eindverantwoordelijke functionaris voor de KPI binnen NS de kwaliteit monitort.
- Evalueer de meetpraktijk halfjaarlijks met alle betrokken om het lerend vermogen te vergroten en formaliseer deze evaluatie.
- Breng de onzekerheden in het meetsysteem ook kwantitatief in beeld en leg deze in een document vast. Leg daarnaast ook de belangrijkste risico's en de bijbehorende beheersmaatregelen vast.

Bij deze aanbevelingen geldt dat NS medio 2016 een risicoanalyse heeft opgesteld van de kwetsbare delen van de processen om de prestatie-indicatoren te bepalen. Op basis hiervan is een plan van aanpak geschreven met benodigde activiteiten en te behalen resultaten. In de kern gaat het hierbij om de aandachtsgebieden die zijn opgenomen in de onderstaande tabel.

Aandachtsgebied	Verbetermaatregelen
Aansluiting interne sturing en eigenaarschap	<ul style="list-style-type: none"> <li>— Borgen van eigenaarschap van prestatie-indicatoren</li> <li>— Expliciet vastleggen hoe prestatie-indicatoren kunnen worden beïnvloed</li> <li>— Verbeteren duiding en analyse van uitkomsten van prestatie-indicatoren</li> </ul>

Aandachtsgebied	Verbetermaatregelen
Kwetsbaar proces door gebrek aan interne beheersmaatregelen en IT controls	<ul style="list-style-type: none"> <li>– Proceseigenaren benoemen per prestatie-indicator om risico's in het ketenproces te beheersen</li> <li>– Verder automatiseren van rapportageproces en verbeteren van interne beheersing van proces</li> <li>– Start van implementatie van <i>information security management system</i></li> </ul>
Onvoldoende inzicht in controls bij data leverende derden	<ul style="list-style-type: none"> <li>– Met derde partijen afspraken maken over:                             <ul style="list-style-type: none"> <li>- Inzichtelijk maken van huidige staat van beheersing</li> <li>- Verbetering van die beheersing in de toekomst</li> </ul> </li> </ul>
Onduidelijkheid over en veranderlijkheid van publieksdefinitie	<ul style="list-style-type: none"> <li>– Expliciete en centrale vastlegging van publieksdefinities en gedetailleerder afstemmen hiervan met ministerie van lenM</li> <li>– Proces afspraken maken intern en met het ministerie van lenM over wijzigingen in publieksdefinities. (publieksdefinities zijn een 'contract')</li> </ul>

De aandachtsgebieden en de verbetermaatregelen geïdentificeerd door NS overlappen voor een groot deel met voornoemde aanbevelingen die BTB/AEF doen. De aanbeveling inzake de halfjaarlijkse evaluatie van de meetpraktijk zien we niet direct terug in het verbetertraject. Deze aanbeveling kan echter ook buiten het verbetertraject worden georganiseerd.

NS heeft als doel om voor rapportagejaar 2017 de grootste onzekerheden in de processen aanwijsbaar te verbeteren. Het beeld van BTB/AEF is dat NS hiervoor op schema ligt. Voor wat betreft de kwaliteit van de brondata is er onlangs bijvoorbeeld een risicoanalyse per leverancier uitgevoerd. Wel moeten er nog afspraken worden gemaakt met de betreffende leveranciers om (aanvullende) beheersmaatregelen in te richten. Per saldo heeft NS de verbeterpunten in beeld en werkt aan de implementatie hiervan.

#### 2.1.4 Informatie

##### Bevindingen

De steekproef die is gebruikt, is een voldoende afspiegeling van de totale populatie reizigers in de spits. Hierin worden de meest frequente herkomst-bestemmingen meegenomen. Dit betreft meer dan 97% van de reizigers. Voor de prestatie-indicator wordt een selectie gemaakt van reizigers die in de spits hebben ingecheckt (het maakt daarbij niet uit of deze reiziger binnen of buiten de spits weer uitcheckt). Reizigers die buiten de spits inchecken, maar in de spits uitchecken worden niet meegenomen in de selectie van reizigers. Deze reizigers zijn niet nodig om de zitplaatskans te bepalen. De populatie is groot genoeg. Ze tellen overigens wel mee in de meting van de drukte van de treinen in de spits.

Er is geen aanleiding om te twijfelen aan de betrouwbaarheid van de informatie uit het meetsysteem. NS heeft met het meetsysteem verschillende tests gedraaid, waarmee fouten



zijn opgespoord en hersteld. Het meetsysteem blijft ook continu in ontwikkeling door de toevoeging van stations en het opstellen van nieuwe rapportages. Dit vraagt om frequente controles, die eveneens worden uitgevoerd door NS.

Trendbreuken (zoals in het aantal reizigers per trein) worden geanalyseerd en gerapporteerd middels interne dashboards. De persoon/afdeling verantwoordelijk voor de data of deelapplicatie meetsysteem is ook voor deze analyses verantwoordelijk.

**Aanbevelingen**

- Gelet op het voorgaande hebben wij geen aanbevelingen.

**2.1.5 Bodem- en streefwaarde**

**Bevindingen**

In de onderstaande tabel is het voorstel van lenM opgenomen voor de aanpassing van de bodem- en streefwaarde.

Situatie	Bodemwaarde	Streefwaarde
Huidig	98,7%	99,2%
Nieuw	94,3%	95,5%

Voor de berekening van deze waarden is gebruikgemaakt van het gegeven, dat de realisatiewaarde in 2015 volgens de oude meting gelijk was aan de bodemwaarde. Op basis hiervan is de bodemwaarde voor de nieuwe methode gelijk gesteld aan de realisatiewaarde volgens de nieuwe methode en is de streefwaarde vervolgens bepaald door het uitvoeren van een regressieanalyse.

Voor de bepaling van zowel de bodemwaarde als de streefwaarde is gebruik maakt van een set gegevens die betrekking heeft op de periode Q1 2015 - Q4 2015<sup>2</sup>. Dit is niet consistent in vergelijking met de dataset die is gekozen voor de overige prestatie-indicatoren (zie vervolg).

Ondanks deze inconsistentie biedt het gegeven dat de realisatiewaarde en bodemwaarde over 2015 gelijk zijn, de zekerheid dat de vertaling van de bodemwaarde ambitie-neutraal plaatsvindt. Deze zekerheid bieden alternatieve statistische methoden niet. Gezien het feit dat de nieuwe bodemwaarde is bepaald aan de hand van de set gegevens Q1 2015 - Q4 2015 is het logisch om deze set ook te gebruiken voor het bepalen van de nieuwe streefwaarde.

Alternatieve methoden voor de berekening van de bodem- en streefwaarde zijn de deltamethode en de methode van onder- en overschrijding. Deze methoden worden door zowel NS als Q-consult<sup>3</sup>, een adviesbureau dat de statistische methoden voor de omrekening van de huidige bodem- en streefwaarde voor vervoercapaciteit van NS eerder in opdracht van het ministerie van lenM heeft beoordeeld, als minder geschikt geacht dan de gebruikte methode. Ondanks de grote spreiding kan vanwege de vele meetwaarden op dagbasis voor de prestatie-indicator zitplaatskans toch een vrij nauwkeurige omrekening worden bepaald.

<sup>2</sup> Hierbij is gebruik gemaakt van het per 1/10/2015 herijkte fractiemodel.

<sup>3</sup> Q-Consult beoordeling omrekening prestatie-indicatoren NS V2 d.d. 28 juli 2016.

BTB/AEF onderschrijft de wijze waarop de nieuwe bodem- en streefwaarde is bepaald en bevestigt dat de uitkomsten een ambitieneutrale doorvertaling zijn.

**Aanbevelingen**

- Gelet op het voorgaande hebben wij geen aanbevelingen.

## 2.2 Top 10 reizen met de slechtste vervoerplaatskans in de spits

### 2.2.1 Publieksdefinitie

**Bevindingen**

De publieksdefinitie is op enkele aandachtspunten na helder, compleet en consistent gedefinieerd. In aanvulling op de bevindingen bij de prestatie-indicator Zitplaatskans merken we op, dat pas in de beschrijving van het meetsysteem staat toegelicht hoe de selectie van de top 10 tot stand komt. Daarnaast bevat de publieksdefinitie in tegenstelling tot de beschrijving van de oude prestatie-indicator geen lijst met de reizen met de slechtste vervoerplaatskans die deel uitmaken van de top 10. Onze inschatting is dat de belangstellende lezer wil weten hoe de top 10 is samengesteld en uit welke reizen deze bestaat.

**Aanbevelingen<sup>4</sup>**

- Neem in het eerste deel van de publieksdefinitie op hoe de top-10 wordt samengesteld.
- Neem de top 10 van slechtste reizen op in een op het betreffende jaar toegespitste publieksdefinitie in het Vervoerplan.<sup>5</sup>

### 2.2.2 Meetmethode

**Bevindingen**

De meetmethode is nauwkeurig, compleet en consistent beschreven. Hierbij geldt dat helder is aangegeven hoe de top 10 wordt samengesteld. Hiervoor selecteert NS allereerst de 1.000 reistrajecten met de meeste reizigers. Vervolgens stelt NS de tien reistrajecten vast met de laagste kans op een vervoerplaats over de basisperiode (periode 1 juni jaar x-2 en 31 mei jaar x-1). Belangrijkste reden hiervoor is dat een recentere basisperiode volgens NS niet haalbaar is. Indien deze datum naar voren wordt geschoven, wordt het lastiger voor NS om tijdig (op materieel) te sturen op de top 10 van drukste trajecten. Het beeld van BTB/AEF is dat er ook bij een latere vaststelling van de top 10 voldoende sturingsmogelijkheden overblijven, onder meer via de band van de wijzigingsbladen.

**Aanbevelingen<sup>6</sup>**

- Onderzoek nader of het mogelijk is om voor het bepalen van de top 10 uit te gaan van een meer recente basisperiode.

<sup>4</sup> In aanvulling op aanbevelingen onder KPI zitplaatskans.

<sup>5</sup> In dit verband is het aan te bevelen om in de generieke publieksdefinitie op te nemen dat in de op het betreffende jaar toegespitste publieksdefinitie in het vervoerplan de geselecteerde top-10 van dat jaar wordt vermeld.

<sup>6</sup> In aanvulling op aanbevelingen onder prestatie-indicator Zitplaatskans.

### 2.2.3 Meetpraktijk

De meetpraktijk van deze prestatie-indicator komt overeen met die voor de prestatie-indicator Zitplaatskans. BTB/AEF heeft geen aanvullende bevindingen en aanbevelingen.

### 2.2.4 Informatie

De informatie uit de meetsystemen (waaronder representativiteit van de steekproef en betrouwbaarheid van het meetstelsel) van deze prestatie-indicator komt overeen met die voor de prestatie-indicator Zitplaatskans. We hebben geen aanvullende bevindingen en aanbevelingen.

### 2.2.5 Bodem- en streefwaarde

#### Bevindingen

In de onderstaande tabel is het voorstel van NS opgenomen voor de aanpassing van de bodem- en streefwaarde.

Situatie	Bodemwaarde	Streefwaarde
Huidig	96,5%	97,5%
Nieuw	95,1%	96,0%

Het ministerie van IenM wil deze nieuwe waarden in de toekomst omgedraaid presenteren. In dat geval wordt de bodemwaarde 4,9% en de streefwaarde 4,0%. Deze presentatie is niet consistent met de presentatie van de zitplaatskans. Belangrijker is echter dat deze manier van presenteren aansluit bij de definitie en de meetmethodiek. Het betreft immers een indicatie van de kans dat een reiziger die tijdens de spits op de tien slechtste reizen reist op een deel van zijn reis in een overvolle trein zit of staat. Omdat deze presentatie de inhoud beter dekt, onderschrijven wij de aanpassing. Het is wel van belang om deze consequent door te voeren in alle interne- en externe documentatie.

Voor de doorvertaling van de bodem- en streefwaarde is de regressiemethode toegepast. Consistent met de overige prestatie-indicatoren (m.u.v. zitplaatskans) is Q2 2015 - Q1 2016<sup>7</sup> als meetperiode genomen. De berekeningen conform deze methode waren in de praktijk lastig te maken, omdat er geen gegevens op dagbasis beschikbaar waren om de bodemwaarde opnieuw te berekenen. NS beschikte alleen over cijfers op maandbasis, waardoor de set aan gegevens (12 waarnemingen) feitelijk te klein is voor een lineaire regressie met een acceptabele onzekerheidsmarge. Ook is de correlatie tussen de cijfers relatief laag. Dit maakt dat BTB/AEF, in lijn met het rapport van Q-Consult<sup>8</sup>, van mening is dat de methode van regressie voor de berekening van de nieuwe bodem- en streefwaarde toch de voorkeur verdient boven alternatieve methoden, zoals de deltamethode en de methode van onder- en overschrijding.

#### Aanbevelingen

- Gelet op het voorgaande hebben wij geen aanbevelingen.

<sup>7</sup> Hierbij is gebruik gemaakt van het per 1/10/2015 herijkte fractiemodel.

<sup>8</sup> Zie paragraaf 2.1.4.

## 2.3 Zitplaatskans reizigers in spits HSL

### 2.3.1 Publieksdefinitie

#### Bevindingen

De publieksdefinitie van de prestatie-indicator is specifiek en compleet. De beknopte beschrijving van het meetsysteem draagt bij aan het begrip van de publieksdefinitie.

Evenals de prestatie-indicator Zitplaatskans HRN beperkt deze prestatie-indicator zich tot de spits. Het exacte tijdsframe is correct geformuleerd: een rit is in de spits als de check-in op een werkdag tussen 06:35-08:55 uur en 16:05-18:25 uur heeft plaatsgevonden.

De scope is helder afgebakend. Het betreft alle reizigers die in- en uitchecken op een HRN-station en met NS reizen. Voor de HSL-Zuid diensten wordt vervolgens alleen gekeken naar de 12 meest gemaakte reizen over de HSL-Zuid. Dit zijn de reizen in beide richtingen tussen de volgende stations:

1. Amsterdam Centraal - Rotterdam Centraal
2. Amsterdam Centraal - Breda
3. Amsterdam Zuid - Rotterdam Centraal
4. Schiphol Airport - Rotterdam Centraal
5. Schiphol Airport - Breda
6. Rotterdam Centraal - Breda.

#### Aanbevelingen

- Gelet op het voorgaande hebben wij geen aanbevelingen.

### 2.3.2 Meetmethode

#### Bevindingen

De meetmethode en de onderliggende aannames voor de onderhavige prestatie-indicator komt overeen met de prestatie-indicator Zitplaatskans reizigers in de spits HRN 2017.

Het enige verschil zit in de laatste component van de meetmethode, de bepaling van de indicator. Het gaat hierbij om een selectie van reizigers die vanaf een hoofdrailnetstation naar een hoofdrailnetstation reizen en daarbinnen de twaalf meest gemaakte reizen over de HSL-Zuid. Binnen de methodiek van herkomst-bestemmingsrelaties (HB-relaties) heeft NS 12 HB-relaties gekozen die een representatieve weergave zijn van de prestaties van de binnenlandse HSL-Zuid diensten. Hierbij baseert NS zich op feitelijke gegevens: NS weet dat voor elk van die HB-relaties geldt dat de reizigers voor circa 90% of meer in de IC direct (of IC Brussel of Den Haag-Eindhoven) zitten. Het hoge aandeel reizigers in de IC direct was ook het criterium om deze 12 HB-relaties te selecteren.

#### Aanbevelingen

- Gelet op het voorgaande hebben wij geen aanbevelingen.

### 2.3.3 Meetpraktijk

De meetpraktijk van de prestatie-indicator zitplaatskans HSL komt overeen met die voor de prestatie-indicator Zitplaatskans HRN. BTB/AEF heeft geen aanvullende bevindingen en aanbevelingen.

### 2.3.4 Informatie

De informatie uit meetsystemen van de prestatie-indicator zitplaatskans HSL komt overeen met die voor de prestatie-indicator Zitplaatskans HRN. BTB/AEF heeft geen aanvullende bevindingen en aanbevelingen.

### 2.3.5 Sturingsmogelijkheden

#### Bevindingen

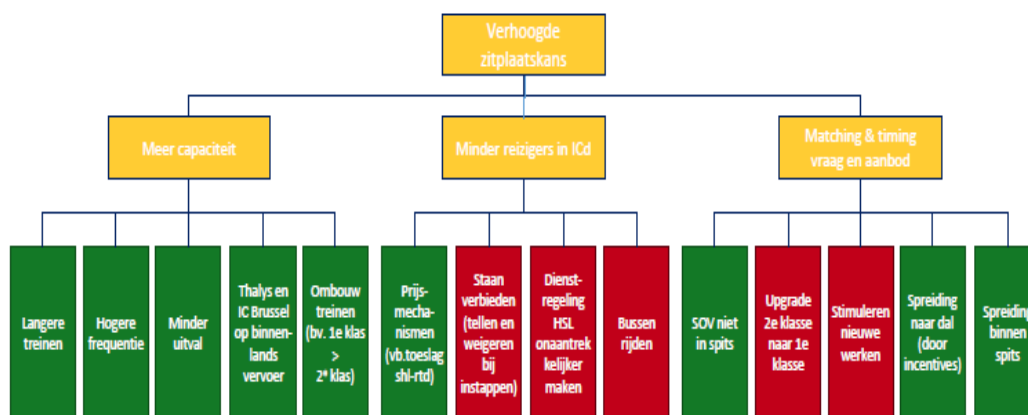
##### Beperking uitval

Een manier om de netto zitplaatscapaciteit te vergroten is het terugdringen van uitval van materieel. Hiertoe hebben NS en ProRail verbeterplannen opgesteld<sup>9</sup> die nu in uitvoering zijn. Volgens inschatting van NS kan realisatie van bijvoorbeeld 6% uitval een positieve impact van 0,5% kunnen hebben op de prestatie-indicator in 2019. De plausibiliteit van de verbeterplannen wordt momenteel in een review beoordeeld en is geen onderdeel van deze audit.

##### Overige sturingsmogelijkheden

##### *Compleetheid*

Om de zitplaatskans op de HSL te verhogen heeft NS de volgende potentiële maatregelen geïdentificeerd:



De groen gemarkeerde maatregelen zijn door NS als meest kansrijk beoordeeld en verder uitgewerkt. De rood gemarkeerde maatregelen zijn in overleg met het ministerie van IenM niet verder onderzocht, omdat deze niet wenselijk zijn, nauwelijks beïnvloedbaar zijn door NS of geen effect op de prestatie-indicator hebben.

