

# DECISIO

## **Prestatie-indicatoren nieuwe concessie HRN**

### **Auditrapport**

Ministerie van Infrastructuur en Waterstaat

15 november 2023

## TITEL

Prestatie-indicatoren nieuwe concessie HRN, auditrapport

## DATUM

15 november 2023

## OP VERZOEK VAN

Ministerie van Infrastructuur en Waterstaat

## PROJECTTEAM DECISIO

Niels Hoefsloot

Gerwin van der Meulen

Floortje Cieraad

Freek Lier

Met medewerking van Gijsbert van Eck (SWECO, senior adviseur railtechniek)

## CONTACTGEGEVENS DECISIO | ECONOMISCH ONDERZOEK EN ADVIES

Valkenburgstraat 212

1011 ND Amsterdam

T 020 – 67 00 562

E [info@decisio.nl](mailto:info@decisio.nl)

I [www.decisio.nl](http://www.decisio.nl)

## Inhoudsopgave

1.	Inleiding	1
1.1	Aanleiding	1
1.2	Ontwerpconcessie voor het Hoofdrailnet 2025 – 2033	1
1.3	Vraagstelling	1
1.4	Leeswijzer	2
2.	Prestatie-indicatoren	3
2.1	Doel van prestatie-indicatoren	3
2.2	Definities van prestatie-indicatoren	3
3.	Beoordelingsmethode/normenkader	6
3.1	Normenkader	6
3.2	Scenario's en (externe) ontwikkelingen	7
4.	Toets PI's aan normenkader	10
4.1	Algemeen klantoordeel	10
4.2	Klantoordeel sociale veiligheid	12
4.3	Aantal drukke treinen per week in de spits	13
4.4	Zitplaatskans in de Spits/in het Dal tweede klas	15
4.5	Reizigerspunctualiteit 3 min/ 10 min HRN	17
4.6	Kwaliteit van aansluitingen van NS op andere vervoerders	19
4.7	Impactvolle verstoringen veroorzaakt door NS	20
4.8	Kwaliteit reisinformatie inclusief verstoringen	23
4.9	Vermeden CO <sub>2</sub> -uitstoot	24
5.	Bevindingen	26
5.1	Bevindingen op hoofdlijnen	26
5.2	Bevindingen per prestatie-indicator	27
5.3	Conclusies en aanbevelingen	29
	Bijlage 1: Bronnen	32

# 1. Inleiding

## 1.1 Aanleiding

De nieuwe concessie voor het rijden van treinen op het Hoofdrailnet (HRN) zal onderhands aan NS worden gegund. NS krijgt daarmee voor de periode 2025 - 2033 een exclusief recht op het rijden van treinen op het Hoofdrailnet<sup>1</sup>. Om een kwalitatief hoogwaardige en doelmatige afwikkeling van het spoorvervoer te borgen worden prestatieafspraken gemaakt. In de (ontwerp)concessie voor het Hoofdrailnet 2025 - 2033 zijn prestatie-indicatoren (PI's) afgesproken waarmee de prestaties worden gemeten. In voorliggend rapport doen we verslag van onze toets van de verschillende prestatie-indicatoren zoals opgenomen in de ontwerpconcessie voor het Hoofdrailnet 2025 – 2033.

## 1.2 Ontwerpconcessie voor het Hoofdrailnet 2025 – 2033

Bij de toets zijn we uitgegaan van de ontwerp-concessie voor het Hoofdrailnet 2025 – 2033 (versie 14 juli 2023, hierna ontwerpconcessie). In de ontwerpconcessie is een aantal definities opgenomen die ook in dit toetsrapport van belang zijn.

**Bodemwaarde (BW):** waarde voor het jaarlijks minimaal te realiseren prestatieniveau op een prestatie-indicator.

**Dal(uren):** de periode op maandag tot en met vrijdag tussen 09:00 en 16:00 uur en tussen 18:30 en 06:30 uur en op zaterdagen, zondagen en feestdagen.

**HRN:** het Hoofdrailnet bestaande uit de spoorvervoerdiensten die als zodanig zijn aangewezen bij het Besluit Hoofdrailnet.

**Informatie-indicator (II):** maatstaf voor een niet bindende prestatie van NS.

**Prestatie-indicator (PI):** maatstaf voor een bindende prestatie met een bodemwaarde en een streefwaarde.

**Spits(uren):** de periode op werkdagen van 06:30 tot 09:00 uur (ochtendspits) en van 16:00 tot 18:30 uur (avondspits).

**Streefwaarde (SW):** waarde voor na te streven prestatieniveau op een prestatie-indicator in 2029 en 2033.

## 1.3 Vraagstelling

Bij de toets van de prestatie-indicatoren is gewerkt met de volgende door het ministerie van Infrastructuur en Waterstaat geformuleerde onderzoeksvragen over de PI's in de ontwerpconcessie:

1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?
2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?

<sup>1</sup> Voor de beschrijving van de scope van de concessie verwijzen wij naar ontwerpconcessie voor het Hoofdrailnet 2023 – 2033, Artikel 2. Reikwijdte en opgave voor NS

3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk.
4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?
5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties.
6. Zijn de bodem- en streefwaarden realistisch en (redelijkerwijs) haalbaar?

In dit onderzoek zijn deze onderzoeksvragen verder geoperationaliseerd en beantwoord voor de elf prestatie-indicatoren die in de ontwerpconcessie 2025 – 2033 zijn opgenomen

## 1.4 Leeswijzer

In hoofdstuk twee introduceren we de PI's. In hoofdstuk drie presenteren we het normenkader dat is gebruikt voor de toetsing van de PI's. In hoofdstuk vier wordt per PI verslag gedaan van de toetsing aan de normen. Hoofdstuk vijf vat de bevindingen samen en presenteert enkele aanbevelingen naar aanleiding van onze bevindingen.

## 2. Prestatie-indicatoren

### 2.1 Doel van prestatie-indicatoren

In de ontwerpconcessie zijn afspraken vastgelegd over de te leveren prestaties, die worden gemeten met prestatie-indicatoren (PI's). In Artikel 52 staat beschreven dat NS in elk kalenderjaar minimaal de bij de PI's horende bodemwaarden dient te realiseren en daarnaast streeft naar het behalen van de streefwaarden in kalenderjaar 2029 en in 2033. De resultaten voor alle PI's moeten controleerbaar tot stand komen, op statistisch verantwoorde wijze en zodanig dat een vergelijking tussen verschillende kalenderjaren mogelijk is. In het geval van het niet realiseren van bodemwaarden en het behalen van streefwaarden is een bonus/malus regeling opgenomen (Artikel 57).

### 2.2 Definities van prestatie-indicatoren

De volgende tabel geeft een overzicht van de PI's. Hierbij is de naamgeving uit de ontwerpconcessie gebruikt. Voor de PI 'Impactvolle Verstoringen veroorzaakt door NS' en 'Aantal drukke treinen per week in de Spits' geldt dat de genoemde bodem- en streefwaarden gehaald worden als op of onder de norm wordt gepresteerd (minder is beter). Voor de overige prestatie-indicatoren geldt dat genoemde bodem- en streefwaarden gehaald worden als op of boven het niveau gepresteerd wordt (meer is beter). Twee van de elf PI's zijn van NS en ProRail gezamenlijk.

Tabel 2.1 Overzicht prestatie-indicatoren ontwerpconcessie

Prestatie-indicator	Bodemwaarde 2025-2029	Streefwaarde 2029
Algemeen Klantoordeel	7,5	7,7
Klantoordeel Sociale veiligheid	7,7	7,9
Aantal drukke treinen per week in de Spits	110	95
Zitplaatskans in de Spits tweede klas <sup>2</sup>	94,3 %	95,5 %
Zitplaatskans in het Dal tweede klas <sup>3</sup>	98,0 %	98,7 %
Reizigerspunctualiteit 3 min. HRN <sup>4</sup>	84,4 %	86,0 %
Reizigerspunctualiteit 10 min. HRN <sup>5</sup>	94,5 %	95,1 %
Kwaliteit van aansluitingen van NS op andere vervoerders	90,3 %	91,2 %
Impactvolle verstoringen veroorzaakt door NS	240	190
Kwaliteit reisinformatie inclusief verstoringen	75,5 %	77,5 %
Vermeden CO <sub>2</sub> -uitstoot	534	620

Bron: ontwerpconcessie voor het Hoofdrailnet

<sup>2</sup> Deze waarde betreft zitplaatskans spits en moet nog ambitieneutraal omgezet worden naar zitplaatskans tweede klas spits

<sup>3</sup> Deze waarde betreft zitplaatskans dal en moet nog ambitieneutraal omgezet worden naar zitplaatskans tweede klas dal

<sup>4</sup> Dit betreft een gezamenlijke prestatie-indicator met ProRail

<sup>5</sup> Dit betreft een gezamenlijke prestatie-indicator met ProRail

De volgende definities van de PI's zijn gehanteerd in de ontwerpconcessie voor het HRN:

## *Algemeen klantoordeel*

De prestatie-indicator 'Algemeen Klantoordeel' brengt de algehele tevredenheid van reizigers over het geboden product in beeld met een rapportcijfer.

## *Klantoordeel Sociale veiligheid*

De prestatie-indicator 'Klantoordeel Sociale veiligheid' geeft aan welk gemiddeld rapportcijfer reizigers toekennen aan de mate van ervaren veiligheid op stations en in de treinen.

## *Aantal drukke treinen per week in de Spits*

De prestatie-indicator 'Aantal drukke treinen per week in de Spits' geeft het gemiddeld aantal treinen weer per werkweek met een vertrektijd tijdens de spits waarvoor het aantal reizigers groter is geweest dan het aantal gerealiseerde vervoerplaatsen.

## *Zitplaatskans in de Spits tweede klas<sup>6</sup>*

De prestatie-indicator 'Zitplaatskans in de Spits tweede klas' geeft een indicatie van de kans dat een reiziger in de tweede klas die tijdens de Spits in de trein stapt de volledige reis heeft kunnen zitten.

## *Zitplaatskans in het Dal tweede klas<sup>7</sup>*

De prestatie-indicator 'Zitplaatskans in het Dal tweede klas' geeft een indicatie van de kans dat een reiziger in de tweede klas die tijdens het dal in de trein stapt de volledige reis heeft kunnen zitten.'

## *Reizigerspunctualiteit 3 minuten HRN<sup>8</sup>*

De prestatie-indicator 'Reizigerspunctualiteit 3 minuten' geeft aan welk percentage van de reizigers met minder dan drie minuten vertraging is aangekomen op het eindstation van de reis met NS.

