



Rijksinstituut voor Volksgezondheid
en Milieu
*Ministerie van Volksgezondheid,
Welzijn en Sport*

Programmatische Aanpak Meten **Vliegtuiggeluid (PAMV):**

Validatie Doc.29-model voor Schiphol

**Programmatiese Aanpak Meten
Vliegtuiggeluid (PAMV):**
Validatie Doc.29-model voor Schiphol

RIVM-rapport 2023-0429



Colofon

© RIVM 2024

Delen uit deze publicatie mogen worden overgenomen op voorwaarde van bronvermelding: Rijksinstituut voor Volksgezondheid en Milieu (RIVM), de titel van de publicatie en het jaar van uitgave.

DOI 10.21945/RIVM-2023-0429

A. Sahai (auteur, coördinator), RIVM
T. Wartenberg (auteur), RIVM
N. Mabjaia (auteur), RIVM
R. Hogenhuis (auteur), NLR
S. Heblj (auteur), NLR
K. Vinkx (auteur), To70

Contact:

Abhishek Sahai
Centrum Milieukwaliteit
abhishek.sahai@rivm.nl

Dit onderzoek werd verricht in opdracht van het ministerie Infrastructuur en Waterstaat, directie Luchtvaart in het kader van de Programmatische Aanpak Meten Vliegtuiggeluid (PAMV).

Dit is een uitgave van:
**Rijksinstituut voor Volksgezondheid
en Milieu**
Postbus 1 | 3720 BA Bilthoven
Nederland
www.rivm.nl

Publiekssamenvatting

Programmatische Aanpak Meten Vliegtuiggeluid (PAMV):

Validatie Doc.29-model voor Schiphol

In Nederland zijn alle normen over vliegtuiggeluid gebaseerd op berekeningen. Omwonenden willen dat de overheid meer gebruikmaakt van metingen, omdat zij die meer vertrouwen. Dit vertrouwen kan worden vergroot door met metingen te controleren of de rekenmodellen kloppen (valideren). Om te kijken hoe zo'n systeem ingericht kan worden, zijn in dit onderzoek de uitkomsten van berekeningen en metingen van vliegtuiggeluid rond Schiphol vergeleken.

Bij een hoge geluidbelasting blijken de berekeningen en metingen vrij goed overeen te komen. De geluidbelasting is het geluidniveau van vliegtuigen keer het aantal overvliegende vluchten. De geluidbelasting in L_{den} is de gemiddelde geluidbelasting over een jaar gewogen over een periode van 24 uur. Bij een hoge geluidbelasting bleek het verschil tussen meten en rekenen per meetpost minder dan 2 decibel L_{den} . Dat maakt de berekeningen voor een hoge geluidbelasting betrouwbaar. De L_{den} is een geluidmaat voor de jaargemiddelde geluidbelasting. Het geeft in principe een waarde voor de gewogen jaargemiddelde etmaal belasting.

Bij lage geluidbelasting is het verschil tussen meten en rekenen groter. Dat komt onder andere omdat er weinig betrouwbare meetdata van lage geluidniveaus zijn. De meetpunten rond Schiphol meten veel vluchten met een laag geluidniveau namelijk niet. Hierdoor wordt een flink aantal vliegbewegingen met een laag geluid niet geregistreerd. Dit geldt bijvoorbeeld voor plaatsen die verder weg van Schiphol liggen.

In dit onderzoek zijn de berekeningen gedaan met het zogeheten Doc.29-rekenmodel. De resultaten van de metingen zijn verzameld via meetposten van Schiphol (het NOMOS-meetnetwerk). Van alle meetposten van dit netwerk is gekeken welke geschikt zijn voor de validatie. Hiervoor zijn criteria gebruikt die eerder zijn bepaald voor de nationale meetstrategie PAMV (Programmatische Aanpak Meten Vliegtuiggeluid). Meer onderzoek is nodig om de oorzaken van de gevonden verschillen te verklaren. Aanbevolen wordt te onderzoeken welk effect de beperkte meetdata van lage geluidniveaus heeft op de resultaten. Ook adviseren we uit te zoeken hoe deze geluidniveaus beter kunnen worden gemeten.

Een consortium van het RIVM, het Nederlands Lucht- en Ruimtevaartcentrum (NLR) en adviesbureau To70 heeft dit onderzoek gedaan. Aanleiding was een aanbeveling uit 2019 om de kwaliteit van de berekeningen van vliegtuiggeluid te beoordelen. Met de uitkomsten van dit onderzoek is ook de eerste stap gezet om een validatieprogramma op te zetten dat lange tijd en structureel de rekenresultaten met metingen controleert.

Kernwoorden: vliegtuiggeluid, meten, rekenen, validatie, PAMV, Doc.29

Synopsis

Programme-Based Approach to Measuring Aircraft Noise (PAMV)

Validation of Doc.29 model for Schiphol

In the Netherlands, all aircraft noise limit values are based on modelling. Those who live near airports would like the government to make more use of measurements, because they have more confidence in them. This confidence can be enhanced by conducting measurements to verify the models used (validation). In order to investigate what such a validation system could look like, researchers have compared the results of modelling and measurements of aircraft noise near Amsterdam Airport Schiphol.

For high noise exposures, the modelling and the measurements were reasonably consistent. The noise exposure (L_{den}) is the annually averaged weighted noise level of aircraft. For high noise exposures, the difference between the results of measurements and modelling was less than 2 dB L_{den} per measurement location. This means that modelling is reliable for high noise exposures.

For low noise exposures, the difference between the results of measurements and modelling was more pronounced. One of the reasons for this is the lack of reliable low noise level measurement data. This is because the measurement locations near Schiphol do not measure many of the flights that produce low noise levels. As a consequence, a great many flights that produce low noise levels are not registered. This is the case, for example, for many measurement locations further away from Schiphol.

For this research, the Doc.29 model was used for modelling. The measurement results were gathered from Schiphol measurement locations (the NOMOS measurement network). All measurement locations that form part of this network were assessed for their suitability for validation, using criteria drawn up earlier for the Programme-Based Approach to Measuring Aircraft Noise (*Programmatische Aanpak Meten Vliegtuiggeluid*, PAMV) measurement strategy. The researchers noted that further research is needed to explain the reasons for the differences. One of their recommendations is to study the effect of the lack of low noise level measurement data on the results. A second recommendation is to investigate how the measurement of low noise levels could be improved.

This research was carried out by a consortium made up of RIVM, the National Aerospace Laboratory (NLR) and the consultancy To70. It was prompted by a recommendation from 2019 to review the quality of aircraft noise modelling. The results of this research are the first step towards setting up a validation programme that uses measurements to verify modelling results consistently and for the long term.

Keywords: aircraft noise, measuring, modelling, validation, PAMV, Doc.29

Inhoudsopgave

Samenvatting – 9

1 Inleiding – 15

- 1.1 Programmatische Aanpak Meten Vliegtuiggeluid (PAMV) – 15
 - 1.1.1 Aanleiding PAMV – 15
 - 1.1.2 Aanbevelingen PAMV-verkenningfase – 16
- 1.2 Validatie – 17
 - 1.2.1 Aanleiding en behoefte – 17
 - 1.2.2 Relevante informatie uit de literatuur – 20
 - 1.2.3 PAMV-validatieaanpak – 25
 - 1.2.4 Relatie met andere onderzoeken – 27
- 1.3 Opdrachtverlening – 29
- 1.4 Leeswijzer – 29

2 Vergelijking meten en rekenen – 31

- 2.1 Aanpak en methodologie – 31
- 2.2 Gebruikte invoergegevens – 31
 - 2.2.1 Meetgegevens – 31
 - 2.2.2 Berekeningsgegevens – 33
- 2.3 Resultaten vergelijking meten en rekenen – 34
 - 2.3.1 Resultaten zonder gebruik van subselecties van meetdata – 35
 - 2.3.2 Resultaten voor subselecties van meetdata – 47
 - 2.3.3 Discussie: in hoeverre is er sprake van een systematisch verschil? – 55

3 Conclusies en aanbevelingen – 61

- 3.1 Aanbevelingen – 63

Referenties – 65

Samenvatting

Aanleiding en behoefte

Rekenmodellen spelen een belangrijke rol binnen de huidige systematiek rondom vliegtuiggeluid. Alle normen binnen het Nederlandse beleid zijn gebaseerd op berekende waarden en de handhaving van de jaarlijkse geluidbelasting wordt ook op basis van berekeningen gedaan. Lokaal uitgevoerde metingen worden momenteel voornamelijk voor informatievoorziening aan omwonenden gebruikt en hebben geen formele rol in de wet- en regelgeving.

Tijdens de verkenning die is uitgevoerd voor de Programmatische Aanpak Meten Vliegtuiggeluid (PAMV) (Smetsers, Siegmund et al. 2019) is vanuit de omgeving duidelijk de wens naar voren gekomen om metingen een belangrijkere rol te geven binnen het luchtvaartdossier. Een van de mogelijkheden om de rol van metingen te vergroten, is het gebruik van metingen om rekenmodellen te valideren. Door een vergelijking van rekenresultaten met metingen kan onderzocht worden of de rekenmodellen verbetering behoeven en kunnen beide methoden elkaar versterken.

Om de verschillen tussen meet- en rekenresultaten te onderzoeken en te duiden, is tijdens de PAMV-verkenningsfase aanbevolen om een systeem op te zetten dat met metingen controleert of de resultaten van geluidberekeningen van voldoende kwaliteit zijn. Hiermee kan onderzocht worden of de modelinvoergegevens een goede weergave van de werkelijkheid geven, maar ook of het wenselijk is om verbeteringen door te voeren aan de geluidmodellering zelf. Onderdeel van de ontwikkeling van dit systeem is het bepalen van criteria om de kwaliteit van de rekenresultaten te beoordelen en om te bepalen wanneer nader onderzoek naar verbeteringen nodig of wenselijk is. Deze criteria kunnen gebaseerd zijn op verschillen in geluidbelastingniveaus, maar ook bijvoorbeeld op statistische correlatiecoëfficiënten tussen berekende en gemeten data. Door het toepassen van zo'n validatiesysteem en de uitkomsten daarvan openbaar en transparant te delen, zou het vertrouwen van burgers in de rekenresultaten verhoogd kunnen worden.

Tijdens de verkenning is verder aanbevolen om de focus van het validatietraject op het Doc.29-model te leggen, aangezien dit model naast Schiphol op termijn ook voor de regionale luchthavens van nationale betekenis wordt gebruikt. Er is op dit moment beperkt inzicht in hoe betrouwbaar het Doc.29-rekenmodel voor de Nederlandse situatie is. Met de Nederlandse situatie wordt niet alleen de intensieve luchtvaartoperatie in Nederland bedoeld, maar ook de Nederlandse weersomstandigheden (veel voorkomende harde wind en neerslag, combinatie van lage temperaturen en hoge luchtvochtigheid).

Doel

Dit onderzoek heeft als doel om het Doc.29-model, specifiek de implementatie van dit model voor Schiphol, te valideren. Hiervoor is het jaargemiddelde, berekende geluidbelasting bepaald met het Doc.29-

model vergeleken met de gemeten geluidbelasting op meetposten van het NOMOS-meetsysteem rondom Schiphol. De geluidbelasting is uitgedrukt in de L_{den} -geluidmaat.

Het onderzoek is uitgevoerd in opvolging van twee aanbevelingen uit de PAMV-verkenningsfase:

1. Signaalfunctie: Het ontwikkelen van een methodiek om de kwaliteit van geluidberekeningen te beoordelen.
2. Modelontwikkeling: Het starten van een structureel en langjarig modelvalidatieprogramma met de focus op het Doc.29-rekenmodel.

De term 'structureel' in aanbeveling 2 geeft aan dat het proces van het vergelijken van rekenresultaten met metingen met een bepaalde frequentie dient te worden herhaald. De frequentie waarmee de validatie structureel kan worden uitgevoerd, moet nog worden bepaald.

Aanpak

De validatie van het Doc.29-model voor Schiphol is door een consortium van RIVM-NLR-To70 uitgevoerd in twee fasen – een validatie-opstartfase en een uitwerkfase. De opstartfase is in 2022 uitgevoerd en heeft medio 2022 tot een validatieaanpak geleid (Sahai 2022). Voor de validatieaanpak zijn criteria bepaald die voor de validatie van het Doc.29- model gehanteerd moeten worden en stappen die tijdens de uitwerkfase gevolgd moeten worden. Vervolgens is begin 2023 begonnen aan de validatie-uitwerkfase dat tot het huidige rapport heeft geleid. De belangrijkste punten uit de validatieaanpak staan hieronder samengevat:

Validatiescope

Op basis van een literatuurstudie en eerdere ervaring van het consortium met validatie van rekenmodellen is de volgende scope voor de validatie voorgesteld:

- De validatie richt zich op geluidberekeningen volgens de ECAC Doc.29 (4th edition)-methode, specifiek de implementatie voor Schiphol.
- De validatie richt zich op het verbeteren van specifieke invoergegevens van geluidberekeningen (dat wil zeggen: niet de modellering zelf, wat op Europees niveau wordt bepaald).
- De validatie richt zich op de jaargemiddelde geluidbelasting, uitgedrukt in de L_{den} -geluidmaat. Wel zal ook een analyse op SEL-niveau (dat wil zeggen: vluchtniveau) worden gepresenteerd om inzicht te krijgen in de L_{den} -verschillen. Dit wordt gedaan omdat de L_{den} -geluidbelastingmaat wordt opgebouwd door het optellen van geluidniveaus per vliegbeweging uitgedrukt in de SEL-geluidmaat.
- De validatie ondersteunt het onderzoek naar het toepassingsbereik van metingen en berekeningen. Dit onderzoek wordt apart gerapporteerd en heeft het doel om te onderzoeken tot waar metingen en berekeningen van vliegtuiggeluid betrouwbaar kunnen worden uitgevoerd.

Aanvullende meeteisen voor validatie

De gebruikte meetgegevens voor validatie moeten in eerste instantie zijn verzameld met meetsystemen die geschikt zijn voor validatie en op locaties die voldoen aan de validatie-eisen, zoals eerder vastgelegd in de PAMV Nationale Meetstrategie (Sahai 2021). Er is voor de validatieaanpak ook over aanvullende meeteisen nagedacht. Hierbij wordt voorgesteld om metingen alleen te gebruiken voor validatie als ze zijn uitgevoerd onder de volgende condities:

- Windsnelheid: 5 m/s of minder, tenzij er voor de specifieke meetpost ruimere limieten zijn vastgesteld.
- Neerslag: geen neerslag.
- Elevatiehoek: 60 graden of meer.

De implicatie van het toepassen van deze aanvullende meeteisen is dat alleen een deel van alle metingen uiteindelijk voor de validatie wordt gebruikt. Afhankelijk van de locatie van de meetpost in relatie tot de vliegroutes kan dit betekenen dat een significant aantal metingen niet aan deze strenge eisen voldoet. Tegelijkertijd zorgt het toepassen van deze eisen voor een eerlijke vergelijking tussen metingen en berekeningen.

Validatiecriterium

Voor de validatieaanpak is een streefwaarde voorgesteld voor het maximale verschil tussen meten en rekenen. Een maximaal verschil van 2 dB L_{den} per meetpost wordt hierbij als een *initiële streefwaarde* beschouwd. Bij verschillen tussen gemeten en berekende geluidbelasting van groter dan 2 dB L_{den} , zal gezocht moeten worden naar mogelijkheden voor modelverbetering of naar een verklaring voor het verschil in de meetwaarden. Op basis van bevindingen tijdens de validatie kan deze streefwaarde, waar nodig, worden aangescherpt.

De overeenkomst tussen metingen en berekeningen van vliegtuiggeluid is in het voorliggende rapport op basis van een statistische analyse bepaald. Hiermee is een poging gedaan om statistisch te duiden hoe goed de correlatie tussen metingen en berekeningen is, hoe groot het absolute verschil tussen metingen en berekeningen is en hoe hoog de spreiding in dit verschil is. Hiernaast had de in dit rapport gepresenteerde analyse ook als doel om nader inzicht te verkrijgen in de verschillen op totaal niveau, op vluchtsoortniveau (starts/landingen) en op vliegtuigtypeniveau. Deze informatie kan in een eventueel vervolgotraject voor modelverbetering worden gebruikt.

Raakvlakken met het onderzoek naar het toepassingsbereik

Delen van het onderzoek naar het toepassingsbereik van metingen en berekeningen (Sahai 2024) zijn op dezelfde vergelijking van berekende en gemeten geluidbelasting gebaseerd als het validatieonderzoek. In het toepassingsbereikonderzoek is onderzocht of het verschil tussen meten en rekenen verder weg van de luchthaven en in gebieden met een lage geluidbelasting zodanig toeneemt dat het niet meer als acceptabel kan worden beschouwd (meer dan 2 dB L_{den}). Het punt vanaf waar dit gebeurt, bepaalt in principe het toepassingsbereik van het rekenmodel.

Voor het toepassingsbereikonderzoek is verder het effect van de verstorende factoren, zoals wind, neerslag en lage elevatiehoek, op het

verschil tussen meten en rekenen in kaart gebracht. Inzicht in deze aspecten is direct ook voor het validatieonderzoek van belang, omdat het duidelijk maakt of de versturende factoren de vergelijking van berekeningen met metingen verslechteren.

Hiernaast zijn de volgende punten niet in het toepassingsbereik onderzoek geadresseerd en worden in het voorliggende rapport over validatie gerapporteerd:

- De vergelijking van de berekende en gemeten geluidbelasting wordt in het validatierapport uitgebreid met een meer uitvoerige statistische analyse.
- De analyse wordt verder verdiept door de vergelijking ook op vliegtuigtype-niveau te presenteren naast een analyse op vluchtsoort-niveau.
- Er wordt het effect onderzocht op de vergelijking tussen metingen en berekeningen van het gebruik van de drempelwaarde en de vliegtuigbewegingen die daardoor in het meetdataset ontbreken. Dit effect wordt geanalyseerd door een vergelijking te maken van de berekende geluidbelasting indien alle vliegbewegingen van en naar Schiphol worden meegenomen, versus de berekende geluidbelasting als alleen vliegbewegingen worden beschouwd die met een gemeten vlucht kon worden gekoppeld.
- Er wordt voor het validatieonderzoek ook een vergelijking op SEL-niveau per meetpost gepresenteerd voor het verkrijgen van meer inzicht in de verschillen waargenomen op L_{den} -niveau.