BTB/AEF beschouwt het aantal geïdentificeerde en nader onderzochte maatregelen als compleet. NS heeft de maatregelen bovendien overzichtelijk gedocumenteerd. Wij kunnen ons vinden in de keuze om de rood gemarkeerde maatregelen niet verder te onderzoeken. Zo kent de inzet van bussen te veel afhankelijkheden en vormt dit geen realistisch alternatief voor de IC direct. Voor de upgrade van tweede naar eerste klasse geldt dat in de zitplaatskans berekening geen onderscheid wordt gemaakt tussen eerste en tweede klas, waardoor deze actie geen effect heeft op de prestatie-indicator.

<sup>9</sup> Zie voor de verbeterplannen NS de Brief van R. van Boxtel aan Staatssecretaris Dijkzma d.d. 29 april 2016.

#### *Afhankelijkheden en haalbaarheid*

Voor elk van de als kansrijk beoordeelde maatregelen is door NS onderzocht of de betreffende maatregel technisch, juridisch, operationeel, financieel en maatschappelijk haalbaar zou zijn om in 2019 effect te kunnen sorteren. De uitkomsten van dit onderzoek zijn vastgelegd in een uitgebreide rapportage.

NS concludeert verder dat de overige onderzochte maatregelen gelet op de verschillende afhankelijkheden niet haalbaar, te onzeker en/of niet wenselijk zijn. Twee voorbeelden:

- De onderzochte maatregel om langere treinen in te zetten is volgens NS te onzeker en/of niet maakbaar. De verwerving en ombouw (naar het HSL-netwerk) van het benodigde additionele materieel zal circa vier jaar kosten. Aanpassing van het bestaande rijkspark kost (in theorie) minimaal 1,5 - 2 jaar vanaf het besluit om dit te doen. Hier zitten echter risico's aan verbonden, onder meer vanwege kritische ombouwcapaciteit bij NedTrain en opstelcapaciteit bij ProRail. NS heeft concreet de afhankelijkheden in kaart gebracht. BTB/AEF onderschrijft in dit verband de risico's en concludeert dat de maatregel te laat in gang kan worden gezet om effect te sorteren in 2019.
- Een ander voorbeeld betreft de onderzochte maatregel om naar een hogere frequentie in de dienstregeling te gaan (van 4 naar 5 treinen op Rotterdam-Breda en Amsterdam - Rotterdam). NS geeft aan dat het qua dienstregeling in 2017 niet haalbaar is om de frequentie te verhogen. Niet duidelijk is echter of dit voor 2018 dan wel mogelijk zal zijn. BTB/AEF begrijpt echter dat ook voor deze maatregel extra treinstellen nodig zijn, waarbij bovengenoemde afhankelijkheden optreden.

Wat beperkt inzichtelijk is gemaakt zijn de financiële implicaties van de onderzochte maatregelen (voor zover het geen bedrijfsvertrouwelijke informatie betreft). De reden hiervoor is dat financiële overwegingen niet leidend zijn geweest voor NS. Desalniettemin geven we in overweging om toch een kosteninschatting per maatregel te geven, zodat er meer zicht is op de bedrijfseconomische afwegingen die NS maakt (dan wel dient te maken).

Tot slot merkt BTB/AEF op dat de focus van het onderzoek naar de haalbaarheid (maar ook de effectiviteit, zie vervolg) van de maatregelen ligt op 2019, het jaar van de midterm review. In het kader van het vaststellen van de bodem- en streefwaarde is dit ook begrijpelijk. De maatregelen kunnen echter ook een effect hebben op de langere termijn, dat wil zeggen na 2019. Ons beeld uit de gesprekken is, dat dit besef er bij NS wel degelijk is. Maatregelen die na eerste analyse niet haalbaar lijken en/of geen effect lijken te sorteren in 2019, worden wel degelijk nader onderzocht voor de langere termijn.

Al met al zijn de kansrijke maatregelen door NS grondig getoetst op haalbaarheid. BTB/AEF ziet geen aanleiding om de conclusies van NS in twijfel te trekken.

#### *Effectiviteit*

Voor ieder van de als kansrijk beoordeelde maatregelen is door NS onderzocht welke bijdrage de maatregel (oftewel de kwantitatieve impact) zou kunnen leveren in het resultaat op de prestatie-indicator in 2019:

- Meer capaciteit:
  - langere treinen (impact 1,9%)
  - hogere frequentie (impact 1,5%)
  - minder uitval (Bijvoorbeeld impact uitval 6% = 0,5%)
  - ombouw treinen (impact 0,7%)
- Minder reizigers:
  - prijsmechanismen (impact max 0,1%)

- Matching en timing vraag en aanbod:
  - StudentenOV niet in spits (impact 2,1%)
  - spreiding naar dal (impact 0,0%)
  - spreiding binnen Spits (impact max 0,1%).

Onderzochte maatregelen zoals prijsmechanismen, spreiding binnen spits en naar dal en NS medewerkers uit de spits hebben volgens NS (vrijwel) geen kwantitatieve impact in 2019 en verdienen daarmee geen verdere aandacht. Deze uitkomsten zijn geen verassing voor BTB/AEF. HSL-Zuid kenmerkt zich bijvoorbeeld door de beperkte prijselasticiteit. De impact van zowel positieve als negatieve incentives lijkt daarom klein.

NS heeft voldoende inzichtelijk gemaakt hoe de kwantitatieve impact van de maatregelen is berekend. Naar het oordeel van BTB/AEF is de berekening van de kwantitatieve impact van de maatregelen op het resultaat van de prestatie-indicator in 2019 op een logische wijze uitgevoerd.

**Aanbevelingen**

- Bepaal op basis van de conclusies van de review op het Verbeterplan IC direct in samenspraak met het ministerie van IenM in welke mate de door NS en ProRail voorgestelde verbetermaatregelen haalbaar zijn en aanleiding geven tot aanscherping van de streefwaarde.
- Beperk het onderzoek naar de haalbaarheid van sturingsmogelijkheden niet tot 2019. Blijf aandacht houden voor maatregelen die ook in de periode daarna impact kunnen genereren.

**2.3.6 Bodem- en streefwaarde**

**Bevindingen**

In de onderstaande tabel is het voorstel van zowel NS als het ministerie van IenM opgenomen voor de aanpassing van de bodem- en streefwaarde.

	<b>Bodemwaarde</b>	<b>Streefwaarde</b>
Voorstel IenM	91,2%	96,4%
Voorstel NS	91,2%	94,4%
Groeiscenario	Hoog	Midden

De bodem- en streefwaarde van NS is opgebouwd zoals weergegeven in de onderstaande tabel. De toelichting op de opbouw volgt op de volgende pagina's.

	<b>Bodemwaarde</b>	<b>Streefwaarde</b>
Component 1: Baseline groeiscenario	92,2% (zeer hoog IC direct)	94,4% (midden)
Component 2: Netto onzekerheid Sedna	-0,5%	0% (geen marge gerekend)
Component 3: Marge dagelijkse variatie KPI	-0,5%	0% (geen marge gerekend)

	Bodemwaarde	Streefwaarde
Voorstel NS	91,2%	94,4%

#### Navolgbaarheid berekening bodem- en streefwaarde

NS heeft de berekening en opbouw van de bodem- en streefwaarde uitgebreid vastgelegd. Ook de onderliggende aannames, keuzemogelijkheden en de afwegingen zijn inzichtelijk gemaakt. De berekening van zowel de bodemwaarde en de streefwaarde bestaat uit de volgende componenten:

#### *Component 1: Reizigersgroei*

Als startpunt is genomen de zitplaatskans zoals verwacht op basis van een berekening van de reizigersgroei in de vorm van scenario's.

Jaarlijks worden er door NS in het voorjaar drie groeiscenario's (bandbreedteinschattingen van de reizigersgroei) opgesteld. De groeiscenario's dienen als basis voor de strategie van NS en worden bijvoorbeeld ook verwerkt in het vervoerplan. Dit is een extern geauditeerd proces.

De voorjaarsprognose 2016 maakt onderscheid in drie reizigersgroei-scenario's: een zeer laag scenario, een middenscenario en een zeer hoog scenario. Verschillende factoren zijn van invloed op de ontwikkeling van een groeiscenario. Denk met name aan (niet door NS beïnvloedbare exogene) factoren zoals economische ontwikkeling, demografische ontwikkeling en de benzineprijs (waardoor mensen meer of minder de auto gebruiken als vervoermiddel). Maar ook endogene factoren zoals de opbouw van de dienstregeling en de tarifiering (de prijs van een treinkaartje) spelen een rol.

De groeiscenario's zijn gebaseerd op een normaalverdeling. Dit is een kansverdeling met twee parameters: de verwachtingswaarde en de standaardafwijking. De kans (c.q. de verwachtingswaarde) dat het middenscenario zich ontwikkelt is het grootst. Het middenscenario bevat voor de exogene en endogene factoren de meest waarschijnlijke aannames. In de normaalverdeling is gekozen voor een standaarddeviatie<sup>10</sup> van 2 sigma. Aan de onderkant en bovenkant van de normaalverdeling is er derhalve een 2,5% kans is dat reizigersaantallen respectievelijk zeer veel lager dan wel zeer veel hoger zijn dan het middenscenario. Dit zijn de zeer laag en zeer hoog scenario's die aannames bevatten die met een kleinere kans op zullen treden. De kans dat het vervoer zich volgens deze scenario's ontwikkelt is dan dus ook kleiner dan bij het middenscenario.

Omdat de voorziene groei op de HSL-Zuid sterker is dan de voorziene groei op het HRN is er een groeiscenario toegevoegd welke de bandbreedte corrigeert voor deze groei ten opzichte van de groei op het HRN voor het zeer hoog scenario: groeiscenario 'Zeer hoog IC direct'.<sup>11</sup> Per HB-relatie zijn hierbij de specifieke groeiverwachtingen gebruikt.<sup>12</sup> De sterkere groeioprognose voor de HSL-Zuid komt voor NS overigens niet als een verrassing. Een belangrijke oorzaak van de (verwachte) reizigersgroei is de (toekomstige) uitbreiding van de dienstregeling. Voorts

<sup>10</sup> Dit is een statistische maat om de spreiding te berekenen.

<sup>11</sup> Ook de bandbreedte voor het zeer lage groeiscenario HRN is voor de HSL gecorrigeerd naar een zeer laag IC Direct scenario.

<sup>12</sup> Voor de reizigers die gebruik maken van de IC direct maar niet in één van de zes HB-relaties vallen is het groeipercentage van het hele netwerk als uitgangspunt genomen (+7%).



was al bekend dat de reizigersgroei op specifiek de HSL-Zuid hoger was dan het gemiddelde op het HRN.

De groeiscenario's zijn berekend op basis van de realisatiecijfers van de eerste 5 maanden van 2016. De (statistische) gevoeligheid van deze constructie is door NS (statistisch) getest en vastgelegd. Op basis van de cijfers heeft NS het basisjaar 2016 geconstrueerd (zitplaatskans 96,9%). Hiervoor is het volgende gedaan:

- Juni tot en met augustus zijn gelijk aan maart tot en met mei, waarbij de capaciteit van de treinen 5% wordt verhoogd om te simuleren dat er minder reizigers zijn (dus meer plaatsen).
- September tot en met november zijn gelijk aan maart tot en met mei, waarbij de capaciteit van de treinen 5% wordt verlaagd om de drukte in de herfst te simuleren.
- December is gelijk aan januari.

Uitgangspunten bij de groeiscenario's zijn:

- Een uitval van 8% (gebaseerd op het feitelijke uitvalpercentage over de eerste vijf maanden van 2016).
- De huidige dienstregeling van de IC Brussel (waarbij geen gebruik wordt gemaakt van het HSL-Zuid traject).
- De reizigersprognose uit de voornoemde voorjaarsprognose.

Aannames voor de groeiscenario's zijn:

- Op de HB-relaties Rotterdam - Breda en Breda - Rotterdam krimpt het aantal reizigers in de IC direct, dat wordt veroorzaakt doordat die reizigers vanaf de dienstregeling 2017 de Intercity 1100 kunnen nemen tussen Den Haag, Rotterdam en Breda. De totale vervoersstroom op deze relatie groeit echter wel. NS heeft de aanname gedaan dat de groei op de hele HB-relatie een goede indicatie geeft voor de zitplaatskans, door deze reizigers in de huidige IC direct-treinen mee te nemen. Dit omdat de KPI op HB-niveau wordt bepaald, en niet op treinserieniveau.
- Totaal aantal reizigers is gelijk aan het aantal instappers.
- De capaciteit op traject Rotterdam - Breda wordt verdubbeld om te simuleren dat er 2 keer zo veel IC direct-treinen gaan rijden. Aanname is dat de aanvoer van reizigers zich 50/50 verdeelt over de twee al rijdende treinen in 2016 en de twee 'nieuwe treinen'.

In onderstaand tabel zijn de verschillende groeiscenario's weergegeven.

Scenario	% groei HSL Zuid in 2019 t.o.v. 2015	Verwachte zitplaatskans 2019 (zonder maatregelen)
Zeer laag	17%	95,3%
Midden	19%	94,4%
Zeer hoog	20%	93,6%
Zeer hoog IC direct	21%	92,2%

De manier waarop de groeiscenario's t.o.v. het basisjaar tot stand zijn gekomen, is voor BTB/AEF begrijpelijk:

- De uitgangspunten en aannames zijn helder. De keuze om uit te gaan van 8% uitval is begrijpelijk. Hierin zijn de meest recente realisaties verwerkt. Zoals toegelicht in het kader van de sturingsmogelijkheden wordt via het verbeterplan toegewerkt naar een reductie van de uitval. Het mogelijke (positieve) effect van deze maatregelen op de zitplaatskans kan worden verdisconteerd in de streefwaarde.
- Meer in het bijzonder komt het onderliggende proces voor de berekening van het groeiscenario 'Zeer hoog IC direct' overeen met het (door een externe partij geauditeerde) proces voor de berekening van de algemene groeiscenario's. Het gehanteerde model is bovendien vaker gebruikt, bijvoorbeeld in geval van de Hanzelijn. BTB/AEF heeft derhalve geen aanleiding om te twijfelen aan de betrouwbaarheid van de berekening van de groeiscenario's.

Voor de bodemwaarde heeft NS voor het specifieke groeiscenario Zeer hoog IC direct gekozen (21%-92,2%). Reizigersgroei op de IC direct is volgens NS in dit kader maatschappelijk wenselijk. Niet alleen omdat de reizigers behoefte heeft aan een snelle verbinding via de HSL-Zuid (reiziger op 1,2 en 3), maar ook om de investering in de aanleg van de HSL-Zuid terug te verdienen. In het algemeen bestelt NS haar materieel gebaseerd op het hoogste groeiscenario. Om deze redenen kiest NS ervoor geen lager groeiscenario te gebruiken (indien wordt gekozen voor een van de algemene scenario's). BTB/AEF onderschrijft deze keuze. De bodemwaarde is het minimumniveau waar NS niet onder mag presteren. Deze bodemwaarde moet ons inziens ook te halen zijn indien het hoogste groeiscenario zich ontwikkelt. Gelet op het feit dat de groeiscenario's in grote mate exogeen worden bepaald, is de invloed van NS op deze ontwikkeling zeer beperkt c.q. nihil.

Voor de streefwaarde is door NS voor een lager groeiscenario gekozen, namelijk het middenscenario (19%-94,4%). Hierbij is door NS geen extra veiligheidsmarge genomen voor de verwachte reizigersgroei op de HSL-Zuid. Volgens BTB/AEF een logische keuze omdat het middenscenario de meest waarschijnlijke aannames bevat en de kans daarmee ook het grootst is dat dit middenscenario wordt bewaarheid.

#### *Component 2: Effecten van project Sedna*

Vervolgens heeft NS in de tweede stap de effecten van project Sedna meegenomen. Project Sedna betreft de verandering van de dienstregeling van de IC-Brussel. Met als doel een betere dienstverlening aan de reiziger door een reistijdverkorting streeft project Sedna naar een ander opzet van het product IC-Brussel. In de nu voorliggende scenario's wordt de IC-Brussel verder geïntegreerd met de IC direct. Hierover zijn het ministerie van IenM, NS en ProRail nog in gesprek.

Het project Sedna heeft drie (mogelijke) gevolgen voor de zitplaatskans in de IC direct:

- Een van de vier IC direct treinen (dat wordt dan een IC-Brussel) kent een langere samenstelling (7 bakken).
- Ook internationale reizigers gaan zitplaatsen innemen op de IC direct.
- Tussen Amsterdam en Schiphol verdwijnt één reismogelijkheid per uur. Eerder genoemde vervanging van één: IC direct voor IC Brussel (die nu naast de vier IC directs rijdt).

Ter bepaling van de bodemwaarde heeft NS de effecten van deze gevolgen berekend:

- Het effect van de langere IC direct-trein is op basis van het Zeer hoog IC direct scenario vastgesteld op 0,2% (AUD2.3, AUD2.4 en nadere toelichting per email).
- Het effect van de internationale reizigers moet volgens NS worden gekwantificeerd op +/- 1,5%. Dit laatste effect lijkt een overreactie van de scenarioberekening te zijn. Er is

daarom gekozen om een afgeronde risicomarge van netto -0,5% in te rekenen waarin beide effecten zijn gecombineerd.

NS heeft de kwantificering van dit effect (de berekening) goed inzichtelijk gemaakt. BTB/AEF onderschrijft de gemaakte berekeningen.

Volgens NS is in de streefwaarde 'de ambitie' ingebouwd dat er geen risicomarge is gerekend voor de onzekerheid van het effect Sedna. Voor deze onzekerheid geldt dat in het middenscenario het effect van de langere IC direct-trein afgerond +0,1% is en het effect van de internationale reizigers in de trein -0,1%. In totaal is hier dus geen effect. In dit geval kan dit derhalve ook niet als 'ambitie in de streefwaarde' worden gepresenteerd. Wij geven in overweging de onderbouwing van de bodem- en streefwaarde hier op aan te passen.

#### *Component 3: Effecten dagelijkse variatie prestatie-indicator*

De prestatie-indicator wordt beïnvloed door dagelijkse variatie in de spreiding van reizigers en het aantal reizigers. Deze variatie wordt bepaald door factoren zoals de weersomstandigheden, gedrag reizigers en evenementen.