## *Reizigerspunctualiteit 10 minuten HRN<sup>9</sup>*

De prestatie-indicator 'Reizigerspunctualiteit 10 minuten' geeft aan welk percentage van de reizigers met minder dan tien minuten vertraging is aangekomen op het eindstation van de reis met NS.

## *Kwaliteit van aansluitingen van NS op andere vervoerders*

De prestatie-indicator 'Kwaliteit van aansluitingen van NS op andere vervoerders' geeft op basis van aankomstpunctualiteit 3 minuten met een percentage weer hoe goed NS een overstap mogelijk maakt op treinen van andere vervoerders.

---

<sup>6</sup> Deze PI wordt op een later moment omgezet in een PI 'zitplaatstijd in de spits tweede klas'. Het moment van deze overgang is afhankelijk van de ervaring die wordt opgedaan met de informatie-indicator 'zitplaatstijd in de spits tweede klas', maar vindt ultimo bij de midterm review plaats.

<sup>7</sup> Deze PI wordt op een later moment omgezet in een prestatie-indicator 'zitplaatstijd in het Dal tweede klas'. Het moment van deze overgang is afhankelijk van de ervaring die wordt opgedaan met de informatie-indicator 'zitplaatstijd in het Dal tweede klas', maar vindt ultimo bij de midterm review plaats.

<sup>8</sup> Dit betreft een gezamenlijke PI met ProRail

<sup>9</sup> Dit betreft een gezamenlijke PI met ProRail

## *Impactvolle verstoringen veroorzaakt door NS*

De prestatie-indicator 'Impactvolle Verstoringen veroorzaakt door NS' geeft het aantal impactvolle verstoringen van de treindienst weer die zijn veroorzaakt door NS.

## *Kwaliteit Reisinformatie inclusief Verstoringen*

De prestatie-indicator 'Kwaliteit Reisinformatie inclusief verstoringen' geeft weer in hoeverre vertragingen, spoorwijzigingen, opgeheven stoppen tijdig en juist aan de reizigers zijn gecommuniceerd en er eveneens volledig en tijdig handelingsperspectief is geboden in geval van ernstige verstoringen, met een juiste prognose van de duur van de verstoring. Het betreft de communicatie over treinen van alle OV-vervoerders via NS Reisplanners, borden/omroep stations en in (NS) treinen en via de NDOV-loketten met hun afnemers.

## *Vermeden CO<sub>2</sub>-uitstoot*

De prestatie-indicator 'Vermeden CO<sub>2</sub>-uitstoot' geeft weer hoeveel CO<sub>2</sub>-uitstoot in totaal vermeden is door reizigers die de trein nemen en de auto laten staan, minus de CO<sub>2</sub>-uitstoot in de operatie van NS en in de keten.

De PI's worden ook op de website van NS en in achterliggende documenten benoemd. NS heeft een actueel overzicht van prestaties op de website<sup>10</sup>. Hierbij staan de prestatie-indicatoren verdeeld over zes thema's<sup>11</sup> (betrouwbaarheid, deur-tot-deur reis, reisgemak, reisinformatie, veiligheid en algemeen) weergegeven. Wie wil kan hier per PI terugvinden wat de definitie van de PI is en hoe NS historisch gezien gepresteerd heeft. Onder de nieuwe concessie zal deze informatie per PI ook worden gedeeld op de website van NS.

Per prestatie-indicator heeft NS ook een definitiedocument opgesteld. Dit document bevat zowel de publieke omschrijving als een gedetailleerde definitie van de prestatie-indicator. Het definitiedocument beschrijft daarnaast hoe de indicator berekend wordt, de statistische verantwoording aan de hand van de representativiteit, betrouwbaarheid, beperkingen en onzekerheden van meetsystemen en ten slotte de verantwoordelijkheden en versiebeheer. Onderwerp van deze audit is onder meer de onderlinge consistentie van de definities en (beschrijvingen van) meetmethoden.

---

<sup>10</sup> <https://dashboards.nsjaarverslag.nl/prestaties>

<sup>11</sup> Conform de huidige concessie



### 3. Beoordelingsmethode/normenkader

Het doel van de PI's is om de kwaliteit van de geboden diensten te monitoren. Het beoordelen van de PI's is door ons gedaan op basis van een normenkader, waarin de onderzoeksvragen verder geconcretiseerd zijn. Het normenkader sluit zo veel mogelijk aan bij eerdere audits.

#### 3.1 Normenkader

Onderstaand overzicht geeft het normenkader weer. Het normenkader is onderverdeeld in zes onderzoeksvragen over de definitie, meetmethode, meetpraktijk, betrouwbaarheid van brongegevens en het ambitieniveau en de haalbaarheid van bodem- en streefwaarden. Per onderzoeksvraag is onderzocht of aan de normen voldaan wordt. Deze normen hebben we onder elke onderzoeksvraag weergegeven. Een voorbeeld is de mate waarin een definitie *specifiek* is. Hierbij is getoetst of er sprake is van een nauwkeurige en duidelijke definitie die geen ruimte laat voor verwarring. Een ander voorbeeld is de mate waarin de meetmethode *valide* is. Hierbij is getoetst of de methode daadwerkelijk meet wat het beoogt te meten.

##### 1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?

*Normen*

- a. Specifiek
- b. Compleet
- c. Meetbaar
- d. In zichzelf consistent
- e. Voor één uitleg vatbaar
- f. Afgebakend in tijd
- g. Afgebakend in scope (breedte)

##### 2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?

*Normen*

- a. Specifiek
- b. Compleet
- c. Valide
- d. In zichzelf consistent
- e. Voor één uitleg vatbaar
- f. Sluit aan op definitie

##### 3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?

*Normen*

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

**4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?**

*Normen*

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

**5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties.**

*Normen*

- a. Is alle relevante actuele informatie gebruikt bij het opstellen van scenario's?
- b. Zijn gebruikte scenario's transparant beschreven?
- c. Zijn gebruikte methoden en modellen gevalideerd?
- d. Zijn ontwikkelingen en scenario's, BW en SW op een evenwichtige manier gewogen?
- e. Mogelijkheden van NS en ProRail<sup>12</sup> om sturing te geven aan realisatie
- f. Hoe verhouden BW en SW zich tot prestaties in het verleden en voorbeelden elders?

**6. Zijn de bodem- en streefwaarden realistisch en (redelijkerwijs) haalbaar?<sup>13</sup>**

*Normen*

Zelfde normen als onder 5.

## 3.2 Scenario's en (externe) ontwikkelingen

Onderzoeksvraag 5 en 6 gaan over de mate waarin bodem- en streefwaarden van de verschillende PI's realistisch en ambitieus zijn. Dit is niet zonder meer objectief meetbaar en toetsbaar. Een mogelijke indicator om prestaties van NS te beoordelen vormen prestaties uit het verleden. Voor zover daar al gegevens over zijn, speelt de vraag in welke mate eerdere prestaties ook vergelijkbaar zijn met de toekomstige situatie in de nieuwe concessie voor het Hoofdrailnet. Prestaties in het verleden dienen daarom in het licht van veranderingen in de omgeving gezien te worden.

NS heeft de afgelopen jaren een methodiek ontwikkeld om tot bandbreedtes te komen om op basis daarvan het gesprek over bodem- en streefwaarden te voeren<sup>14</sup>. Deze methodiek gaat uit van vijf stappen:

1. Bepalen relevante ontwikkelingen voor NS op het HRN (en waar relevant voor ProRail);
2. Berekenen impact op PI's;
3. Uitwerken impact op bandbreedte;
4. Intern valideren;
5. Gesprek tussen lenW en NS (en waar relevant met ProRail).

<sup>12</sup> Voor de indicatoren van NS en ProRail gezamenlijk.

<sup>13</sup> 'Realistisch' en '(redelijkerwijs) haalbaar' zijn niet verder geoperationaliseerd. Deze termen, al dan niet in combinatie, worden in dit rapport door elkaar gebruikt.

<sup>14</sup> Deze methode is in 2022 gevalideerd door AEF.

De ontwikkelingen en impacts op de PI's worden expertmatig (en deels modelmatig) door NS en ProRail uitgewerkt. Achterliggende aannames zijn met de directeurs (van zowel NS als ProRail) afgestemd en gedeeld met IenW. Bij de midterm review worden aannames geactualiseerd en vindt opnieuw een gesprek tussen IenW en NS (en waar relevant met ProRail) over bodem- en streefwaarden plaats.

Een voorloper van de methodiek is in 2020 getoetst (Rebel 2020), waarna verbeteringen zijn doorgevoerd. De methodiek is in 2022 opnieuw getoetst (AEF, 2022) en valide bevonden. De methodiek is in het kader van voorliggend onderzoek daarom niet opnieuw getoetst. Wel is met een onafhankelijke, expertmatige blik (spoorespecialist SWECO) gekeken naar de ontwikkelingen en de (logica van de) doorwerking daarvan op de PI's. De belangrijkste bevinding is dat in algemene zin de lijst met ontwikkelingen herkenbaar is, en er geen opmerkelijke onderwerpen in staan. Het zijn allemaal thema's en ontwikkelingen die in meer of mindere mate relevant zijn voor de PI's. De wijze waarop de impacts zijn doorvertaald naar de verschillende PI's is voor zover te beoordelen plausibel. Daarbij signaleren we de volgende aandachtspunten:

- Een behoorlijk aantal werkzaamheden aan het spoor wordt gecombineerd tot een langdurige ontwikkeling die wordt aangeduid met een “boeggolf” aan werkzaamheden. De impact hiervan op de PI's is onder meer gebaseerd op de door ProRail berekende Ervaren extra reizigersminuten. Het nadeel van de gecombineerde impact is dat het effect van individuele ontwikkelingen niet zichtbaar is en (binnen de scope van deze audit) dus ook niet verder kon worden getoetst.
- Enkele verwijzingen naar de duur van werkzaamheden wijken op verschillende plaatsen in de ontwikkelingenlijst van elkaar af. Hoewel dit naar alle waarschijnlijkheid geen invloed heeft op de ontwikkelingen en de PI's, bevelen we, omwille van de interne consistentie en transparantie, aan om planningen van werkzaamheden consistent volgens de meest recente inzichten op te nemen.
- Bij sommige ontwikkelingen wordt uitgegaan van een impact op de PI's terwijl de ontwikkelingen nog onzeker zijn. Dit speelt vooral in relatie tot de issues met baanstabieleit waar op verschillende plekken naar wordt verwezen. Dit wordt nu geotechnisch uitgebreid onderzocht door ProRail. Omdat nog niet duidelijk is wat de effecten zullen zijn kan er niet op voorhand worden gezegd of de toekomstige intensivering van dienstregeling (sneller rijden of meer treinen) wel mogelijk is. Daarbij moet men bedenken dat oude baanlichamen niet gebouwd zijn voor dit soort intensief gebruik. De risico's en onzekerheden hierover worden nog eens vergroot door klimatologische veranderingen. Dit is een belangrijk aandachtspunt waarvan het de vraag is of dit, gezien de onzekerheden, redelijkerwijs kan worden meegewogen in de scenario's voor de PI's.