Bevindingen

Algemene conclusies

De analyse in het voorliggende validatieonderzoek is uitgevoerd voor een dataset waarin iedere gemeten vliegtuigbeweging aan een berekende vliegtuigbeweging is gekoppeld (de zogenoemde gekoppelde dataset). Op basis van de analyse kunnen de volgende algemene conclusies over de vergelijking tussen metingen en berekeningen worden getrokken:

- Er is een sterke correlatie tussen metingen en berekeningen en allebei volgen dezelfde trends. De correlatie wordt sterk verbeterd door gebruik te maken van validatiemeetposten.
- Het toepassen van de aanvullende meeteisen die zijn bepaald voor de validatieaanpak voor wind, neerslag en elevatiehoek verkleint het verschil tussen metingen en berekeningen en elimineert grotendeels de spreiding in het verschil.
- De waarde van de geluidbelasting op basis van metingen komt in het algemeen hoger uit dan de berekende geluidbelasting.
- Het verschil tussen metingen en berekeningen is relatief klein voor hoge geluidbelasting en neemt met afnemende geluidbelasting toe. Een regressieanalyse van het verschil laat zien dat het verschil bij hoge geluidbelasting 1.5 dB L_{den} bedraagt en neemt tot circa 2.5 dB toe bij de laagste gemeten geluidbelasting.
- Op de NOMOS-meetposten van Schiphol wordt een zogeheten 'drempelwaarde' toegepast. In de meetdatabase van NOMOS worden alleen die vliegbewegingen geregistreerd waarvan het

maximale geluidniveau boven de drempelwaarde uitkomt. Door het gebruik van een drempelwaarde op de NOMOS-meetposten worden lage geluidniveaus minder vaak gemeten, terwijl deze geluidniveaus wel worden berekend. Dit heeft met name een effect op de vergelijking tussen meten en rekenen op meetposten die verder van de luchthaven en in gebieden met een lage geluidbelasting liggen. De vergelijking bij lage geluidbelasting is hierdoor minder goed mogelijk.

- De (medianen van de) verschillen op SEL-niveau na het toepassen van validatiefilters (aanvullende eisen aan wind, neerslag en elevatiehoek) liggen voor de meerderheid van meetposten binnen een bereik van -2 tot +4 dBA SEL. Dit geeft aan dat er ook op vluchtniveau in het algemeen een vrij goede overeenkomst tussen meten en rekenen te zien is.

Conclusies op basis van vergelijking op vluchtsoort- en vliegtuigtype-niveau

Bij de analyse op vluchtsoort-niveau zijn losse analyses uitgevoerd voor starts en landingen. De analyses laten zien dat het verschil tussen metingen en berekeningen bij hoge geluidbelasting voor beide vluchtsoorten klein is, met een iets groter verschil voor landingen dan voor starts. De analyse op vliegtuigtype-niveau laat zien dat er geen specifiek vliegtuigtype aan te merken is waarvoor het verschil tussen metingen en berekeningen structureel hoger uitkomt. De locatie waar de vergelijking wordt uitgevoerd lijkt bepalender te zijn voor de mate van de verschillen.

Effect van ontbrekende vliegtuigbewegingen op de vergelijking meten en rekenen

Om het effect van ontbrekende vliegtuigbewegingen door het gebruik van de drempelwaarde nader te analyseren, is de vergelijking tussen meten en rekenen gemaakt voor een dataset waarin geen koppeling is gemaakt tussen de gemeten en berekende vliegtuigbewegingen. De berekeningen in deze dataset zijn uitgevoerd voor alle vliegtuigbewegingen van en naar Schiphol in het gebruiksjaar 2019. De vergelijking van de geluidbelasting voor alle berekende vliegtuigbewegingen en alle gemeten vliegtuigbewegingen (zonder koppeling van beiden) laat zien dat vooral de berekende waarden bij lage geluidbelasting omhoog gaan indien alle vliegtuigbewegingen worden berekend. De gemeten waarden blijven daarbij onveranderd. De correlatie tussen meten en rekenen is voor deze dataset in het algemeen minder goed dan voor de gekoppelde dataset.

Vervolgens is een vergelijking gemaakt van de berekende geluidbelasting voor alle vliegtuigbewegingen en voor alleen gekoppelde vliegtuigbewegingen. Deze vergelijking laat zien dat het verschil tussen beiden tot een geluidbelasting van ongeveer 55 dB L_{den} verwaarloosbaar is, maar neemt met afnemende geluidbelasting duidelijk toe. Bij een geluidbelasting van 45 dB bedraagt het verschil tussen beide datasets gemiddeld 4 dB L_{den} . Deze constatering is op basis van alle meetposten bepaald, zowel validatie- als niet-validatiemeetposten, en bevestigt dat het NOMOS-meetsysteem een aanzienlijk aantal relevante vliegtuigbewegingen bij lage geluidbelasting niet kan meten.

Aanbevelingen

Toekomstige vergelijking meten en rekenen

Voor toekomstige vergelijkingen tussen berekeningen en metingen van vliegtuiggeluid kan op basis van het validatieonderzoek een aantal aanbevelingen worden gedaan:

- Bij vergelijkingen tussen meten en rekenen dient rekening gehouden te worden met het effect van de drempelwaarde op de gemeten geluidbelasting. Als blijkt dat op bepaalde locaties geen goede vergelijking gemaakt kan worden door de drempelwaarde, kan onderzocht worden of de drempelwaarde op specifieke locaties kan worden verlaagd. Hierbij is het van belang om vast te stellen in hoeverre de lagere drempelwaarde niet leidt tot minder nauwkeurige meetresultaten doordat metingen van lage geluidniveaus gevoeliger zijn voor stoorgeluid. Nog een mogelijkheid is het verplaatsen van een aantal meetposten in gebieden met een lage geluidbelasting onder of dichterbij de vliegroutes. Op deze manier komen ook meer metingen voor de vergelijking beschikbaar.
- Het wordt aanbevolen om de vergelijking in ieder geval voor het jaar 2022 en later te herhalen om het effect van een nieuw detectie algoritme voor metingen op de NOMOS-metposten te analyseren (Schiphol 2022). Indien door toepassing van dit algoritme meer (betrouwbare) data voor de vergelijking bij lage geluidbelasting beschikbaar komen, zal dat de statistische vergelijking tussen meten en rekenen versterken. Tevens zouden meer data bij lage geluidbelasting helpen om het toepassingsbereik van het model concreter te formuleren.

Ontwikkeling signaalfunctie

Het validatieonderzoek heeft de basis gelegd voor de ontwikkeling van het validatiesysteem (de zogenoemde signaalfunctie) waarmee de verschillen tussen meten en rekenen (structureel) gemonitord kunnen worden. Ook kan met dit systeem een eventuele toekomstige toename van de verschillen worden gesignaleerd. Het wordt aanbevolen om het werk uitgevoerd voor het validatieonderzoek te gebruiken voor de ontwikkeling van de signaalfunctie.

Uitvoering modelverbetering

In het voorliggende onderzoek zijn de oorzaken van het verschil tussen meten en rekenen niet onderzocht. Conform de PAMV-validatieaanpak dient bij verschillen tussen gemeten en berekende geluidbelasting die groter zijn dan 2 dB L_{den} gezocht te worden naar mogelijkheden voor modelverbetering of naar een verklaring voor het verschil in de meetwaarden. Het wordt aanbevolen om de verschillen in een vervolgotraject voor modelverbetering nader te onderzoeken.

1 Inleiding

1.1 Programmatische Aanpak Meten Vliegtuiggeluid (PAMV)

1.1.1 *Aanleiding PAMV*

Rondom Nederlandse luchthavens worden al vele jaren metingen van vliegtuiggeluid uitgevoerd. Op basis daarvan worden omwonenden geïnformeerd over de geluidniveaus en de geluidbelasting in hun omgeving. In de regelgeving en handhaving van de luchtvaart wordt géén gebruik gemaakt van (lokaal uitgevoerde) geluidmetingen.

De ondergeschikte rol van geluidmetingen in het luchtvaartdossier leidde in de afgelopen jaren tot veel discussie en maatschappelijke onvrede. Voor de minister van Infrastructuur en Waterstaat was dat aanleiding om een Programmatische Aanpak voor het Meten van Vliegtuiggeluid (PAMV) op te starten. Met de PAMV wil de minister zowel berekeningen als metingen van vliegtuiggeluid verbeteren en beide methoden onderling versterken. Gegevens over de geluidbelasting door vliegverkeer moeten in de nieuwe situatie voor iedereen herkenbaar en betrouwbaar zijn en een solide basis vormen voor informatie voor omwonenden en beleidsbeslissingen (ministerie van Infrastructuur en Waterstaat 2018).

Een belangrijke eerste stap bij het verbeteren van meet- en rekengegevens is om te bepalen hoe goed ze met elkaar overeenkomen en op welke punten er behoefte is aan verbetering van rekengegevens, respectievelijk meetgegevens. Door een validatie van het rekenmodel op basis van geschikte meetgegevens uit te voeren, kan vastgesteld worden of en in welk bereik de rekenresultaten betrouwbaar zijn en of er verbetering van het model wenselijk of nodig is om de overeenkomst tussen beiden te verbeteren. Een validatie van het rekenmodel kan een belangrijke bijdrage leveren aan het behalen van het overkoepelende doel van de PAMV – het terugwinnen van het vertrouwen van de omwonenden in de huidige meet- en rekensystematiek.

De PAMV is medio 2019 gestart met een verkenningsfase. De verkenningsfase is door de kennisinstututen RIVM, NLR en KNMI uitgevoerd. Hierbij zijn mogelijkheden onderzocht om het stelsel van meten, berekenen en beleven van vliegtuiggeluid te verbeteren. De resultaten van deze verkenning zijn in december 2019 gerapporteerd (Smetsers, Siegmund et al., 2019). In de beleidsreactie op dat rapport, van 18 december 2019, heeft de minister toegezegd dat zij alle aanbevelingen uit dat rapport zal uitvoeren (ministerie van Infrastructuur en Waterstaat 2019). Sinds medio 2020 zit de PAMV in de uitwerkfase, waarbij de diverse aanbevelingen worden uitgewerkt. Het wordt beoogd dat de uitwerkfase uiteindelijk zal leiden tot een structurele fase, waarbij een aantal activiteiten structureel wordt uitgevoerd. Hiermee kan vervolg worden gegeven aan de kennis en inzichten die de PAMV heeft opgeleverd en kan richting de omgeving worden aangetoond dat het doel niet alleen het terugwinnen van het vertrouwen in de meet- en rekengegevens was, maar ook het behouden van het vertrouwen.

1.1.2 *Aanbevelingen PAMV-verkenningsfase*

Het rapport van de verkenningsfase doet zeven aanbevelingen, die hier kort worden herhaald en samengevat. Voor meer informatie over de aanbevelingen en hoe ze tot stand kwamen, wordt verwezen naar het PAMV-verkenningsrapport uit 2019 (Smetsers, Siegmund et al., 2019). De aanbevelingen die daarin zijn gedaan, waren:

1. *Meetstrategie*
Ontwikkel een nationale meetstrategie, met visies over de verschillende meetdoelen. Leg per meetdoel vast waaraan metingen moeten voldoen. Werk de nationale meetstrategie op regionaal niveau uit.
2. *Signaalfunctie*
Ontwikkel een methodiek die op basis van meetgegevens kan signaleren of de resultaten van geluidberekeningen van de gewenste kwaliteit zijn.
3. *Modelontwikkeling*
 - a. Leg voor verschillende doeleinden de criteria vast waaraan modelberekeningen moeten voldoen. Start een structureel en langjarig modelvalidatieprogramma, met de focus op berekeningen volgens de Doc.29-methode (ECAC.CEAC 2016a). Gebruik zo mogelijk bestaande meetsets en pas de signaalfunctie methodiek toe om de kwaliteit van het rekenmodel te beoordelen.
 - b. Om inzicht te krijgen in de toepasbaarheid van berekeningen en metingen wordt aanbevolen om voor één of voor een aantal luchthavens de grenzen waarbinnen geluidberekeningen en metingen betrouwbaar uitgevoerd kunnen worden te analyseren.
4. *Publiekscommunicatie*
Verbeter de publiekscommunicatie over meten, rekenen, beleven en regelgeving van vliegtuiggeluid. Bied algemene informatie gecoördineerd aan, vanuit één goed vindbare weblocatie.
5. *Hinder en gezondheid*
Ga geluidhinder en slaapverstoring rond luchthavens systematisch monitoren volgens een wetenschappelijk verantwoorde standaardmethode. Gebruik deze gegevens onder meer om te onderzoeken of er aanvullende geluidindicatoren zijn die beter aansluiten bij de manier waarop omwonenden de effecten van vliegtuiggeluid ervaren.
6. *Citizen Science*
Betrek de omgeving op een gestructureerde manier bij het opzetten van een aanvullend meetprogramma. Faciliteer citizen science-projecten voor hoogrisicogroepen, zoals burgers die dicht bij een start- of landingsbaan wonen of een verhoogde gevoeligheid voor omgevingsgeluid hebben.
7. *Wetenschappelijke aansturing*
Zorg voor onafhankelijke deskundigheid rond vliegtuiggeluid om de kwaliteit van de structurele werkzaamheden en het verbetertraject te toetsen en te borgen.

De basis voor het werk dat is uitgevoerd voor de validatie in dit onderzoek werd gelegd bij de opvolging van aanbeveling 1, dat wil zeggen tijdens de ontwikkeling van de kaders van de nationale meetstrategie (Sahai 2021). In het rapport over de nationale

meetstrategie uit 2021 zijn kaders bepaald voor het meten van vliegtuiggeluid voor het validatiedoel. Alleen metingen van meetposten die geschikt zijn voor validatie dienen voor de vergelijking te worden gebruikt. Vervolgens heeft de regionale uitwerking van de nationale meetstrategie voor Schiphol de meetposten geïdentificeerd die voor validatie geschikt zijn en de meetdata die voor de vergelijking met rekenresultaten kunnen worden gebruikt.

Het validatierapport gaat in op aanbevelingen 2 en 3 voor de PAMV. Er wordt een begin gemaakt met het ontwikkelen van een methodiek waarmee berekende en gemeten gegevens met elkaar kunnen worden vergeleken en waarmee inzicht kan worden verkregen in waar de verbeterpunten potentieel liggen. Deze methodiek wordt vervolgens toegepast om de rekenresultaten van het Doc.29-model voor Schiphol te vergelijken met metingen rondom Schiphol. De in dit rapport uitgevoerde analyse wordt als de eerste stap gezien richting een structurele validatie die is aanbevolen in aanbeveling 3.

1.2 Validatie

1.2.1

Aanleiding en behoefte

Het gebruik van rekenmodellen speelt een belangrijke rol binnen de huidige systematiek rondom vliegtuiggeluid. Alle normen die binnen het Nederlandse beleid worden gehanteerd, zijn gebaseerd op berekende waarden. Het maken van jaarprognoses van de verwachte geluidbelasting en de handhaving van de geproduceerde geluidbelasting op basis van het gerealiseerde vliegverkeer wordt ook op basis van berekeningen gedaan. Hoewel de rekenmodellen zelf wel gebruikmaken van eerder uitgevoerde metingen, hebben actuele en lokaal uitgevoerde metingen geen rol. Dergelijke metingen worden momenteel voornamelijk voor informatievoorziening aan omwonenden gebruikt.

Tijdens de voor de PAMV uitgevoerde verkenning is vanuit de omgeving de duidelijke wens naar voren gekomen om metingen een belangrijkere rol te geven binnen het luchtvaartdossier. Een van de mogelijkheden om de rol van metingen te vergroten, is om metingen te gebruiken om de rekenmodellen te valideren (Smetsers, Siegmund et al. 2019). Hiermee kunnen de rekenresultaten aan de hand van metingen worden verbeterd en kunnen beide methoden onderling worden versterkt.

In de praktijk zullen er vrijwel altijd verschillen optreden tussen metingen en berekeningen. Dit komt doordat bij berekeningen aannames worden gedaan die kunnen leiden tot verschillen met de werkelijkheid. Ook is bij metingen sprake van meeton nauwkeurigheden en geluid afkomstig van andere bronnen dat ongemerkt meegenomen kan worden en aan het vliegtuiggeluid worden toegewezen. Alhoewel er altijd een verschil tussen metingen en berekeningen zal blijven, is het wel van belang dat het verschil binnen een acceptabele marge blijft. Als het verschil boven een acceptabele marge stijgt, dient nader onderzoek te worden gedaan naar de mogelijke verschillen. Hierbij is het dan belangrijk om te beseffen dat:

- zowel berekeningen als metingen nooit 100 procent nauwkeurig zullen zijn;

- de gewenste nauwkeurigheid van berekeningen en metingen afhangt van de toepassing waar ze voor worden gebruikt.

Om de verschillen tussen meet- en rekenresultaten te onderzoeken en te duiden is in het PAMV-verkenningrapport (Smetsers, Siegmund et al., 2019) aanbevolen om een systeem op te zetten dat met metingen controleert of de resultaten van geluidberekeningen van voldoende kwaliteit zijn. Hiermee kan onderzocht worden of de modelinvoergegevens een goede weergave van de werkelijkheid geven, maar ook of er verbeteringen doorgevoerd kunnen worden aan de geluidmodellering zelf. Door het toepassen van zo'n systeem en de uitkomsten daarvan openbaar en transparant te delen, zou het vertrouwen van burgers in de rekenresultaten verhoogd kunnen worden.

Tijdens de verkenning is verder aanbevolen om de focus van het validatietraject op het Doc.29-model te leggen, gezien dit model naast Schiphol op termijn ook voor de regionale luchthavens van nationale betekenis zal worden gebruikt. Er is op dit moment beperkt inzicht in hoe betrouwbaar het Doc.29-rekenmodel voor de Nederlandse situatie is. Met de Nederlandse situatie wordt niet alleen de intensieve luchtvaart operatie in Nederland bedoeld, maar ook de Nederlandse weersomstandigheden (bijvoorbeeld veel voorkomende harde wind en neerslag, combinatie van lage temperaturen en hoge luchtvochtigheid).