Om te voorkomen dat NS de bodemwaarde niet realiseert als gevolg van deze dagelijkse variatie, heeft NS berekend wat het effect van deze factor is. In dit verband is - om de verdeling te bepalen van de gewogen gemiddelde zitplaatskans op jaarniveau - gebruik gemaakt van de zogenaamde bootstrapping-methode. NS *resampled* (met teruglegging) 100 maal een jaar (261 dagen) aan dag-spitsgegevens. Dit resulteert in 100 nieuwe jaar cijfers en een gemiddelde score van 93,67%. Vervolgens heeft NS ook de standaarddeviatie bepaald op 0,233%. Het uiteindelijke effect is vastgesteld op basis van twee keer de standaarddeviatie, oftewel -0,5%. BTB/AEF ondersteunt de gemaakte berekening.

Wederom om 'voldoende ambitie' te houden in de streefwaarde is door NS geen marge gerekend voor de dagelijkse variatie van de berekening van de prestatie-indicator. Voor de dagelijkse variatie zou de marge eveneens -0,5% bedragen. Dit omdat alle cijfers, voor de hele dataset, iets hoger worden gemaakt en daarmee dus vooral het gemiddelde van de cijfers hoger wordt en niet per se de variatie groter.

#### Ambitieniveau bodem- en streefwaarde

##### *Historische- en recente realisaties*

NS heeft niet naar historische en recente realisaties gekeken om de bodem- en streefwaarde te bepalen. De dienstregeling tot augustus 2016 verschilt van de huidige dienstregeling en zal per 2017 zeer waarschijnlijk ook weer veranderen, bijvoorbeeld door de toevoeging van de Intercity Den Haag - Eindhoven en de Intercity Amsterdam - Brussel. Mede gelet op de verwachte reizigersgroei op de HSL-Zuid en de beperkte sturingsmogelijkheden is deze insteek begrijpelijk. Overigens geldt dat de groeiscenario's zijn berekend op basis van de realisatiecijfers van de eerste 5 maanden van 2016. Indirect is deze recente historie derhalve wel meegenomen.

##### *Verwachte toekomstige ontwikkelingen*

Zoals toegelicht heeft NS op zorgvuldige wijze vier groeiscenario's opgesteld, waarvan er een specifiek voor de HSL-Zuid.

Dat de bodemwaarde op basis van het groeiscenario Zeer hoog IC direct is vastgesteld, is naar inschatting van BTB/AEF realistisch. De bodemwaarde moet immers realistisch haalbaar zijn

en niet in sterke mate afhankelijk zijn van parameters (zoals reizigersgroei door exogene factoren, het project Sedna en de gevoeligheid van de prestatie-indicator) die NS niet kan beïnvloeden. Hierbij merken we op dat NS ook het materieel plant op het hoogste groeiscenario.

Volgens NS wordt in de streefwaarde ambitie getoond door het middenscenario als uitgangspunt te nemen en geen marge te rekenen voor dagelijkse variatie en onzekerheid Sedna. Op basis van de normaalverdeling (de kansverdeling met twee parameters: de verwachtingswaarde en de standaardafwijking) is de keuze voor het middenscenario logisch. Het middenscenario bevat de meest waarschijnlijke aannames en de kans is daarmee ook het grootst is dat dit scenario wordt bewaarheid. Hier plaatsen we de volgende kanttekeningen bij:

- De keuze om uit te gaan van het middenscenario betekent ook dat indien de reizigersgroei lager uitpakt, NS de streefwaarde voor 2019 waarschijnlijk betrekkelijk eenvoudig kan behalen. Het tegenovergestelde is echter ook waar: is de reizigersgroei hoger, dan wordt het een enorme uitdaging voor NS om de streefwaarde te behalen.
- Voor de onzekerheid volgend uit Sedna is het effect op basis van het middenscenario nul. Op dit punt kan dus ook niet gesproken worden van enige ambitie.

Deze kanttekeningen leiden wat BTB/AEF betreft niet tot een aanpassing van de streefwaarde, omdat het optreden van het middenscenario voor het grootste deel afhankelijk is van externe factoren, die niet door NS zijn te beïnvloeden.

#### *Sturingsmogelijkheden*

In de opbouw van de bodemwaarde is geen rekening gehouden met het effect van verbetermaatregelen. Dit is begrijpelijk. De kansrijke maatregelen zijn door NS grondig getoetst op haalbaarheid en effectiviteit. De onderzochte maatregelen bevatten veel risico's en afhankelijkheden en zijn niet op korte termijn (max. 1 jaar) te realiseren. De door NS (en het ministerie van IenM) voorgestelde bodemwaarde van 91,2% is derhalve realistisch.

BTB/AEF trekt - met uitzondering van de potentiële positieve impact van de reductie van uitval - voor de streefwaarde eenzelfde conclusie. De streefwaarde moet haalbaar en voldoende ambitieus zijn. In dit verband heeft NS tijdens de audit concreet in beeld gebracht waarom de verschillende geïdentificeerde maatregelen niet tijdig maakbaar zijn om in 2019 een positief effect te kunnen genereren op de streefwaarde. BTB/AEF heeft hier een toets op uitgevoerd en geen aanleiding gevonden om te twifelen aan de conclusies van NS.

Afhankelijk van de uitkomsten van de review van het Verbeterplan IC direct is eventueel meer ambitie mogelijk in de streefwaarde. Volgens inschatting van NS kan de streefwaarde bijvoorbeeld met 0,5% worden opgehoogd indien de uitval komende jaren tot 6% wordt gereduceerd. Op basis van de uitkomsten van de review kunnen NS en het ministerie van IenM publieksdefinitie bepalen in welke mate de voorgestelde verbetermaatregelen aanleiding geven tot aanscherping van de streefwaarde.

#### **Aanbevelingen**

- Bepaal op basis van de conclusies van de review op het Verbeterplan IC direct met het ministerie van IenM in welke mate de door NS en ProRail voorgestelde verbetermaatregelen haalbaar zijn en aanleiding geven tot aanscherping van de door NS voorgestelde streefwaarde van 94,4%. Is de conclusie van de review bijvoorbeeld dat 6% uitval haalbaar is, verhoog de streefwaarde dan met 0,5%.

## 3 Prestatie-indicatoren ProRail

Voor ProRail heeft de audit betrekking op drie prestatie-indicatoren, te weten Klantoordeel reizigersvervoerders, Klantoordeel goederenvervoerders en Klantoordeel klanthinder als gevolg van storingen aan de infrastructuur. De prestatie-indicatoren Klantoordeel reizigersvervoerders en Klantoordeel goederenvervoerders worden voor een belangrijk deel volgens dezelfde methodiek gemeten. Ook de opzet van de publieksdefinities en de voorgestelde wijzigingen van voornoemde prestatie-indicatoren verschillen weinig van elkaar. De bevindingen en aanbevelingen die gelden voor de prestatie-indicator Klantoordeel reizigersvervoerders zijn daarom ook van toepassing op de prestatie-indicator Klantoordeel goederenvervoerders. Punten die specifiek van toepassing zijn op het klantoordeel van goederenvervoerders zijn opgenomen in de betreffende paragraaf.

### 3.1 Klantoordeel reizigersvervoerders

#### 3.1.1 Publieksdefinitie

##### Bevindingen

De publieksdefinitie is niet op alle onderdelen helder, concreet en eenduidig interpreteerbaar beschreven:

- Een eerste aandachtspunt is het feit dat de vragenlijst volgens de publieksdefinitie alleen gericht is op vervoerders met een volwaardige toegangsovereenkomst die openbaar vervoer van reizigers verzorgen. Vervoerders met een beperkte toegangsovereenkomst participeren niet in het onderzoek. Dit zijn bijvoorbeeld vervoerders bij evenementen. Hun aandeel in het totale vervoersvolume is betrekkelijk gering.
- Een tweede aandachtspunt is dat vervoerders met een volwaardige toegangsovereenkomst die zeer beperkt gebruik maken van de spoorinfrastructuur, eveneens niet worden meegenomen in het onderzoek. Het blijkt te gaan om vervoerders die vanuit het buitenland een aantal honderd meter de grens over rijden om Nederland te bereiken. Dit staat in de publieksdefinitie niet beschreven. Er is ook geen ondergrens voor het gebruik van de infrastructuur gedefinieerd.
- Een derde aandachtspunt betreft de beschrijving van de scope van de prestatie-indicator. Deze beschrijving bevat een beperkte tijdsafbakening. De publieksdefinitie vermeldt niet dat het oordeel dat aan respondenten gevraagd wordt, betrekking heeft op het afgelopen kalenderjaar. Dit wordt wel gecommuniceerd naar de respondenten via de uitnodiging die zij krijgen.

##### Aanbevelingen

- Maak helder om welke reden de vragenlijst alleen wordt uitgezet onder vervoerders met een volwaardige toegangsovereenkomst en wat het aandeel van deze vervoerders in het totale vervoer is.

- Maak helder wat de ondergrens is voor de vervoerders die meegenomen worden in het onderzoek en verwerk deze in de publieksdefinitie. Of voeg aan de publieksdefinitie toe: '*alle vervoerders wiens hoofdactiviteit het is om reizigers **binnen Nederland** te vervoeren*'.
- Neem in de publieksdefinitie op dat in de vragenlijst gevraagd wordt om een oordeel over het afgelopen kalenderjaar.

### 3.1.2 Meetmethode en meetpraktijk

#### Bevindingen

Voor het klantonderzoek liggen de meetmethode en de meetpraktijk zodanig in elkaars verlengde, dat we deze in een paragraaf beschrijven.

De meetmethode betreft ten eerste het gebruik van de vragenlijst en ten tweede de manier waarop aan de hand van de ingevulde vragenlijsten de score op de prestatie-indicator berekend wordt.

Aan de ingevulde vragenlijsten valt op dat de kwantitatieve scores en kwalitatieve toelichting niet altijd consistent zijn. Een cijfer 7 gaat in sommige vragenlijsten gepaard met een positieve toelichting, terwijl hier in andere vragenlijsten juist een negatief oordeel aan is toegevoegd. Een mogelijke reden kan zijn dat de cijfers die gegeven kunnen worden niet per vraag zijn toegelicht. De verschillen roepen de vraag op hoe betrouwbaar de scores van de respondenten zijn en welke conclusies hieruit getrokken kunnen worden.

De weging van de vragen voor het berekenen van een score is deels gespecificeerd. Het generieke en specifieke deel tellen beiden voor 50% mee in de score. Het generieke deel bestaat uit drie aspecten met een gelijk gewicht. In de meetmethode maakt ProRail niet expliciet hoe de aspecten in het specifieke deel wegen, namelijk allemaal gelijk.

In de meetmethode is niet beschreven hoe ProRail omgaat met vragenlijsten die niet volledig zijn ingevuld. Dit is ook niet vermeld in een achterliggende rapportage van TNS NIPO, de organisatie die het onderzoek voor ProRail uitvoert. Uit gevoerde gesprekken blijkt dat de niet volledig ingevulde vragenlijsten wel worden meegenomen in de berekening van de score.

Een belangrijk onderdeel van de berekening van de score voor de prestatie-indicator is de weging van de vervoerders. Tot op heden wegen alle reizigersvervoerders even zwaar mee in de berekening van de score. NS zal vanaf 2017 voor 50% gaan meetellen in de berekening. De regionale vervoerders zullen eveneens voor 50% meetellen, waarbij elke vervoerder een gelijk gewicht zal hebben. Bij een beperkte deelname van de regionale vervoerders aan het onderzoek, blijft hun aandeel opgeteld 50%. Dit kan betekenen dat een klein aantal vervoerders veel gewicht in de schaal leggen in het bepalen van de score van de prestatie-indicator. Hierdoor bestaat het risico op een vertekening van het beeld dat in de score tot uitdrukking komt.

De weging van de vervoerders is niet conform het marktaandeel. ProRail geeft aan dat het gewicht van NS, mocht wel van deze weging sprake zijn, zou toenemen van circa 11% in de oude situatie naar circa 90% in de nieuwe situatie. Omdat deze stijging als te fors werd gezien, is besloten NS een gewicht van 50% te geven. In de volgende paragraaf rekenen we uit wat het effect is op de score als NS wel conform het marktaandeel meegenomen zou worden.

ProRail bepaalt op voorhand welke respondenten uitgenodigd worden voor het beantwoorden van de generieke en specifieke vragen. Voor de specifieke vragen worden de

relevante individuen benaderd. Bij te lage respons worden respondenten nog telefonisch benaderd door het onderzoeksbureau dat ProRail gebruikt. Met het onderzoeksbureau is een minimale respons afgesproken. Deze kwaliteitsborging is niet gedocumenteerd.

Het aantal respondenten kan per vervoerder sterk verschillen, afhankelijk van het aantal speerpunten en afspraken dat relevant is, het aantal benaderde respondenten en het aantal respondenten dat de vragenlijst invult. De kans dat één medewerker van NS de score voor 50% kan bepalen, is dankzij de afspraken over minimale respons zeer gering. Het risico op aantasting van de representativiteit geldt vooral voor het specifieke deel van de vragenlijst, dat in de praktijk door een beperkt aantal respondenten wordt ingevuld (veelal een per specifieke vraag per vervoerder). Het aantal respondenten per vraag is niet opgenomen in de rapportage over het onderzoek, zodat hier geen inzicht in bestaat.

De wijze waarop ProRail de score op de prestatie-indicator uiteindelijk berekent, is wel op hoofdlijnen, maar niet in detail beschreven. Hierdoor is uit de beschrijving niet op te maken in welke volgorde de scores van verschillende respondenten gemiddeld worden. Verschillende volgordes leveren in de praktijk uiteenlopende scores op. In de praktijk blijkt ProRail eerste een score per vraag per vervoerder te bepalen, dan een score per vervoerder en vervolgens een score over alle vervoerders. Deze volgorde is tot stand gekomen in samenwerking met TNS NIPO. BTB/AEF onderschrijft deze volgorde.

Wel is duidelijk dat in de verwerking van de vragenlijsten en het berekenen van de score voor de prestatie-indicator gewerkt wordt met het vier-ogen principe. Hoe dit gebeurt is echter ook niet helder beschreven.

### **Aanbevelingen**

- Voeg per vraag in de vragenlijst een korte beschrijving toe van de betekenis van de cijfers, zodat dit voor de respondenten helder is, bijvoorbeeld van ‘zeer ontevreden’ tot ‘zeer tevreden’.
- Borg de representativiteit van het klantoordeel door in de instructie voor het invullen van de vragenlijst meer richtlijnen op te nemen en de weging van vervoerders te herijken
- Leg beter vast hoe de ingevulde vragenlijsten verwerkt worden tot een score op de prestatie-indicator en geef daarbij aan:
  - hoe wordt omgegaan met niet volledig ingevulde vragenlijsten
  - op basis van welke methode (volgorde van middeling) de score berekend wordt
  - hoe de kwaliteit van de berekening om de score te bepalen, wordt geborgd
- Maak in de rapportage van het onderzoek inzichtelijk door hoeveel respondenten de vragen (generiek en specifiek) beantwoord zijn
- Stel een risicomatrix op waarin de risico's van het onderzoek beschreven zijn, evenals de maatregelen om deze te beheersen, bijvoorbeeld:
  - risico: beperkte respons van (een categorie van) vervoerders
  - maatregel: na x aantal dagen een herinnering per mail en na y dagen een telefonisch contact met de vraag om de vragenlijst alsnog in te vullen.

### **3.1.3 Informatie**

De wijze waarop de informatie verzameld wordt om de prestatie-indicator te berekenen, is in het voorgaande beschreven. Hierop zijn er geen toevoegingen.

### 3.1.4 Bodem- en streefwaarde

#### Bevindingen

Het plan is om de methodiek waarmee de klanttevredenheid wordt bepaald in 2017 op twee punten te veranderen:

- De weging van de vervoerders in het bepalen van de score wordt aangepast. NS (NS Reizigers en NS Internationaal) telt in de toekomst voor 50% mee en de overige vervoerders gezamenlijk eveneens voor 50%. Voorheen was het gewicht van elke vervoerder gelijk.
- Het onderwerp ‘prijs-kwaliteitverhouding’ van treinpaden in de vragenlijst vervalt.

Het voorstel van het ministerie van IenM is om de bodem- en streefwaarde, die in gehele getallen (zonder decimalen) worden uitgedrukt, niet aan te passen. De bodemwaarde blijft 6 en de streefwaarde blijft 7.

BTB/AEF is nagegaan wat de effecten zijn van de wijzigingen in de vragenlijst:

- De aanpassing van het gewicht van NS in de weging van vervoerders blijkt weinig uit te maken voor de score op de prestatie-indicator. De score neemt met 0,1 toe van 6,42 naar 6,52.
- Het verwijderen van de vraag over de prijs-kwaliteitverhouding van de treinpaden verandert de score van 6,42 naar 6,56.

Opgeteld verandert de score voor 2015 door de optelsom van veranderingen met 0,22 van 6,42 naar 6,64.

In het geval het marktaandeel als leidraad voor de weging genomen zou worden (zie voorgaande paragraaf), verandert de score van 6,42 naar 6,44. In de praktijk is het effect van deze aanpassing dus gering. Inclusief de verwijdering van de vraag over de prijs-kwaliteitverhouding van de treinpaden komt de score uit op 6,54.

Door de afronding op hele cijfers heeft deze kleine aanpassing in de score een substantieel effect op de uiteindelijke score. Het cijfer slaat namelijk om van 6 naar 7. Hiermee zou de streefwaarde van 7 nu al bereikt zijn en ontbreekt het dus aan een prikkel (voortvloeiend uit de prestatie-indicator) om de klanttevredenheid verder te verbeteren. Tegelijk is het beeld van BTB/AEF dat een bijstelling van de streefwaarde naar 8 te ambitieus is. Deze score is voor ProRail naar verwachting niet te realiseren.

Om een oplossing te bieden voor de beschreven situatie, stelt BTB/AEF voor de cijfers ‘af te ronden’ op halve punten. Dit wil zeggen dat de streefwaarde niet 7 of 8 hoeft te zijn, maar ook 7,5 kan zijn. Het gebruik van halve punten biedt voor de streefwaarde meer mogelijkheden. In de onderstaande tabel is aangegeven hoe de afronding bij het gebruik van halve punten kan geschieden.