Aan de hand van prognoses van het volume en diverse andere (sector brede) ontwikkelingen is een beeld (bandbreedte) geschetst van hoe prestaties zich naar verwachting ontwikkelen en zich verhouden tot prestaties in het verleden. Een deel van de afgesproken bodem- en streefwaarden in de ontwerpconcessie ligt lager en een deel ligt hoger dan de in de lopende concessie overeengekomen bodem- en streefwaarden of in het verleden gerealiseerde waarden.

Aangezien de methode voor de totstandkoming van de bandbreedtes eerder als valide is beoordeeld en we de ontwikkelingen en impacts als plausibel beschouwen gaan we ervan uit dat de geschetste bandbreedte een haalbare range is voor de prestatie van NS. Voor zover inzichtelijk zijn de bodem- en streefwaarden daarmee realistisch te noemen wanneer deze binnen de bandbreedte vallen.

Voor de mate waarin de bodem- en streefwaarden ambitieus zijn, is het niet mogelijk een beoordeling te geven. Tussen IenW en NS is niet afgesproken wanneer een streefwaarde ambitieus is en hier is binnen de scope van de opdracht geen norm voor te bepalen<sup>15</sup>. Daarmee ontbreekt de mogelijkheid tot een onderbouwde beoordeling.

---

<sup>15</sup> Lagere waarden dan in het verleden of in de huidige concessie kunnen nog steeds ambitieus zijn. Dat is het geval wanneer er inspanningen worden gedaan om de negatieve impact van (externe) ontwikkelingen zo veel mogelijk te niet te doen of te compenseren.

## 4. Toets PI's aan normenkader

In dit hoofdstuk beschrijven we alle PI's en toetsen we ze volgens het normenkader. Hiervoor gaan we, per PI, steeds in op de definitie(s), meetmethode en -praktijk (inclusief brongegevens) en de bodem- en streefwaarden.

### 4.1 Algemeen klantoordeel

De PI 'Algemeen klantoordeel' voldoet aan de gestelde normen<sup>16</sup>. Voor de *definitie* van deze PI geldt echter een aandachtspunt.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓*	✓	✓	✓	?
				6. Zijn de bodem- en streefwaarden realistisch en (redelijkerwijs) haalbaar?

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

#### 4.1.1 Definitie

De PI voldoet aan alle eisen die we als norm hanteren voor de definitie. De reiziger wordt specifiek gevraagd om een rapportcijfer te geven op een schaal van 1 tot en met 10 (meetbaar) als waardering voor de treinrit op dat moment (afgebakend in tijd en scope). We signaleren een beperkte inconsistentie in het gebruik van de definitie voor deze PI, namelijk de 'algehele reizigerstevredenheid over het geboden product' (ontwerpconcessie) versus de reizigerswaardering over het reizen per trein (definitiedocument) versus het totaaloordeel over de rit (vraag enquête OV-Klantenbarometer). Het is aan te bevelen deze beperkte inconsistentie aan te passen.

#### 4.1.2 Meetmethode en -praktijk

De PI voldoet aan alle eisen die we als norm hanteren voor de meetmethode en -praktijk. De meetmethode en -praktijk is niet de verantwoordelijkheid van NS, maar van de uitvoerder en de opdrachtgevers van OV-Klantenbarometer. Deze partijen maken onderling afspraken over meetmethode en -praktijk (waaronder kwaliteitsborging, representativiteit, steekproef,

<sup>16</sup> Voor enkele normen van onder onderzoeksvraag 3, 4 en 5 geldt dat deze niet volledig toetsbaar zijn binnen de scope van dit onderzoek. Dit geldt voor alle PI's.

betrouwbaarheid, etc.) van de OV-Klantenbarometer. NS is afnemer van de resultaten van de OV-Klantenbarometer. NS voert een herberekening en controle uit op de resultaten.

#### 4.1.3 Bodem- en streefwaarden

Onderstaande tabel laat de historische prestatie (2019) en de bodem- en streefwaarde voor de huidige en ontwerpconcessie zien.

*Tabel 4.1 Historische prestatie en bodem- en streefwaarden voor de PI 'Algemeen klantoordeel'*

Algemeen klantoordeel	Gemeten waarde (2019)	Lopende concessie (2020 – 2024)	Ontwerpconcessie (2025 – 2029)
Historische prestatie	7,8		
Bodemwaarde		7,3	7,5
Streefwaarde		7,5	7,7

De bodem- en streefwaarden zijn *lager* dan de prestatie in 2019, maar *hoger* dan de bodem- en streefwaarden in de huidige concessie. Deze veranderingen zijn het gevolg van de geïnventariseerde ontwikkelingen en de impact daarvan op de vastgestelde bandbreedtes.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen, zijn die realistisch te noemen. Zoals in hoofdstuk 3 beschreven is, is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn<sup>17</sup>.

<sup>17</sup> Ondanks dat waarden lager zijn kunnen ze in het licht van de verwachte ontwikkelingen nog steeds ambitieus zijn. Uitgaande van de bandbreedtes zijn inspanningen nodig om bodem- en streefwaarden te halen. De mate van ambitie is niet te beoordelen.

## 4.2 Klantoordeel sociale veiligheid

De PI 'Klantoordeel sociale veiligheid' voldoet aan de gestelde normen.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓	✓	✓	✓	?

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

### 4.2.1 Definitie

De definitie uit het definitiedocument voldoet aan de gestelde normen. De definitie is specifiek en compleet. Het is mogelijk om deze PI te meten en kan maar op één manier worden uitgelegd. Er wordt op meerdere momenten in het jaar (behalve tussen 23.30 en 6.30 uur) gemeten in HRN-treinen door het hele land.

### 4.2.2 Meetmethode- en praktijk

Het meten van deze PI is onderdeel van de OV-klientenbarometer en gebeurt door een onderzoeksbureau in opdracht van CROW. De verantwoordelijkheid voor het meten ligt bij het geselecteerde onderzoeksbureau. In de OV-klientenbarometer wordt het onderzoek en de meetpraktijk toegelicht en dit sluit logisch aan op de beschreven onderzoeksmethode. De methode is specifiek, compleet en valide. Daarnaast is deze in zichzelf consistent en voor één uitleg vatbaar. De methode sluit logisch aan op de gehanteerde definitie. Kwaliteitsborging en validatie van data ligt hoofdzakelijk bij externe onderzoekers en is in het kader van deze audit niet verder onderzocht.

### 4.2.3 Bodem- en streefwaarden

De volgende tabel geeft de historische prestatie (2019) en voor de huidige- en ontwerpconcessie de bodem- en streefwaarde weer.

Tabel 4.2 Historische prestatie en bodem- en streefwaarden voor de PI 'Klantoordeel sociale veiligheid'

Klantoordeel sociale veiligheid	Gemeten waarde (2019)	Lopende concessie (2020 – 2024)	Ontwerpconcessie (2025 – 2029)
Historische prestatie	8,3		
Bodemwaarde		7,7	7,7
Streefwaarde		7,9	7,9

De bodem- en streefwaarden zijn *lager* dan de prestatie in 2019. Dit is het gevolg van de geïnventariseerde ontwikkelingen en de impact daarvan op de vastgestelde bandbreedtes. De bodem- en streefwaarden voor de ontwerpconcessie zijn *gelijk* aan de bodem- en streefwaarden van de lopende concessie.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen zijn die realistisch te noemen. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn<sup>18</sup>.

## 4.3 Aantal drukke treinen per week in de spits

De PI 'Aantal drukke treinen in de spits' voldoet aan de gestelde normen. Voor de *definitie* van deze PI geldt echter een aandachtspunt.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓*	✓	✓	✓	?

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

<sup>18</sup> Ondanks dat waarden lager zijn kunnen ze in het licht van de verwachte ontwikkelingen nog steeds ambitieus zijn. Uitgaande van de bandbreedtes zijn inspanningen nodig om bodem- en streefwaarden te halen. De mate van ambitie is niet te beoordelen..



### 4.3.1 Definitie

De PI voldoet aan de eisen die we als norm hanteren voor de definitie. Aandachtspunt is dat de definitie in de ontwerpconcessie minder specifiek is dan die in het PI Definitiedocument. Bij de definitie in de ontwerpconcessie ontbreekt de afbakening in scope en de exacte afbakening in de tijd. Wij adviseren om in de ontwerpconcessie een uitgebreide definitie op te nemen met de afbakening in tijd en de treinen die in scope zijn beschreven; of een verwijzing naar deze informatie op te nemen.

### 4.3.2 Meetmethode en -praktijk

De beschreven meetmethode is specifiek, compleet, valide, in zichzelf consistent, voor één uitleg vatbaar en sluit aan op de definitie van de PI uit het PI Definitiedocument. De meetpraktijk komt, voor zover te toetsen, overeen met de beschreven meetmethode. (Kwantitatieve) onzekerheden in het huidige systeem zitten in het toedelingsmodel (welke treinen hebben reizigers volgens het model genomen, versus welke in werkelijkheid) en in de ophoging met historisch veldwerk. Voor ontbrekende gegevens en meetfouten wordt gecorrigeerd onder andere met steekproefgegevens. Het is niet de verwachting dat de aanstaande wijziging in het meetsysteem leidt tot een (grote) trendbreuk. Op het nieuwe systeem wordt in het najaar nog een audit uitgevoerd.

### 4.3.3 Bodem- en streefwaarden

De volgende tabel geeft de historische prestatie (2021) en voor de huidige- en ontwerpconcessie de bodem- en streefwaarde weer.