In 2018 is door het NLR een trendvalidatie uitgevoerd, waarbij de berekende geluidbelasting met gemeten geluidbelasting is vergeleken. Dit was gedaan voor acht meetposten van het NOMOS-meetnet rondom Schiphol voor een periode van tien maanden van het gebruiksjaar 2014-2015. De uitkomst van deze trendvalidatie was dat de berekende geluidbelasting en gemeten geluidbelasting in het algemeen dezelfde trend volgden.

De trendvalidatie was echter beperkt in scope en omvang en zou als een proof-of-concept kunnen worden beschouwd voor het validatiesysteem waarmee in het voorliggende onderzoek een begin wordt gemaakt. Een aantal beperkingen van de trendvalidatie waarop kan worden voortgebouwd is:

- Er werd tijdens de trendvalidatie slechts beperkt gekeken naar de exacte oorzaken van het verschil en of de trends ook voor de resterende 33 NOMOS-meetposten gelden.
- Omdat de kaders van de nationale meetstrategie (Sahai 2021) in 2018 nog niet bepaald waren, werd er geen onderscheid gemaakt tussen meetposten die wel of niet geschikt zijn voor validatie.
- Er werd geen uitvoerige statistische analyse uitgevoerd om de overeenkomst tussen de berekende en gemeten geluidbelasting uit te drukken. De spreiding in het verschil tussen berekende en gemeten waarden werd hierbij niet in kaart gebracht.
- De invloed van versturende factoren, zoals wind, neerslag en elevatiehoek, op het verschil tussen meten en rekenen werd niet in detail geanalyseerd (alleen indicatief voor wind).
- Het was op basis van de trendvalidatie niet vast te stellen of de positieve trend tussen berekeningen en metingen ook voor de jaren na de onderzochte periode zal gelden.

- Het werd niet bepaald wanneer het model als van voldoende kwaliteit kan worden beschouwd, dat wil zeggen welk verschil acceptabel is en wanneer een verbetering van het model aan de orde is.

Door onder andere de bovengenoemde redenen werd tijdens de PAMV-verkenningfase aanbevolen om de trendvalidatie als basis te gebruiken voor de ontwikkeling van het validatiesysteem waarmee (op termijn) een structurele validatie kan worden uitgevoerd. Onderdeel van de ontwikkeling van dit systeem zou het bepalen van criteria zijn om de kwaliteit van de rekenresultaten te beoordelen en om te bepalen wanneer nader onderzoek naar modelverbetering nodig of wenselijk is. Deze criteria kunnen gebaseerd zijn op verschillen in geluidniveaus of geluidbelastingniveaus, maar ook bijvoorbeeld op statistische correlatiecoëfficiënten tussen berekende en gemeten data.

De term 'structureel' in structurele validatie houdt een herhaling in van het proces van vergelijken van rekenresultaten met metingen met een bepaalde frequentie. Er kunnen diverse aanleidingen zijn om een validatie opnieuw uit te voeren, zoals een aanpassing van de rekenmethodiek of actualisatie van de geluidtabellen op Europees niveau, operationele veranderingen bij een luchthaven, introductie van nieuwe vliegtuigtypes in de vloot, behoefte om inzicht te verkrijgen in de rekenkwaliteit over tijd, et cetera. De frequentie waarmee de validatie en eventuele modelverbetering zal worden uitgevoerd, dient nog te worden bepaald. In het Verenigde Koninkrijk (VK) wordt bijvoorbeeld door de CAA-UK (Britse civiele luchtvaartautoriteit) jaarlijks een vergelijking van het ANCON-model met metingen uitgevoerd, al is daar meer sprake van kalibratie van het model dan van validatie. Voor andere sectoren in Nederland, zoals rail- en wegverkeer, wordt een jaarlijkse vergelijking van berekende en gemeten geluidbelasting voor de geluidmonitor gemaakt (Joosten 2020). Een jaarlijkse vergelijking zou de luchtvaart dan meer in lijn brengen met de andere transportsectoren. Afhankelijk van de beleidsbehoefte kunnen voor de validatiefrequentie ook andere keuzes worden gemaakt.

Het kan ook besloten worden om het beoogde validatiesysteem te automatiseren, waarbij het verschil tussen meten en rekenen automatisch kan worden gemonitord. Indien het verschil groter wordt dan een vooraf bepaalde waarde of de correlatie slechter wordt dan voorgesteld in de validatiecriteria, kan nader onderzoek worden uitgevoerd naar de oorzaak van het verschil om het model eventueel te verbeteren. Op hoofdlijnen kunnen er drie mogelijke uitkomsten zijn indien verschillen tussen meten en rekenen nader onderzocht worden:

- De invoergegevens van het rekenmodel (deels Nederlandse, deels Europese invloedsfeer) en/of de rekenmethodiek (Europese invloedsfeer) kunnen worden aangepast om de verschillen te verkleinen.
- Er wordt geen duidelijke oorzaak gevonden of er wordt geconstateerd dat het verkleinen van de verschillen in een specifiek geval niet nodig, wenselijk of mogelijk is.
- Er wordt geconstateerd dat in bepaalde gevallen aanpassingen van het vlieggedrag wenselijk zijn, waardoor de operaties worden aangepast (dit is alleen aan de orde daar waar dit mogelijk is).

Het doen van aanpassingen aan het rekenmodel heeft gevolgen voor beleid en mogelijk voor de mate van bescherming van omwonenden. Deze gevolgen dienen goed meegewogen te worden bij het kiezen voor een validatie en een eventuele verbetering van het model. Indien het model of de invoer daarvan verandert, betekent dit meestal dat zowel de jaarlijks berekende geluidbelasting verandert als de methode waarmee de normen worden vastgesteld. Voordat gestart wordt met het structureel verbeteren van het model, dient daarom ook vast te staan wat dit betekent voor de handhaving.

1.2.2 *Relevante informatie uit de literatuur*

Voordat met de uitvoering van de validatie is begonnen, is door het PAMV-consortium een literatuuronderzoek uitgevoerd wat betreft de validatiescope, aanvullende meeteisen voor de validatie en een mogelijke streefwaarde voor een acceptabel verschil tussen meten en rekenen. De uitkomsten van dit literatuuronderzoek staan hieronder kort samengevat.

1.2.2.1 Validatiescope

De Europese Richtlijn 2015/996 (EU 2015) raadt het gebruik van rekenmodellen van geluidbelasting aan die gebaseerd zijn op de ECAC Doc.29-methode (ECAC.CEAC 2016a). Doc.29 is een in internationale samenwerking ontwikkeld model voor het berekenen van vliegtuiggeluidbelasting en baseert zich op de meest actuele inzichten over het modelleren van vliegtuiggeluid. Ook de ICAO via Doc.9911 (ICAO 2018) raadt het gebruik van methoden aan die grotendeels overeenkomen met de ECAC Doc.29-methodologie. Als gevolg hiervan, worden ook buiten Europa rekenmodellen voor vliegtuiggeluid gebruikt die op vergelijkbare richtlijnen zijn gebaseerd en een vergelijkbare globale scope kennen. Enkele voorbeelden van modellen die gebaseerd zijn op Doc.29 zijn het ANCON-model in het Verenigd Koninkrijk, AEDT in de Verenigde Staten (vroeger INM (FAA 2008)) en het door EUROCONTROL ontwikkelde IMPACT-model. Aangezien er weinig referenties bekend zijn over een validatie van Doc.29 zelf, richtte het literatuuronderzoek zich ook op de validatie van vergelijkbare modellen als Doc.29.

Voor het vaststellen van de scope van de validatiestudie is onder andere gekeken naar de geluidmaten die voor validatie worden gebruikt en welke mogelijkheden tot modelverbetering door de diverse organisaties worden toegepast.

1. *Geluidmaten*: De geluidmaten L_{den} en L_{night} zijn beleidsmatig gezien de belangrijkste maten voor de uiteindelijke validatie van een rekenmodel zoals Doc.29. Een vergelijking alleen op L_{den} -niveau heeft echter als beperking dat verschillen voor bepaalde vliegtuigtypes door het optellen van alle vliegtuigbewegingen mogelijk niet gevonden worden (Trow 2018). Een vergelijking van gemeten en berekende geluidniveaus per vliegtuigtype helpt hierbij en kan waardevolle inzichten leveren. Zo kunnen er bijvoorbeeld afwijkende vliegtuigtypes en de daaraan gerelateerde vliegprocedures voor nadere analyse en verbetering worden geïdentificeerd. Voor een dergelijke vergelijking wordt de geluidmaat SEL (Sound Exposure Level) het vaakst gebruikt, met een beperkt aantal studies die ook zijn uitgevoerd voor het

maximale geluidniveau L_{Amax} naast SEL (Giladi 2020), (BAC-Australia 2005), (Rhodes 2002). Om vervolgens weer een koppeling te maken met de L_{den} - en L_{night} -geluidmaten worden de SEL-waarden van alle vliegtuigbewegingen gedurende het jaar opgeteld (Rhodes D. P. 2018).

2. *Mogelijkheden tot modelverbetering*: Er zijn verschillende mogelijkheden om het rekenmodel aan te passen en tot een betere overeenkomst tussen gemeten en berekende waarden te komen:
 - Aanpassing van (specifieke) invoergegevens¹.
 - Aanpassing van de geluidmodellering.
 - Aanpassing van de onderliggende geluid- (Noise-Power-Distance, NPD) tabellen.

Het PAMV-verkenningrapport (Smetsers, Siegmund et al. 2019) adviseert om eerst met het verbeteren van invoergegevens te beginnen alvorens onderzoek te doen naar mogelijke verbeteringen van de geluidmodellering zelf. Uit het literatuuronderzoek blijkt dat ook in andere studies een vergelijkbare aanpak voor modelvalidatie wordt gevolgd. De meeste studies beginnen met een aanpassing van invoergegevens, zoals de gemodelleerde vliegprocedures en vliegroutes om deze beter overeen te laten komen met de in de praktijk gevlogen procedures en routes. Een aantal parameters zoals het startgewicht van het vliegtuig en de motorsetting moet dan alsnog worden geschat. Daarnaast heeft het moment waarop de kleppen en het landingsgestel worden in- of uitgeklaapt invloed op de vliegprestaties en dus op de geluidproductie van een vliegtuig.

Alhoewel het aanpassen en verbeteren van model-invoergegevens ertoe kan leiden dat een rekenmodel meer representatief wordt voor de praktijk, zullen ook na aanpassing van de invoergegevens verschillen blijven bestaan tussen gemeten en berekende waarden. Een andere mogelijke methode om het verschil te verkleinen is het aanpassen van de geluidtabellen (ook wel NPD-tabellen genoemd). Hierbij worden geluidniveaus voor specifieke motorsettings via een decibelcorrectie aangepast om de berekende geluidniveaus dichter bij gemeten geluidniveaus te brengen. Het verkenningrapport van de PAMV raadt deze aanpak af, maar een dergelijke correctie is wel door diverse organisaties in het buitenland toegepast (Trow 2018), (BAC-Australia 2005), (Rhodes 2002). Deze aanpak leidt uiteindelijk tot een kleiner verschil tussen gemeten en berekende waarden, maar kent het nadeel dat de resultaten na toepassing van een decibelcorrectie alleen representatief zijn voor de meetlocaties waar werd gemeten. Ook kunnen deze correcties per luchthaven verschillen en worden eventuele afwijkingen in de metingen overgenomen in de geluidberekening.

1.2.2.2 Aanvullende meeteisen

Weerscondities kunnen een verstrend effect hebben op het gemeten geluid. De aanwezigheid van neerslag, sneeuw of wind van hoge snelheden kan grote invloed hebben op de gemeten waarde van vliegtuiggeluid. In sommige gevallen kunnen de betreffende metingen

¹ Met *specifieke invoergegevens* wordt bedoeld op invoergegevens die specifiek zijn voor een luchthaven, zoals het gebruik of de temperatuur. Daarnaast heb je de *algemene invoergegevens* zoals de geluidtabellen.

door ongunstige weersomstandigheden ongeschikt zijn voor het vergelijken met berekende waarden. Sinds de jaren zeventig is er expliciet aandacht voor de invloed van weerscondities op het meten van vliegtuiggeluid, onder andere via ISO 3891 (Noise 1978). In deze norm werden voor de eerste keer eisen aan de weerscondities gesteld waarbij metingen van vliegtuiggeluid werden uitgesloten als ze niet aan deze eisen voldeden. Deze weerscondities waren:

- Aanwezigheid van neerslag.
- Aanwezigheid van wind met een snelheid van 5 m/s (10 knopen) of hoger.
- Een combinatie van luchtvochtigheid en temperatuur die tot een demping van 10 dB/100 m of hoger leidt in de derde-octafband rondom 8 kHz.

Deze norm is vervolgens in 2009 vervangen door ISO 20906 voor het onbemand meten van vliegtuiggeluid (ISO-20906/TC 2009), waarbij de eisen over de weerscondities minder streng zijn gemaakt. De nieuwe norm beveelt aan om metingen van vliegtuiggeluid uit te sluiten:

- Bij aanwezigheid van neerslag.
- Bij aanwezigheid van wind met een snelheid van 10 m/s (20 knopen) of hoger.

De eis die rekening houdt met de hoeveelheid demping bij een gegeven luchtvochtigheid en temperatuur is daarmee komen te vervallen. Ook laat de nieuwe norm hogere windsnelheden toe tot 10 m/s. In de trendvalidatie Doc.29 voor Schiphol (Hogehuis and Heblj 2018) is ook gekeken naar het effect van wind. Op basis van een steekproef, waarbij opnames bij harde wind werden teruggeluisterd, bleek dat in de meeste gevallen de harde wind ervoor zorgde dat het gemeten geluidniveau veel hoger werd. In sommige gevallen was het vliegtuig zelfs helemaal niet hoorbaar. Vervolgens werden analyses uitgevoerd naar het effect van het wel of niet meenemen van meetgegevens die bij een windsnelheid van boven de 10 m/s zijn gemeten. Hierbij bleek dat de verschillen tussen gemeten en berekende gegevens kleiner werden op het moment dat metingen bij harde wind niet werden meegenomen. Dit effect speelde in de onderzochte periode met name in de wintermaanden. Ook bleek te correlatie tussen de gemeten en berekende waarden na het uitsluiten van de betreffende metingen hoger.

Naast meetcondities die te maken hebben met het weer zijn er ook andere meetcondities waarmee rekening dient te worden gehouden (denk bijvoorbeeld aan achtergrondgeluid, afscherming en reflecties door objecten et cetera). Hierbij wordt in de literatuur aanbevolen om zoveel mogelijk aan te sluiten bij de ISO-norm 20906 (ISO-20906/TC 2009), die ook voor de ontwikkeling van de PAMV nationale meetstrategie als basis heeft gediend (Sahai 2021).

- 1.2.2.3 Effect van verschillende paramaters op het gemiddelde verschil tussen meten en rekenen en de spreiding in het verschil
In principe kunnen alle gemodelleerde invoergegevens die niet overeenkomen met waarden uit de praktijk tot een verschil leiden tussen metingen en berekeningen (naast een aantal andere factoren).

In de VS heeft het bedrijf Wyle voor de luchthaven Denver International Airport in 1998 een validatie uitgevoerd voor het verbeteren van het rekenmodel INM (Plotkin 2013). Zoals genoemd in paragraaf 1.2.2.1, is het INM-model gebaseerd op een vergelijkbare methodologie als Doc.29. Voor deze studie zijn starts gemeten voor de vliegtuigtypes Boeing 727, 737 en 757 vergeleken met gemodelleerde starts voor dezelfde vliegtuigtypes. In totaal is de vergelijking gemaakt voor 635 vliegtuigbewegingen op een aantal meetlocaties recht onder de vliegroutes. Vervolgens is gekeken of het verschil tussen gemeten en berekende SEL-waarden toeneemt of afneemt bij een variatie van invoergegevens die relevant zijn tijdens het starten. De parameters die in de studie aan bod komen zijn aangegeven in de eerste kolom van Tabel 1.

Tabel 1 Invloed van toename van verschillende parameters op het gemiddelde verschil tussen gemeten en berekende SEL-waarden en op de spreiding in dit verschil uit (Plotkin 2013).

Parameter	Effect van toename parameter op <i>gemiddeld verschil</i> tussen gemeten en berekende SEL	Effect van toename parameter op <i>spreiding</i> in verschil tussen gemeten en berekende SEL
Afstand tot meetpost	↑	↑
Afstand tot de luchthaven	↑	— (Geen trend te zien)
Vlieghoogte	↑	— (Geen trend te zien)
Vliegsnelheid	↑	— (Geen trend te zien)
Startgewicht (Weight-margin ²)	— (Geen trend te zien)	— (Geen trend te zien)
Elevatiehoek	↓	↓ ³
Luchttemperatuur	↓	— (Geen trend te zien)
Relatieve luchtvochtigheid	↑	— (Geen trend te zien)
Luchtdruk	↑	— (Geen trend te zien)
Windsnelheid	↑	— (Geen trend te zien)

In Tabel 1 wordt het effect samengevat van een toename van de betreffende parameter op het *gemiddeld verschil* tussen gemeten en berekende SEL-waarden en op de *spreiding* in het verschil. Een goed werkend en gevalideerd model zal niet alleen een beperkt gemiddeld verschil tussen gemeten en berekende waarden laten zien, maar zal ook een lage spreiding in het verschil geven bij een vergelijking met metingen.

Hieronder staat een korte toelichting op de trends uit Tabel 1:

- Volgens de analyse voor luchthaven Denver wordt het verschil in SEL-waarden groter naarmate de afstand tussen het vliegtuig en de meetpost toeneemt. Bij grote afstanden tussen het vliegtuig en de meetpost wordt de kans groot dat wind en andere weersomstandigheden niet constant blijven over de

² Weight margin = (Maximaal toelaatbaar startgewicht – startgewicht)/startgewicht.