Score	'Afgerond cijfer'
...	
6,25 – 6,74	6,5
6,75 – 7,24	7,0
7,25 – 7,74	7,5
7,75 – 8,24	8,0
...	

Door de streefwaarde op (afgerond) 7,5 vast te stellen wordt een oplossing gevonden voor het feit dat de realisatiewaarde door de aanpassingen in de meetmethode stijgt van 6 en 7, wat maakt dat er voor ProRail geen prikkel meer bestaat om de klanttevredenheid verder te verhogen. Het nieuwe streefcijfer betekent dat de ambitie zoals die ook in de huidige situatie bestaat, behouden blijft.

#### Aanbevelingen

- Verander de bodemwaarde niet en houd deze op hetzelfde niveau als in 2015.
- Werk voor de streefwaarde met een 'afronding' op halve punten en stel de streefwaarde vast op 7,5 (feitelijk dus 7,25 – 7,74) om daarmee voldoende ambitie in te bouwen.
- Hanteer in de toekomst schaduwberekeningen om het effect van aanpassingen in de inhoud van de vragenlijst en de wijze van verwerking te bepalen.

## 3.2 Klantoordeel goederenvervoerders

### 3.2.1 Publieksdefinitie

#### Bevindingen

De publieksdefinitie voor deze prestatie-indicator is gelijk aan die van de voorgaande, afgezien van de beschreven berekeningsmethodiek.

De scope is zodanig gedefinieerd, dat helder is welke goederenvervoerders wel en niet meegenomen worden in het klantoordeel. Hierbij valt op dat ook een brancheorganisatie wordt meegenomen. Deze wordt beschouwd als middelgrote vervoerder. BTB/AEF is hier geen voorstander van. Een brancheorganisatie heeft zelf geen praktische ervaring met de dienstverlening van ProRail en kan daardoor alleen een indirect oordeel geven (afgaand op hun leden). Daarnaast bestaat het risico dat sprake is van dubbeltellingen, omdat ook de vervoerders die vertegenwoordigd worden door de brancheorganisaties gevraagd wordt aan het onderzoek deel te nemen.

Voor het klantoordeel van goederenvervoerders is het voorstel van ProRail om de grote, middelgrote en kleine vervoerders elk voor een derde deel mee te laten wegen in de score. Het marktaandeel van een vervoerder wordt bepaald aan de hand van gereden kilometers in het voorgaande jaar. Kleine goederenvervoerders definieert ProRail als vervoerders met een marktaandeel tussen 0% en 5,0% van het marktaandeel, middelgrote als die met een marktaandeel van 5,0% en 24,9% en grote vervoerders als die met een marktaandeel tussen 25,0% en 100%. Binnen de drie categorieën hebben alle vervoerders een gelijk gewicht. Voor

de consistentie in de indeling in categorieën acht BTB/AEF het raadzaam de bovengrens van de categorie van kleine vervoerders vast te stellen op 4,9% in plaats van 5,0%.

**Aanbevelingen**

- Pas de bovengrens van de categorie van kleine vervoerders aan van 5,0% naar 4,9% om daarmee de consistentie in de indeling te borgen.
- Heroverweeg de deelname van brancheorganisaties aan het onderzoek naar het klantoordeel of voorkom dubbelstellingen die expliciet te maken namens welke vervoerders zij aan het onderzoek meedoen.

**3.2.2 Meetmethode en meetpraktijk**

Zoals hierboven beschreven volgt uit de publieksdefinitie dat de categorieën van kleine, middelgrote en grote vervoerders alle voor een derde deel meewegen in de score. Het wordt hierbij niet duidelijk om hoeveel partijen het per categorie gaat en wat het marktaandeel per categorie is. BTB/AEF heeft dit bij ProRail opgevraagd. De cijfers zijn opgenomen in de onderstaande tabel.

Categorie	Aantal vervoerders (absoluut)	Aantal vervoerders (procentueel)	Marktaandeel
Klein	12	75%	23%
Middelgroot	3	19%	23%
Groot	1	6%	54%
<b>Totaal</b>	<b>16</b>	<b>100%</b>	<b>100%</b>

In de gepresenteerde cijfers hebben we de brancheorganisatie buiten beschouwing gelaten.

Daar waar het gaat om de aantallen vervoerders zijn de grote vervoerders in het onderzoek oververtegenwoordigd. In termen van marktaandeel zijn de grote vervoerders juist ondervertegenwoordigd. Voor de kleine en middelgrote vervoerders geldt het omgekeerde.

In lijn met de benadering van het klantoordeel van de reizigersvervoerders is het voorstel van BTB/AEF om de aandelen van de categorieën meer in lijn te brengen met de marktaandelen. Dit doet beter recht aan de marktposities. Voor het gewicht van de categorieën betekent dit het volgende:

- Kleine vervoerders 25%
- Middelgrote vervoerders 25%
- Grote vervoerders 50%.

De bovenstaande cijfers gelden voor het jaar 2015.

**Aanbevelingen**

- Herijk de gewichten van de verschillende categorieën van goederenvervoerders, waardoor deze beter in lijn zijn met de marktaandelen.
- Monitor veranderingen in marktaandeel van de goederenvervoerders en pas hierop indien nodig jaarlijks de weging aan.
- Neem in de beschrijving van de meetmethode op wat per categorie het aantal vervoerders en hun opgetelde marktaandeel is.

### 3.2.3 Informatie

De wijze waarop de informatie verzameld wordt om de prestatie-indicator te berekenen, is in het voorgaande beschreven. Hierop zijn er geen toevoegingen.

### 3.2.4 Bodem- en streefwaarde

#### Bevindingen

De methodiek waarmee de klanttevredenheid wordt bepaald, verandert doordat:

- De weging van de vervoerders in de berekening anders wordt.
- Het onderwerp ‘prijs-kwaliteitverhouding’ van treinpaden in de vragenlijst vervalt.

Deze aanpassingen leiden wat betreft ProRail niet tot een wijziging van de bodem- en streefwaarde. Als bodemwaarde blijft een score van 6 en als streefwaarde een score van 7 gelden.

BTB/AEF is nagegaan wat de effecten van de wijzigingen zijn:

- De wijziging in de weging van vervoerders blijkt amper invloed te hebben op de score op de prestatie-indicator. De score stijgt met 0,01 van 5,68 naar 5,69.
- Door het verwijderen van de vraag over de prijs-kwaliteitverhouding van de treinpaden verandert de score van 5,68 naar 5,91.

Per saldo leiden de wijzigingen tot een aanpassing van de score van 5,68 naar 5,91. Deze aanpassing is zodanig beperkt, dat deze niet leidt tot een aanpassing van de bodem- en streefwaarde. BTB/AEF adviseert om vast te houden aan de huidige bodem- en streefwaarde.

BTB/AEF heeft ook berekend wat het effect is van de andere verdeling die ter overweging is voorgelegd (50-25-25):

- De wijziging in de weging van vervoerders blijkt amper invloed te hebben op de score op de prestatie-indicator. De score stijgt met 0,02 van 5,68 naar 5,70.
- Door het verwijderen van de vraag over de prijs-kwaliteitverhouding van de treinpaden verandert de score zoals gezegd van 5,68 naar 5,91.

Beide maatregelen gecombineerd hebben een gezamenlijke stijging van 0,25 punten. De score verandert van 5,68 naar 5,93.

De streefwaarde voor 2019 is in 2016 bijgesteld van een 6 naar een 7. De streefwaarde van een score van 7 geeft naar het oordeel van BTB/AEF voldoende ambitie voor ProRail. Het verder verhogen van het streefcijfer is wat ons betreft te ambitieus. Een streefwaarde van 6 geeft juist weer te weinig ambitie, omdat ProRail dan feitelijk geen aanvullende inspanningen ten opzichte van de huidige situatie hoeft te plegen.

ProRail scoort op dit moment in de ogen van de goederenvervoerders minder goed dan in die van de reizigersvervoerders. Wat BTB/AEF zou de inspanning van ProRail erop gericht moeten zijn om dit verschil op te heffen en net als voor de reizigersvervoerders op het streefcijfer van 7 uit te komen.

#### Aanbevelingen

- Verander de bodem- en streefwaarde niet en houd deze op hetzelfde niveau als in 2016.
- Hanteer in de toekomst schaduwberekeningen om het effect van aanpassingen in de inhoud van de vragenlijst en de wijze van verwerking te bepalen.

### 3.3 Klanthinder als gevolg van storingen infra

#### 3.3.1 Publieksdefinitie

##### Bevindingen

De publieksdefinitie van de prestatie-indicator is helder, concreet en eenduidig beschreven, behoudens een beperkt aantal aandachtspunten.

Ten opzichte van de audit van BTB/AEF in 2015 is in de publieksdefinitie nu wel helder welke storingen wel en niet onderdeel zijn van de prestatie-indicator en ook op welke wijze deze gemeten wordt.

In de publieksdefinitie van de prestatie-indicator worden termen gebruikt die niet nader gedefinieerd worden. Het gaat hier bijvoorbeeld om functieherstel en be- en bijsturing. Onze suggestie is om deze termen nader te beschrijven, zodat helder is wat hieronder wordt verstaan. Verder doen we de suggestie om in de publieksdefinitie ook aan te geven op welke tijdsbasis ProRail de prestatie-indicator rapporteert. Intern gebeurt dit maandelijks, waarbij gekeken wordt naar de score van de afgelopen maand, de score *year to date* en de score *moving annual total*. Aan het ministerie van IenM rapporteert ProRail per halfjaar en per jaar over de klanthinder. De bovenstaande punten zijn suggesties om de communicatie over de prestatie-indicator te verbeteren, maar staan los van de kwaliteit van de publieksdefinitie.

##### Aanbevelingen

- Gelet op het voorgaande hebben wij geen aanbevelingen.

#### 3.3.2 Meetmethode

##### Bevindingen

De voorgestelde meetmethode is naar het oordeel van BTB/AEF op dit moment de meest valide methode om de hinder voor klanten te meten.

De naam klanthinder suggereert dat de prestatie-indicator meet hoeveel hinder klanten in de praktijk ervaren. Wat echter gemeten wordt, is het aantal storingen aan de infrastructuur en het aantal treinvertragsminuten dat deze storingen opleveren. Omdat de correlatie tussen de indicator klanthinder en de samengestelde waarde van punctualiteit en uitval hoog is ( $>0,7$ ), geeft de prestatie-indicator wel een valide beeld van hinder die reizigers ervaren.

Een toekomstige stap kan zijn om de klanthinder te berekenen aan de hand van gegevens van de OV-chipkaart. Deze geven een nog betrouwbaarder beeld van de klanthinder. ProRail geeft aan dat het op dit moment niet mogelijk is om deze gegevens in te zetten, onder meer omdat het niet te bepalen is hoeveel reizigers in vertraagde, omgeleide of opgeheven treinen zitten. BTB/AEF doet voor de langere termijn de suggestie om in beeld te brengen op welke wijze OV-chipkaart gegevens wel te gebruiken zijn en de reizigershinder directer in beeld te brengen.

In de nieuwe meetmethode werkt ProRail niet meer met een indeling van vertragingen in tijdscategorieën om de klanthinder te meten, maar maakt men gebruik van de feitelijke vertraging van treinen. Dit is mogelijk door de beschikbaarheid van betere meetgegevens. Het gebruik van feitelijke vertragingen biedt naar het oordeel van BTB/AEF een betrouwbaarder beeld van de klanthinder dan in het verleden.

Ten opzichte van de huidige situatie worden ook leeg materieel ritten in de toekomst niet meer meegenomen in de prestatie-indicator. BTB/AEF onderschrijft deze aanpassing, omdat deze voor reizigers niet (direct) relevant zijn en hierdoor ook dubbeltellingen worden voorkomen.

De bijgestelde maximering van vertraging per treinnummer op 30 minuten (voorheen 60 minuten) is ingegeven door een analyse van gegevens afkomstig uit het gebruik van de OV-chipkaart de zogenaamde A2-corridor. Het beeld hieruit is dat reizigers bij vertragingen maximaal 30 minuten extra reistijd hebben. Omdat de treinseries die op de A2-corridor rijden 25% van het totaal aantal vertraagde reizen veroorzaken, is sprake van een representatieve steekproef. BTB/AEF onderschrijft de bijstelling van de maximale vertraging.

Ook de keuze om voor omgeleide treinen uit te gaan van 15 minuten vertraging is helder onderbouwd en plausibel. Dit geldt tot ook voor de vertraging van 30 minuten die wordt aangehouden voor opgeheven treinen.<sup>13</sup>

#### **Aanbevelingen**

- Gelet op het voorgaande hebben wij geen aanbevelingen.

### **3.3.3 Meetpraktijk**

#### **Bevindingen**

De storingen die plaatsvinden deelt ProRail zoveel mogelijk toe aan categorieën van oorzaken. Dit proces verloopt in de praktijk conform de procedure. Het is (in de praktijk) echter niet altijd duidelijk wat de oorzaak is van een storing. In voorkomend geval voert ProRail nader onderzoek uit. Het komt voor dat een oorzaak uiteindelijk niet meer te achterhalen is. De storing komt dan in de categorie 'onbekend' terecht. Het aantal storingen in deze categorie is dermate klein, dat dit de berekening van de prestatie-indicator maar zeer beperkt beïnvloedt.

Naast een toewijzing aan een oorzaak, worden storingen ook toegewezen aan een bedrijfsonderdeel van ProRail of vervoerders. Na een voorlopige toedeling krijgt het betreffende onderdeel 28 dagen de tijd om deze te fiatteren. Gebeurt dit niet, dan wordt de toewijzing publieksdefinitie en kan hier niet meer van afgeweken worden. Het voordeel hiervan is dat alle storingen toegewezen worden aan een bedrijfsonderdeel. Het risico dat een onderdeel niet reageert op een onterechte toewijzing, is in de praktijk beperkt.

Voor de registratie van de vertragingen, opheffingen, omleidingen en verleggingen worden er periodiek steekproeven gedaan om de kwaliteit van de registratie te bepalen. De uitkomsten van deze steekproeven gebruikt ProRail om het vastleggingsproces daar waar nodig verder te verbeteren. Er zijn rapportages (Dashboard Monitoring) beschikbaar die inzicht geven in deze uitkomsten en daarmee in de kwaliteit van de registratie van de vertragingen. In deze rapportages worden ook conclusies en verbeteracties verbonden aan de waarnemingen omtrent kwaliteit. Uit de registratie blijkt dat circa 15% van de registraties een fout bevat. ProRail stuurt erop om deze fouten te herstellen.

Het proces van toewijzing van vertragingen aan storingen is ten opzichte van de voorgaande audit verbeterd, in de zin dat het proces nu helder in beeld is. Verder is door het uitvoeren van

<sup>13</sup> In het consultatiedocument dat door ProRail is voorgelegd aan Locov staat dat voor zowel omgeleide als opgeheven treinen met 30 minuten vertraging wordt gerekend. Dit is niet in lijn met het feitelijke voorstel. Dit is reeds onder de aandacht gebracht van ProRail.

interne checks vastgesteld dat slechts een klein percentage van de storingen niet wordt toegedeeld aan een vertraging. Binnen ProRail loopt momenteel een actie om alle vertragingen toe te delen aan een storing. Deze actie wordt naar verwachting van ProRail eind van het jaar afgerond.

ProRail monitort de kwaliteit van de meetgegevens voor de vertragingen op een adequate manier. De aanpassingen in de wijze van berekening, waaronder het feit dat de feitelijke vertragingen worden meegenomen en het instellen van een maximum aan een vertraging van 30 minuten, maakt dat de meetmethodiek en daarmee ook de meetpraktijk verandert. Een van deze veranderingen is, dat de gegevens in de toekomst afkomstig zijn uit meerdere systemen. Hierdoor moet ook het proces van kwaliteitsborging herijkt worden. Omdat dit nog plaats moet vinden, kan BTB/AEF nog geen oordeel geven over de toekomstige borging van de kwaliteit.

Als onderdeel van de kwaliteitsborging geldt dit ook voor de wijze van risicobeheersing. Er dient nog een nieuwe inventarisatie van risico's plaats te vinden en een vertaling hiervan naar beheersmaatregelen.

#### **Aanbevelingen**

- Herijk gezien de wijzigingen in de meetpraktijk de wijze waarop de kwaliteit van de meetpraktijk wordt gemonitord en pas ook de risicobeheersing hierop aan.

### **3.3.4 Informatie**

#### **Bevindingen**

Voor de berekening van de prestatie-indicator worden alle storingen en hieraan gekoppelde treinvertragingen meegenomen. Er spelen dan ook geen vragen over de representativiteit en daarmee statistische betrouwbaarheid. Alle gegevens uit de meetsystemen worden meegenomen. In de voorgaande paragrafen is beschreven hoe ProRail hiermee omgaat.

#### **Aanbevelingen**

- Gelet op voorgaande hebben wij geen aanbevelingen.

### **3.3.5 Bodem- en streefwaarde**

#### **Bevindingen**

Tot voorheen was de klanthinder een informatie-indicator en geen prestatie-indicator. Er was dan ook geen sprake van een bodem- en streefwaarde. Intern binnen ProRail waren deze wel al bepaald. Deze hebben formeel echter geen status. De bodem- en streefwaarde die in de huidige situatie door ProRail gehanteerd worden, zijn:

- Bodemwaarde 610 storingen in klasse 1 en 2 gezamenlijk
- Streefwaarde 546 storingen in klasse 1 en 2 gezamenlijk.

De vertaling van de (informele) huidige naar de (formele) nieuwe bodem- en streefwaarde is gebeurd door de grenzen van de hinderklassen die onderscheiden worden aan te passen, zodanig dat het aantal storingen in de hindercategorieën gelijk blijft. In de onderstaande tabel zijn de huidige en nieuwe bandbreedtes opgenomen. Hier veranderen de bodem- en streefwaarde in het beeld van ProRail niet.

Hindercategorie	Huidige bandbreedte (treinvertragsminuten)	Huidig aantal storingen	Nieuwe bandbreedte (treinvertragsminuten)	Nieuw aantal storingen
Categorie 1	> 6.000	84	> 2.400	85
Categorie 2	1.500 – 6.000	546	680 – 2.400	547
Categorie 3	60 – 1.500	2.934	40 – 680	2.963
Categorie 4	< 60	7.353	< 40	7.322

De vraag aan BTB/AEF is of de bodem- en streefwaarde voldoende ambitie bevatten ten opzichte van de tot nu toe gerealiseerde data en of beide waarden logisch volgen uit de sturingsmogelijkheden van ProRail.