Tabel 4.3 Historische prestatie en bodem- en streefwaarden voor de PI 'Aantal drukke treinen in de spits'

Aantal drukke treinen per week in de spits	Gemeten waarde (2021) <sup>19</sup>	Lopende concessie (2020 – 2024)	Ontwerpconcessie (2025 – 2029)
Historische prestatie	4		
Bodemwaarde		110	110
Streefwaarde		100	95

Deze PI wordt gemeten sinds 2021. Dit was ten tijde van covid. Het is daarom niet zinvol om de bandbreedte en bodem- en streefwaarden te relateren aan deze historische prestatie. Ten opzichte van de lopende concessie is de bodemwaarde *gelijk* gebleven en is de streefwaarde *hoger* (oftewel *minder* drukke treinen per week in de spits).

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen zijn die realistisch te

<sup>19</sup> Metingen van de jaren voor 2021 zijn niet beschikbaar doordat de PI pas sinds 2021 wordt gemeten.

noemen. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn.

## 4.4 Zitplaatskans in de Spits/in het Dal tweede klas

De PI's 'Zitplaatskans in de Spits/ in het Dal tweede klas' voldoen aan de gestelde normen. Voor de *definitie* van deze PI's en de *haalbaarheid* van de bodem- en streefwaarden geldt echter een aandachtspunt.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓*	✓	✓	✓	?
✓	6. Zijn de bodem- en streefwaarden realistisch en (redelijkerwijs) haalbaar?			
✓*	✓*			

✓ Voldoet aan normen

✓\* Voldoet met een of enkele aandachtspunten

Niet volledig objectief toetsbaar binnen scope

? Geen oordeel

- Voldoet niet aan normen

### 4.4.1 Definitie

De definities in de ontwerpconcessie voldoen op hoofdlijnen aan de normen. Wel zijn ze minder specifiek dan de definities in de definitiedocumenten. Wij adviseren om in de ontwerpconcessie in de definitie ook de afbakening in tijd en de treinen die in scope zijn te beschrijven óf een verwijzing naar de specifieke afbakening elders op te nemen. Onze aanbeveling is om bij de onderliggende stukken ook aan te geven waarom de spijtijden met/zonder 5 minuten marge gehanteerd worden.

### 4.4.2 Meetmethode en -praktijk

De beschreven meetmethode is specifiek, compleet, valide, in zichzelf consistent, voor één uitleg vatbaar en sluit aan op de definitie van de PI uit het PI Definitiedocument. De meetpraktijk komt voor zover te toetsen overeen met de beschreven meetmethode. (Kwantitatieve) onzekerheden in het huidige systeem zitten in het toedelingsmodel (welke treinen hebben reizigers volgens het model genomen, versus welke in werkelijkheid) en in de ophoging met historisch veldwerk. Voor ontbrekende gegevens en meetfouten wordt gecorrigeerd onder andere aan de hand van steekproefgegevens. Op het nieuwe systeem wordt in het najaar een audit uitgevoerd.

### 4.4.3 Bodem- en streefwaarden

Onderstaande tabel geeft de historische prestatie (2019/ 2021) en voor de huidige- en ontwerpconcessie de bodem- en streefwaarde weer. Vanaf 2023 kan NS onderscheid maken tussen zitplaatskans en zitplaatskans tweede klas. De KPI wordt eind 2023 ambitieneutraal omgezet met de toevoeging van 2<sup>de</sup> klas.

*Tabel 4.4 Historische prestatie en bodem- en streefwaarden voor de PI 'Zitplaatskans in de spits'*

Zitplaatskans in de spits	Gemeten waarde (2019)	Lopende concessie <sup>20</sup> (2020 – 2024)	Ontwerpconcessie <sup>21</sup> (2025 – 2029)
<b>Historische prestatie</b>	94,9%		
<b>Bodemwaarde</b>		94,3%	94,3%
<b>Streefwaarde</b>		95,5%	95,5%

*Tabel 4.5 Historische prestatie en bodem- en streefwaarden voor de PI 'Zitplaatskans in het dal'*

Zitplaatskans in het dal	Gemeten waarde (2021) <sup>22</sup>	Lopende concessie (2020 – 2024)	Ontwerpconcessie <sup>23</sup> (2025 – 2029)
<b>Historische prestatie</b>	99,8%		
<b>Bodemwaarde</b>		n.v.t.	98,0%
<b>Streefwaarde</b>		n.v.t.	98,7%

De bodemwaarde voor de PI 'Zitplaatskans in de spits tweede klas' is lager dan de waarde gemeten in 2019. De streefwaarde is daarentegen hoger dan deze waarde. Voor de PI 'Zitplaatskans in het dal tweede klas' geldt dat deze pas wordt gemeten sinds 2021. Dit was ten tijde van covid dus het is daarom niet zinvol om de bandbreedte en bodem- en streefwaarden te relateren aan deze historische prestatie. De bandbreedtes zijn grotendeels bepaald door geïnventariseerde ontwikkelingen en de impact daarvan.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodemwaarden binnen de bandbreedte liggen zijn die realistisch te noemen. De streefwaarden liggen boven de bandbreedte. Dit roept vragen op over de haalbaarheid ervan. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn<sup>24</sup>.

<sup>20</sup> Onderscheid tussen eerste en tweede klas is vanaf eind 2023 mogelijk. Deze bodem- en streefwaarden betreffen dus 1<sup>e</sup> en 2<sup>de</sup> klas samen.

<sup>21</sup> Onderscheid tussen eerste en tweede klas is vanaf eind 2023 mogelijk. De bandbreedte wordt eind 2023 ambitieneutraal omgezet.

<sup>22</sup> Metingen van de jaren voor 2021 zijn niet beschikbaar

<sup>23</sup> Onderscheid tussen eerste en tweede klas is vanaf eind 2023 mogelijk. De bandbreedte wordt eind 2023 ambitieneutraal omgezet.

<sup>24</sup> Ondanks dat waarden lager zijn kunnen ze in het licht van de verwachte ontwikkelingen nog steeds ambitieus zijn. Uitgaande van de bandbreedtes zijn inspanningen nodig om bodem- en streefwaarden te halen. De mate van ambitie is niet te beoordelen.

## 4.5 Reizigerspunctualiteit 3 min/ 10 min HRN<sup>25</sup>

De PI's 'Reizigerspunctualiteit 3 minuten/ 10 minuten' voldoen aan de gestelde normen.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓	✓	✓	✓	?

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

### 4.5.1 Definitie

De definitie uit het definitiedocument voldoet aan de gestelde normen. De definitie is specifiek en compleet. Het is mogelijk om deze PI te meten en kan maar op één manier worden uitgelegd. Daarnaast is de afbakening van de scope van de definitie duidelijk: alle treinen in de HRN-concessie op ieder moment van het jaar. Een aandachtspunt is dat in de ontwerpconcessie is besloten om de IC Berlijn trein niet op te nemen. In het huidige Definitiedocument staat de zin “Dus wél IC Berlijn en IC Brussel, maar níét Thalys, Eurostar, ICE en Nightjet.” Na behandeling van de ontwerpconcessie in de Kamer en wanneer duidelijkheid is geboden over de internationale scope in de nieuwe concessie dient het definitiedocument hierop aangepast te worden. NS heeft aangegeven het definitiedocument indien nodig vóór ondertekening van de concessie aan te passen.

### 4.5.2 Meetmethode en -praktijk

De scope en beschikbare tijd van deze audit maakt het niet mogelijk mee te lopen in de praktijk en de meetpraktijk uit eigen ervaring te beoordelen. De beoordeling van de meetpraktijk ten opzichte van de beschreven meetmethode is gebaseerd op beschrijvingen van de praktijk, procesflows, kwaliteitscontroles en monitoringsrapporten. Voor zover inzichtelijk komt de meetpraktijk overeen met de beschreven meetmethode en is geborgd dat de meetpraktijk ook gerapporteerd wordt. Op basis van de documenten en toegelichte processen wordt voldaan aan de gestelde normen. De methode is daarmee specifiek, compleet en valide. Daarnaast is deze in zichzelf consistent en voor één uitleg vatbaar. De methode sluit tenslotte logisch aan op de gehanteerde definitie.

<sup>25</sup> PI gezamenlijk met ProRail.

### 4.5.3 Bodem- en streefwaarden

Onderstaande tabel geeft de historische prestatie (2019/ 2021) en voor de huidige- en ontwerpconcessie de bodem- en streefwaarde weer.

Tabel 4.6 Historische prestatie en bodem- en streefwaarden voor de PI 'Reizigerspunctualiteit 3 minuten'

Reizigerspunctualiteit 3 minuten	Gemeten waarde (2019)	Lopende concessie (2020 – 2024)	Ontwerpconcessie (2025 – 2029)
<b>Historische prestatie</b>	88,1%		
<b>Bodemwaarde</b>		84,4% <sup>26</sup>	84,4%
<b>Streefwaarde</b>		87,0% <sup>27</sup>	86,0%

Tabel 4.7 Historische prestatie en bodem- en streefwaarden voor de PI 'Reizigerspunctualiteit 10 minuten'

Reizigerspunctualiteit 10 minuten	Gemeten waarde (2019)	Lopende concessie (2020 – 2024)	Ontwerpconcessie (2025 – 2029)
<b>Historische prestatie</b>	96,1%		
<b>Bodemwaarde</b>		95,1% <sup>28</sup>	94,5%
<b>Streefwaarde</b>		95,8% <sup>29</sup>	95,1%

De bodem- en streefwaarden zijn *lager* dan de prestaties van reizigerspunctualiteit 3 minuten in 2019 en reizigerspunctualiteit 10 minuten in 2021. Daarnaast liggen deze waarden *lager* dan de bodem- en streefwaarden in de lopende concessie (een uitzondering geldt voor de bodemwaarde voor 'Reizigerspunctualiteit 3 minuten', deze is gelijk gebleven). Deze veranderingen zijn het gevolg van de geïnventariseerde ontwikkelingen en de impact daarvan op de vastgestelde bandbreedtes.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen zijn die realistisch te noemen. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn<sup>30</sup>.

<sup>26</sup> De huidige bodemwaarde omgerekend naar de normtijd van 3 minuten.

<sup>27</sup> De huidige bodemwaarde omgerekend naar de normtijd van 3 minuten.

<sup>28</sup> De huidige bodemwaarde omgerekend naar de normtijd van 10 minuten.

<sup>29</sup> De huidige bodemwaarde omgerekend naar de normtijd van 10 minuten.

<sup>30</sup> Ondanks dat waarden lager zijn kunnen ze in het licht van de verwachte ontwikkelingen nog steeds ambitieus zijn. Uitgaande van de bandbreedtes zijn inspanningen nodig om bodem- en streefwaarden te halen. De mate van ambitie is niet te beoordelen..