³ Lagere elevatiehoeken leiden tot minder betrouwbare meetresultaten en daarmee tot grotere verschillen tussen gemeten en berekende SEL-niveaus.

atmosferische lagen tussen het vliegtuig en de meetpost. Hierdoor kunnen diffractie- en refractie-effecten optreden, die het pad van het geluid buigen en tot andere geluidsniveaus leiden dan wordt verwacht bij constante weerscondities. Hiernaast kunnen grote afstanden gepaard gaan met lage waarden van de elevatiehoek, die ook tot onzekerheden leiden in de laterale spreiding van het geluid.

- In Tabel 1 is ook te zien dat op grote afstanden van het vliegtuig tot de luchthaven (ongeveer 20 km en verder) het gemiddelde verschil toeneemt. Op grotere afstanden van de luchthaven is het vliegtuig meestal op grotere vlieghoogtes en vliegt met een hogere snelheid.
- Zowel het gemiddelde verschil als de spreiding nemen bij lage elevatiehoeken toe in vergelijking met hoge elevatiehoeken. Deze constatering komt overeen met de aanbeveling uit ISO 20906 en van de Britse luchtvaartautoriteit CAA om alleen die vliegtuigbewegingen te gebruiken die binnen een elevatiehoek van minimaal 45 of minimaal 60 graden vallen.
- In Tabel 1 is ook het effect van meteorologische invoergegevens te zien op het verschil tussen gemeten en berekende waarden. Een groter verschil wordt gerapporteerd bij hoge relatieve luchtvochtigheid waarden (vanaf circa 70%) en hogere windsnelheden (vanaf ongeveer 15 knopen, oftewel 7.7 m/s). Ook worden grotere verschillen geconstateerd voor hoge luchtdruk waarden en bij lage luchttemperaturen.
- Voor de afzonderlijke invloed van het startgewicht op het verschil tussen metingen en berekeningen zijn in de studie geen duidelijke trends gevonden. Wel is uit andere literatuur bekend dat het gemodelleerde startgewicht (voor starts) tot grote verschillen met gemeten geluidsniveaus kan leiden indien het startgewicht niet correct is geschat. Ook in de trendvalidatie voor Schiphol (Hogehuis and Heblj 2018) wordt het modelleren van het startgewicht als een mogelijke optie voor modelverbetering aangemerkt.

De trends in Tabel 1 zijn vooral representatief voor de vliegtuigen, vliegprocedures en vliegroutes die voor de Wyle-studie in Denver zijn geanalyseerd. Het is niet zeker dat dezelfde trends voor Schiphol of een andere Nederlandse luchthaven zullen worden geconstateerd. Zo verschilt de luchthaven van Denver van Nederlandse luchthavens doordat deze ver boven zeeniveau ligt en omdat het klimaat verschilt van het Nederlandse klimaat. Daarnaast is gebruikgemaakt van een ouder geluidmodel dan de Doc.29-implementatie voor Schiphol. Voor de situatie in Nederland is daarom nader onderzoek nodig. Daarnaast is de aangegeven studie alleen voor starts uitgevoerd, voor landingen is er op dit moment geen vergelijkbare studie aan het PAMV-consortium bekend.

- 1.2.2.4 Streefwaarde voor het verschil tussen gemeten en berekende gegevens
- In de literatuur is gezocht naar richtlijnen voor welke verschillen tussen gemeten en berekende gegevens als 'acceptabel' kunnen worden beschouwd. Er worden in de literatuur geen concrete grenzen gesteld aan de definitie van een 'acceptabel' verschil tussen gemeten en berekende geluidsniveaus of geluidbelasting. Zoals ook eerder toegelicht, kunnen voor individuele vluchten grote verschillen in geluidsniveaus

ontstaan die door diverse redenen niet te vermijden zijn (Giladi 2020), (Simons, Snellen et al. 2015), (Merino-Martínez, Heblíj et al. 2019). De Europese Richtlijn 2015/996 stelt voor alle model-invoergegevens met een nauwkeurigheid te bepalen waarmee de *geluidemissie* binnen een onzekerheidsmarge van ± 2 dBA kan worden berekend. Op basis van deze richtlijn is in het Verenigd Koninkrijk een richtwaarde gebruikt van ± 2 dBA SEL voor de spreiding tussen gemeten en berekende geluidniveaus (Trow 2018). Aangezien de validatie voor de PAMV voor de L_{den} -geluidmaat (dat wil zeggen geluidbelasting) zal worden uitgevoerd, is ook naar andere sectoren gekeken waar eveneens een validatie wordt uitgevoerd, bijvoorbeeld weg- en spoorverkeer. In de geluidmonitor voor de jaren 2017-2019 (Joosten 2020) is een gemiddeld verschil van 2 dB L_{den} tussen berekende en gemeten geluidbelasting voor weg- en spoorverkeer gehaald, gemiddeld over alle meetposten. Deze waarde van 2 dB L_{den} is gemiddeld en wordt niet op alle meetlocaties gehaald, dat wil zeggen op sommige meetlocaties het verschil (veel) hoger ligt en op sommige meetposten lager.

Door gebrek aan andere referentiewaarden en om de validatie voor de PAMV consistent te houden met andere sectoren, zal voor de validatie in de PAMV een maximaal verschil van 2 dB L_{den} als een *initiële* streefwaarde worden beschouwd. Hierbij zal het streven zijn om het verschil per meetpost onder 2 dB L_{den} te houden. Indien de verschillen tussen gemeten en berekende geluidbelasting bij een gegeven meetpost groter zijn dan 2 dB L_{den} , zal gezocht moeten worden naar mogelijkheden voor modelverbetering of naar een verklaring voor het verschil in de meetwaarden. Op basis van bevindingen in de PAMV in de uitwerkfase kan de streefwaarde worden bijgesteld.

1.2.3 PAMV-validatieaanpak

Het PAMV-consortium heeft vervolgens onder andere op basis van het literatuuronderzoek een validatieaanpak bepaald. Hierbij is voorgesteld om de validatie van het Doc.29-model voor Schiphol in twee fasen uit te voeren – een validatie opstartfase en een uitwerkfase. De opstartfase is in 2022 uitgevoerd en heeft medio 2022 tot een validatieaanpak geleid (Sahai 2022). Deze aanpak is door de onafhankelijke PAMV-expertgroep bekeken en de expertgroep heeft in het derde kwartaal van 2022 een advies over de voorgestelde aanpak uitgebracht. Dit advies is door het consortium in de aanpak grotendeels verwerkt. Vervolgens is begin 2023 werk begonnen aan de validatie-uitwerkfase dat tot het huidige rapport heeft geleid. Voor de validatieaanpak zijn criteria bepaald die voor de validatie gehanteerd moeten worden en de stappen die voor de validatie gevolgd dienen te worden. De analyse die hoofdstuk 2 van dit rapport wordt gepresenteerd, volgt grotendeels de stappen die in de validatieaanpak zijn geformuleerd. De belangrijkste punten uit de validatieaanpak worden hieronder samengevat.

Validatiescope

Op basis van de uitgevoerde literatuurstudie en eerdere ervaring van het consortium met validatie van rekenmodellen is de volgende scope voor de validatie voorgesteld:

- De validatie richt zich op geluidberekeningen volgens de ECAC Doc.29 (4th edition)-methode, specifiek de implementatie daarvan voor Schiphol.

- De validatie richt zich op het verbeteren van (specifieke) invoergegevens van geluidberekeningen (dat wil zeggen niet de modellering zelf).
- De validatie richt zich op de L_{den} -geluidmaat. Wel zal ook een analyse op SEL-niveau worden gepresenteerd om inzicht te krijgen in de L_{den} -verschillen.
- De validatie ondersteunt het onderzoek naar het toepassingsbereik van berekeningen en metingen. In paragraaf 1.2.4.1 staat een toelichting op de overlappen tussen het validatieonderzoek en het onderzoek naar het toepassingsbereik.

Aanvullende meeteisen voor validatie

De voor validatie gebruikte meetgegevens moeten zijn verzameld met voor validatie geschikte meetsystemen en op locaties die voldoen aan de validatie-eisen, zoals eerder vastgelegd in de PAMV nationale meetstrategie (Sahai 2021). Er is voor de validatieaanpak naast de eisen aan de meetlocatie ook over aanvullende meeteisen nagedacht (zie 1.2.2.2). Onder andere op basis van de literatuur en de eerdere ervaring van het consortium wordt hierbij voorgesteld om metingen van dergelijke systemen alleen te gebruiken voor validatie op het moment dat deze metingen zijn uitgevoerd onder de volgende condities:

- Windsnelheid: 5 m/s of minder, tenzij er voor de specifieke meetpost ruimere limieten zijn vastgesteld.
- Neerslag: geen neerslag.
- Elevatiehoek: 60 graden of meer.

Validatiecriteria en aanvullende overwegingen voor de validatie

Voor de validatieaanpak is een streefwaarde voorgesteld voor het verschil tussen meten en rekenen van 2 dB L_{den} per meetpost. Deze waarde wordt hierbij als een initiële streefwaarde beschouwd. Bij verschillen tussen gemeten en berekende L_{den} -waarden die groter zijn dan 2 dB L_{den} moet gezocht worden naar mogelijkheden voor modelverbetering of naar een verklaring voor het verschil in de meetwaarden. Op basis van de uitkomsten van de validatie kan deze streefwaarde, waar nodig, worden bijgesteld.

Hiernaast is in de validatieaanpak ook een aantal aanvullende overwegingen voor de validatie beschreven:

1. Ligging van een meetlocatie ten opzichte van de luchthaven: Hiervoor wordt geadviseerd om voor de validatie zowel meetposten dichtbij als verder weg van de luchthaven te gebruiken. Ook de kaders van de nationale meetstrategie geven een vergelijkbare aanbeveling met de eis aan de dekkinggraad van het meetstelsel. Tijdens de validatie-uitwerkfase kan onderzocht worden hoe de resultaten op meetposten dichtbij en verder weg onderling verschillen en wat dit voor het toepassingsbereik van het model betekent.
2. Percentage vliegtuigpassages dat gemeten kan worden: Het kan voorkomen dat een deel van de vliegtuigpassages door een meetpost niet gemeten kan worden. Dit kan samenhangen met situaties waarin geluidniveaus van vliegtuigen niet boven de drempelwaarde van een meetpost uitkomen. Als een aanzienlijk deel van de vliegtuigpassages door de meetpost niet kan worden gemeten, worden vooral de luidste geluidniveaus gemeten en

lagere geluidniveaus niet. Dit kan een invloed hebben op het geconstateerde verschil tussen gemeten en berekende waarden. Indien blijkt dat het verschil tussen gemeten en berekende geluidbelasting meer dan 2 dB bedraagt, kan gecontroleerd worden in hoeverre dit mede komt doordat een relevant deel van de passages niet kon worden gemeten.

3. De beschouwde validatieperiode:
Om te zorgen dat voldoende variaties in weersgegevens meegenomen worden, is het wenselijk om gegevens te gebruiken van minstens één en bij voorkeur twee of meer gehele jaren. In de literatuur wordt aanbevolen om de validatie voor tenminste een periode van twee jaar uit te voeren (Rhodes D. P. 2018), wat tot een robuustere validatie van het rekenmodel leidt. Als gegevens van één jaar beschikbaar zijn of worden gebruikt, kan op een later moment de validatie alsnog voor meerdere jaren worden uitgevoerd.
4. Het beschouwde aantal metingen:
Een laag aantal metingen kan ertoe leiden dat een klein aantal afwijkende gemeten of berekende geluidniveaus resulteren in een overschrijding van het criterium van 2 dBA. Zodra dit aan de orde is, kan bij het onderzoek naar de oorzaken van verschillen tussen metingen en berekeningen vastgesteld worden of het relatief lage aantal beschikbare metingen een mogelijke oorzaak van dit verschil is. Het analyseren van de oorzaken op zichzelf zal ook worden bemoeilijkt als voor de te onderzoeken situaties weinig metingen beschikbaar zijn. Als in de praktijk blijkt dat dit regelmatig voorkomt of zodra een validatie uitgevoerd gaat worden op een luchthaven met een beperkt aantal vliegtuigbewegingen, kan overwogen worden om een richtlijn op te stellen voor een minimaal aantal metingen dat nodig is om een goede validatie uit te kunnen voeren.

Veel van de hierboven genoemde overwegingen komen in hoofdstuk 2 van het voorliggende rapport terug.

1.2.4 *Relatie met andere onderzoeken*

In de volgende paragrafen staat een toelichting op de link van het voorliggende validatieonderzoek met andere gerelateerde onderzoeken.

1.2.4.1 PAMV-toepassingsbereik metingen en berekeningen

Het onderzoek naar het toepassingsbereik van metingen en berekeningen had als doel om de grenzen te bepalen waarbinnen metingen en berekeningen van vliegtuiggeluid betrouwbaar kunnen worden uitgevoerd. Voor het bepalen van het toepassingsbereik van berekeningen is eerst een gevoeligheidsanalyse uitgevoerd van de modelinvoergegevens. Vervolgens is een vergelijking gemaakt van berekende en gemeten geluidbelasting voor Schiphol. Het verschil hiertussen is vervolgens als functie van afstand tot de luchthaven en geluidbelasting per meetpost geanalyseerd.

Delen van het onderzoek naar het toepassingsbereik van metingen en berekeningen zijn op dezelfde vergelijking van berekende en gemeten geluidbelasting gebaseerd als het validatieonderzoek. De vergelijking voor het toepassingsbereik-onderzoek was nodig om de

onderzoeksvraag te beantwoorden en te weten tot waar het model nog betrouwbare resultaten geeft. Hiermee is onderzocht of het verschil tussen meten en rekenen verder weg van de luchthaven en in gebieden met een lage geluidbelasting toeneemt en tot waar de rekenmodellen nog een goede overeenstemming met gemeten waarden laten zien. Het punt vanaf waar het verschil zodanig toeneemt dat het niet meer als acceptabel kan worden beschouwd (meer dan 2 dB L_{den}), bepaalt in principe het toepassingsbereik van het rekenmodel. Een vergelijking van metingen en berekeningen is dus voor beide onderzoeken nodig geweest om tot de benodigde antwoorden op de onderzoeksvragen te komen.

Voor het toepassingsbereik-onderzoek is verder het effect van versturende factoren zoals wind, neerslag en lage elevatiehoek op het verschil tussen meten en rekenen in kaart gebracht. Inzicht in deze aspecten is direct ook voor het validatieonderzoek van belang, omdat het duidelijk maakt of de versturende factoren de vergelijking van berekeningen met metingen verslechteren.

Hiernaast zijn de volgende punten niet in het toepassingsbereik onderzoek geadresseerd en worden in het voorliggende rapport over validatie Doc.29 voor Schiphol gerapporteerd:

- De vergelijking van de geluidbelasting berekend met het Doc.29-model en de gemeten geluidbelasting op de NOMOS-metposten wordt in dit rapport uitgebreid met een meer uitvoerige statistische analyse.
- De analyse wordt verder verdiept door de vergelijking ook op vliegtuigtype-niveau te presenteren naast de analyse op vluchtsoort-niveau.
- Dit onderzoek kijkt ook naar het effect op de vergelijking tussen metingen en berekeningen van het gebruik van de drempelwaarde en de vliegtuigbewegingen die daardoor in het meetdataset ontbreken. Dit effect wordt geanalyseerd door een vergelijking te maken van de berekende geluidbelasting als alle vliegbewegingen van en naar Schiphol worden meegenomen versus de berekende geluidbelasting die alleen op basis van vliegbewegingen die met een gemeten vlucht kon worden gekoppeld is bepaald.
- Er wordt voor het validatieonderzoek ook een vergelijking op SEL-niveau per meetpost gepresenteerd voor het verkrijgen van meer inzicht in de verschillen waargenomen op L_{den} -niveau.

1.2.4.2 Ontwikkeling signaalfunctie

Om invulling te geven aan aanbevelingen 2 en 3 van het PAMV-verkenningsrapport is op langere termijn het doel om toe te werken naar een (geautomatiseerd) validatiesysteem. Dit systeem zal signaleren in hoeverre dusdanige verschillen tussen gemeten en berekende gegevens optreden die aanleiding geven voor nader onderzoek en eventuele modelverbetering. Het validatiesysteem zal dus een 'signaalfunctie' rol vervullen en kan worden opgezet op basis van de in het voorliggende validatieonderzoek gepresenteerde validatiemethodiek. In het voorliggende onderzoek is een begin gemaakt met de ontwikkeling van zo'n validatiesysteem. Verdere ontwikkeling van de signaalfunctie is in een vervolgotraject beoogd. Het consortium heeft voor het ministerie van IenW over de signaalfunctie en de

benodigde validatiefrequentie een apart advies opgesteld, zodat hierover een besluit kan worden genomen.

1.3 Opdrachtverlening

Dit validatieonderzoek is uitgevoerd als onderdeel van de uitwerkfase van de PAMV. Het ministerie van IenW heeft in 2022 het consortium van RIVM, NLR en To70 binnen de PAMV de opdracht gegeven om een validatie uit te voeren van het Doc.29-rekenmodel voor Schiphol. Medio 2022 is de validatieaanpak bepaald en vervolgens is het consortium begin 2023 met de uitvoering van de validatie aan de slag gegaan met als eerste stap het verzamelen van de reken- en meetgegevens. Hierna is de vergelijking uitgevoerd conform de voorgestelde validatieaanpak. In de zomer van 2023 zijn de eerste resultaten van de validatie door het consortium gereed gemaakt.

1.4 Leeswijzer

Hoofdstuk 2 van dit rapport presenteert de vergelijking van met het DOC.29-model uitgevoerde berekeningen met metingen die zijn uitgevoerd op de NOMOS-metposten rondom Schiphol. Hierbij beschrijft paragraaf 2.2 eerst hoe en welke meet- en rekengegevens zijn verzameld. Daarna worden in paragraaf 2.3 de resultaten van de vergelijking gepresenteerd en geanalyseerd. De analyse is eerst gepresenteerd voor de gekoppelde dataset waarbij iedere gemeten vlucht met een berekende vlucht is gekoppeld. Het effect van de drempelwaarde is hierna onderzocht en hoe dit de vergelijking van berekeningen met metingen heeft beïnvloed. Hierna is de analyse verder verdiept tot vluchtsoort-niveau en vliegtuigtype-niveau met een beschrijving van de meest relevante bevindingen. Als laatste is een verdiepende regressieanalyse gepresenteerd om te onderzoeken in hoeverre er sprake is van een systematisch verschil tussen berekeningen en metingen.