Om deze vragen te beantwoorden hebben we een analyse gemaakt van de ontwikkeling van het aantal beïnvloedbare en niet-beïnvloedbare storingen per hindercategorie in de periode vanaf 2011. De beïnvloedbare storingen zijn die door processen en techniek, terwijl de niet-beïnvloedbare storingen betrekking hebben op weer en derden. De ontwikkeling van het aantal storingen is opgenomen in de onderstaande tabel. Per jaar wordt het aantal storingen getoond en de delta ten opzichte van het voorgaande jaar (de delta wordt weergegeven vanaf 2012, en is niet beschikbaar voor 2011 omdat de gegevens van 2010 bij ons niet bekend zijn).

Categorie	2011	2012	Δ	2013	Δ	2014	Δ	2015	Δ	Gem. Δ
<b>Hindercat. 1</b>	<b>7</b>	<b>26</b>	<b>271%</b>	<b>62</b>	<b>138%</b>	<b>69</b>	<b>11%</b>	<b>84</b>	<b>22%</b>	<b>111%</b>
- Beïnvloedbaar	5	21	320%	35	67%	42	20%	37	-12%	99%
- Niet-beïnvloedbaar	2	5	150%	27	440%	27	0%	47	74%	166%
<b>Hindercat. 2</b>	<b>138</b>	<b>244</b>	<b>77%</b>	<b>470</b>	<b>93%</b>	<b>492</b>	<b>5%</b>	<b>541</b>	<b>10%</b>	<b>46%</b>
- Beïnvloedbaar	67	118	76%	226	92%	262	16%	242	-8%	44%
- Niet-beïnvloedbaar	71	126	77%	244	94%	230	-6%	299	30%	49%
<b>Totaal</b>	<b>145</b>	<b>270</b>	<b>86%</b>	<b>532</b>	<b>97%</b>	<b>567</b>	<b>7%</b>	<b>625</b>	<b>10%</b>	<b>50%</b>
- Beïnvloedbaar	72	139	93%	261	88%	304	16%	279	-8%	47%
- Niet-beïnvloedbaar	73	131	79%	271	107%	257	-5%	346	35%	54%

Uit de tabel komt naar voren dat het aantal storingen in de hindercategorieën 1 en 2 sinds 2011 sterk is toegenomen. De beïnvloedbare storingen zijn na een jarenlange stijging voor het eerst afgenomen. Uit de cijfers valt niet op te maken of hier sprake is van een trendbreuk. ProRail geeft wel aan dat gericht gestuurd wordt op het beperken van de beïnvloedbare storingen en dat deze sturing zijn vruchten afwerpt. De sturing bestaat uit het beperken van het aantal storingen in absolute zin en het verminderen van de vertraging per storing, zodat deze in een andere categorie gaan vallen. Deze sturing is mogelijk op zowel de beïnvloedbare als de niet-beïnvloedbare storingen. Voor de niet-beïnvloedbare storingen is het bijvoorbeeld

mogelijk om de gevolgen van weersinvloeden te beperken. Per publieksdefinitie zijn de beïnvloedbare storingen beter te beheersen dan de niet-beïnvloedbare storingen.

De bodemwaarde voor de prestatie-indicator ligt onder de waarde die in 2015 gerealiseerd is. Om de bodemwaarde te halen is een afname van het aantal storingen nodig van 2,4%. Deze daling dient binnen twee jaar te worden gerealiseerd. Deze daling staat haaks op de trend van de afgelopen jaren, waarin het aantal storingen is toegenomen. Zoals beschreven is deze trend in 2015 voor de beïnvloedbare storingen wel gekeerd.

De bodemwaarde zoals door ProRail voorgesteld is bepaald op basis van de storingsgegevens van Q2 2015 t/m Q1 2016. De voorgestelde bodemwaarde is gelijk aan de in deze periode gerealiseerde waarde. Vanuit deze optiek is de bodemwaarde helder onderbouwd. Omdat het gaat om een ondergrens voor de te leveren prestatie is deze ook realistisch en haalbaar.

Voor de streefwaarde geldt dat deze is gebaseerd op een afname van het aantal storingen in de hindercategorieën 1 en 2 van 2,5% per jaar in de periode tot 2019. Gerelateerd aan de ontwikkeling van het aantal storingen en de omvang van de storingen in de afgelopen jaren, is dit een ambitieus doel. Binnen ProRail is echter sprake van een groot aantal initiatieven om de voorgestelde daling te realiseren, waaronder het programma Beter en Meer, het reductieprogramma TAO's (met een focus op reductie van storingen door derden), het programma Prestatiegericht Onderhoud (PGO) en het Programma Meldkamer Spoor. Daarnaast zijn er technische verbeteringen (onder meer LED-seinen en een nieuw type lassen) en is er een separaat project voor het verminderen van de beïnvloedbare storingen door processen. De initiatieven richten zich op zowel een afname van het aantal storingen als een daling van de vertraging per storing.

ProRail heeft geen zicht op de effecten van deze initiatieven op het aantal storingen en de vertraging per storing. De afname van het aantal storingen van 2,5% per jaar is gebaseerd op een schatting. Door het ontbreken van een berekening, is het voor BTB/AEF niet mogelijk om deze afname te verifiëren. De beoordeling van de streefwaarde dient dus op andere gronden plaats te vinden.

Als we de aanname doen dat ProRail het aantal niet-beïnvloedbare storingen in de hindercategorieën 1 en 2 ten opzichte van 2015 weet te stabiliseren, dan dient de voorgestelde daling volledig gerealiseerd te worden door een afname van het aantal beïnvloedbare storingen. Dit is een afname van circa 28% over drie jaar (2017-2019), ofwel circa 10% per jaar.

Afgelopen jaar bedroeg de afname van het aantal beïnvloedbare storingen 8%. Gezien de eerder beschreven initiatieven die ProRail uitvoert, lijkt de afname met circa 10% per jaar haalbaar en is de streefwaarde binnen bereik. Zoals aangegeven geldt hier de aanname dat het aantal niet-beïnvloedbare storingen constant blijft. De afgelopen jaren laten deze echter nog een stijgende trend zien. Vanuit deze optiek adviseert BTB/AEF de streefwaarde niet lager vast te stellen en de ambitie dus niet verder op te voeren. Onze beeld is dat dit niet realistisch is.

Overall heeft BTB/AEF het beeld dat de bodemwaarde en streefwaarde realistisch en tegelijk ambitieus zijn vastgesteld.

#### **Aanbevelingen**

- Gelet op voorgaande hebben wij geen aanbevelingen.



### 3.3.6 Opvolging eerdere aanbevelingen

In de voorgaande audit van BTB/AEF zijn voor deze prestatie-indicator vier aanbevelingen gedaan. Hieronder beschrijven we hoe ProRail hiermee is opgegaan.

***Aanbeveling 1. Scherp in de procesbeschrijving de publieksdefinitie van de prestatie-indicator aan, waarbij wordt aangegeven dat twee categorieën van hinder worden meegenomen, dat de eenheid waarin de indicator wordt uitgedrukt het aantal storingen per categorie is en dat alle geplande treinritten (ook goederen en leeg materieel) meegenomen worden.***

ProRail heeft de publieksdefinitie van de prestatie-indicator aangepast. De treinritten van leeg materieel zijn uit de scope verwijderd (zie tekst hierboven). De aanbeveling is opgevolgd.

***Aanbeveling 2. Werk voor deze prestatie-indicator met een risicodossier, waarin de geconstateerde risico's rond de meetpraktijk en getroffen risicobeheersmaatregelen staan vastgelegd; in het auditrapport wordt verwezen naar een beschrijving van het monitoringsysteem waarin risico's die spelen inzichtelijk zijn gemaakt.***

Het ontbreekt nog aan een uitgeschreven risicodossier. De risico's in de berekening van de prestatie-indicator zijn binnen ProRail wel bekend en er wordt erop gestuurd deze zo goed mogelijk te beheersen. Dit is echter niet vastgelegd.

#### **Aanbeveling**

- Stel alsnog een risicodossier op dat intern gebruikt kan worden om te sturen op de kwaliteit van de prestatie-indicator.

***Aanbeveling 3. Stel als onderdeel van de procesbeschrijving een procedure op voor het koppelen van TVTA's aan TAO's.***

Deze beschrijving is inmiddels beschikbaar.

***Aanbeveling 4. Voer opnieuw een check in op het proces van het toewijzen van treinvertragingen aan storingen aan de infrastructuur.***

Deze check wordt nu periodiek uitgevoerd, waardoor de kwaliteit van de toewijzing in beeld is en beter op deze kwaliteit gestuurd kan worden. Uit interne rapportages van ProRail blijkt dat dit ook gebeurt.

## 4 Gezamenlijke prestatie-indicatoren

De prestatie-indicatoren Reizigerspunctualiteit < 5 minuten HRN, Reizigerspunctualiteit < 15 minuten HRN, Reizigerspunctualiteit < 5 minuten HSL en Reizigerspunctualiteit < 5 minuten HSL komen voor wat betreft de publieksdefinitie, de meetmethode en de meetpraktijk sterk overeen. De bevindingen en aanbevelingen die gelden voor 'Reizigerspunctualiteit < 5 minuten HRN' zijn daarom ook van toepassing op de overige prestatie-indicatoren. Bevindingen en aanbevelingen die specifiek zijn voor deze prestatie-indicatoren zijn opgenomen in de betreffende paragrafen.

### 4.1 Reizigerspunctualiteit < 5 minuten HRN

#### 4.1.1 Publieksdefinitie

##### Bevindingen

De publieksdefinitie van de prestatie-indicator is compleet en specifiek geformuleerd. De beknopte beschrijving van de meetsystematiek draagt bij aan de duidelijkheid van de publieksdefinitie. Hierbij merken we (net als bij de prestatie-indicatoren Vervoercapaciteit) op, dat de beschreven stappen voor een relatieve buitenstaander naar verwachting moeilijk te volgen zijn. Wij geven om die reden in overweging om de beschrijving van het meetsysteem in de publieksdefinitie te versimpelen. Dit is bijvoorbeeld mogelijk door het gebruik van een stroomdiagram, waarin een schematische voorstelling van de beschreven processtappen is opgenomen.

De tekst van de publieksdefinitie beschouwen we, op een aantal aandachtspunten na, als eenduidig. In de beschrijving van het meetsysteem in de publieksdefinitie staat het volgende:

*'Op de belangrijkste herkomst-bestemmingen van het vorige kwartaal (met meer dan 100 reizigers per 100 dagen en reizen op 20 verschillende dagen) wordt in de reisplanner op twee dagen van tevoren gekeken wat de reismogelijkheden zijn op elke dag.'*

Dit roept de vraag op of hiermee wordt bedoeld 20 verschillende dagen per jaar of per 100 dagen. Uit de onderliggende documentatie en de gesprekken concluderen we dat het om 20 dagen per 100 dagen gaat. Dit kan helderder in de publieksdefinitie vermeld worden.

De voorlaatste zinnen in de publieksdefinitie zijn bovendien multi-interpretabel. De zinsnede 'exclusief Thalys en Eurostar' kan zowel worden geïnterpreteerd als uitzondering op de zinsnede 'behalve in een internationale trein met zitplaatsreservering of toeslag' als een beperking ten opzichte van de gehele zin 'In scope zijn alle reizigers die inchecken én uitchecken op een Hoofdrailnet station en met NS reizen behalve in een internationale trein met

*zitplaatsreservering of toeslag*<sup>1</sup>. De publieksdefinitie moet dusdanig worden geformuleerd dat duidelijk is dat de Thalys en de Eurostar buiten scope zijn, omdat reizigers die in- en uitchecken (via de poortjes en paaltjes) niet met deze treinen kunnen reizen. In de ICE, IC Brussel en IC Berlin is in- en uitchecken wel mogelijk, wat maakt dat deze treinen wel onderdeel zijn van de scope.

De publieksdefinitie is helder afgebakend in tijd en scope. Voor het bepalen van de reisbelofte is een keuze gemaakt is voor een termijn van twee dagen. De reisbelofte van twee dagen voor de dag waarop een reis wordt gemaakt, is de norm ten opzichte waarvan de reizigerspunctualiteit wordt gemeten. Een kortere termijn van bijvoorbeeld een uur maakt dat verstoringen die plaatsvinden net voor een geplande reis in de reisbelofte worden verwerkt en derhalve geen (negatief) effect hebben op de reizigerspunctualiteit. Met de termijn van twee dagen is ook de consistentie ten opzichte van de huidige prestatie-indicator geborgd. Meer in het bijzonder zijn geplande werkzaamheden zijn verwerkt in de reisbelofte. Zij kunnen daarom geen afwijking veroorzaken. De situatie van een landelijk uitgedunde dienstregeling (bijvoorbeeld ingeval van winterse omstandigheden) is enkel verwerkt in de reisbelofte indien deze meer dan twee dagen van tevoren is aangekondigd. Volgens NS is dat echter zelden tot nooit het geval, en heeft een landelijk uitgedunde dienstregeling een negatieve impact op de reizigerspunctualiteit.

#### **Aanbevelingen**

- Maak in de publieksdefinitie duidelijk dat het om 20 dagen per 100 dagen gaat.
- Scherp de scope zodanig aan, dat helder is welke treinen hier wel/niet in vallen.

### **4.1.2 Meetmethode**

#### **Bevindingen**

De meetmethode is over het geheel genomen nauwkeurig, compleet en consistent beschreven. Intern is deze gevisualiseerd in een stroomschema.

De nauwkeurigheid van de meetmethode is verbeterd in vergelijking met de huidige prestatie-indicator Reizigerspunctualiteit. Er is sprake van een fundamenteel andere manier van meten. Kort gezegd geeft de huidige prestatie-indicator Reizigerspunctualiteit een gewogen gemiddelde weer van de treinpunctualiteit aan de hand van meetpunten 35 meetpunten bij stations. De weging vindt plaats op basis van het percentage (geprognoseerde) reizigers. Hierbij wordt een beperkt aantal vooraf gedefinieerde aansluitingen meegenomen. In de nieuwe prestatie-indicator tellen (in beginsel) alle reizigers mee die in- en uitchecken. De reizigerspunctualiteit wordt berekend op basis van de OV-chipkaart data, de bijbehorende reisbeloftes en de uitvoering daarvan. De reisbelofte bepaalt het beloofde aankomsttijdstip en de realisatie geeft vervolgens aan of de reis is verlopen zoals beloofd, of dat er vertraging is opgelopen.

Voor de berekening van de prestatie-indicator Reizigerspunctualiteit in het meetsysteem RPUN moesten enkele aannames gedaan worden. NS heeft ervoor gekozen om een zogenaamde *Failure mode and effect analysis* (FMEA) uit te voeren op de aannames. In deze analyse zijn alle aannames in een lijst gezet en voor elke aanname foutsituaties besproken en ingeschat wat de score is op de volgende punten:

- kans op voorkomen
- detecteerbaarheid
- impact op prestatie-indicatorenniveau
- impact op detailniveau.

Via deze methode is gekomen tot een score per aannames. NS heeft de meetmethode om te komen tot de indicator opgedeeld in zeven stappen c.q. componenten. De aannames met de hoogste scores heeft NS voor vijf van de zeven componenten (opvragen reisbelofte, bepalen realisatie, herplannen, koppelen OV-chipkaart data aan reisbeloftes en genereren datamart) van de meetmethode geclusterd en uitgebreid getoetst. De wegingsmethodiek en de toetsing zijn helder beschreven en vastgelegd in een centraal toegankelijk rapport.

De belangrijkste observaties in dit verband zijn:

- Een van de belangrijkste aannames in het bepalen van de reizigerspunctualiteit is, dat de reiziger wil reizen van z'n check-in naar z'n check-out station op het tijdstip van inchecken. In uitzonderlijke gevallen zal deze aannames niet opgaan. NS heeft geen informatie over reizigers die hebben afgezien van een beoogde verplaatsing per trein. Doordat deze reizigers afzien van hun reis checken ze niet in en uit.
- Voorts wordt er voor reizigerspunctualiteit uitgegaan van de beloofde reis volgens de reisplanner direct na check-in. Als dit reisadvies gerealiseerd kan worden (d.w.z. alle treinen in dit reisadvies rijden en alle aansluitingen worden gehaald), dan wordt de vertraging van de reizigers gemeten als de aankomstvertraging van de laatste trein in het reisadvies. Alleen in situaties waarbij de reis niet gerealiseerd kan worden, en derhalve de reizigers ook vertraging oplopen, wordt er gekeken naar de check-out tijd. Alleen in dit geval leidt een langer verblijf in het station tot een grotere vertraging dan feitelijk opgelopen door de reiziger.
- Een ander belangrijk punt is dat als een reiziger te laat op zijn bestemming aankomt, doordat een trein is uitgevallen, dan zal dit gemeten worden. Uitval telt dus mee in de bepaling van reizigerspunctualiteit. Dit was in de huidige indicator HRN reizigerspunctualiteit ook het geval.

Gelet op nauwkeurigheid van de wegingsmethodiek voor de aannames en het grote aantal aannames dat is getoetst, is het beeld van BTB/AEF dat de methodiek voldoende valide is. Aandachtspunt is echter wel dat voor de componenten c.q. stappen in de meetmethode 'verzamen van CICO-data' en 'bepaling indicator' de aannames niet zijn getoetst en vastgelegd door NS.<sup>14</sup>

#### Aanbeveling

- Maak inzichtelijk wat de onderliggende aannames zijn voor de componenten c.q. stappen in de meetmethode 'verzamen van CICO-data' en 'bepaling indicator', en hoe deze zijn getoetst. Indien dit niet is gebeurd, voer deze toetsing uit.

### 4.1.3 Meetpraktijk

#### Bevindingen

De meetpraktijk voor de prestatie-indicatoren Reizigerspunctualiteit verschilt niet wezenlijk van de meetpraktijk voor de prestatie-indicator Vervoercapaciteit.

In de meetpraktijk worden gegevens vanuit verschillende (primaire) bronnen verzameld, verwerkt en samengevoegd. De OV-chipkaart data en de reisbelofte worden bijvoorbeeld gekoppeld aan de uitvoering van het reisplan. Dit gebeurt door middel van het draaien van

<sup>14</sup> De laatste stap c.q. component 'bepaling indicator' is het laatste onderdeel van de meetmethode, en ziet dus niet op de gehele meetmethode. Mogelijk liggen ook aan de laatste stap aannames ten grondslag die nog in beeld moeten worden gebracht en vastgelegd door NS.

query's vanuit de verschillende gegevensbronnen in een daarvoor ingerichte digitale omgeving RPUN. Aan het einde van dit proces wordt de prestatie-indicator berekend.