## 4.6 Kwaliteit van aansluitingen van NS op andere vervoerders

De PI 'Kwaliteit van aansluitingen NS op andere vervoerders' voldoet aan de gestelde normen. Voor de *definitie* van deze PI geldt echter een aandachtspunt.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓*	✓	✓	✓	?

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

### 4.6.1 Definitie

De definitie uit het definitiedocument voldoet aan de gestelde normen, behalve dat de definitie in het definitiedocument niet voldoende specifiek is. De naam en de definitie (in het definitiedocument) van de PI doen vermoeden dat het meet of de overstap op andere concessieovervoerders gehaald wordt, maar dat is niet aan de orde. Specifiek gaat het om aankomstpunctualiteit op stations waar ook andere vervoerders actief zijn. Het advies is daarom om de naam en definitie op zo'n manier te formuleren dat deze overeenkomt met hetgeen wat daadwerkelijk wordt gemeten (namelijk de aankomstpunctualiteit op stations waar een overstap op een andere vervoerder mogelijk is). De definitie uit het Definitiedocument is wel compleet en meetbaar. Daarnaast is de afbakening van de scope van de definitie duidelijk: alle treinen in de HRN-concessie die aankomen op een station waar overstappen op andere concessieovervoerders mogelijk zijn op ieder moment van het jaar.

### 4.6.2 Meetmethode- en praktijk

De meetmethode voldoet aan de gestelde normen. De methode is specifiek, compleet en valide. Daarnaast is deze in zichzelf consistent en voor één uitleg vatbaar. De methode sluit tenslotte logisch aan op de gehanteerde definitie. De beoordeling van de meetpraktijk ten opzichte van de beschreven meetmethode is gebaseerd op beschrijvingen van de praktijk, procesflows, kwaliteitscontroles en monitoringsrapporten. De scope en beschikbare tijd van deze audit maakt het niet mogelijk mee te lopen in de praktijk en de meetpraktijk uit eigen ervaring te beoordelen. Voor zover inzichtelijk komt de meetpraktijk overeen met de beschreven meetmethode en is geborgd dat de meetpraktijk ook gerapporteerd wordt.

### 4.6.3 Bodem- en streefwaarden

Onderstaande tabel laat de historische prestatie (2019) en de bodem- en streefwaarde voor de huidige en ontwerpconcessie zien.

Tabel 4.8 Historische prestatie en bodem- en streefwaarden voor de PI 'Kwaliteit aansluitingen van NS op andere vervoerders'

Kwaliteit van aansluitingen NS op andere vervoerders	Gemeten waarde (2019)	Lopende concessie (2020 – 2024)	Ontwerpconcessie (2025 – 2029)
Historische prestatie	92,2%		
Bodemwaarde		90,2% <sup>31</sup>	90,3%
Streefwaarde		91,8% <sup>32</sup>	91,2%

De bodem- en streefwaarden zijn *lager* dan de prestatie in 2019. De streefwaarde in de lopende concessie is eveneens *hoger* dan de streefwaarde in de ontwerpconcessie. Dit geldt niet voor de bodemwaarde: deze ligt iets *lager* in de ontwerpconcessie ten opzichte van de lopende concessie. Deze veranderingen zijn het gevolg van de geïnventariseerde ontwikkelingen en de impact daarvan op de vastgestelde bandbreedtes.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen zijn die realistisch te noemen. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn<sup>33</sup>.

## 4.7 Impactvolle verstoringen veroorzaakt door NS

De PI 'Impactvolle verstoringen veroorzaakt door NS' voldoet aan het grootste deel van de gestelde normen<sup>34</sup>. Voor de *definitie* van deze PI geldt echter een aandachtspunt. Daarnaast is het binnen de scope van dit onderzoek niet mogelijk om te beoordelen of de meetpraktijk overeenkomt met de beschrijving van de meetmethode.

<sup>31</sup> De huidige bodemwaarde omgerekend naar de normtijd van 3 minuten

<sup>32</sup> De huidige bodemwaarde omgerekend naar de normtijd van 3 minuten

<sup>33</sup> Ondanks dat waarden lager zijn kunnen ze in het licht van de verwachte ontwikkelingen nog steeds ambitieus zijn. Uitgaande van de bandbreedtes zijn inspanningen nodig om bodem- en streefwaarden te halen. De mate van ambitie is niet te beoordelen.

<sup>34</sup> Voor enkele normen van onder onderzoeksvraag 3, 4 en 5 geldt dat deze niet volledig toetsbaar zijn binnen de scope van dit onderzoek. Dit geldt voor alle PI's.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓*	✓	?	✓	?

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

## 4.7.1 Definitie

De definitie uit het definitiedocument voldoet aan de gestelde normen, met uitzondering van de norm “*in zichzelf consistent*”<sup>35</sup>. In de ontwerpconcessie heeft deze PI de naam ‘Impactvolle verstoringen veroorzaakt door NS’ en is daarmee specifiek en anders dan (dus niet consistent met) de naam in het definitiedocument: ‘aantal verstoringen veroorzaakt door NS’<sup>36</sup>. Het advies is daarom om deze definitie in het definitiedocument ook specifiek te beschrijven door aan te geven dat het gaat om ‘impactvolle’ verstoringen. De definitie uit het Definitiedocument is verder wel specifiek, compleet en meetbaar. Daarnaast is de afbakening van de scope van de definitie duidelijk: alle beladen treinen samen die gehinderd worden door de verstoring gedurende het hele jaar.

## 4.7.2 Meetmethode en -praktijk

De meetmethode voldoet aan de gestelde normen. De methode is specifiek, compleet en valide. Daarnaast is deze in zichzelf consistent en voor één uitleg vatbaar. De methode sluit tenslotte logisch aan op de gehanteerde definitie. Het is niet mogelijk te oordelen of de PI voldoet aan de normen ten aanzien van de meetpraktijk. Onderbouwing hiervan is nog niet beschikbaar omdat de PI nog in ontwikkeling is. Op basis van de statistische verantwoording concluderen we dat de brongegevens voldoende betrouwbaar zijn en de kwaliteitscontrole op verschillende manieren geborgd is.

<sup>35</sup> Zie paragraaf 3.1 voor toelichting op deze norm

<sup>36</sup> Het ministerie van Infrastructuur en Waterstaat heeft al het verzoek bij NS gedaan om de naam in het Definitiedocument te wijzigen naar Impactvolle verstoringen veroorzaakt door NS en het zo gelijk te trekken met de naam die ProRail hanteert: impactvolle verstoringen veroorzaakt door de infrastructuurbeheerder. Tussen NS en IenW is afgesproken om eind 2023 het Definitiedocument te updaten.



### 4.7.3 Bodem- en streefwaarden

Onderstaande tabel laat de historische prestatie (2019) en de bodem- en streefwaarde voor de huidige en ontwerpconcessie zien. In de huidige concessie is de PI een informatie indicator, dus er is geen bodem- en streefwaarde opgenomen in de huidige concessie.

*Tabel 4.9 Historische prestatie en bodem- en streefwaarden voor de PI 'Impactvolle verstoringen veroorzaakt door NS'*

Impactvolle verstoringen veroorzaakt door NS	Gemeten waarde (2019)	Ontwerpconcessie (2025 – 2029)
<b>Historische prestatie</b>	172	
<b>Bodemwaarde</b>		240
<b>Streefwaarde</b>		190

De bodem- en streefwaarden zijn *lager* dan de prestatie voor 2019. Dit is het gevolg van de geïnterpreteerde ontwikkelingen en de impact daarvan op de vastgestelde bandbreedtes.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen zijn die realistisch te noemen. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn<sup>37</sup>.

<sup>37</sup> Ondanks dat waarden lager zijn kunnen ze in het licht van de verwachte ontwikkelingen nog steeds ambitieus zijn. Uitgaande van de bandbreedtes zijn inspanningen nodig om bodem- en streefwaarden te halen. De mate van ambitie is niet te beoordelen.

#### 4.8 Kwaliteit reisinformatie inclusief verstoringen

De PI “Kwaliteit reisinformatie inclusief verstoringen” voldoet aan de gestelde normen<sup>38</sup>.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
✓	✓	✓	✓	?
				✓

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

### 4.8.1 Definitie

De PI voldoet aan alle eisen die we als norm hanteren voor de definitie. De kwaliteit van reisinformatie wordt per station en per trein (idem voor wat betreft eventueel vervangende treinen) gemeten. Reisinformatie bestaat uit vertrektijden en afwijkingen van de dienstregeling. De PI meet of de afwijking van de dienstregeling juist en tijdig wordt gecommuniceerd.

#### 4.8.2 Meetmethode en – praktijk

De procesbeschrijving, het verzamelen van data en berekenen van de scores om tot de Kwaliteit Reisinformatie incl. verstoringen te komen is specifiek, compleet, valide, in zichzelf consistent, voor één uitleg vatbaar en aansluitend op definitie beschreven. De beoordeling van de meetpraktijk is gebaseerd op beschrijvingen van de praktijk, procesflows, kwaliteitscontroles en monitoringsrapporten. Voor zover inzichtelijk komt de meetpraktijk overeen met de beschreven meetmethode en is geborgd dat de meetpraktijk ook gerapporteerd wordt. Het PI Definitiedocument geeft aan dat de tabel met betrouwbaarheid van meetsystemen op onderdelen nog geactualiseerd dient te worden. Actualisatie van het definitiedocument gebeurt jaarlijks en staat op de planning om voor de ondertekening van de ontwerpconcessie in 2023 te zijn uitgevoerd.

#### 4.8.3 Bodem- en streefwaarden

Onderstaande tabel geeft de historische prestatie (2021) en voor de huidige en ontwerpconcessie de bodem- en streefwaarde weer.

<sup>38</sup> Voor enkele normen van onder onderzoeksvraag 3, 4 en 5 geldt dat deze niet volledig toetsbaar zijn binnen de scope van dit onderzoek. Dit geldt voor alle PI's.

Tabel 4.10 Historische prestatie en bodem- en streefwaarden voor de PI 'Kwaliteit reisinformatie inclusief verstoringen'

Kwaliteit reisinformatie inclusief verstoringen	Gemeten waarde (2021)	Lopende concessie (2020 – 2024)	Ontwerpconcessie (2025 – 2029)
Historische prestatie	74,6 %		
Bodemwaarde		65,0%	75,5%
Streefwaarde		68,5%	77,5%

Deze PI wordt pas gemeten sinds 2021. Dit was ten tijde van Covid. Daarnaast zijn na 2021 verbeteringen in het algoritme doorgevoerd wat invloed heeft op de hoogte van de gemeten waarden. Het is daarom niet zinvol om de bandbreedte en bodem- en streefwaarden te relateren aan de historische prestatie en de bodem- en bandbreedte van de lopende concessie.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen zijn die realistisch te noemen. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn.