Hoofdstuk 3 beschrijft vervolgens de belangrijkste conclusies van de uitgevoerde validatie en aanbevelingen voor toekomstige validaties en eventuele vervolgonderzoeken.

2 Vergelijking meten en rekenen

Dit hoofdstuk beschrijft de vergelijking van rekenresultaten van het Doc.29-rekenmodel met metingen op de NOMOS-geluidmeetposten rondom Schiphol. Zoals genoemd in paragraaf 1.2.3 is een deel van deze vergelijking ook in het PAMV-rapport over het toepassingsbereik van metingen en berekeningen gepresenteerd. In het voorliggende rapport wordt de vergelijking uitgebreid met een meer uitvoerige statistische analyse en de vergelijking wordt ook nog verdiept op vliegtuigtype-niveau. De vergelijking in dit hoofdstuk wordt verder ook naast op L_{den} -niveau op SEL-niveau gepresenteerd om meer inzicht te verkrijgen in het waargenomen verschil per meetpost op L_{den} -niveau.

Het hoofdstuk begint met een beschrijving van de aanpak en methodologie gevolgd voor de vergelijking in paragraaf 2.1, gevolgd door een beschrijving van de gebruikte meet- en rekengegevens in paragraaf 2.2. Paragraaf 2.3 beschrijft de analyse en uitkomsten van de vergelijking en sluit af met een verdiepende regressieanalyse om te duiden in hoeverre er sprake is van een systematisch verschil tussen meten en rekenen.

2.1 Aanpak en methodologie

Voor de vergelijking van berekeningen met metingen is gebruikgemaakt van metingen van het NOMOS-meetsysteem en berekeningen uitgevoerd met het Doc.29-rekenmodel. Dit is gedaan voor het vliegverkeer van en naar Schiphol voor het gebruiksjaar 2019. De gebruikte gegevens zijn beschreven in paragraaf 2.2. De gegevens bevatten naast de meet- en rekenwaarden per vliegtuigpassage ook informatie over kenmerken van de vlucht (zoals vluchtsoort en vliegtuigtype), datum en tijdstip van de vliegtuigpassage. De meet- en berekeningsgegevens zijn aan elkaar gekoppeld en zijn verrijkt met meteorologische gegevens. Mede op basis van de meteorologische gegevens is vastgesteld of de meting voldoet aan de gestelde aanvullende meeteisen uit de PAMV-validatieaanpak. Via een aantal statistische vergelijkingen zijn vervolgens diverse (deel)vergelijkingen tussen metingen en berekeningen uitgevoerd en geanalyseerd in paragraaf 2.3.

2.2 Gebruikte invoergegevens

2.2.1 Meetgegevens

Voor de meetgegevens is gebruikgemaakt van meetdata die afkomstig is van de meetposten van het NOMOS-geluidmeetsysteem. Hiervoor zijn alle metingen van civiel vliegverkeer in het gebruiksjaar 2019 voor Schiphol geanalyseerd.

De leverancier van het meetsysteem is het bedrijf Envirosuite (voorheen Bruel en Kjaer). Voor het ophalen van meetgegevens van beide luchthavens is gebruikgemaakt van de Airport Noise and Operations Monitoring System (ANOMS) database van Envirosuite. In de ANOMS database worden meetdata van alle door de luchthaven beheerde meetposten opgeslagen, inclusief aanvullende informatie over de vlucht,

zoals het vluchtsoort en het vliegtuigtype. Schiphol heeft RIVM voor dit onderzoek voor het ophalen van de Schiphol-meetdata toegang verleend tot de ANOMS-database.

Wat betreft de meetdata kunnen twee opmerkingen worden gemaakt die voor de uitgevoerde vergelijking van belang zijn.

1. Meteo:

Voor de analyse van meteocondities is gebruikgemaakt van publiek beschikbare urregelingen van het KNMI (KNMI 2023). De reden hiervoor was dat de meteogegevens uit ANOMS niet met de KNMI-gegevens overeenkwamen. Ze waren bovendien niet gebaseerd op weergegevens van windmeters en andere op de meetposten geïnstalleerde meteoapparatuur. De meerderheid van de NOMOS-meetposten heeft geen meteoapparatuur beschikbaar, waardoor de KNMI-gegevens als betrouwbaarder werden beschouwd. Hiervoor is gebruikgemaakt van het meest dichtbij liggende KNMI-weerstation per meetpost.

2. SEL versus L_{Amax} :

Een geluidmeting van een vliegbeweging wordt uitgedrukt in een geluidmaat zoals de L_{Amax} (het maximale geluidniveau) of SEL. Voor het doelvalidatie is de SEL het meest belangrijk omdat de geluidbelasting in L_{den} wordt bepaald door een optelling van de SEL-waarden van individuele vliegtuigpassages. Het voorliggende rapport focust daarom op de SEL-geluidmaat wat betreft de verschillen op vluchtniveau en er worden geen resultaten gepresenteerd voor de L_{Amax} -geluidmaat.

Voor de analyse zijn enkele bewerkingen op de totale meetdataset uitgevoerd om ervoor te zorgen dat verstoorde of onbetrouwbare metingen de analyse en de conclusies van het onderzoek niet beïnvloeden. Hiernaast zijn de volgende basisfilters op de meetdataset toegepast:

- Alleen startende of landende vliegtuigen van of naar Schiphol worden geanalyseerd (overvluchten van andere luchthavens zijn buiten beschouwing gelaten).
- De berekeningen zijn voor het grootste deel van de analyse gekoppeld aan metingen. Wat deze keuze van het koppelen van vliegbewegingen voor de vergelijking tussen meten en rekenen betekent, wordt in paragraaf 2.3 nader toegelicht.

Vervolgens zijn er op de overgebleven metingen de validatie-eisen toegepast. Deze eisen gelden voor de beschouwde meetlocatie – wel of niet geschikt voor validatie conform de kaders van de PAMV nationale meetstrategie (Sahai 2021) en aanvullende meeteisen bepaald als onderdeel van de PAMV-validatieaanpak (Sahai 2022). De aanvullende eisen worden als 'filters' toegepast op de meetdata voor:

- Windsnelheid (max 5 m/s);
- Neerslag (geen regen of sneeuw);
- Elevatiehoek (minimaal 60 graden).

De metingen die op deze wijze overblijven, worden geschikt geacht om een betrouwbare vergelijking met berekeningen te maken. Uit deze metingen wordt per meetpost een L_{den} -waarde bepaald. De L_{den} is een geluidmaat voor de jaargemiddelde geluidbelasting. Het geeft in principe

een waarde voor de gewogen jaargemiddelde etmaal belasting. Vliegtuigbewegingen in de avond en nacht krijgen hierbij een straffactor van respectievelijk 5 en 10 dB. Het optellen van de vliegtuigbewegingen komt neer op het optellen van SEL-waarden per vliegtuigbeweging.

Na het toepassen van de validatiefilters ontstaan enkele uitdagingen door het wegvallen van metingen. Zoals toegelicht in het rapport over het toepassingsbereik van metingen en berekeningen (Sahai 2024) blijkt een deel van de metingen na het toepassen van de validatie-eisen ongeschikt voor validatie. Daardoor ligt het aantal bruikbare metingen voor de validatie lager dan het totale aantal metingen. Dit heeft als gevolg dat een L_{den} die bepaald is op basis van de gefilterde set metingen op sommige locaties lager uitvalt. Ook betekent dit dat de L_{den} voor sommige locaties bepaald moet worden op basis van een relatief laag aantal metingen. Het effect hiervan op de vergelijking van meten en rekenen voor Schiphol wordt in de komende paragrafen van dit hoofdstuk nader onderzocht en toegelicht.

2.2.1.1 Afwijkende meetlocaties

Twee validatiemeetlocaties lieten in dit onderzoek opvallend grote verschillen zien tussen meten en rekenen. Deze locaties worden gezien als 'outliers' en zijn daarom niet meegenomen in verdere conclusies.

Meetlocatie 18 is een validatielocatie waar de meetresultaten van NOMOS sterk afwijken van de Doc.29-berekeningen. Nader onderzoek heeft laten zien dat tijdens de inventarisatie in 2020 een onjuist adres voor deze meetpost is overgenomen, waardoor de berekeningen zijn uitgevoerd voor de verkeerde locatie. Gezien de benodigde doorlooptijd is dit niet meer gecorrigeerd, maar zijn de resultaten voor meetlocatie 18 verder niet meer meegenomen in de analyse.

Meetlocatie 26 is ook een validatielocatie, maar desondanks zijn de verschillen groter dan op andere locaties. Op deze meetlocatie blijken veel vliegtuigbewegingen te zijn gemeten en berekend met zeer lage elevatiehoeken op relatief grote afstand (3 km). Mogelijk zijn hierbij stoorgeluiden opgepikt en gekoppeld aan vliegtuigen die worden overstemd door het stoorgeluid. Nader onderzoek moet dit verder uitwijzen. Meetpost 26 is voorlopig ook niet in de analyse meegenomen.

2.2.2 Berekeningsgegevens

Om de gemeten en berekende geluidbelastingniveaus op de meetpostlocaties te kunnen vergelijken, zijn voor alle NOMOS-locaties geluidberekeningen uitgevoerd met het ECAC Doc.29-model, vierde editie (ECAC.CEAC 2016b). Meer specifiek is gebruikgemaakt van de Doc.29 implementatie voor Schiphol. Naast het rekenmodel zelf dat is vastgelegd in (ECAC.CEAC 2016b) omvat deze implementatie een totaal aan werkafspraken, de vaste invoergegevens gebaseerd op de Aircraft Noise and Performance Database versie 2.2 (Eurocontrol 2019), en de benodigde nabewerkingstappen. De implementatie zelf is beschreven in het rapport over de toepassing van ECAC Doc.29 voor het bepalen van de geluidbelasting van het vliegverkeer van Schiphol (Heblij, Derei and Hogenhuis 2017).

De geluidberekeningen zijn uitgevoerd voor alle vliegtuigbewegingen van en naar Schiphol in het gebruiksjaar 2019 (1 november 2018 tot en met 31 oktober 2019). De vluchtgegevens, inclusief de vliegbanen, zijn afkomstig uit FANOMOS. Alleen vliegtuigbewegingen waarvoor essentiële gegevens ontbreken of waarvan vastgesteld kan worden dat gegevens niet juist zijn, zijn niet doorgerekend. Waar mogelijk is General Aviation (GA) verkeer ook doorgerekend. Een uitzondering is helikopterverkeer: dit wordt niet ondersteund binnen de Doc.29-implementatie voor Schiphol.

De berekeningen zijn uitgevoerd per vliegtuigbeweging: voor alle vliegtuigbewegingen zijn de geluidblootstellingsniveaus (SEL) bepaald. Dit maakt het mogelijk om de vergelijking meten en rekenen ook voor specifieke doorsnedes te maken, zoals alleen startend verkeer of voor specifieke vliegtuigtypes. Door optelling van de geluidblootstellingsniveaus van alle vliegtuigbewegingen kan, rekening houdend met de straffactoren, ook de geluidbelasting (L_{den}) worden berekend.

Het is doorgaans gebruikelijk om geluidbelastingberekeningen uit te voeren op een raster van rekenpunten op regelmatige afstand van elkaar. Een dergelijk raster wordt vervolgens gebruikt voor het vaststellen van geluidscontouren. Voor het voorliggende onderzoek is echter geen gebruik gemaakt van een raster, maar zijn puntberekeningen uitgevoerd op de exacte locaties van de NOMOS-meetposten. De puntberekeningen zijn uitgevoerd voor alle meetpostlocaties, inclusief de locaties waarop laagfrequent geluid wordt gemeten (in dBC). Omdat het rekenmodel echter alleen geluidniveaus in dBA kan bepalen, zijn de berekende niveaus voor de meetposten 78 en 80 niet geschikt voor de vergelijking van meten en rekenen. Deze twee meetposten worden daarom ook buiten de huidige analyse gelaten, naast meetposten 18 en 26 die afwijkende metingen lieten zien.

De set met berekeningsgegevens is relatief groot ten opzichte van de set met meetgegevens. In de berekeningen zijn ruim een half miljoen vliegtuigbewegingen doorgerekend op *elk* van de 41 meetlocaties. Ter vergelijking: het totaal aantal metingen bedraagt minder dan 3 miljoen voor alle locaties samen. Het grote verschil kan worden verklaard door het feit dat een meting alleen kan worden uitgevoerd voor vliegtuigen die relatief dicht in de buurt van een meetpost komen en daar een voldoende hoog geluidniveau boven de meedrempelwaarde veroorzaken. Berekeningen kunnen praktisch altijd worden uitgevoerd. Ook voor vliegtuigbewegingen met een laag geluidniveau en voor vliegtuigbewegingen die niet in de buurt van de meetpost komen. Bij de vergelijking van meten en rekenen en de analyse in de komende paragrafen dient rekening te worden gehouden met dit fundamentele verschil.

2.3 Resultaten vergelijking meten en rekenen

Deze paragraaf presenteert de vergelijking van berekeningen en metingen voor het gebruiksjaar 2019 voor Schiphol. De resultaten van de vergelijking worden eerst gepresenteerd voor de gekoppelde dataset, waarbij de L_{den} - en SEL-waarden van alleen die vliegtuigbewegingen met

elkaar worden vergeleken die zowel gemeten als berekend zijn. Vervolgens wordt geanalyseerd wat het effect van de ontbrekende vliegtuigbewegingen op de geluidbelasting is. Vliegtuigbewegingen kunnen om verschillende redenen in één van beide datasets ontbreken, maar ontbreken veelal in de meetgegevens door het gebruik van de drempelwaarde in het NOMOS-meetsysteem. Deze tweede vergelijking wordt gedaan door de geluidbelasting voor alle vliegtuigbewegingen van en naar Schiphol in 2019 te berekenen en die te vergelijken met de voor de gekoppelde dataset berekende geluidbelasting. Alhoewel het effect van de ontbrekende vliegtuigbewegingen op de *gemeten* L_{den} dan nog steeds onbekend blijft, kan niettemin een indruk worden verkregen van het deel van de geluidbelasting dat door het gebruik van de drempelwaarde ontbreekt.

Paragraaf 2.3.2 presenteert vervolgens een verdiepende analyse op vluchtsoort-niveau en vliegtuigtype-niveau voor de gekoppelde dataset. Het doel daarbij is om te bepalen of de verschillen groter zijn voor starts of landingen en of er bepaalde vliegtuigtypen zijn waarvoor het verschil (structureel) hoger ligt dan de gemiddelde verschillen. Paragraaf 2.3.3 presenteert als laatst een verdiepende regressieanalyse om te bepalen in hoeverre er sprake is van een systematisch verschil tussen metingen en berekeningen.

2.3.1 *Resultaten zonder gebruik van subselecties van meetdata*

2.3.1.1 Analyse gekoppelde dataset

Voordat de analyse wordt gepresenteerd, wordt een aantal punten benoemd waar rekening mee dient te worden gehouden bij de vergelijking tussen meten en rekenen:

- De vergelijking tussen berekeningen en metingen in dit rapport is uitgevoerd op basis van de *huidige* Doc.29-implementatie voor Schiphol. In een later traject, na deze fase van de PAMV, is ook een modelverbetering beoogd met focus op punten waarop het rekenmodel kan worden verbeterd. De eventuele modelverbetering zal in dat geval afzonderlijk worden gerapporteerd. De uitkomsten van modelverbetering kunnen potentieel een invloed hebben op de vergelijking tussen meten en rekenen.
- Aangezien de vergelijking in dit onderzoek alleen voor de luchthaven Schiphol wordt gemaakt, zijn de gerapporteerde uitkomsten dan ook vooral van toepassing voor Schiphol. Hoe de vergelijking voor de andere (regionale) luchthavens eruitziet, is voorlopig onbekend. Hierover is door het PAMV-consortium een afzonderlijk advies opgesteld.
- De analyse in dit hoofdstuk is voornamelijk uitgevoerd voor vliegtuigbewegingen waarvoor zowel een gemeten als berekende SEL-waarde beschikbaar waren. Het effect van het gebruik van alle beschikbare gegevens op de vergelijking wordt in paragraaf 2.3.1.2 nog nader toegelicht.
- Omdat door het toepassen van de validatiefilters (aanvullende meeteisen) een deel van de metingen wegvalt, is een ondergrens toegepast op een minimaal aantal gekoppelde vliegtuigbewegingen per meetpost. Indien het aantal gekoppelde vliegtuigbewegingen lager is dan 50, wordt de betreffende meetpost niet in de analyse meegenomen. Voor de huidige

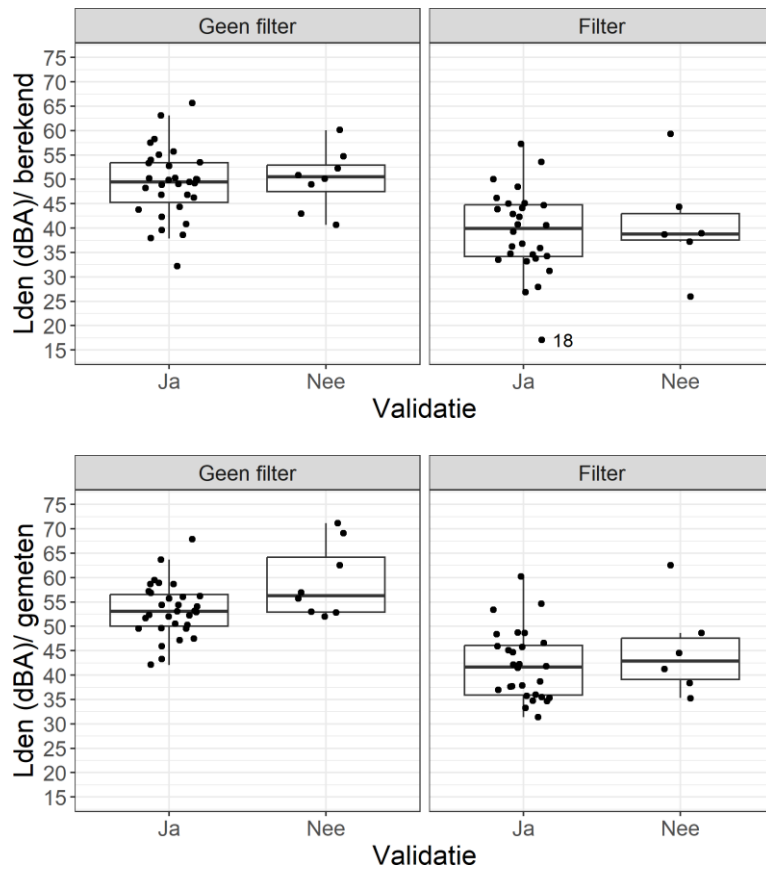
analyse betrof dit in totaal twee meetposten waarvoor het aantal koppelingen op minder dan 50 uitkwam. Deze ondergrens is voorgesteld door onderzoekers in het Verenigd Koninkrijk (VK) om een minimaal niveau van statistische betrouwbaarheid te behalen (Rhodes D. P. 2018), (Trow 2018).