Voor de sturing en de organisatie van de meetpraktijk zijn de taken, verantwoordelijkheden en bevoegdheden van de verschillende actoren in het proces vastgelegd<sup>15</sup> en is de rapportagecyclus beschreven:

- De berekening van de aantallen reizigers totaal en binnen normtijd per belofte datum is belegd bij het NSR kenniscentrum. Binnen het Kenniscentrum zijn geen specifieke verantwoordelijken aangewezen voor de berekening.
- De afdeling Business Intelligence & Analytics - verantwoordelijk voor het verzamelen en verwerken van de (primaire) data - ontbreekt in de beschrijving van het organisatiemodel. Gelet op de belangrijke rol, vindt BTB/AEF het van belang om deze afdeling hierin wel op te nemen.
- De aanlevering en verwerking van (rapportages over de kwaliteit van de) interne en externe brondata is niet vastgelegd in de rapportagecyclus. Dit terwijl het kenniscentrum wel afhankelijk is van interne en externe aanlevering van, de (kwaliteit)controle op en de verwerking van de brondata.

Op de meetpraktijk voert NS voldoende controles uit, waaronder controles op de koppeling van reizen aan reisbeloftes. Ook is het beeld dat er sprake is van voldoende controle op de kwaliteit van de brondata. De taken en verantwoordelijkheden alsook de rapportages zijn op dit punt niet vastgelegd. Wij geven in overweging om dit wel te doen.

Voor wat betreft het niveau van de eindverantwoordelijken voor de prestatie-indicator heeft NS niet inzichtelijk gemaakt in hoeverre zij actief de gehele meetpraktijk van de prestatie-indicatoren reizigerspunctualiteit (op basis van de beschikbare (kwaliteits)rapportages) monitoren. Daarnaast ontbreekt het aan vaste momenten, waarop de gehele meetpraktijk met betrokkenen (waaronder ProRail) onder leiding van de eindverantwoordelijken wordt geëvalueerd.

NS heeft de marge van onzekerheden in het meetsysteem kwalitatief inzichtelijk gemaakt door de beperkingen en afwegingen van het meetsysteem en de meetpraktijk in kaart te brengen. Een voorbeeld van een onzekerheid is dat er geen weging naar reisafstand plaatsvindt. De onderliggende afweging is dat elke reiziger voor NS en ProRail even belangrijk is. Uit de audit is naar voren gekomen dat de marge van onzekerheden in het meetsysteem ook kwantitatief in beeld kan worden gebracht. Denk bijvoorbeeld aan het aantal reizigers dat uit de (reguliere) dataset valt door bijvoorbeeld een verkeerde waarneming (zoals reizigers die vergeten om uit te checken). Vaak blijken deze foutmarges zeer beperkt te zijn. Dat betekent echter niet dat deze niet (structureel) in beeld kunnen worden gebracht en worden vastgelegd.

Tot slot valt op dat er niet wordt gewerkt met een (algemene) matrix waarin de risico's en beheersmaatregelen zijn vastgelegd. Een groot risico is bijvoorbeeld dat reisadviezen niet op tijd worden opgevraagd. De data inzake de reisadviezen wordt slechts een week lang opgeslagen. Daarna verdwijnt de data. Er bestaat geen beheersmaatregel voor deze situatie.

<sup>15</sup> Zie voor een overzicht van de sturing & organisatie het format zitplaatskans HRN onder meetpraktijk.

### Aanbevelingen

- Maak medewerkers binnen het NSR Kenniscentrum verantwoordelijk voor het verzamelen van de benodigde data en de berekening van de prestatie-indicator om daarmee de kwaliteit en continuïteit van de berekening en de resultaten beter te borgen.
- Neem de (interne en externe) leveranciers van de brondata op basis waarvan de indicator wordt bepaald en de rol van Business Intelligence & Analytics op in het organisatiemodel. Hierdoor is makkelijker te sturen op tijdigheid en kwaliteit van de brondata.
- Leg de interne en externe aanlevering van, de (kwaliteit)controle op en de verwerking van de brondata vast in een rapportagecyclus.
- Maak inzichtelijk hoe de eindverantwoordelijke functionaris voor de KPI binnen NS de kwaliteit monitort.
- Evalueer de meetpraktijk halfjaarlijks met alle betrokken om het lerend vermogen te vergroten en formaliseer deze evaluatie.
- Breng de onzekerheden in het meetsysteem ook kwantitatief in beeld en leg deze in een document vast. Leg daarnaast ook de belangrijkste risico's en de bijbehorende beheersmaatregelen vast.

### Verbetering KPI-rapportageproces

NS heeft medio 2016 een risicoanalyse opgesteld van de kwetsbare delen van de KPI-processen. Op basis hiervan is een plan van aanpak geschreven met benodigde activiteiten en te behalen resultaten. In de kern gaat het om de volgende aandachtsgebieden:

Aandachtsgebied	Verbetermaatregelen
Aansluiting interne sturing en eigenaarschap	<ul style="list-style-type: none"> <li>– Borgen eigenaarschap</li> <li>– Expliciet vastleggen hoe PI's kunnen worden beïnvloed</li> <li>– Verbeteren duiding en analyse van PI-uitkomsten</li> </ul>
Kwetsbaar proces door gebrek aan interne beheersmaatregelen en IT controls	<ul style="list-style-type: none"> <li>– Proceseigenaren benoemen per KPI om risico's in het ketenproces te beheersen</li> <li>– Het rapportageproces verder automatiseren</li> <li>– Verbeteren interne beheersing</li> <li>– Start implementatie van information security management system</li> </ul>
Onvoldoende inzicht in controls bij data leverende derden	<p>Met derde partijen afspraken maken over:</p> <ul style="list-style-type: none"> <li>– Inzichtelijk maken van huidige staat van beheersing</li> <li>– Verbetering van die beheersing</li> </ul>
Onduidelijkheid over en veranderlijkheid van publieksdefinitie	<ul style="list-style-type: none"> <li>– Expliciete en centrale vastlegging van publieksdefinities, gedetailleerder afstemmen met lenM</li> <li>– Proces afspraken maken intern en met lenM over wijzigingen in publieksdefinities. (publieksdefinities zijn een 'contract')</li> </ul>

De aandachtsgebieden en de verbetermaatregelen geïdentificeerd door NS overlappen voor een groot deel met voornoemde aanbevelingen die BTB/AEF doet. De aanbeveling inzake de halfjaarlijkse evaluatie van de meetpraktijk zien we niet direct terug in het verbetertraject. Dit kan echter ook buiten het verbetertraject worden georganiseerd.

NS heeft als doel om voor rapportagejaar 2017 de grootste onzekerheden in de processen aanwijsbaar te verbeteren. Beeld is dat NS op schema ligt. Voor wat betreft de kwaliteit van de brondata is er onlangs bijvoorbeeld een risico-analyse per leverancier uitgevoerd. Wel moeten er nog afspraken worden gemaakt met de betreffende leveranciers om (aanvullende) beheersmaatregelen in te richten. Kortom, NS heeft de verbeterpunten in beeld en werkt aan implementatie.

#### 4.1.4 Informatie

##### Bevindingen

Er is sprake van een representatieve steekproef. Meer dan 80% van de reizigers gebruikt de check-in en check-out. Reizigers die niet uitchecken, worden automatisch uitgesloten. Dit is van belang om vervuiling van de data te voorkomen. Dit betreft een klein aantal. De reizen tussen HRN-stations waarbij ook een andere vervoerder betrokken is, worden automatisch uitgesloten. Dit is belangrijk voor een goede afbakening, omdat het cijfers anders zou worden beïnvloed door andere vervoerders.

Er is geen aanleiding om te twifelen aan de betrouwbaarheid van de informatie uit het meetsysteem. NS heeft met het meetsysteem verschillende tests gedraaid, waarmee fouten zijn opgespoord en hersteld. Het meetsysteem blijft ook continu in ontwikkeling door de toevoeging van stations en het opstellen van nieuwe rapportages. Dit vraagt om frequente controles, die eveneens worden uitgevoerd door NS. Door middel van schaduwdraaien kan worden gekeken of de cijfers correleren. Kleinere technische aanpassingen die geen significante verschillen opleveren, worden na controle doorgevoerd. De afweging en de onderliggende criteria zijn niet vastgelegd.

Trendbreuken (zoals reizigers in de trein) worden geanalyseerd en gerapporteerd middels interne dashboards. De persoon/afdeling verantwoordelijk voor de data of deelapplicatie meetsysteem is ook voor deze analyses verantwoordelijk.

##### Aanbevelingen

- Leg de afweging voor het al dan niet doorvoeren van technische aanpassingen na overleg met het ministerie van IenM vast.

#### 4.1.5 Bodem- en streefwaarde

In de onderstaande tabel is het voorstel van NS en ProRail opgenomen voor de aanpassing van de bodem- en streefwaarde.

Situatie	Bodemwaarde	Streefwaarde
Huidig	90,0%	92,3%
Nieuw	89,2%	91,3%

De nieuwe bodem- en streefwaarde zijn bepaald door toepassing van de methode van lineaire regressie. Hiervoor is gekozen op basis van de hoge correlatie ( $R > 0,7$ ). Deze hoge correlatie maakt statistisch gezien dat regressie de geëigende methode is om de nieuwe bodem- en streefwaarde te bepalen. BTB/AEF ondersteunt de keuze voor deze methode.

Voor de herijking van de bodem- en streefwaarde is gebruikgemaakt van een set van gegevens die loopt van Q2 2015 t/m Q1 2016. Het alternatief was geweest om gegevens te gebruiken over de periode Q1 2015 t/m Q4 2015. De nu gebruikte gegevens zijn meer recent en consistent in de gekozen periode ten opzichte van de overige prestatie-indicatoren (m.u.v. zitplaatscapaciteit). Daarnaast is ook de correlatie tussen de oude en de nieuwe gegevens voor de gebruikte set hoger dan die voor de set van gegevens over 2015. BTB/AEF kan zich vinden in de keuze voor de gebruikte set van gegevens.

Overall heeft BTB/AEF het beeld dat de omzetting van de bodem- en streefwaarde ambitieneutraal is uitgevoerd.

#### **Aanbevelingen**

- Gelet op het voorgaande hebben wij geen aanbevelingen.

## **4.2 Reizigerspunctualiteit < 15 minuten HRN**

### **4.2.1 Publieksdefinitie**

De opzet en inhoud van de publieksdefinitie van de KPI reizigerspunctualiteit < 15min HRN is vrijwel gelijk aan de KPI reizigerspunctualiteit < 5min HRN. De prestatie-indicatoren < 5min en < 15 min samen bieden de mogelijkheid zowel op korte als langere vertragingen te sturen. Onderhavige indicator geeft een goed beeld van de negatieve uitschieters in reizigerspunctualiteit. Een reden om voor de grens van < 15 min te kiezen is dat uit het klantoordeel naar voren komt dat vanaf 15 minuten de klantwaardering sterk afneemt. Voorts zorgt een verruiming van de grens naar bijvoorbeeld 30 minuten ervoor dat het percentage van de reizigers voor wie de treinreis qua reistijd met maximaal 15 minuten vertraging is verlopen, zeer groot is. Dit is voor sturing van beperkte waarde.

### **4.2.2 Meetmethode**

Meetmethoden voor de prestatie-indicatoren reizigerspunctualiteit HRN zijn identiek. BTB/AEF heeft geen aanvullende bevindingen en aanbevelingen.

### **4.2.3 Meetpraktijk**

De meetpraktijk van deze prestatie-indicator komt overeen met die voor de prestatie-indicator Reizigerspunctualiteit < 5 min HRN. BTB/AEF heeft geen aanvullende bevindingen en aanbevelingen.

### **4.2.4 Informatie uit meetsystemen**

Voor informatie uit meetsystemen - waaronder de statistische betrouwbaarheid - verwijzen we naar de beschrijving van de prestatie-indicator reizigerspunctualiteit < 5min HRN op dit punt. Wederom zijn er voor < 15 min HRN geen andere bevindingen of aanbevelingen.



#### 4.2.5 Bodem- en streefwaarde

In de onderstaande tabel is het voorstel van NS en ProRail opgenomen voor de aanpassing van de bodem- en streefwaarde.

Situatie	Bodemwaarde	Streefwaarde
Huidig	93,7%	95,6%
Nieuw	96,7%	97,3%

De nieuwe bodem- en streefwaarde zijn bepaald door toepassing van de methode van lineaire regressie. De correlatie is voor deze prestatie-indicator lager dan die voor de prestatie-indicator Reizigerspunctualiteit < 5 min. HRN en ligt ook onder de statistische ondergrens ( $R < 0,7$ ). Dit door de relatief grote variantie in de vertragingen.

Ondanks dit gebrek is deze gebruikte methode toch de meest adequate voor de ambitieneutrale doorvertaling van de bodem- en streefwaarde. De alternatieven, zoals de deltamethode en de methode van over- en onderschrijding zijn statistisch gezien minder betrouwbaar. BTB/AEF onderschrijft het gebruik van de regressiemethode.

Voor de aanpassing van de bodem- en streefwaarde is gebruikgemaakt van een set van gegevens die loopt van Q2 2015 t/m Q1 2016. Het alternatief was geweest om gegevens te gebruiken over de periode Q1 2015 t/w Q4 2015. De nu gebruikte gegevens zijn meer recent. Daarentegen is de correlatie tussen de oude en de nieuwe gegevens voor de gebruikte set lager dan die voor de set van gegevens over 2015. Dit verschil is echter beperkt.

BTB/AEF kan zich vinden in de keuze om gebruik te maken van de set van gegevens van Q2 2015 t/m Q1 2016. Deze keuze is in lijn met de gegevens die gebruikt zijn voor de berekening van de prestatie-indicator Reizigerspunctualiteit < 5 min HRN. Daarnaast is het verschil in bodem- en streefwaarde door het gebruik van de gekozen set ten opzichte van de alternatieve set relatief klein (0,1%).

Overall heeft BTB/AEF het beeld dat de omzetting van de bodem- en streefwaarde ambitieneutraal is uitgevoerd.

#### Aanbevelingen

- Gelet op het voorgaande hebben wij geen aanbevelingen.

### 4.3 Reizigerspunctualiteit < 5 minuten HSL

De prestatie-indicator vervangt de huidige prestatie-indicatoren Reizigerspunctualiteit HSL-Zuid 5 minuten voor zover toerekenbaar aan NS (NS) en de Punctualiteit HSL-producten 3 minuten (ProRail). De nieuwe indicator wordt ingevoerd, omdat een gezamenlijke prestatie-indicator beter aansluit bij de reizigersbeleving. Alle situaties die vertraging veroorzaken tellen in de berekening van de prestatie-indicator mee. Voor de reiziger maakt het immers niet uit wie de veroorzaker is. De samenvoeging van de oude prestatie-indicatoren bovendien in lijn met de eerdere aanbeveling van BTB/AEF om de consistentie en harmonisatie in de set van prestatie-indicatoren te vergroten.<sup>16</sup>

<sup>16</sup> Audit prestatie-indicatoren vervoer- en beheerconcessie d.d. 16 oktober 2015, p. 28.

### 4.3.1 Publieksdefinitie

#### Bevindingen

De publieksdefinitie van de prestatie-indicator is voldoende specifiek, compleet en consistent. Twee formuleringen dienen echter te worden verduidelijkt. Volgens de publieksdefinitie vindt de bepaling van de aankomsttijd en de omvang van vertraging plaats op basis van uitchecktijd, wanneer een trein onder meer 'te veel te laat' vertrekt. Dit moet echter 'veel te laat' zijn en bovendien is een duiding van de tijd wenselijk, in dit geval meer dan 5 minuten te laat. Daarnaast bevelen wij aan om de term 'normtijd' zoals gehanteerd in de beschrijving van het meetsysteem, nader te omschrijven. Bij eerste lezing is niet direct duidelijk, dat hiermee gerefereerd wordt aan < 5min vertraging.

Verder mist in de publieksdefinitie, meer in het bijzonder in de beschrijving van het meetsysteem, de vermelding dat treinen die te laat aankomen (met meer dan 5 minuten vertraging) ook meetellen in de prestatie-indicator. Deze vertragingscategorie ontbreekt nu.

De tijdsafbakening van de prestatie-indicator is hetzelfde als voor de prestatie-indicator Reizigerspunctualiteit HRN. Hierover hebben we geen aanvullende opmerkingen.

De scope wijkt wel af van de prestatie-indicator Reizigerspunctualiteit HRN. Deze afwijkende scope is helder afgebakend. In scope zijn alle reizigers die in- en uitchecken tussen de geselecteerde HSL-stations. De geselecteerde relaties zijn in overleg met het ministerie van IenM vastgesteld en representeren de meest bereisde HB-relaties over de HSL.

Reizigers die in- of uitchecken op Amsterdam- Centraal en Amsterdam-Zuid en een deel van hun reis over het HRN maken, zijn hierbij inbegrepen. Hiermee wordt een klein deel van de prestatie-indicator dus mede bepaald door de HRN-punctualiteit tussen Amsterdam Centraal, Amsterdam Zuid en Schiphol. Dit is niet te voorkomen, omdat er geen gegevens beschikbaar zijn van het overstaptijdstip van een reiziger tussen treinen op het HRN en de HSL. Er is niet inzichtelijk wat voor effect de toevoeging van dit deel van het HRN heeft op de realisatiewaarde. Vanuit het oogpunt van transparantie wordt aanbevolen om als beperking van de meetmethode op te nemen dat binnen de scope van de prestatie-indicator ook HB-relaties worden meegenomen, die deels gebruik maken van het HRN. In dit verband wordt eveneens aanbevolen om het besluitvormingsproces en de overwegingen rondom de selectie van HB-relaties inzichtelijk te maken in de publieksdefinitie.