## 4.9 Vermeden CO<sub>2</sub>-uitstoot

De PI 'Vermeden CO<sub>2</sub>-uitstoot' voldoet aan het grootste deel van de gestelde normen over de definitie en de bodem- en streefwaarden<sup>39</sup>. Voor de *definitie* van deze PI geldt echter een aandachtspunt. Daarnaast is het binnen de scope van dit onderzoek niet mogelijk om vast te stellen of aan de normen over *meetmethode*, *meetpraktijk* en *brongegevens* is voldaan.

Definitie	Meetmethode, meetpraktijk, brongegevens		Bodem- en streefwaarden	
1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?
6. Zijn de bodem- en streefwaarden realistisch en (redelijkerwijs) haalbaar?				
✓*	?	?	?	✓

✓	Voldoet aan normen
✓*	Voldoet met een of enkele aandachtspunten
	Niet volledig objectief toetsbaar binnen scope
?	Geen oordeel
-	Voldoet niet aan normen

<sup>39</sup> Voor enkele normen van onder onderzoeksvraag 3, 4 en 5 geldt dat deze niet volledig toetsbaar zijn binnen de scope van dit onderzoek. Dit geldt voor alle PI's.

#### 4.9.1 Definitie

De definitie in de ontwerpconcessie is smaller dan in het definitiedocument van NS, maar voldoet in principe aan de normen specifiek, compleet, meetbaar, in zichzelf consistent en voor één uitleg vatbaar. De afbakening in tijd (per jaar) en afbakening in scope (wat behoort precies tot de operatie van NS en de keten) ontbreken in de definitie in de ontwerpconcessie. Deze elementen zijn wel onderdeel van de uitgewerkte definitie in het definitiedocument van NS. Wij adviseren om in de ontwerpconcessie een verwijzing naar deze informatie op te nemen

#### 4.9.2 Meetmethode en -praktijk

De meetmethode (en -praktijk) voor deze PI is in ontwikkeling en moet nog verder door NS worden uitgewerkt. Daarom kunnen we binnen de scope van deze audit de normen over de meetmethode – en praktijk niet beoordelen.

#### 4.9.3 Bodem- en streefwaarden

Onderstaande tabel laat de historische prestatie (2019) en de bodem- en streefwaarde voor de huidige en ontwerpconcessie zien.

*Tabel 4.11 Historische prestatie en bodem- en streefwaarden voor de PI 'Vermeden CO<sub>2</sub>-uitstoot'*

Vermeden CO <sub>2</sub> -uitstoot	Gemeten waarde (2019)	Ontwerpconcessie (2025 – 2029)
Historische prestatie	498	
Bodemwaarde		534
Streefwaarde		620

De bodemwaarde voor de PI is lager dan de waarde gemeten in 2019. De streefwaarde is daarentegen hoger dan deze waarde. De bandbreedtes zijn grotendeels bepaald door geïnterviewde ontwikkelingen en de impact daarvan.

Op basis van de audit op de methode waarmee de bandbreedte van de prestatie op de PI's is vastgesteld, nemen we aan dat de bandbreedte een haalbare range is voor de prestatie van NS. Gegeven dat de bodem- en streefwaarden binnen de bandbreedte liggen zijn die realistisch te noemen. Zoals in hoofdstuk 3 beschreven is het binnen de scope van deze audit niet mogelijk om te beoordelen in welke mate bodem- en streefwaarden ambitieus zijn<sup>40</sup>.

<sup>40</sup> Ondanks dat waarden lager zijn kunnen ze in het licht van de verwachte ontwikkelingen nog steeds ambitieus zijn. Uitgaande van de bandbreedtes zijn inspanningen nodig om bodem- en streefwaarden te halen. De mate van ambitie is niet te beoordelen.

## 5. Bevindingen

In dit hoofdstuk lichten we de belangrijkste bevindingen en aandachtspunten toe en presenteren we onze conclusies en aanbevelingen.

### 5.1 Bevindingen op hoofdlijnen

Onze belangrijkste bevindingen op hoofdlijnen zijn:

- Behoudens enkele verschillen tussen de **definities** in de *ontwerpconcessie* en de *PI-definitie documenten*, zijn de definities helder, concreet en consistent. Het advies is om de definities van PI's waarbij aandachtspunten naar voren komen, aan te scherpen of aan te passen. Dit gaat om de PI's '*Algemeen klantoordeel*', '*Kwaliteit van aansluitingen van NS op andere vervoerders*', '*Zitplaatskans spits en dal*' en '*Aantal drukke treinen per week in de spits*'. Deze aandachtspunten staan nader toegelicht in paragraaf 5.2.
- De toegepaste **meetmethoden** sluiten goed aan op de definities.
- Voor de meeste PI's geldt dat de beoordeling van de **meetpraktijk** ten opzichte van de beschreven meetmethode is gebaseerd op beschrijvingen van de praktijk, procesflows, kwaliteitscontroles en monitoringsrapporten. De scope en beschikbare tijd van deze audit maakt het niet mogelijk mee te lopen in de praktijk en de meetpraktijk uit eigen ervaring te beoordelen. Voor zover inzichtelijk komt de meetpraktijk overeen met de beschreven meetmethode en is geborgd dat de meetpraktijk ook gerapporteerd wordt. Niet beantwoord zijn enkele vragen over de aansluiting van de meetmethode/meetpraktijk bij de definitie van '*Algemeen klantoordeel*' en '*Klantoordeel sociale veiligheid*'. Dit omdat hiervoor de OV-klantenbarometer wordt gebruikt. Dit is een apart onderzoek uitgevoerd in opdracht van CROW door een extern onderzoeksbureau en valt buiten de scope van deze audit.
- De **betrouwbaarheid** van **brongegevens** is niet volledig te toetsen binnen de scope van deze audit. NS en ProRail hebben echter afdoende interne validatie en toetsprocedures en eerdere audits die het aannemelijk maken dat het meten van de PI's op betrouwbare wijze mogelijk is.
- De **ontwikkelingen** en **impacts** op de PI's worden door experts van NS en ProRail uitgewerkt. De methode die hiervoor in gebruik is, is in 2022 getoetst en valide bevonden. De methodiek is in het kader van voorliggend onderzoek niet opnieuw getoetst. Wel is met een onafhankelijke expertmatige blik (spooexpert SWECO) gekeken naar de ontwikkelingen en de (logica van de) doorwerking daarvan op de PI's. De belangrijkste bevinding is dat in algemene zin de lijst herkenbaar is er geen opmerkelijke onderwerpen in staan. Het zijn allemaal thema's en ontwikkelingen die in meer of mindere mate relevant zijn voor de PI's. De wijze waarop de impacts zijn doorvertaald naar de verschillende PI's is voor zover te beoordelen plausibel.

Enkele aandachtspunten:

- Een behoorlijk aantal werkzaamheden aan het spoor wordt gecombineerd tot een langdurige ontwikkeling die wordt aangeduid met een "boeggolf" aan werkzaamheden. De impact hiervan op de PI's is onder meer gebaseerd op de door ProRail berekende Ervaren extra reizigersminuten. Het nadeel van de gecombineerde impact is dat het effect van

- individuele ontwikkelingen niet zichtbaar is en (binnen de scope van deze audit) dus ook niet verder kon worden getoetst.
- Enkele verwijzingen naar de duur van werkzaamheden wijken op verschillende plaatsen in de ontwikkelingenlijst van elkaar af. Hoewel dit naar alle waarschijnlijk geen invloed heeft op de ontwikkelingen en de PI's, bevelen we omwille van de interne consistentie en transparantie aan om planningen consistent volgens de meest recente inzichten op te nemen.
  - Een deel van de afgesproken bodem- en streefwaarden in de ontwerpconcessie ligt lager en een deel ligt hoger dan de in de lopende concessie overeengekomen bodem- en streefwaarden of in het verleden gerealiseerde waarden. Aan de hand van ontwikkelingen (zie voorgaand punt) is een plausibel beeld geschetst hoe prestaties zich naar verwachting ontwikkelen en zich verhouden tot prestaties in het verleden. Aangezien de methode voor de totstandkoming van de bandbreedtes eerder als valide is beoordeeld en we de ontwikkelingen en impacts als plausibel beschouwen gaan we ervan uit dat de geschetste bandbreedte een haalbare range is voor de prestatie van NS (en ProRail voor de gezamenlijke PI's). Gegeven dat de bodem- en streefwaarden (met uitzondering van zitplaatskans) binnen de bandbreedte liggen zijn die daarmee **realistisch** te noemen. Voor de PI's over zitplaatskans geldt dat de streefwaarden buiten de berekende bandbreedtes vallen. Dit roept vragen op over de haalbaarheid ervan.
  - De mate waarin bodem- en streefwaarden **ambitieuw** zijn, is binnen de scope van deze audit niet objectief vast te stellen. Ten aanzien van de mate waarin de bodem- en streefwaarden ambitieus zijn is het niet mogelijk een beoordeling te geven. Tussen IenW en NS is niet afgesproken wanneer een streefwaarde ambitieus is en hier is binnen scope van de opdracht geen norm voor te bepalen. Daarmee ontbreekt de mogelijkheid tot een onderbouwde beoordeling.

## 5.2 Bevindingen per prestatie-indicator

In deze paragraaf beschrijven we de belangrijkste aandachtspunten per PI. Dit betreft alleen aandachtspunten bij de gestelde normen. Ook geven we aan wanneer er geen oordeel geveld kon worden<sup>41</sup> over een bepaalde norm doordat relevante informatie niet of beperkt voorhanden was.

### Algemeen klantoordeel

We signaleren een beperkte inconsistentie in het gebruik van de definitie voor deze prestatie-indicator. In het definitiedocument is zowel onder 2) Publieke omschrijving als onder 3) Gedetailleerde definitie steeds gesproken over de waardering van de reiziger over het *reizen per trein*, in plaats van de waardering van de *reiziger voor de betreffende rit*. Ook in de ontwerpconcessie wordt een net iets gewijzigde definitie gehanteerd, namelijk: “de algehele tevredenheid van reizigers over het geboden product”. De aanbeveling is om de gehanteerde definities gelijk te trekken.

<sup>41</sup> Zoals toegelicht geldt dit voor onderzoeksvraag 5 bij alle PI's.

## **Klantoordeel sociale veiligheid**

Geen aandachtspunten.