- Conform de PAMV-validatieaanpak (Sahai 2022) wordt een verschil van 2 dB L_{den} per meetpost tussen de gemeten en berekende geluidbelasting als de bovengrens beschouwd wat betreft een aanvaardbaar verschil. Een verschil hoger dan deze waarde geeft aanleiding voor nader onderzoek naar de oorzaken van het verschil.

2.3.1.1.1 Analyse berekende en gemeten waarden

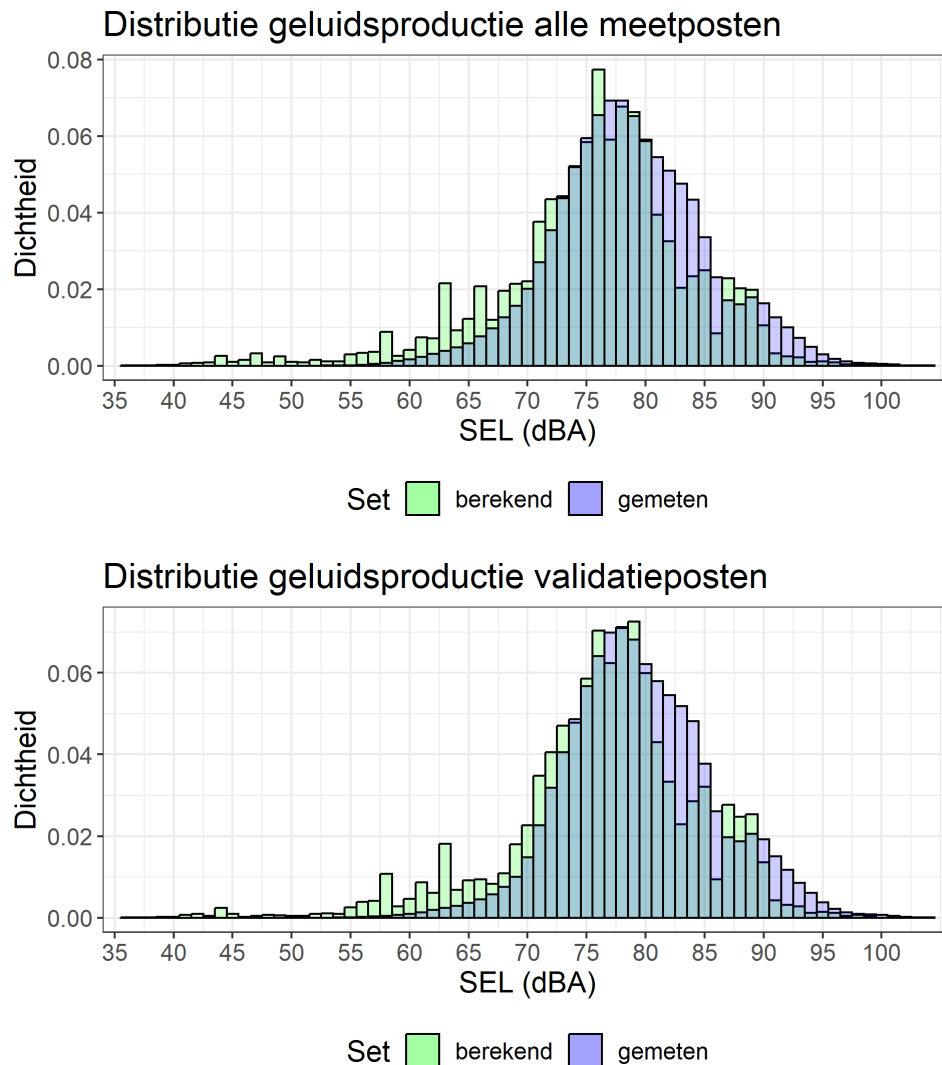
De analyse begint met een presentatie van de berekende en gemeten geluidbelasting in Figuur 1 (tevens gepresenteerd in (Sahai 2024)). Figuur 1 laat zien dat voor de gekoppelde dataset de berekende geluidbelasting (bovenste grafiek) lager uitvalt dan de gemeten geluidbelasting (onderste grafiek). Zowel de mediaan-waarde als de minimaal en maximaal gemeten L_{den} 's liggen hoger dan de berekende L_{den} 's. Dit is het geval zowel vóór het toepassen van de validatiefilters voor de elevatiehoek, wind en neerslag, als na het toepassen van de validatiefilters.

Een verdere opsplitsing is gemaakt in Figuur 1 op basis van meetposten die volgens de kaders van de nationale meetstrategie (Sahai 2021) geschikt zijn voor validatie en meetposten die niet geschikt zijn voor validatie (in de figuur is dat aangegeven als validatie ja/nee). Door het gebruik van de box-plots wordt ook de spreiding in beide waarden (berekend en gemeten) zichtbaar. In Figuur 1 is te zien dat de geluidbelasting na het toepassen van de validatiefilters lager wordt dan voor het toepassen van de filters (zoals eerder toegelicht). Ook de berekende L_{den} -waarde wordt dan lager, doordat de L_{den} bepaald wordt voor vliegtuigbewegingen waarvoor zowel een gemeten als een berekende SEL-waarde beschikbaar is. De betreffende berekende vliegtuigbewegingen worden samen met de gemeten vliegtuigbewegingen uitgefilterd.



Figuur 1 Vergelijking van berekende (boven) en gemeten (onder) geluidbelasting op de NOMOS meetposten opgesplitst over meetposten wel en niet geschikt voor validatie, voor en na toepassen van validatiefilters.

Figuur 2 laat zien dat de metingen ook op SEL-niveau doorgaans hoger uitkomen dan berekeningen door de distributies van de berekende en gemeten geluidniveaus te laten zien. De vergelijking van de berekende en gemeten distributies is in Figuur 2 weergegeven voor alle meetposten (bovenste grafiek) en voor alleen de validatiemeetposten (onderste grafiek). Hieruit blijkt dat de berekende distributie (groen) geluidniveaus van hoger dan 80 dBA SEL minder vaak laat zien in vergelijking met de gemeten distributie (paars). Dit geldt zowel voor alle meetposten als voor alleen de validatiemeetposten. Daarnaast worden geluidniveaus van minder dan 70 dBA SEL met een relatief hoge frequentie berekend, terwijl deze geluidniveaus veel minder vaak worden gemeten. Dit is een gevolg van het gebruik van de drempelwaarde, waarbij vliegtuigbewegingen onder de meetdrempel niet in het NOMOS-meetsysteem worden geregistreerd.

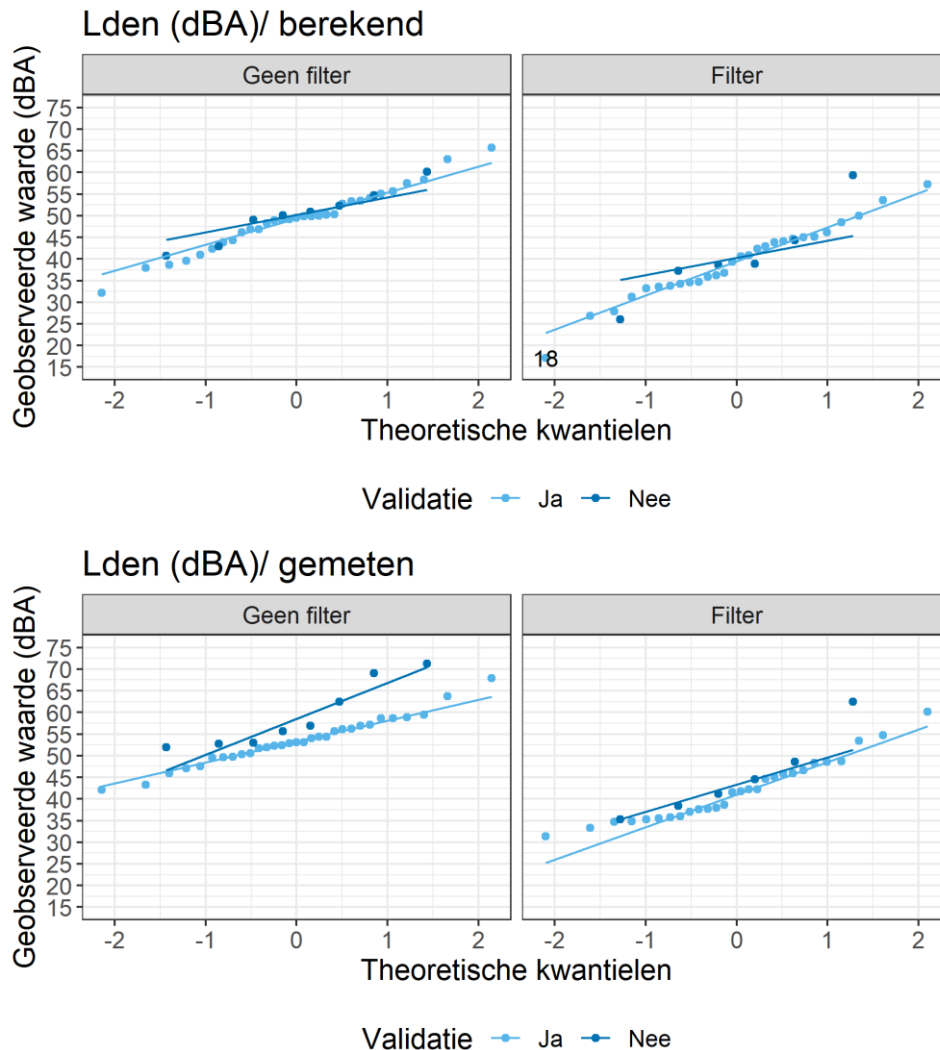


Figuur 2 Vergelijking van de distributies voor de berekende en gemeten geluidsniveaus (SEL) voor alle NOMOS-metposten (boven) en alleen validatiemetposten (onder).

Figuur 3 breidt de statistische analyse uit met een zogeheten Q-Q (quantile-quantile)-plot van de berekende en gemeten geluidbelasting. In een Q-Q-plot wordt de hypothese getoetst dat de waarden onderdeel uitmaken van dezelfde normaalverdeling of distributie met één gemiddelde en één variantie. Voor de berekende en gemeten geluidbelasting kan met een Q-Q-plot worden getoetst of beide sets van data voor en na filtering dezelfde verdeling volgen en of er punten zijn die afwijkend gedrag tonen en daardoor niet in dezelfde verdeling passen. Op de y-as in Figuur 3 worden de geobserveerde waarden getoond en op de x-as de standaardafwijking van iedere waarde in relatie tot het gemiddelde. De steilheid van de curve geeft de grootte van de variantie van de groep weer.

In Figuur 3 is te zien dat het toepassen van validatiefilters een groter effect heeft op de gemeten dataset (onderste grafiek) dan op de berekende dataset (bovenste grafiek). De uitbijeters (punten met

afwijkend gedrag) worden door het gebruik van validatiemeetposten en filtering grotendeels verwijderd. Voor de gemeten geluidbelasting is een duidelijk verschil zichtbaar tussen validatieposten en niet validatieposten wanneer er geen validatiefilters worden toegepast. Na het toepassen van de validatiefilters volgen de datasets een beter vergelijkbare verdeling.



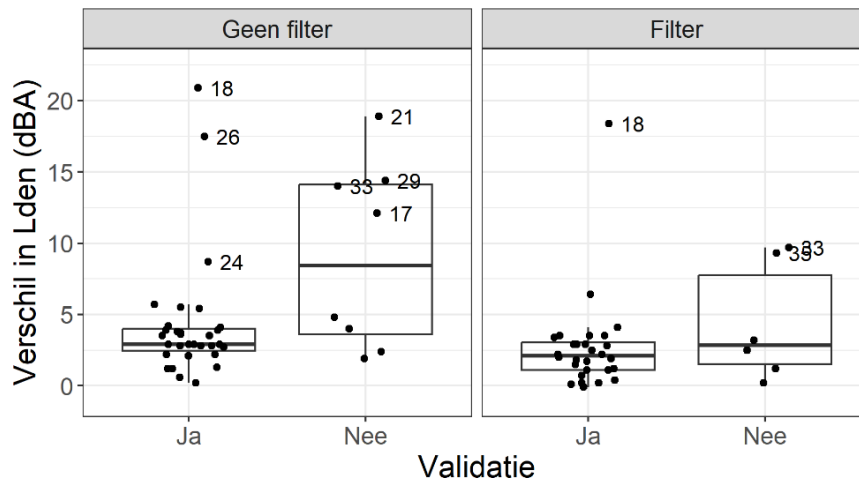
Figuur 3 Q-Q-plots voor de berekende (boven) en gemeten (onder) geluidbelasting.

2.3.1.1.2 Analyse verschil meten en rekenen

Figuur 4 toont het verschil tussen de gemeten en berekende geluidbelasting en de spreiding in het verschil via box-plots. Bij een verschil dat groter dan nul is, is de gemeten waarde hoger dan de berekende waarde. Figuur 4 laat zien dat het verschil voor meetposten die volgens de kaders van de nationale meetstrategie geschikt zijn voor validatie kleiner is dan voor meetposten die niet voor validatie geschikt zijn. Verder valt het op dat het toepassen van validatiefilters het verschil tussen meten en rekenen verder verkleint. De box-plots laten zien dat

niet alleen het verschil tussen meten en rekenen door het toepassen van validatie-eisen wordt verkleind, maar ook de spreiding in het verschil.

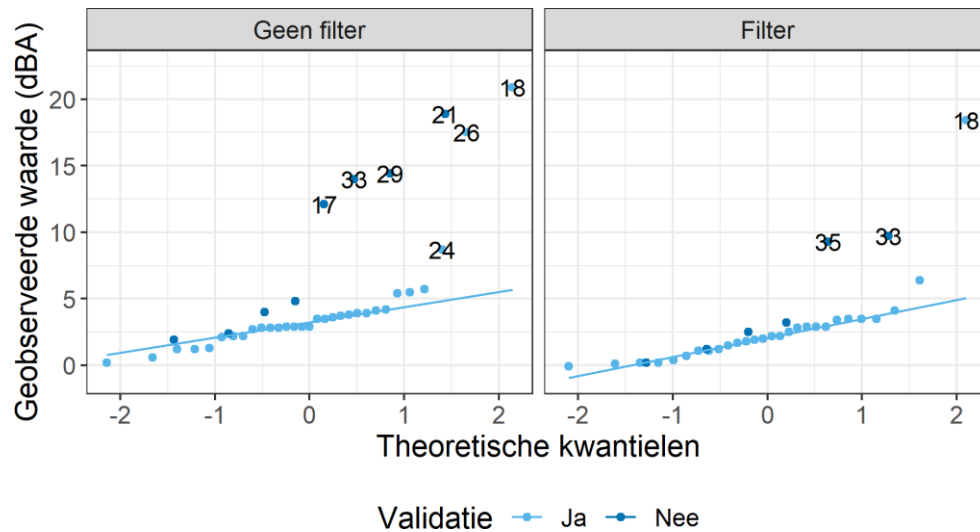
De linker box-plot in Figuur 4 laat zien dat meetposten 18 en 26 uitbijters zijn. Beide meetposten zijn gemarkeerd als 'geschikt voor validatie', maar laten toch hoge verschillen tussen gemeten en berekende geluidbelasting zien. Zoals toegelicht in paragraaf 2.2.1, zijn de metingen op meetpost 26 mogelijk verstoord door een lage elevatiehoek. Voor deze meetpost verkleint filtering het verschil aanzienlijk, maar er blijken nog onzekerheden te zitten aan de meetkant. Voor meetpost 18 bleek uit een nadere analyse dat de berekende niveaus niet juist zijn door de onjuist geregistreerde locatie voor deze meetpost. De berekening is hierdoor voor een locatie in Amsterdam uitgevoerd, terwijl de meetpost in werkelijkheid in Haarlem staat.



Figuur 4 Box-plot voor het verschil tussen gemeten en berekende geluidbelasting zonder toepassing van validatiefilters (links) en na toepassing van validatiefilters (rechts).

Met een gepaarde t-test is er verder getoetst of er statistisch gezien een significant verschil is voor de validatiemeetposten voor en na filtering. Hieruit blijkt dat wanneer meetposten 18 en 26 niet mee worden gewogen, het verschil tussen beide groepen inderdaad statistisch significant is (p-waarde < 0.05).

In Figuur 5 wordt de Q-Q-plot voor het verschil tussen gemeten en berekende geluidbelasting gepresenteerd. De Q-Q-plot bevestigt dat wat betreft validatiemeetposten, meetposten 18 en 26 niet tot dezelfde distributie als de resterende validatiemeetposten behoren. Na filtering blijft slechts één validatiemeetpost (18) over, die ver buiten de verwachte verdeling valt. Vanwege het afwijkende gedrag en de (mogelijke) oorzaken daarvoor, zijn meetposten 18 en 26 buiten de validatie-analyse gelaten die in de volgende paragrafen wordt gepresenteerd.