#### Aanbevelingen

- Herformuleer het criterium 'te veel te laat' naar 'veel te laat (> 15 min.)'.
- Noem specifiek de normtijd die bij de KPI hoort (in onderhavig geval dus 5 minuten).
- Neem in stap 5 van het meetsysteem in de publieksdefinitie ook de vertragingencategorie 'treinen die te laat aankomen' op. Deze categorie telt immers ook mee in de (berekening van de) prestatie-indicator.
- Leg als aanname vast dat binnen de scope van de prestatie-indicator ook HB-relaties worden meegenomen, die deels gebruik maken van het HRN. Vermeld daarbij ook dat dit effect kan hebben op de realisatiewaarde.
- Neem in de afbakening of als een FAQ op basis van welke criteria en processen gekomen is tot de selectie van zes HB-relaties en welk percentage van de HSL-reizigers hiermee in- en uitgesloten wordt van de berekening van de prestatie-indicator.

#### 4.3.2 Meetmethode

##### Bevindingen

De meetmethode komt grotendeels overeen met die van de prestatie-indicator Reizigerspunctualiteit < 5 min HRN. Het betreft immers een uitsnede van die prestatie-indicator. Het verschil met de prestatie-indicator Reizigerspunctualiteit HRN is, dat er in onderhavige indicator een selectie van 6 HB-relaties is gemaakt die de meeste reizigers vervoeren. Dit is voldoende beschreven en getoetst. BTB/AEF heeft ten opzichte van de prestatie-indicator Reizigerspunctualiteit HRN geen aanvullende bevindingen en aanbevelingen.

##### Aanbevelingen

- Gelet op het voorgaande hebben wij geen (aanvullende) aanbevelingen.

#### 4.3.3 Meetpraktijk

De meetpraktijk komt overeen met die van de prestatie-indicatoren Reizigerspunctualiteit HRN. Het betreft namelijk een uitsnede van de data voor deze indicator. Hierover zijn geen aanvullende opmerkingen.

#### 4.3.4 Informatie

De informatie uit de meetsystemen (waaronder representativiteit van de steekproef en betrouwbaarheid van het meetsysteem) van deze prestatie-indicator komt overeen met die voor de prestatie-indicatoren Reizigerspunctualiteit HRN. De geselecteerde relaties zijn in overleg met het ministerie van IenM vastgesteld en representeren de meest bereide HB-relaties over de HSL. De basisgegevens zijn derhalve representatief. We hebben verder geen aanvullende bevindingen en aanbevelingen.

#### 4.3.5 Bodem- en streefwaarde

##### Bevindingen

In de onderstaande tabel is het voorstel van NS en ProRail opgenomen voor de aanpassing van de bodem- en streefwaarde. In de tabel staan ook de bodem- en streefwaarde van de huidige prestatie-indicatoren van NS en ProRail vermeld.

Situatie	Bodemwaarde	Streefwaarde
Huidig NS	94,0%	96,0%
Huidig ProRail	80,0%	82,0%
Nieuw	82,5%	84,3%

NS en ProRail hebben de lineaire regressiemethode toegepast in doorvertaling bodem- en streefwaarde. Hiervoor baseren zij zich op de waarden van de huidige KPI van ProRail. Er is gebruikgemaakt van een set van gegevens die loopt van Q2 2015 t/m Q1 2016.

Reden om voor de waarden van de huidige prestatie-indicator van ProRail te kiezen, is de hogere correlatie tussen deze prestatie-indicator (gebaseerd op aankomstpunctualiteit) en de nieuwe prestatie-indicator ( $R > 0,7$ ), waar de huidige prestatie-indicator van NS zich

baseerde op slechts een subset van de data, te weten 'toerekenbaar aan NS'. Deze had een lagere correlatie met de nieuwe prestatie-indicator ( $R = 0,59$ ).

De keuze voor het tijdvak Q2 2015 t/m Q1 2016 is begrijpelijk en consistent met de keuze van tijdvak voor Reizigerspunctualiteit HRN. Deze gegevens zijn het meest recent en ook de correlatie tussen de oude en de nieuwe gegevens voor de gebruikte set hoger dan die voor de set van gegevens over 2015.

Gelet op de hogere correlatie is vanuit statistisch oogpunt regressie de meest geëigende methode om de nieuwe bodem- en streefwaarde te bepalen. De omzetting van de bodem- en streefwaarde is ambitieneutraal uitgevoerd aan de hand van regressieparameters. BTB/AEF ondersteunt de keuze voor deze methode.

Wel merken wij nog op dat bij de 'Brussel-treinen' de huidige gemiddelde punctualiteit op het Belgische deel van het traject lager is dan de huidige gemiddelde punctualiteit op de HSL (het Nederlandse deel van het traject). Op het moment dat ervoor wordt gekozen om deze treinen ook mee te laten wegen in de (scope van de) prestatie-indicator, dan kan dit een negatief effect hebben op de realisatiewaarde, terwijl de sturingsmogelijkheden voor NS voor wat betreft het Belgische deel van het traject beperkt tot nihil zijn. Weliswaar wordt gerekend vanaf een Nederlandse HB-relatie (i.e. Breda of Rotterdam), de vertraging die de trein oploopt in het voortraject (en te laat aankomt in bijvoorbeeld Rotterdam) telt ook mee in de prestatie-indicator. Hierop hebben NS en ProRail slechts beperkte invloed, omdat het vertraging betreft die in België al is opgelopen. Dit kan in de toekomst een effect hebben op de haalbaarheid van de streefwaarde.

#### **Aanbevelingen**

- Monitor de geïmporteerde punctualiteit uit Brussel en stel vast wanneer deze een te negatieve invloed heeft op de realisatiewaarde. Bepaal regelmatig in samenspraak met het ministerie van IenM welke consequenties dit heeft gehad op de gerealiseerde prestaties.

## **4.4 Reizigerspunctualiteit < 15 minuten HSL**

### **4.4.1 Publieksdefinitie**

De opzet en inhoud van de publieksdefinitie van de KPI reizigerspunctualiteit < 15 min HSL is vrijwel gelijk aan de KPI reizigerspunctualiteit < 5 min HSL.

De gekozen grens van 15 minuten is consistent met de publieksdefinitie van de prestatie-indicator Reizigerspunctualiteit HRN < 15 min. Bovendien blijkt dat voor HSL-reizigers hetzelfde geldt als voor HRN-reizigers: men ervaart vanaf 15 minuten de trein echt als vertraagd. Uitval heeft een relatief groot aandeel in de waarde van de prestatie-indicator Reizigerspunctualiteit HSL gemeten op 15 minuten: 60% van de reizen met meer dan 15 minuten vertraging heeft uitval als oorzaak. Voor de prestatie-indicator Reizigerspunctualiteit HSL gemeten op 5 minuten is dat 35%. BTB/AEF onderschrijft de keuze voor de grens van 15 minuten.

### **4.4.2 Meetmethode**

Meetmethoden voor de prestatie-indicatoren reizigerspunctualiteit HSL zijn identiek. BTB/AEF heeft geen aanvullende bevindingen en aanbevelingen.

#### **4.4.3 Meetpraktijk**

De meetpraktijk van deze prestatie-indicator komt overeen met die voor de prestatie-indicatoren Reizigerspunctualiteit HRN en Reizigerspunctualiteit < 5 min HSL. BTB/AEF heeft geen aanvullende bevindingen en aanbevelingen.

#### **4.4.4 Informatie uit meetsystemen**

Voor informatie uit meetsystemen - waaronder de statistische betrouwbaarheid - verwijzen we naar de beschrijvingen in de prestatie-indicatoren Reizigerspunctualiteit HRN en Reizigerspunctualiteit < 5 min HSL op dit punt. Wederom zijn er voor < 15 min HSL geen andere bevindingen of aanbevelingen.

#### **4.4.5 Bodem- en streefwaarde**

Een voorstel voor een bodem- en streefwaarde voor onderhavige prestatie-indicator was ten tijde van de audit niet beschikbaar. Ook was nog niet helder of hier in 2017 mee gewerkt gaat worden. BTB/AEF heeft deze waarden daarom niet getoetst.

## 5 Aanvullende aandachtspunten

### 5.1 Opvolging eerdere generieke aanbevelingen

In de eerdere audit van BTB/AEF is een aantal generieke aanbevelingen gedaan om de inhoud en het gebruik van de prestatie-indicatoren verder te verbeteren. Deze aanbevelingen hadden betrekking op:

- Het verder verbeteren van de interne monitoring door NS en ProRail op de kwaliteit van de meetpraktijk.
- Het nader expliciteren van de verantwoordelijkheid voor de kwaliteit van meetgegevens die NS gebruikt van ProRail.
- Het verder verbeteren van de interne vastlegging van de prestatie-indicatoren in termen van publieksdefinitie, meetstelsel en meetpraktijk, en veranderingen daarin binnen NS en ProRail.
- Het nader aanscherpen van het detailniveau waarop de afspraken tussen het ministerie van IenM, NS en ProRail over de prestatie-indicatoren worden gemaakt.
- Het verder vergroten van de consistentie tussen de sets van prestatie-indicatoren die gebruikt worden in de vervoer- en beheerconcessie, zodat deze een logisch geheel vormen.

In het vervolg van deze paragraaf wat de stand van zaken is voor deze aanbevelingen.

#### 5.1.1 Interne monitoring van meetpraktijk

Het beeld uit de audit is dat de meetpraktijk binnen NS en ProRail goed gemonitord wordt. Hiermee bedoelen we dat in de processen om de prestatie-indicatoren in algemene zin voldoende *checks & balances* zijn ingebouwd om fouten in de berekeningen zoveel mogelijk te voorkomen. De resultaten van de interne monitoring, in de vorm van bijvoorbeeld interne rapportages dan wel audits, zijn echter beperkt beschikbaar. De beschikbaarheid hiervan kan worden verhoogd.

##### **Aanbeveling**

- Maak de uitkomsten van interne monitoring op de processen om de prestatie-indicatoren te bepalen binnen NS en ProRail beter inzichtelijk.

#### 5.1.2 Expliciteren verantwoordelijkheid kwaliteit meetgegevens

In de onderlinge relatie tussen NS en ProRail is onderkend dat er nog onduidelijkheid bestaat over wie waarvoor verantwoordelijk is in het geval sprake is van de onderlinge doorlevering van gegevens om de score van een prestatie-indicator te bepalen.

Er zijn vooralsnog geen concrete stappen gezet om de duidelijkheid op dit punt te vergroten. In het geval gegevens van de ene aan de andere organisatie geleverd worden, is niet helder of de leverende of de ontvangende partij verantwoordelijk en aanspreekbaar is op de kwaliteit van de deze gegevens. Wel is binnen NS een breder traject gestart om de betrouwbaarheid van de meetgegevens en daarachter de prestatie-indicatoren te verbeteren. Onderdeel van dit traject is het verhelderen van de afspraken met ProRail (als een van de leveranciers van meetgegevens) over de onderlinge levering van gegevens.<sup>17</sup> De verwachting van NS is dat deze afspraken medio 2017 aangescherpt zijn en daarna in praktijk worden gebracht.

#### **Aanbeveling**

- Monitor de voortgang van het traject om tot verheldering van de verdeling van verantwoordelijkheden te komen bij de levering van gegevens tussen NS en ProRail.

### **5.1.3 Interne vastlegging van prestatie-indicatoren**

We constateren dat de interne vastlegging van de prestatie-indicatoren ten opzichte van de voorgaande audit verbeterd is. Dit geldt vooral voor NS. Ook bij ProRail zijn verbeteringen doorgevoerd, maar tegelijk is binnen deze organisatie nog een aanvullende stap te maken.

Een aandachtspunt voor beide organisaties is het beter vastleggen van de omgang met (aanpassingen in) bodem- en streefwaarde. Deze vastlegging is nu beperkt. Het ontbreekt voor een aantal prestatie-indicatoren bijvoorbeeld aan een beschrijving van de wijze waarop de aanpassingen zijn doorgevoerd en wat de overwegingen hierbij zijn. Hierdoor is het in de toekomst moeilijk om de aanpassingen te herleiden. Dit kan vooral een probleem opleveren in de situatie dat wederom behoefte is aan een ambitieneutrale doorvertaling (vertaling op vertaling).

#### **Aanbeveling**

- Verbeter binnen NS en ProRail de vastlegging van de bodem- en streefwaarde en bouw een historie op van achtereenvolgens doorgevoerde aanpassingen.<sup>18</sup>

### **5.1.4 Detailniveau van afspraken met ministerie van IenM**

Het beeld uit de audit is dat over het detailniveau van de afspraken over de prestatie-indicatoren tussen het ministerie van IenM enerzijds en NS en ProRail anderzijds inmiddels heldere afspraken zijn gemaakt. Onderling is bepaald dat de publieksdefinitie de afspraken bevat die door het departement worden goedgekeurd. Wijzingen in deze publieksdefinitie dienen NS en ProRail aan het ministerie van IenM ter accordering voor te leggen. Het beeld van BTB/AEF is dat deze afspraken voldoende gedetailleerd zijn om de continuïteit in het gebruik van de individuele prestatie-indicatoren te borgen. Voorwaarde is wel dat de publieksdefinities op de eerder beschreven aandachtspunten wordt aangepast.

#### **Aanbeveling**

- Gelet op voorgaande hebben wij geen aanbevelingen.

<sup>17</sup> Zie ook pagina 11 en 12 van deze rapportage. NS heeft een risicoanalyse opgesteld van de kwetsbare delen van de KPI-processen en is op basis daarvan een verbetertraject gestart.

<sup>18</sup> Mogelijk kan NS dit nog meenemen in het voornoemde verbetertraject waarin de onduidelijkheid over en veranderlijkheid van de publieksdefinities wordt aangepakt.

### **5.1.5 Consistentie van sets van prestatie-indicatoren**

In de audit zijn geen indicaties naar voren gekomen dat er sprake is van inconsistenties tussen de prestatie-indicatoren voor NS en ProRail gelden. Omdat de prestatie-indicatoren die in de voorgaande audit aan bod kwamen, vrijwel niet binnen de scope van de huidige audit vallen, hebben we geen check gedaan op het feit of de eerder geconstateerde inconsistenties zijn opgeheven.

#### **Aanbeveling**

- Gelet op het voorgaande hebben wij geen aanbevelingen.

## **5.2 Generieke aanbevelingen volgend uit huidige audit**

Er zijn in de huidige audit geen aanvullende bevindingen gedaan, die leiden tot nieuwe generieke aanbevelingen voor de toekomst. De observaties die gedaan zijn, liggen in lijn met de bevindingen in de voorgaande audit. Zie hiervoor de beschrijving in de voorgaande paragraaf.



## Bijlage 1. Overzicht aanbevelingen

Prestatie-indicator	Aanbeveling
NS	
<p>Zitplaatskans reizigers in de spits HRN</p>	<ul style="list-style-type: none"> <li>– Scherp de formulering van de scope zodanig aan dat helder is welke treinen hier wel/niet in vallen.</li> <li>– Pas het tijdframe voor de spits in het niet-publieke deel van de publieksdefinitie aan naar 06:35-08:55 en 16:05-18:25.</li> <li>– Leg de aannames bij de meetmethode in hetzelfde document vast als de beschrijving van de meetmethode of zorg als alternatief voor centrale vastlegging van beide.</li> <li>– Maak inzichtelijk welke validaties zijn uitgevoerd op de deelapplicaties 3 en 4 van SOFA. Indien er geen validaties zijn uitgevoerd, voer deze dan alsnog uit.</li> <li>– Maak medewerkers binnen het Kenniscentrum verantwoordelijk voor het verzamelen van de benodigde data en de berekening van de prestatie-indicator om daarmee de kwaliteit en continuïteit van de berekening en de resultaten beter te borgen.</li> <li>– Neem de (interne en externe) leveranciers van de brondata op basis waarvan de indicator wordt bepaald en de rol van Business Intelligence &amp; Analytics op in het organisatiemodel. Hierdoor is makkelijker te sturen op tijdigheid en kwaliteit van de brondata.</li> <li>– Leg de interne en externe aanlevering van, de (kwaliteit)controle op en de verwerking van de brondata vast in een rapportagecyclus.</li> </ul>

---

	<ul style="list-style-type: none"><li>– Maak inzichtelijk hoe de eindverantwoordelijke functionaris voor de KPI binnen NS de kwaliteit monitort.</li><li>– Evalueer de meetpraktijk halfjaarlijks met alle betrokken om het lerend vermogen te vergroten en formaliseer deze evaluatie.</li><li>– Breng de onzekerheden in het meetsysteem ook kwantitatief in beeld en leg deze in een document vast. Leg daarnaast ook de belangrijkste risico's en de bijbehorende beheersmaatregelen vast.</li></ul>
Top 10 reizen met slechtste vervoerplaatskans in spits <sup>19</sup>	<ul style="list-style-type: none"><li>– Neem in het eerste deel van de publieksdefinitie op hoe de top-10 wordt samengesteld.</li><li>– Neem de top 10 van slechtste reizen op in een op het betreffende jaar toegespitste publieksdefinitie in het Vervoerplan.</li><li>– Onderzoek nader of het mogelijk is om voor het bepalen van de top 10 uit te gaan van een meer recente basisperiode.</li></ul>
Zitplaatskans reizigers in de spits HSL <sup>20</sup>	<ul style="list-style-type: none"><li>– Bepaal op basis van de conclusies van de review op het Verbeterplan IC direct met het ministerie van IenM in welke mate de door NS en ProRail voorgestelde verbeter-maatregelen haalbaar zijn en aanleiding geven tot aanscherping van de door NS voorgestelde streefwaarde van 94,4%. Is de conclusie van de review bijvoorbeeld dat 6% uitval haalbaar is, verhoog de streefwaarde dan met 0,5%.</li><li>– Beperk het onderzoek naar de haalbaarheid van sturingsmogelijkheden niet tot 2019. Blijf aandacht houden voor maatregelen die ook in de periode daarna impact kunnen genereren.</li></ul>

---

ProRail

<sup>19</sup> In aanvulling op zitplaatskans HRN.

<sup>20</sup> In aanvulling op zitplaatskans HRN.