## **Aantal drukke treinen in de spits**

De definitie in de ontwerpconcessie is minder specifiek dan die in het PI Definitiedocument. Bij de definitie in de ontwerpconcessie ontbreekt de afbakening in scope en de exacte afbakening in de tijd. Wij adviseren om in de ontwerpconcessie een uitgebreide definitie op te nemen waarin ook de afbakening in tijd en de treinen die in scope zijn staat beschreven, of een verwijzing naar deze informatie op te nemen.

## **Zitplaatskans (spits/dal) (tweede klas)**

De definities in de ontwerpconcessie voldoen op hoofdlijnen aan de normen. Wel zijn ze minder specifiek dan de definities in de definitiedocumenten. Wij adviseren om in de ontwerpconcessie in de definitie ook de afbakening in tijd en de treinen die in scope zijn te beschrijven óf een verwijzing naar de specifieke afbakening elders op te nemen. Geef bij de onderliggende stukken ook aan waarom de spittijden met/zonder 5 minuten marge gehanteerd worden.

## **Reizigerspunctualiteit 3 min/ 10 min HRN**

Een aandachtspunt is dat in de ontwerpconcessie is besloten om de IC Berlijn trein niet op te nemen. In het huidige Definitiedocument staat de zin “Dus wél IC Berlijn en IC Brussel, maar niét Thalys, Eurostar, ICE en Nightjet.” NS heeft aangegeven het definitiedocument hierop indien nodig aan te passen vóór ondertekening van de concessie.

## **Kwaliteit van aansluitingen van NS op andere vervoerders**

De definitie van deze PI is niet voldoende specifiek. De naam en de definitie (in het definitiedocument) van de PI doen vermoeden dat het meet of de overstap op andere vervoerders gehaald wordt, maar dat is niet aan de orde. Specifiek gaat het om aankomstpunctualiteit op stations waar ook andere vervoerders actief zijn. Het advies is daarom om de definitie op zo'n manier te formuleren dat deze overeenkomt met hetgeen wat daadwerkelijk wordt gemeten (namelijk de aankomstpunctualiteit op stations waar een overstap op een andere vervoerder mogelijk is).

## **Impactvolle verstoringen veroorzaakt door NS**

In de ontwerpconcessie heeft deze PI de naam 'Impactvolle verstoringen veroorzaakt door NS' en is daarmee specifiek en anders dan (dus niet consistent met) de naam in het definitiedocument: 'aantal verstoringen veroorzaakt door NS'. Het advies is daarom om deze definitie in het definitiedocument ook specifiek te beschrijven.

## **Kwaliteit van Reisinformatie inclusief verstoringen**

De beoordeling van de meetpraktijk is gebaseerd op beschrijvingen van de praktijk, procesflows, kwaliteitscontroles en monitoringsrapporten. Voor zover inzichtelijk komt de meetpraktijk overeen met de beschreven meetmethode en is geborgd dat de meetpraktijk ook gerapporteerd wordt. Het PI

Definitiedocument geeft aan dat de tabel met betrouwbaarheid van meetsystemen op onderdelen nog geactualiseerd dient te worden. Actualisatie van het PI document gebeurt jaarlijks en staat op de planning om voor de ondertekening van de ontwerpconcessie in 2023 te zijn uitgevoerd. Het is belangrijk dat deze actualisatie goed in beeld is bij relevante betrokkenen en daadwerkelijk wordt uitgevoerd.

## Vermeden CO<sub>2</sub>-uitstoot

De definitie in de ontwerpconcessie is smaller dan in het definitiedocument van NS, maar voldoet aan de normen specifiek, compleet, meetbaar, in zichzelf consistent en voor één uitleg vatbaar. De afbakening in tijd (per jaar) en afbakening in scope (wat behoort precies tot de operatie van NS en de keten) ontbreken in de definitie in de ontwerpconcessie. Deze elementen zijn wel onderdeel van de uitgewerkte definitie in het definitiedocument van NS. Wij adviseren om in de ontwerpconcessie een verwijzing naar deze informatie op te nemen.

## 5.3 Conclusies en aanbevelingen

Onderstaande tabel laat per prestatie-indicator zien of aan alle normen voor de verschillende onderzoeksvragen wordt voldaan.

Thema	Onderzoeksvraag	Algemeen Klantordeel	Klantoordeel Sociale Veiligheid	Aantal drukke treinen per week in de spits	Zitplaatskans Spits tweede klas	Zitplaatskans Dal tweede klas	Reizigerspunctualiteit 3 minuten HRN	Reizigerspunctualiteit 10 minuten HRN	Kwaliteit van aansluitingen van NS op andere vervoerders	Impactvolle verstoringen veroorzaakt door NS	Kwaliteit Reisinformatie Inclusief Verstoringen	Vermeden CO <sub>2</sub> - uitstoot
Definitie	1. Is de definitie helder en concreet beschreven en niet op meerdere manieren te interpreteren?	✓*	✓	✓*	✓*	✓*	✓	✓	✓*	✓*	✓	✓*
Meetmethode, meetpraktijk, brongegevens	2. Is de meetmethode helder en concreet beschreven en niet op meerdere manieren te interpreteren, en sluit deze aan op de definitie van de PI?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	?
	3. Komt de meetpraktijk overeen met de beschrijving van de meetmethode en hoe wordt geborgd dat ook zal worden gerapporteerd volgens de meetpraktijk?	✓	✓	✓	✓	✓	✓	✓	✓	?	✓	?
	4. Zijn de brongegevens statistisch betrouwbaar en is de kwaliteitscontrole hiervan voldoende geborgd?	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	?
Bodem en streefwaarden	5. Zijn de bodem- en streefwaarde – mede gelet op de huidige prestaties en de voorziene en actuele relevante ontwikkelingen die deze positief of negatief kunnen beïnvloeden – ambitieus genoeg om op betekenisvolle wijze bij te dragen aan het realiseren van ambitieuze prestaties?	?	?	?	?	?	?	?	?	?	?	?
	6. Zijn de bodem- en streefwaarden realistisch en (redelijkerwijs) haalbaar?	✓	✓	✓	✓*	✓*	✓	✓	✓	✓	✓	✓

- ✓ Voldoet aan normen
- ✓\* Voldoet met een of enkele aandachtspunten
- Niet volledig objectief toetsbaar binnen scope
- ? Geen oordeel
- Voldoet niet aan normen



Voor onderzoeksvragen 3 (meetmethoden), 4 (data) en 5 (ambitieuze bodem- en streefwaarden) geldt dat deze slechts ten dele getoetst konden worden (zie de bevindingen op hoofdlijnen in 5.1). Voor onderzoeksvraag 3 en 4 is wel getoetst of het aannemelijk is dat ze aan de normen voldoen *gegeven* deze beperkingen. Voor onderzoeksvraag 5 was dit niet mogelijk binnen de scope van deze audit.

## Conclusies

Het algemene beeld van de toets is dat voor de verschillende PI's aan de meeste normen voldaan wordt:

- De PI's voldoen aan de gestelde normen voor onderzoeksvraag 1 (definities) en 2 (meetmethode). Aandachtspunt bij de definities is dat er soms afwijkingen zitten in beschrijvingen in verschillende documenten. Aandachtspunt bij de onderzoeksvraag 2 (meetmethode) is de PI Vermeden CO<sub>2</sub>-uitstoot. Deze is nog in ontwikkeling kon daarom in deze audit nog niet worden voorzien van een oordeel.
- Voor onderzoeksvraag 3 (meetmethode) en 4 (meetpraktijk) geldt dat de meetpraktijk en de gebruikte data overeenkomen met de beschreven meetmethode, maar dat binnen de scope van deze audit een volledige toetsing van de praktijk en gebruikte data niet haalbaar was (in de overzichtstabel zijn deze velden daarom grijs gearceerd). Gezien eerdere validaties is het echter aannemelijk dat aan alle normen wordt voldaan. De PI Vermeden CO<sub>2</sub>-uitstoot heeft geen beoordeling en bij onderzoeksvraag 4 (meetpraktijk) ook de PI Impactvolle Verstoringen , veroorzaakt door NS niet. De reden hiervoor is dat deze (aspecten van deze) PI's nog in ontwikkeling zijn.
- Voor onderzoeksvraag 5 (voldoende ambitieuze bodem en streefwaarden) geldt dat de eerder gevalideerde methode om afspraken te maken over bodem- en streefwaarden is toegepast. De achterliggende ontwikkelingen hebben we binnen de scope van deze audit niet voldoende kunnen toetsen. Hierdoor en door het ontbreken van een norm voor 'ambitie', is de mate waarin bodem- en streefwaarden ambitieus zijn niet te beoordelen.
- Voor onderzoeksvraag 6 (haalbaarheid bodem en streefwaarden) voldoen de PI's aan de normen, al zijn er bij de PI's die gaan over de zitplaatskans wel vraagtekens te plaatsen. Afgesproken waarden liggen namelijk buiten de geschatte bandbreedtes.

## Aanbevelingen

We bevelen aan om de definities en beschrijvingen in de verschillende documenten waarin de PI's worden behandeld, gelijk te trekken. Zie paragraaf 5.2 voor concrete aandachtspunten per PI.

Om in de toekomst wel een beoordeling te kunnen geven op de mate waarin bodem- en streefwaarden ambitieus zijn (onderzoeksvraag 5) bevelen we aan om:

- Als ministerie actief de ontwikkelingen van de PI's te monitoren.
- Voor de start van de Mid Term Review (MTR) te beoordelen of de ontwikkelingen die zijn meegenomen bij het opstellen van de bandbreedtes daadwerkelijk hebben plaatsgevonden en wat de daadwerkelijke impact hiervan is geweest op PI's.

- Als er in de toekomst een nieuwe ontwikkelingenlijst of bandbreedtes worden opgesteld, een onafhankelijke expert een toets te laten uitvoeren op de inhoud en evenwichtigheid van de ontwikkelingenlijst en bandbreedtes. Hiermee wordt de neutraliteit en de evenwichtigheid van de ontwikkelingenlijst en de bijbehorende bandbreedtes geborgd.
- Een gezamenlijke norm (of schaal) vast te stellen waaraan het ambitieniveau van PI's kan worden getoetst. Zonder versturende externe factoren zouden PI's in principe jaar op jaar verbeteringen moeten laten zien om ambitieus te kunnen worden genoemd (tot een maximum). Externe ontwikkelingen kunnen ertoe leiden dat een absolute verbetering niet haalbaar is. In dat geval moet aannemelijk worden gemaakt dat er maatregelen worden genomen die leiden tot een relatieve verbetering ten opzichte van de situatie zonder maatregelen. Internationale benchmark(s), prestaties in het verleden en lopende en toekomstige ontwikkelingen kunnen een kader bieden voor afspraken over een norm voor het ambitieniveau.