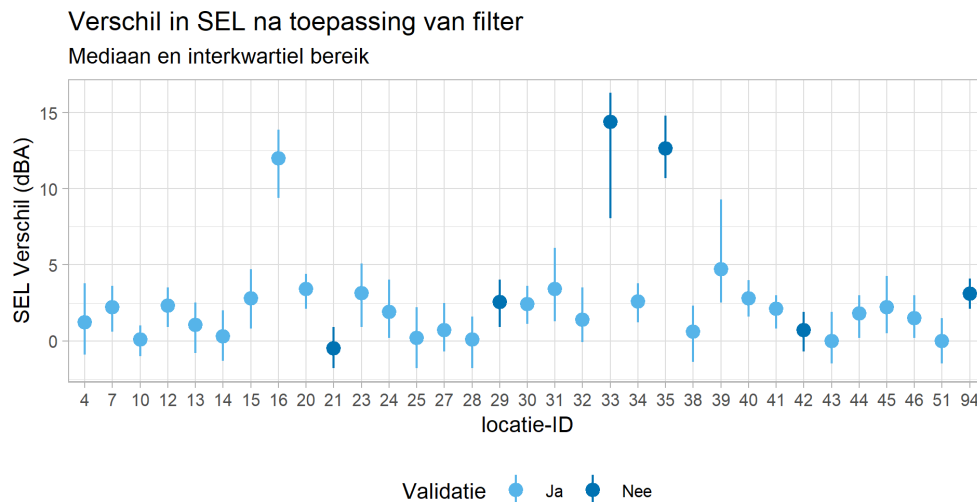
Figuur 5 Q-Q-plot voor het verschil tussen gemeten en berekende geluidbelasting zonder toepassing van validatiefilters (links) en na toepassing van validatiefilters (rechts).

De Q-Q-plots in Figuur 5 bevestigen het beeld dat het verschil tussen meten en rekenen voor de niet-validatiemeetposten anders is dan voor de validatiemeetposten. De spreiding en het gemiddelde van het verschil zijn voor beide groepen ongelijk. Dit is te zien aan de steilheid van de curve. Het afwijkende gedrag voor niet-validatiemeetposten blijft voor een deel ook na filtering aanwezig (meetposten 33 en 35 in de rechterplot in Figuur 5).

Naast de vergelijking op L_{den} -niveau kan ook de vergelijking op SEL-niveau nader inzicht bieden in de waargenomen verschillen per meetpost. Figuur 6 laat de verschillen tussen metingen en berekeningen op SEL-niveau zien, *na* het toepassen van de validatiefilters. Dit wordt gedaan door de mediaan en het interkwartiel-bereik (25%-75%) van het SEL-niveaoverschil per meetpost te presenteren.

Voor bijna alle meetposten ligt het verschil op SEL-niveau na toepassing van de validatiefilters onder 4 dB SEL. Veel van de meetposten in Figuur 4 met een groot verschil op L_{den} -niveau kennen ook een groot verschil op SEL-niveau. Opvallend hierbij is validatiemeetpost 16 die op SEL-niveau een vrij groot verschil laat zien (mediaan van 12 dB SEL). Het verschil voor deze meetpost is ook op L_{den} -niveau relatief groot en bedraagt circa 6 dB L_{den} . Bij deze meetpost gaat het na filtering om een vrij klein aantal metingen dat overblijft en waarvoor de gemeten waarden veel hoger zijn dan de berekende waarden. Aangezien er verder geen aanleiding is dat er op deze meetpost iets is misgegaan, wordt deze meetpost niet als een 'outlier' beschouwd. Daarom wordt deze meetpost in de analyse meegenomen. Meetposten 33 en 35 laten ook grote verschillen zien (zowel op SEL-niveau als op L_{den} -niveau) maar zijn als niet-validatiemeetposten gemarkeerd.

Paragraaf 2.3.2 splitst de vergelijking op SEL-niveau per meetpost verder op basis van vluchtsoort (starts of landingen) en vliegtuigtype op.



Figuur 6 Vershil in SEL-waarden per meetpost met weergave van de mediaan waarde en het interkwartiel-bereik voor de situatie na toepassing van de validatiefilters.

Eén van de doelen van de PAMV is om te bepalen hoe goed de berekeningen met metingen overeenkomen en tot waar het model de meest betrouwbare resultaten geeft. Deze vraag is grotendeels in het onderzoek naar het toepassingsbereik van metingen en berekeningen geadresseerd, maar wordt in het voorliggende rapport nog nader toegelicht. Hiernaast wordt in het voorliggende rapport ook de correlatie gepresenteerd tussen meten en rekenen voor het geval dat alle berekende vliegtuigbewegingen met alle gemeten vliegtuigbewegingen (zonder een koppeling van beiden) worden vergeleken en hoe dit zich verhoudt met de gekoppelde dataset.

Figuur 7 laat de correlatie tussen de gemeten en berekende geluidbelasting voor de gekoppelde dataset zien (zoals ook gepresenteerd in (Sahai 2024)). Dit is gedaan in drie stappen:

1. Door eerst de correlatie voor alle meetposten voor de ongefilterde situatie te laten zien.
2. Gevolgd door de correlatie alleen voor validatiemeetposten (zonder filtering).
3. En als laatst voor validatiemeetposten met de validatiefilters toegepast. N.B. de geluidbelasting na het toepassen van de validatiefilters is lager dan voor het toepassen van de filters. Dit komt doordat na filteren een lager aantal metingen beschikbaar blijft om de L_{den} -waarde te bepalen.

De belangrijkste bevinding op basis van Figuur 7 is dat het toepassen van validatie-eisen, zowel voor de voor de nationale meetstrategie bepaalde meetlocatie als voor de PAMV-validatieaanpak bepaalde aanvullende meeteisen, niet alleen het verschil tussen meten en rekenen verkleinen, maar ook de correlatie tussen deze twee aanzienlijk verhogen. Het toepassen van de validatie-eisen (op de meetlocatie en aanvullende eisen als validatiefilters) verhoogt de R-waarde⁴ van 0.8 tot

⁴ De R-waarde geeft de correlatie weer. Een waarde van 0 geeft aan dat er geen correlatie is; een waarde dicht bij 1 geeft aan dat er een sterke correlatie is. Hierbij is het wel belangrijk om te beseffen dat een hoge R-waarde laat zien dat meten en rekenen elkaar goed volgen, maar niets zegt over hun absolute verschillen.

0.98 met een p-waarde⁵ voor alle drie de gevallen ruim onder het significantieniveau van 0.05. De rode lijn in de drie plots is de lijn van beste fit en de zwarte, gestreepte lijn is de ideale lineaire vergelijking tussen metingen en berekeningen, waarbij een berekende waarde precies met een gemeten waarde overeenkomt.

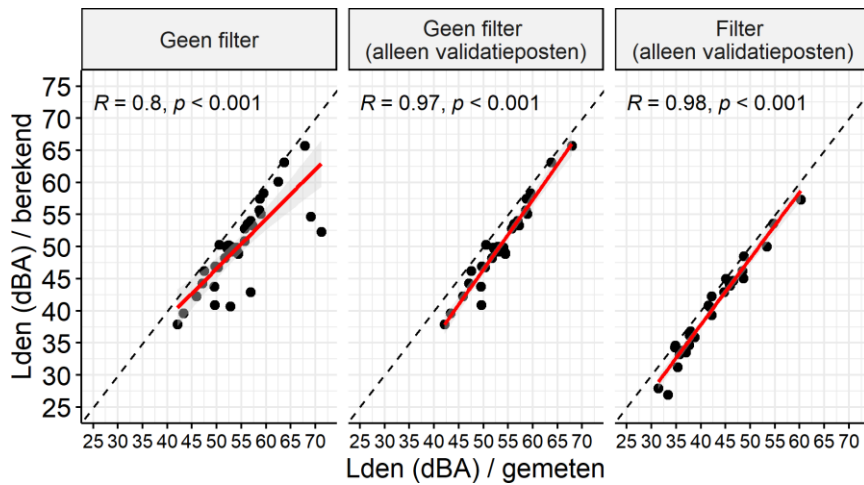
Een andere constatering uit Figuur 7 is dat het verschil tussen meten en rekenen met afnemende geluidbelasting toeneemt. Uit de middelste en rechter plots van Figuur 7 is op te maken dat de berekeningen bij hoge geluidbelasting beter met metingen overeenkomen dan bij lage geluidbelasting. In het onderzoek naar het toepassingsbereik van metingen en berekeningen van vliegtuiggeluid (Sahai 2024) is een ongefilterde waarde van circa 50 dB L_{den} bepaald als de waarde vanaf waar met relatief grote zekerheid kan worden gezegd dat de berekeningen betrouwbaar zijn (dat wil zeggen voor 50 dB L_{den} en hoger). Het verschil tussen meten en rekenen tot deze waarde onder 2 dB blijft en er voor de vergelijking voldoende meetdata beschikbaar zijn.

Bij waarden lager dan 50 dB L_{den} is het verschil tussen meten en rekenen doorgaans groter. Hierbij geeft de vergelijking tussen meten en rekenen door de volgende redenen echter een minder betrouwbaar beeld:

- Het aantal beschikbare metingen is lager door het toepassen van de drempelwaarde.
- De drempelwaarde zorgt ervoor dat metingen van relatief stille passages in sommige gevallen niet meer gemeten worden, terwijl de luidere passages nog wel gemeten worden. Hierdoor lijkt het gemiddelde gemeten geluidniveau hoger dan het in werkelijkheid is.
- Metingen van lagere geluidniveaus kunnen eerder worden verstoord door andere geluidbronnen.

De vergelijking bij geluidbelastingwaarden die hoger dan 50 dB L_{den} zijn, kent een hogere zekerheid, omdat bovenstaande zaken dan in mindere mate spelen. Hierbij kan worden opgemerkt dat de onzekerheid bij zeer hoge geluidbelasting waarden (hoger dan circa 65 dB L_{den} ongefilterd) weer toeneemt, omdat er vanaf deze waarde ook minder meetdata beschikbaar zijn.

⁵ De p-waarde geeft aan of de vergelijking statistisch significant is en heeft ook een waarde tussen 0 en 1. Als de p-waarde kleiner is dan het significantieniveau van 0.05, dan is de vergelijking statistisch relevant.



Figuur 7 Correlatie tussen gemeten en berekende geluidbelasting voor de gekoppelde dataset – zonder toepassing van validatiefilters op alle meetposten (links), alleen voor validatiemeetposten (midden) en na het toepassen van validatiefilters op validatiemeetposten (rechts)

De correlatie uit Figuur 7 en de vergelijking in dit hoofdstuk tot nu toe is gemaakt voor de gekoppelde dataset. De volgende paragraaf presenteert de correlatie als de vergelijking wordt gemaakt voor de geluidbelasting van alle berekende vliegtuigbewegingen met alle gemeten vliegtuigbewegingen (zonder een koppeling van beiden). Hiermee wordt ook onderzocht wat het potentiële effect van de drempelwaarde is op de vergelijking tussen meten en rekenen.

2.3.1.2 Analyse vergelijking gekoppelde dataset en dataset met alle berekende vliegtuigbewegingen

Zoals eerder beschreven, ontbreekt een deel van de metingen door het gebruik van de drempelwaarde waardoor de vergelijking tussen meten en rekenen bij een geluidbelasting lager dan circa 50 dB L_{den} onbetrouwbaarder wordt. Tegelijkertijd is het toepassen van de drempelwaarde ook een internationaal toegepaste manier om ervoor te zorgen dat stoorgeluiden of geluid van andere bronnen in de achtergrond niet te vaak aan een overvliegend vliegtuig worden toegewezen.

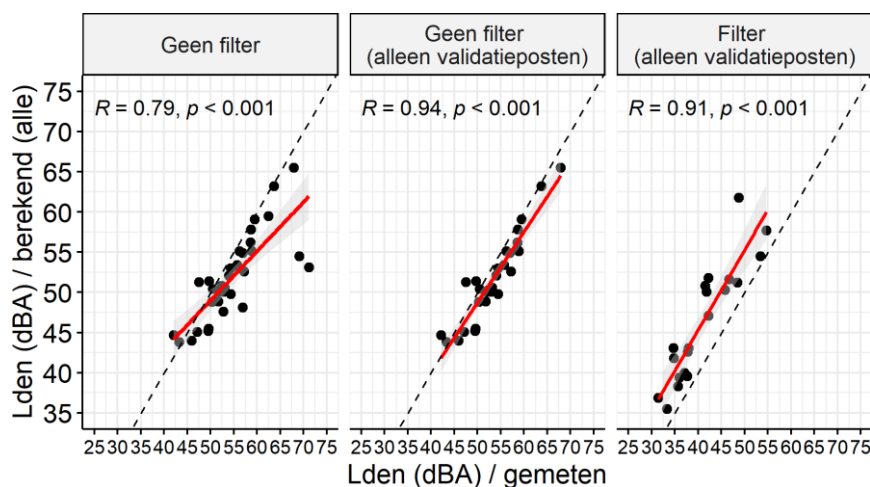
Het niet toepassen van de drempelwaarde kan dus ertoe leiden dat metingen van andere bronnen of stoorgeluiden onterecht als vliegtuiggeluid worden geregistreerd. Recentelijk zijn er nieuwe, meer geavanceerde algoritmes beschikbaar gekomen (ook onder andere voor Schiphol (Schiphol 2022)) waarbij de drempelwaarde door een analyse van het geluid achteraf kan worden bijgesteld. Nieuwe methoden zoals deze kunnen het aantal vliegtuigbewegingen dat kan worden gemeten verhogen. Een dergelijke methode is op Schiphol in 2022 voor de eerste keer getest. Het validatieonderzoek maakt gebruik van meetgegevens van voor 2022, waardoor voor de nu gebruikte meetgegevens nog een vaste drempelwaarde is toegepast. In de toekomst kan onderzocht worden of het gebruik van nieuwe meetmethodes nuttig is bij het verzamelen van meetgegevens voor modelvalidaties.

Deze paragraaf geeft een indruk van het effect van de ontbrekende vliegtuigbewegingen op de berekende geluidbelasting en hoe dit de vergelijking met gemeten geluidbelasting beïnvloedt. Voor de analyse in deze paragraaf is er geen koppeling gemaakt tussen de gemeten en berekende vliegtuigbewegingen. Deze manier van vergelijken heeft het voordeel dat alle vliegtuigbewegingen die in het gebruiksjaar berekend konden worden, zijn meegenomen. Tegelijkertijd heeft deze aanpak twee primaire nadelen:

1. Relatief weinig metingen worden met een groot aantal berekeningen vergeleken (vooral bij meetposten met een lage geluidbelasting).
2. Het is niet exact bekend wat de bijdrage van de niet-gemeten vliegtuigbewegingen op de *gemeten* geluidbelasting is.

Figuur 8 laat de correlatie zien tussen de berekende en gemeten geluidbelasting voor het geval dat alle berekende vliegtuigbewegingen worden vergeleken met alle gemeten vliegtuigbewegingen. Door het vergelijken van Figuur 8 met Figuur 7 kan een aantal constatering worden gedaan:

- Voor de vergelijking zonder toepassing van validatiefilters (linker en middelste plots) gaat de berekende geluidbelasting vooral bij lage geluidbelasting waarden omhoog, terwijl de gemeten geluidbelasting onveranderd blijft.
- De vergelijking van alle berekende vliegtuigbewegingen met alle gemeten vliegtuigbewegingen geeft een minder goede correlatie waarbij de R-waarden lager zijn dan bij de gekoppelde dataset.
- Door het toepassen van validatiefilters (rechter plot) verschuift de trendlijn naar links, maar niet significant naar beneden zoals voor de gekoppelde dataset werd waargenomen. Dit komt grotendeels omdat er veel meer berekeningen zijn dan metingen, ook voor de gefilterde situatie.



Figuur 8 Correlatie tussen gemeten en berekende geluidbelasting als alle vliegbewegingen worden berekend – zonder toepassing van validatiefilters op alle meetposten (links), alleen voor validatiemeetposten (midden) en na het toepassen van validatiefilters op validatiemeetposten (rechts).

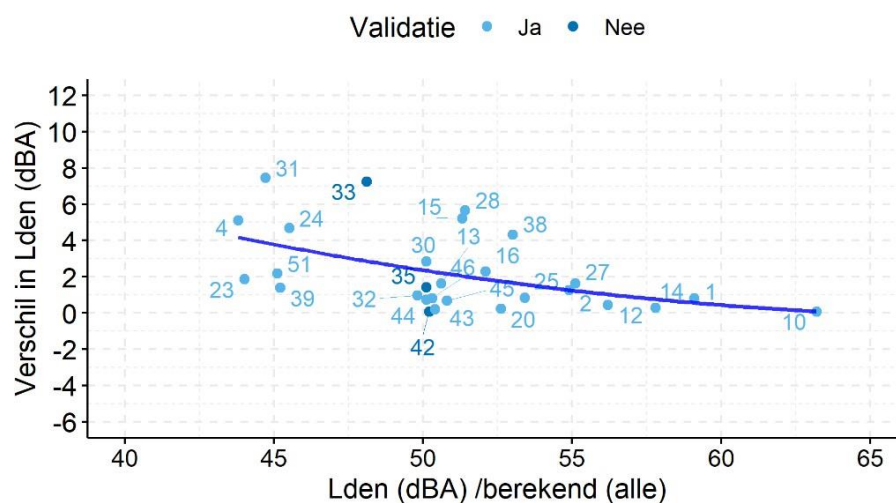
De vergelijking zoals gepresenteerd in Figuur 8 voor *alle* berekende vliegtuigbewegingen en *alle* gemeten vliegtuigbewegingen is (ook

zonder filtering) minder zuiver dan de vergelijking voor de gekoppelde dataset. Voor de gekoppelde dataset wordt hetzelfde aantal berekende en gemeten vliegbewegingen met elkaar vergeleken. Het is te zien in Figuur 8 dat de berekende geluidbelasting bij lage geluidbelasting waarden dichter bij de gemeten geluidbelasting ligt dan bij hoge geluidbelastingwaarden (linkere en middelste plots). Dit hangt mogelijk samen met de verhoogde berekende geluidbelasting bij lage L_{den} -waarden. Indien de bijdrage van de ontbrekende vliegtuigbewegingen aan de *gemeten* geluidbelasting bekend was, zouden de gemeten waarden ook omhoog gaan (verschuiving van de gemeten waarden in Figuur 8 naar rechts). Hierdoor zullen de trends in Figuur 8 meer overeenkomen met de trends gezien voor de gekoppelde dataset in Figuur 7.