## Klantoordeel van reizigersvervoerders

- Maak helder om welke reden de vragenlijst alleen wordt uitgezet onder vervoerders met een volwaardige toegangsovereenkomst en wat het aandeel van deze vervoerders in het totale vervoer is.
- Maak helder wat de ondergrens is voor de vervoerders die meegenomen worden in het onderzoek en verwerk deze in de publieksdefinitie. Of voeg aan de publieksdefinitie toe: '*alle vervoerders wiens hoofdactiviteit het is om reizigers **binnen Nederland** te vervoeren*'.
- Neem in de publieksdefinitie op dat in de vragenlijst gevraagd wordt om een oordeel over het afgelopen kalenderjaar.
- Voeg per vraag in de vragenlijst een korte beschrijving toe van de betekenis van de cijfers, zodat dit voor de respondenten helder is, bijvoorbeeld van 'zeer ontevreden' tot 'zeer tevreden'.
- Borg de representativiteit van het klantoordeel door in de instructie voor het invullen van de vragenlijst meer richtlijnen op te nemen en de weging van vervoerders te herijken.
- Leg beter vast hoe de ingevulde vragenlijsten verwerkt worden tot een score op de prestatie-indicator.
- Maak in de rapportage van het onderzoek inzichtelijk door hoeveel respondenten de vragen (generiek en specifiek) beantwoord zijn.
- Stel een risicomatrix op waarin de risico's van het onderzoek beschreven zijn, evenals de maatregelen om deze te beheersen.
- Verander de bodemwaarde niet en houd deze op hetzelfde niveau als in 2015.
- Werk voor de streefwaarde met een 'afronding' op halve punten en stel de streefwaarde vast op 7,5 (feitelijk dus 7,25 – 7,74) om daarmee voldoende ambitie in te bouwen.
- Hanteer in de toekomst schaduwberendingen om het effect van aanpassingen in de inhoud van de vragenlijst en de wijze van verwerking te bepalen.

Klantoordeel van goederenvervoerders

- Pas de bovengrens van de categorie van kleine vervoerders aan van 5,0% naar 4,9% om daarmee de consistentie in de indeling te borgen.
- Heroverweeg de deelname van brancheorganisaties aan het onderzoek naar het klantoordeel of voorkom dubbeltellingen die expliciet te maken namens welke vervoerders zij aan het onderzoek meedoen.
- Herijk de gewichten van de verschillende categorieën van goederenvervoerders, waardoor deze beter in lijn zijn met de marktaandelen.
- Monitor veranderingen in marktaandeel van de goederenvervoerders en pas hierop indien nodig jaarlijks de weging aan.
- Neem in de beschrijving van de meetmethode op wat per categorie het aantal vervoerders en hun opgetelde marktaandeel is.
- Verander de bodem- en streefwaarde niet en houd deze op hetzelfde niveau als in 2016.
- Hanteer in de toekomst schaduwberekeningen om het effect van aanpassingen in de inhoud van de vragenlijst en de wijze van verwerking te bepalen.

Klanthinder als gevolg van storingen van de infrastructuur

- Herijk - vanwege de wijzigingen in de meetpraktijk - de wijze waarop de kwaliteit van de meetpraktijk wordt gemonitord en pas ook de risicobeheersing hierop aan.

Gezamenlijk

Reizigerspunctualiteit < 5 min. HRN

- Maak in de publieksdefinitie duidelijk dat het om 20 dagen per 100 dagen gaat.
- Scherp de scope zodanig aan, dat helder is welke treinen hier wel / niet in vallen.
- Maak inzichtelijk wat de onderliggende aannames zijn voor de componenten c.q. stappen in de meetmethode 'verzamelen van CICO-data' en 'bepaling indicator', en hoe deze zijn getoetst. Indien dit niet is gebeurd, voer deze toetsing uit.

- Maak medewerkers binnen het NSR Kenniscentrum verantwoordelijk voor het verzamelen van de benodigde data en de berekening van de prestatie-indicator om daarmee de kwaliteit en continuïteit van de berekening en de resultaten beter te borgen.
- Neem de (interne en externe) leveranciers van de brondata op basis waarvan de indicator wordt bepaald en de rol van Business Intelligence & Analytics op in het organisatiemodel. Hierdoor is makkelijker te sturen op tijdigheid en kwaliteit van de brondata.
- Leg de interne en externe aanlevering van, de (kwaliteit)controle op en de verwerking van de brondata vast in een rapportagecyclus.
- Maak inzichtelijk hoe de eindverantwoordelijke functionaris voor de KPI binnen NS de kwaliteit monitort.
- Evalueer de meetpraktijk halfjaarlijks met alle betrokken om het lerend vermogen te vergroten en formaliseer deze evaluatie.
- Breng de onzekerheden in het meetsysteem ook kwantitatief in beeld en leg deze in een document vast. Leg daarnaast ook de belangrijkste risico's en de bijbehorende beheersmaatregelen vast.
- Leg de afweging voor het al dan niet doorvoeren van technische aanpassingen na overleg met het ministerie van IenM vast.

Reizigerspunctualiteit < 15 min. HRN

- Geen aanvullende aanbevelingen ten opzichte van Reizigerspunctualiteit < 5 min. HRN .

Reizigerspunctualiteit < 5 min. HSL<sup>21</sup>

- Herformuleer het criterium 'te veel te laat' naar 'veel te laat (> 15 min.)'.
- Noem specifiek de normtijd die bij de KPI hoort (in onderhavig geval dus 5 minuten).

<sup>21</sup> In aanvulling op Reizigerspunctualiteit <5min HRN.

- Neem in stap 5 van het meetsysteem in de publieksdefinitie ook de verdragingscategorie ‘treinen die te laat aankomen’ op. Deze categorie telt immers ook mee in de (berekening van de) prestatie-indicator.
- Leg als aanname vast dat binnen de scope van de prestatie-indicator ook HB-relaties worden meegenomen, die deels gebruik maken van het HRN. Vermeld daarbij dat dit effect kan hebben op de realisatiewaarde.
- Neem in de afbakening of als een FAQ op basis van welke criteria en processen gekomen is tot de selectie van zes HB-relaties en welk percentage van de HSL-reizigers hiermee in- en uitgesloten wordt van de berekening van de prestatie-indicator.
- Monitor de geïmporteerde punctualiteit uit Brussel en stel vast wanneer deze een negatieve invloed heeft op de realisatiewaarde. Bepaal regelmatig in samenspraak met het ministerie van IenM welke consequenties dit heeft gehad op de gerealiseerde prestaties.

Reizigerspunctualiteit < 15 min. HSL

- Geen aanvullende aanbevelingen ten opzichte van Reizigerspunctualiteit < 5 min. HSL.

## *Bijlage 2. Gesprekspartners*

<b>Nummer</b>	<b>Rol geïnterviewde</b>	<b>Datum</b>
001	NS: Inhoudelijk expert KPI Reizigerspunctualiteit	18-08-2016
002	ProRail: Manager prestatie-analyse	18-08-2016
003	NS: Inhoudelijk expert KPI's vervoercapaciteit	18-08-2016
004	NS: Planning & Control KPI's (waaronder bodem- en streefwaarde)	18-08-2016 15-09-2016
005	NS: Planning & Control KPI's (waaronder bodem- en streefwaarde)	18-08-2016 15-09-2016
006	ProRail: verantwoordelijke KTO reizigers- en goederenvervoerders.	18-08-2016
007	ProRail: Business controller Corporate	22-08-2016
008	ProRail: Beleidsadviseur Infrabeschikbaarheid	24-08-2016
009	ProRail: Vakspecialist performance analyse	24-08-2016 02-09-2016
010	NS: Business Intelligence & Analytics - vergaring, verwerking data en berekening KPI's reizigerspunctualiteit.	23-08-2016
011	NS: Business Intelligence & Analytics - vergaring, verwerking data en berekening KPI's vervoercapaciteit.	23-08-2016 26-08-2016
012	NS: Business Intelligence & Analytics - vergaring, verwerking data en berekening KPI vervoercapaciteit.	26-08-2016
013	ProRail: verantwoordelijke berekeningen Klantoordeel reizigers- en goederenvervoerders	24-08-2016 05-09-2016
014	ProRail: Manager Infrabeschikbaarheid	26-08-2016
015	ProRail: Uitvoerder Klanttevredenheidsonderzoek goederenvervoerders	Week 35 (e-mails)
016	ProRail: Afdelingsmanager Relatiemanagement	29-08-2016
017	NS: Directeur Treindienstontwikkeling en Dienstregeling	22-09-2016
018	NS: Inhoudelijk expert KPI zitplaatskans HSL	22-09-2016





## *Bijlage 3. Brondocumenten*

Omschrijving	Datum	Opsteller
Aannamen voor de berekening van de KPI vervoercapaciteit (en top 10 aandachtstrajecten vervoercapaciteit)	23-08-2016	NS
Audit prestatie-indicatoren vervoer- en beheerconcessie	16-10-2015	BTB/AEF
Audit prestatie-indicatoren voortvloeiende uit vervoer- en beheerconcessie	17-09-2014	KPMG
Beheershandleiding van het systeem	25-11-2014	NS
Beleidsdocument Klanthinder	01-08-2016	ProRail
Berekeningen effect maatregelen ter verbetering KPI HSL Zuid 2019	-	NS
Berekening nieuwe waarde Top-10 drukke treinen	-	NS
Berekening nieuwe waarden Reizigerspunctualiteit	-	NS
Berekening nieuwe waarden zitplaatskans HRN	-	NS
Bijlage 1 voorgenomen besluit lenM voor wijzigingen ten aanzien van prestatie-indicatoren	-	ProRail
Bijlage 2 Publieksdefinities van de prestatie-indicatoren 2016 van de vervoersconcessie 2015-2025	-	ProRail
Bijlage 3 Veelgestelde vragen	-	ProRail
Brondata Realisatie oude en nieuwe methode Q2-15 Q1-16 Reizigerspunctualiteit	-	NS
Brondata Realisatie oude en nieuwe methode Q2-15 Q1-16 Zitplaatskans	-	NS
Consultatiebrief Locov 15 juli 2016 publieksdefinitief.PDF	15-07-2017	ProRail
Publieksdefinitie Reizigerspunctualiteit 15'	27-07-2016	NS
Publieksdefinitie Reizigerspunctualiteit HRN	27-07-2016	NS
Publieksdefinitie PI Reizigerspunctualiteit HSL 5 minuten 2017	01-09-2016	NS
Publieksdefinitie PI Reizigerspunctualiteit HSL 15 minuten 2017	01-09-2016	NS
Publieksdefinitie Top-10 drukke treinen	27-07-2016	NS

Publieksdefinitie Zitplaatskans HRN	27-07-2016	NS
Publieksdefinitie PI Zitplaatskans HSL-Zuid in de spits 2017	01-09-2016	NS
Eindrapportage KPI Zitplaatskans HSL	September 2016	NS
Brief Klanttevredenheidsonderzoeken 2015	16-02-2016	ProRail
Inschatting vervoersvraag IC direct komende jaren	-	NS
Klanthinder Voorbeeld dagoverzicht	05-08-2016	ProRail
Klanthinder Voorbeeld rapportage kwaliteit data klanthinder	05-08-2016	ProRail
Klanthinder 2015-2018 praatsheets	04-06-2016	ProRail
Memo klanthinder aangepast	01-06-2016	ProRail
Memo redenerlijn klanthinder 2017	05-07-2016	ProRail
Omzetting PI Reizigerspunctualiteit HSL 5 minuten	-	NS
Onderbouwing Nieuwe Bodem- en streefwaarde Zitplaatskans spits HSL	-	NS
Onderzoek totstandkoming KPI's	01-12-2014	ProRail
Ontwerpdocument met processtappen voor de berekening	-	NS
Opvolging bevinding klanthinder	30-06-2016	ProRail
Oude publieksdefinitie Aandachtstrajecten Reizigerspunctualiteit	08-12-2015	NS
Oude publieksdefinitie Aandachtstrajecten Vervoercapaciteit	27-07-2016	NS
Oude publieksdefinitie Reizigerspunctualiteit HRN	01-12-2015	NS
Oude publieksdefinitie Vervoercapaciteit HRN	27-07-2016	NS
ProRail halfjaarrapportage 2016 t b v IenM	25-08-2016	ProRail
ProRail klanttevredenheidsonderzoek totaal bestand v7.xlsx	-	ProRail
TNS Nipo rapport KTO ProRail 2015 v8	-	ProRail
ProRail rapportage t.b.v. IenM Q4 2015	04-04-2016	ProRail
Q&A document Consultatie	-	NS
Q-Consult beoordeling omrekening prestatie-indicatoren NS V2 d.d.	28-07-2016	Q-consult
Status van opvolging AEF adviezen	25-07-2016	NS
Testverslag van het systeem	16-12-2014	NS
Toelichting Omzetting PI's zitplaatskansen HRN	-	NS
Toetsdocument Aannames RPUN	-	NS

---

Toetsing van belangrijkste aannamen	-	NS
Uitleg reizigerspunctualiteit	28-04-2016	NS
Voorjaarsprognose 2016 Factsheet	29-03-2016	NS

---

## *Bijlage 4. Beoordelingskader*

### **Vraag a. Zijn de publieksdefinities helder, concreet en eenduidig interpreteerbaar beschreven?**

- Specifiek.
- compleet.
- Meetbaar.
- In zichzelf consistent.
- Voor één uitleg vatbaar.
- Afgebakend in tijd.
- Afgebakend in scope (breedte).

### **Vraag b. Zijn de meetmethoden helder, concreet en eenduidig interpreteerbaar beschreven?**

- Specifiek.
- Compleet.
- Valide.
- In zichzelf consistent.
- Voor één uitleg vatbaar.

### **Vraag c. Zijn de aangepaste bodem- en streefwaarde een ambitieneutrale doorvertaling van de bodem- en streefwaarde in de concessie bijlage 2? Is de gekozen methode van ambitieneutrale omzetting voldoende navolgbaar en statistisch voldoende betrouwbaar? Specifiek ten aanzien van de prestatie-indicatoren Zitplaatskans HRN en negatieve uitschieters wordt naast een beoordeling van de toegepaste methode gevraagd of er een of meerdere betere methoden zijn dan de toegepaste methode en zo ja, welke waarden de toepassing daarvan oplevert.**

- Afwijking tussen bodem- en streefwaarde in plannen 2016 en nu.
- Verklaring voor afwijking tussen bodem- en streefwaarde in plannen 2016 en nu.
- Verhouding tussen historische en recente realisaties:
  - in relatie tot de bodem- en streefwaarde uit plannen 2016
  - in relatie tot de aangepaste bodem- en streefwaarde (toets op ambitieneutraliteit).
- Navolgbaarheid van berekening om tot nieuwe bodem- en streefwaarde te komen.
- Sterkten en zwakten van gebruikte methode (prestatie-indicator Zitplaatskans en negatieve uitschieters).
- Alternatieven voor gebruikte methode (prestatie-indicator Zitplaatskans en negatieve uitschieters) en resultaten van toepassing.

**Vraag d. Voor ProRail is het voornemen de informatie indicator klanthinder als gevolg van storingen aan de infra om te zetten naar prestatie indicator. Bevatten de bodem- en streefwaarde voldoende ambitie t.o.v. de tot nu toe gerealiseerde data?**

**Volgt de totstandkoming van de bodem- en streefwaarde logisch uit de sturingsmogelijkheden van ProRail. Klantoordelen goederenvervoerders: bevat de per 2016 aangepaste streefwaarde voldoende ambitie t.o.v. de tot nu toe gerealiseerde data? Is de gekozen methode m.b.t. de totstandkoming van de streefwaarde voldoende navolgbaar en statistisch betrouwbaar?**

- Relatieve afwijking van bodem- en streefwaarde van historische en recente realisaties.
- Invloed van actuele en toekomstige ontwikkelingen op verwachte toekomstige realisaties.
- Verwachte afwijking van bodem- en streefwaarde van verwachte toekomstige realisaties.

**Vraag e. Komt de meetpraktijk overeen met de beschrijving van de meetsystemen (conformiteit) en wordt er gerapporteerd wat er gemeten wordt (juistheid)?**

- Overeenkomst tussen stappen in de meetpraktijk met de beschrijving van meetsystemen.
- Kwaliteitsborging in meetpraktijk (toetsen/evaluaties/monitors).
- Voldaan aan uitgangspunten en randvoorwaarden van meetsysteem.
- Gebruik van juiste basisgegevens (omvang/tijdvak/et cetera).
- Een op een overname van uitkomsten van meetpraktijk in rapportage.
- In beeld hebben van risico's van meetpraktijk en bijbehorende beheersmaatregelen.

**Vraag f. Is de informatie afkomstig uit de meetsystemen voldoende representatief en statistisch betrouwbaar om een vergelijking tussen verschillende kalenderjaren mogelijk te maken, en is de kwaliteitscontrole van de brongegevens voldoende geborgd? Statistische betrouwbaarheid van basisgegevens (representatieve steekproef).**

- Variantie van basisgegevens.
- Statistische fout in meetsysteem.
- Omgang met tussentijdse aanpassingen in meetsysteem (neutralisatie van afwijkingen/trendbreuken).

**Vraag g. Zijn de - in het kader van deze opdracht relevante - aanbevelingen uit de rapportage van BTB/AEF 'Audit prestatie-indicatoren Vervoer- en Beheerconcessie' van 16 oktober 2015 en uit de rapportage van KPMG 'Audit prestatie-indicatoren voortvloeiende uit vervoer- en beheerconcessie' van 17 september 2014 door NS en ProRail overgenomen of in elk geval stappen in gezet?**

- Bereikte resultaten.
- Uitgevoerde activiteiten.
- Genomen besluiten.
- Aanwezigheid plan van aanpak

**Vraag h. Belangrijk onderdeel van deze auditopdracht is het geven van een expert judgement over de door NS in beeld gebrachte sturingsmogelijkheden en aannames en het beoordelen of NS mogelijkheden over het hoofd heeft gezien om te drukke treinen op de HSL te voorkomen. M.b.t. de prestatie-indicator zitplaatskans HSL dient te worden beoordeeld of de voorgestelde bodem- en streefwaarde zowel vanuit historisch perspectief alsook gebaseerd op recente realisaties, sturingsmogelijkheden van NS en te verwachten toekomstige ontwikkelingen haalbaar en ambitieus zijn. Zo nee, wat is dan een realistische en ambitieuze bodem- en streefwaarde?**

- Compleetheid.

- Uitvoering.
- Planning.
- Effectiviteit.
- Afhankelijkheden.

**Vraag i. Is de voorgestelde bodem- en streefwaarde zowel vanuit historisch perspectief alsook gebaseerd op recente realisaties, sturingsmogelijkheden van NS en ProRail en te verwachten toekomstige ontwikkelingen haalbaar en ambitieus?**

- Relatieve afwijking van bodem- en streefwaarde van historische en recente realisaties.
- Mogelijkheden van NS en ProRail om sturing te geven aan realisatie.
- Verwachte invloed van toekomstige ontwikkelingen.