## Bijlage 1: Bronnen

### *Geraadpleegde personen*

In het kader van dit onderzoek is uitgebreid met betrokkenen van de volgende organisaties gesproken:

- NS
- ProRail

### *Geraadpleegde brondocumenten*

- Andersson Elffers-Felix. (2016). *Audit prestatie-indicatoren vervoer- en beheerconcessie*. Ministerie van Infrastructuur en Milieu.
- Boot, D., & Lustermans, G. (2022). *Methodiek bodem- en streefwaarden PI's NS*. Andersson Elffers Felix. Utrecht: Ministerie van Infrastructuur en Waterstaat.
- Brand, M., & Klein Kranenburg, R. (2023, januari). Definitie PI - Zitplaatstijd in het dal 2de klas. ## *Definitie PI - Zitplaatstijd in het dal 2de klas*.
- CROW. (2023, maart). OV-Klantenbarometer 2022. *Hoofdrapport*.
- Dijkstra, J. (2023, februari 01). Definitie PI - Aantal verstoringen veroorzaakt door NS. 039 *Definitie PI - Aantal verstoringen veroorzaakt door NS*.
- Dijkstra, J. (2023, juni 05). Definitie PI - Kwaliteit van aansluitingen van NS op andere vervoerders. 084 *Definitie PI - Kwaliteit van aansluitingen van NS op andere vervoerders*.
- Klein Kranenburg, R. (2023, januari). Definitie PI - Zitplaatskans in de spits 2e klas. 09X *Definitie PI - Zitplaatskans in de spits 2e klas*.
- Klein Kranenburg, R. (2023, januari). Definitie PI - Zitplaatskans in het dal 2e klas. 09Q *Definitie PI - Zitplaatskans in het dal 2e klas*.
- KPMG. (2017, oktober 31). Audit op prestatie-indicatoren beheerconcessie en vervoerconcessie. Amstelveen.
- Kroeze, C. (2023, januari). Definitie PI - Vermeden CO2 uitstoot. xxx *Definitie PI - Vermeden CO2 uitstoot*.
- Ministerie van Infrastructuur en Waterstaat. (2023, juli). Ontwerp-concessie voor het Hoofdrailnet 2025-2033.
- NS - Werkstroom PI's, in samenwerking met ProRail voor gedeelde PI's. (sd). Ontwikkelingenlijst incl. impact - def (vertrouwelijk).
- NS. (2014, november 25). Toetsing aannames SOFA. (D. Mestrum, & C. Roelofs, Samenstellers)
- NS. (2015). 2015-01-01 Tegels beschrijving oorzaaktypes (vertrouwelijk).
- NS. (2015, november 27). Meetsysteem monitoring C0.1. (B. van den Berg, Samensteller)
- NS. (2019, augustus 16). (17.2) Business Requirements Treinketen 2020 koppelen. (R. van der Straten, Samensteller)
- NS. (2019, augustus 19). (17.3) Business Requirements Treinketen 2020. *Gebruiksvriendelijke Reisinformatie*. (R. van der Straten, Samensteller)
- NS. (2019, april 08). Functioneel Ontwerp Kwaliteit van Aansluitingen Azure V0.5. (A. Rodenburg, Samensteller)

- NS. (2020, 05 25). International Benchmark 2014-2018.
- NS. (2020, 06 30). Klein Onderhoud 2021. *PI's Zitplaatskans*. (R. de Kamer, & E. Wolfswinkel, Samenstellers)
- NS. (2020, november 10). Procesflow PI Kwaliteit van aansluitingen (CDW) (vertrouwelijk).
- NS. (2021, 12 7). Meting PI Zitplaatskans spits 2021/2022.
- NS. (2021, december 16). Procesbeschrijving SOFA. (E. Wolfswinkel, Samensteller)
- NS. (2021, november 24). Veldwerk. (A. d.-v. Linden, Samensteller)
- NS. (2022, november 9). Impactschatting TSB van 80 km/u Rijpwetering.
- NS. (2022, oktober 31). Methodiek voor het komen tot PI-bandbreedtes voor bodem- en streefwaarden (vertrouwelijk).
- NS. (2023, juni 5). 113 Definitie informatie indicator Aankomstpunctualiteit 3 minuten.
- NS. (2023, juni 5). 114 Definitie Informatie Indicator Aankomstpunctualiteit 5 minuten.
- NS. (2023). 230612 Reizigerskilometers Voorjaarsprognose 2023 (vertrouwelijk).
- NS. (2023, januari 19). DOO-Maandrapportage. *December 2022*.
- NS. (2023, augustus 17). PI Algemeen klantoordeel.
- NS. (2023, augustus 24). PI Sociale Veiligheid (vertrouwelijk).
- NS. (2023, mei 9). Kwaliteitsrapport regieteam HRN PI's . *Q1 2023 (intern)*.
- NS. (2023, juli). Procesflow concessie-PI zitplaatskans (vertrouwelijk).
- NS. (2023, 10 8). Slides PI Zitplaatskans 2e klas spits/dal & drukke treinen spits t.b.v. audit Decisio (vertrouwelijk).
- NS. (2023, juli 18 ). Status implementatie Prestatie-indicatoren (vertrouwelijk).
- NS. (2023, januari 10). Toelichting AEF methodiek t.a.v. berekening bandbreedtes PI's (vertrouwelijk). *Werksessie lenW. Utrecht*.
- NS. (2023, 08). Uitgevoerde ARIS controls KvA (08-2023) (vertrouwelijk).
- NS. (2023, 08). Uitgevoerde ARIS controls RPUN 08-2023 (vertrouwelijk).
- NS. (2023, augustus). Uitgevoerde controls Zitplaatskans (testperiode augustus 2023).
- NS. (2023, maart). Vervoerconcessie voor het HRN (vertrouwelijk). *Bijgewerkte PI bandbreedtes t.o.v. financieel kader*.
- NS. (2023, februari 03). Vervoerconcessie voor het HRN (vertrouwelijk). *Dialoogproduct lenW - PI bandbreedtes*.
- NS. (2023, juni 21). Vervoerconcessie voor het HRN deel 2 (vertrouwelijk). *Bijgewerkte PI-bandbreedtes o.b.v. nieuwe reizigersprognoses, dienstregeling, tariefindexatie en volume internationale treinen*.
- NS. (2023, juni 21). Vervoerconcessie voor het HRN\_deel 1 (vertrouwelijk). *Bijgewerkte PI-bandbreedtes o.b.v. nieuwe reizigersprognoses, dienstregeling, tariefindexatie en volume internationale treinen*.
- NS. (2023, januari 30). xxx Definitie Prestatie-indicator Vermeden CO2-uitstoot. (versie 0.1 concept). (C. Kroeze, Samensteller)
- NS. (sd). Concessie PI Vermeden CO2 uitstoot (vertrouwelijk).
- NS. (sd). Controlframework ICF PI- RPUN en KvA.
- NS. (sd). Controlframework ICF PI- ZPK.

- NS Kenniscentrum. (2022, 05 15). FO RPUN. *T.b.v. prestatie-indicator Reizigerspunctualiteit*. (J. Helgers, Samensteller)
- NS. (sd). Meetmethode HRN PI GVR . *Screenshots spoorwijzigingen, vertragingen, opgeheven stops, tijdig handelingsperspectief*.
- NS. (sd). Opvolging aanbeveling Onderzoek Deloitte 20-10-2021. *PI Gebruiksvriendelijke informatie en Aantal drukke treinen per werkweek in de spits HRN* .
- NS. (sd). Procesflow PI Reizigerspunctualiteit (vertrouwelijk). *Kritische Prestatie-indicatoren: RPUN 3 minuten HRN en RPUN 10 minuten HRN*.
- NS. (sd). Rapport voor lenW (vertrouwelijk). *Maandrapport december 2022*.
- NS. (sd). Reizigerspunctualiteit HRN. *Prestatie-indicator van NS & ProRail*.
- NS. (sd). Vrijgave+checks+GVR. *Vergelijking aantallen CDW vs TM1*.
- Oostermeyer, K., & Maassen, P. (2019). *Onderzoek PI Zitplaatskans in de spits HRN en HSL*. Rebel. Rotterdam: Ministerie van Infrastructuur en Waterstaat.
- ProRail. (2022, oktober 3). Spoorkompas. *Interface Requirements Specification Export Punctualiteit naar NSR*.
- ProRail. (2023, 05 19). Keek op de Week 2023 - week 24 (vertrouwelijk).
- ProRail. (2023, 07 11). Maandprestaties ProRail (vertrouwelijk).
- ProRail. (2023). Managementsamenvatting RM Q2 2023 definitief concept (vertrouwelijk). *Belangrijkste risico's Q2 2023: kans en impact* .
- Rodenburg, A. (2023, januari 19). Definitie PI - Kwaliteit Reisinformatie incl. verstoringen. XXX *Definitie PI - Kwaliteit Reisinformatie incl. verstoringen*.
- Schweizer, E., & Egbers, J.-M. (2023, juni 05). Definitie PI - Reizigerspunctualiteit 10 minuten HRN. *088 Definitie PI - Reizigerspunctualiteit 10 minuten HRN*.
- Schweizer, E., & Egbers, J.-M. (2023, juni 05). Definitie PI - Reizigerspunctualiteit 3 minuten HRN. *086 Definitie PI - Reizigerspunctualiteit 3 minuten HRN*.
- Sustainalize. (2023, januari). Advies PI vermeden CO2 uitstoot t.b.v. nieuwe concessie hoofdrailnet.
- Urlings, T. (2020, september 18). Definitie PI - Algemeen klantoordeel. *095 Definitie PI - Algemeen klantoordeel*.
- Urlings, T. (2020, september 17). Definitie PI - Klantoordeel sociale veiligheid. *111 Definitie PI - Klantoordeel sociale veiligheid\_versie 1.1*.
- van Schagen, E., Liu, J., & Veldhuis, W. (2021). *Onderzoek PI's gebruiksvriendelijke reisinformatie | drukke treinen*. Deloitte. Eindhoven: Ministerie van Infrastructuur en Waterstaat.
- Wolfswinkel, E. (2022, november 28). Definitie PI - aantal drukke treinen per week in de spits. *100 Definitie PI - aantal drukke treinen per week in de spits*.