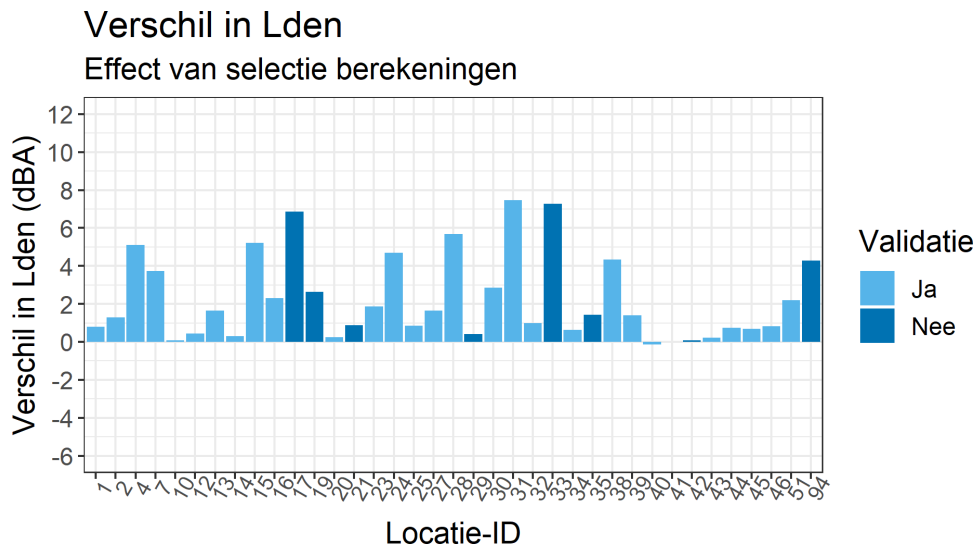
Figuur 9 en Figuur 10 kwantificeren het verschil tussen de datasets met alle berekende vliegtuigbewegingen en alleen berekende gekoppelde vliegtuigbewegingen. Figuur 9 visualiseert dit verschil als een scatter-plot als functie van de berekende geluidbelasting. Het valt in Figuur 9 op dat het verschil tussen beide datasets bij hoge geluidbelasting (vanaf circa 55 dB) minimaal of verwaarloosbaar is en met afnemende geluidbelasting geleidelijk toeneemt. Bij een geluidbelasting van 60 dB is het verschil volgens de blauwe trendlijn (fit) rond de 0 dB terwijl het verschil bij een lage geluidbelasting van 45 dB gemiddeld rond de 4 dB ligt. Deze constatering bevestigt dat het NOMOS-meetsysteem goed in staat is om bij hoge geluidbelasting de meeste vliegbewegingen (of in ieder geval de meeste vliegbewegingen die voor de geluidbelasting relevant zijn) te registreren, maar dat NOMOS bij een lagere geluidbelasting een significant deel van de relevante metingen mist. Figuur 10 presenteert het verschil tussen beide datasets als een staafdiagram met een visualisatie van het verschil per meetpost.

Verskil in L_{den}

Effect van selectie berekeningen



Figuur 9 Verskil tussen de berekende geluidbelasting voor alle vliegbewegingen en alleen de gekoppelde vliegbewegingen als scatter-plot met een lineaire, kwadratische fit. De donker blauwe punten geven de niet-validatiemeetposten aan.



Figuur 10 Verskil tussen de berekende geluidbelasting voor alle vliegbewegingen en alleen gekoppelde vliegbewegingen als staafdiagram met een opsplitsing in validatie en niet-validatiemeetposten.

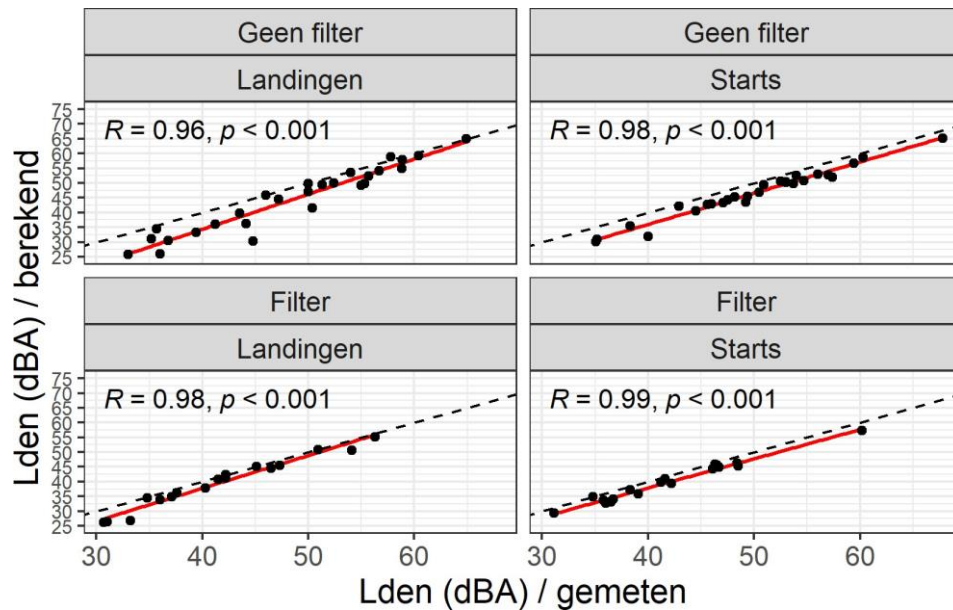
2.3.2 Resultaten voor subselecties van meetdata

In deze paragraaf wordt de analyse voor de vergelijking tussen meten en rekenen verdiept door eerst een opsplitsing te maken op basis van vluchtsoort in sectie 2.3.2.1, gevolgd door een opsplitsing op basis van vliegtuigtype in sectie 2.3.2.2. De analyse in deze paragraaf is uitgevoerd voor de gekoppelde dataset.

2.3.2.1 Analyse op basis van vluchtsoort

Het doel van de analyse op vluchtsoort-niveau is om ten eerste te identificeren of het verschil voor starts en landingen uiteenloopt en of verbetering voor een van de vluchtsoorten tot een betere totale overeenkomst tussen meten en rekenen kan leiden.

Figuur 11 laat de vergelijking op validatiemeetposten afzonderlijk zien voor starts en landingen met de bijbehorende correlatiecoëfficiënten tussen meten en rekenen op basis van vluchtsoort (ook gepresenteerd in (Sahai 2024)). De vergelijking in Figuur 11 laat duidelijk zien dat de correlatie tussen berekeningen en metingen hoog is voor zowel starts als voor landingen. Bij hoge geluidbelasting is het verschil iets groter voor starts dan voor landingen, bij lage geluidbelasting blijft het verschil voor starts vrijwel constant, maar neemt voor landingen met afnemende geluidbelasting toe. Figuur 11 geeft aan dat een verbetering voor zowel starts als voor landingen voordelig kan zijn voor het verbeteren van de overeenkomst tussen metingen en berekeningen op totaal-niveau. Wel dient deze constatering te worden gecombineerd met de vluchtsoort dat op een gegeven meetpost dominant is, dat wil zeggen een verbetering voor starts vooral relevant zal zijn op meetposten die een significant aandeel starts meten en vice-versa voor landingen.



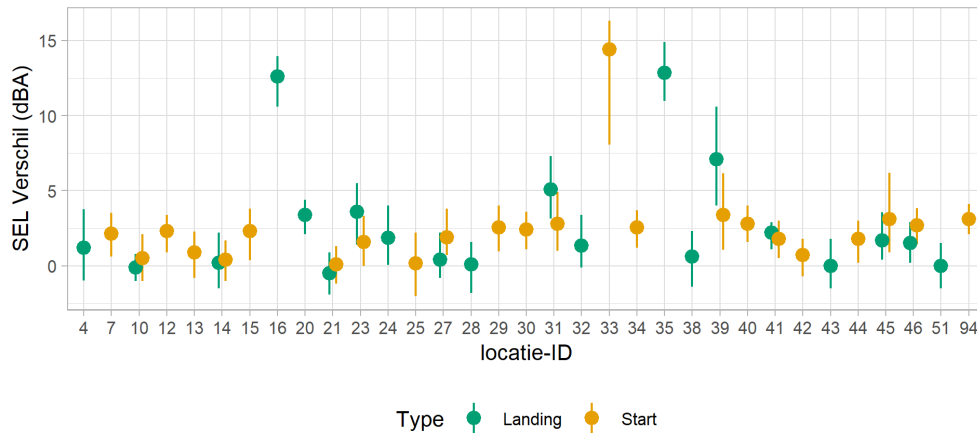
Figuur 11 Vergelijking van gemeten en berekende geluidbelasting op validatiemeetposten, opgesplitst in starts en landingen – na toepassing van validatiefilters (boven) en zonder toepassing van validatiefilters (onder).

Figuur 12 is een uitbreiding van Figuur 6 en laat de verschillen per meetpost op SEL-niveau zien, deze keer opgesplitst voor starts en landingen. In zowel Figuur 6 als Figuur 12 worden de mediaan-waarden en het interkwartiel-bereik aangetoond. Afhankelijk van de groepering die van data is gemaakt (totaal, starts of landingen), komen de mediaan-waarden en het interkwartiel-bereik anders uit. Het aantal meetposten waarop (na filtering) regelmatig zowel starts als landingen wordt gemeten, is beperkt tot tien. Op sommige meetposten zijn de verschillen groter voor starts en op anderen groter voor landingen.

Door een vergelijking van Figuur 6 met Figuur 12 kan worden gezien dat de mediaan en het interkwartiel-bereik voor de totale situatie in Figuur 6 min of meer de mediaan en het interkwartiel-bereik van de *dominerende* vluchtsoort per meetpost volgt.

Vershil in SEL na toepassing van filter

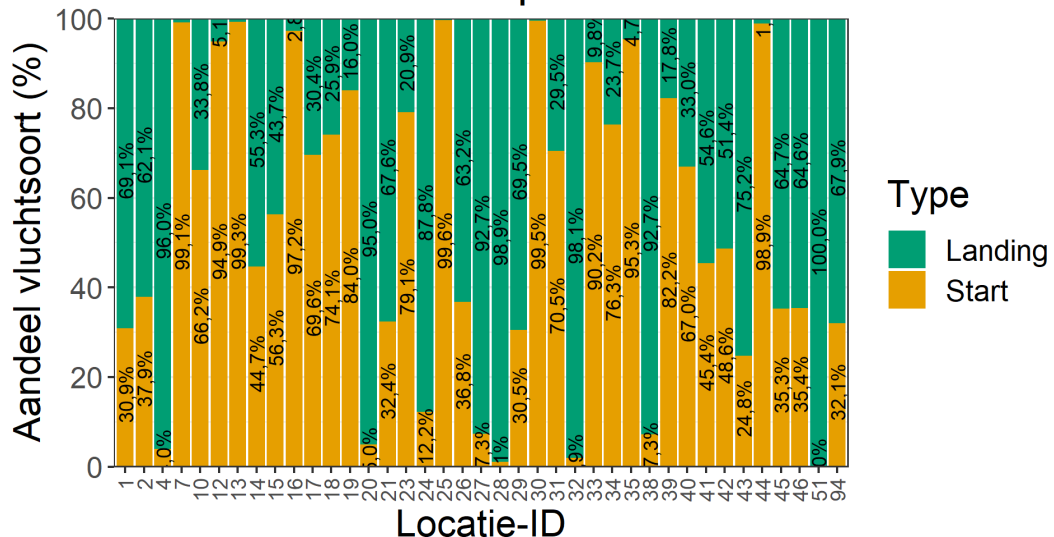
Mediaan en interkwartiel bereik



Figuur 12 Vershil in SEL-waarden per meetpost, opgesplitst op basis van vluchtsoort na toepassing van validatiefilters. De mediaanwaarde en het interkwartiel-bereik zijn weergegeven.

Figuur 13 toont welke vluchtsoort per meetpost het meest voorkomt in de gemeten waarden. Op meetposten 4 en 20 worden bijvoorbeeld voornamelijk landingen gemeten, waardoor de mediaan en het interkwartiel-bereik voor de totale situatie (in Figuur 6) meer met landingen overeenkomt. Voor meetposten 7, 12 en 13 bijvoorbeeld worden vooral starts gemeten, de mediaan en het interkwartiel-bereik in Figuur 6 komen voor deze meetposten dan meer overeen met de betreffende waarden voor starts (zie ook Figuur 12). Door de bevindingen uit Figuur 12 met Figuur 13 te combineren en te kijken naar de bijdrage van de vluchtsoort aan de totale geluidbelasting, kunnen slimme keuzes worden gemaakt in wat betreft het verbeteren van het model op basis van vluchtsoort en in waarop de focus per meetpost moet worden gelegd.

Aandeel vluchtsoort per meetlocatie

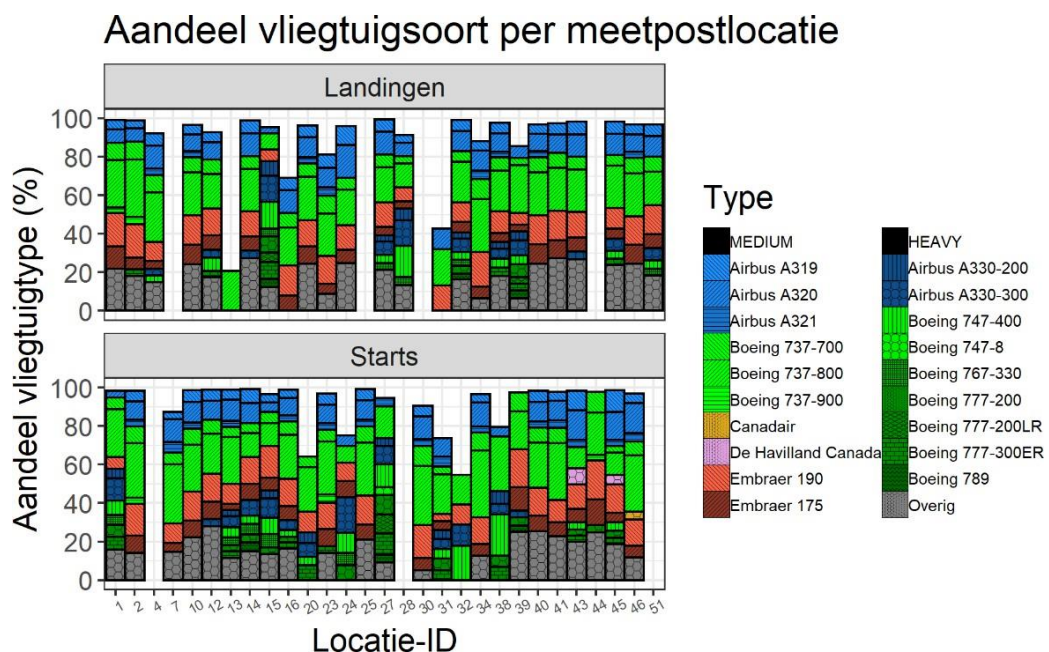


Figuur 13 Aandeel vluchtsoort per meetpost locatie op basis van aantal gemeten vliegbewegingen.

2.3.2.2 Analyse op basis van vliegtuigtype

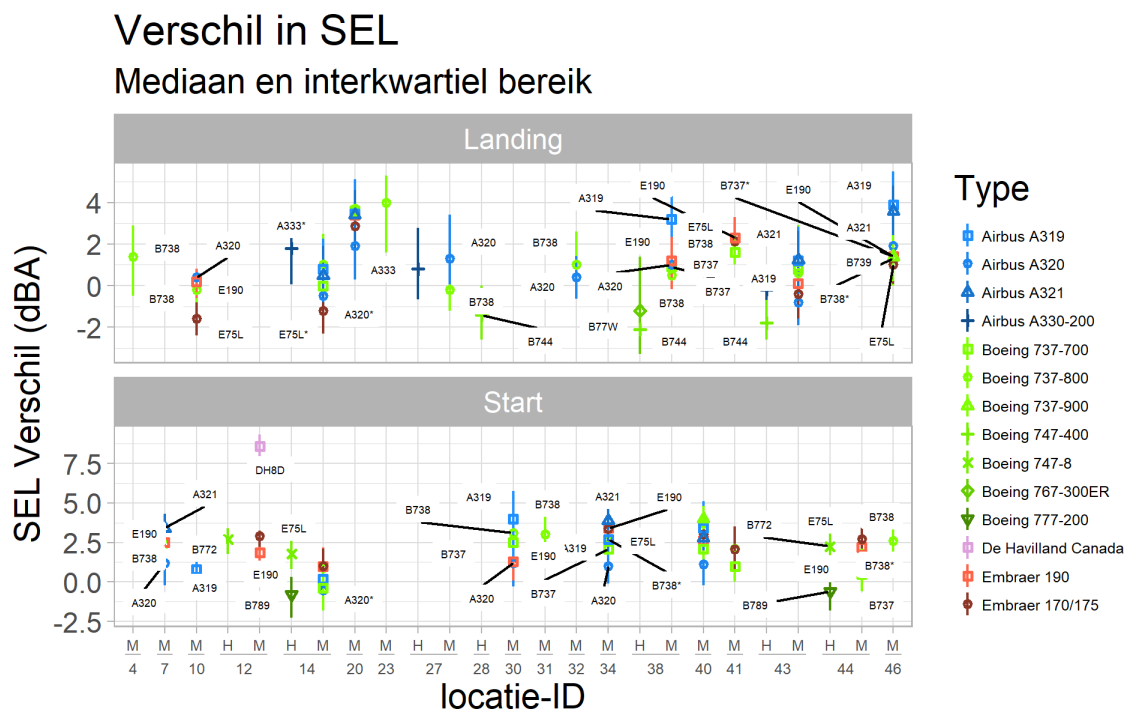
Deze paragraaf wordt de vergelijking van metingen en berekeningen nog een niveau verdiept tot vliegtuigtype-niveau. Het doel van de analyse op vliegtuigtype-niveau is om te identificeren of de verschillen voor een bepaald vliegtuigtype (structureel) groter zijn en meer aandacht verdienen bij een eventuele modelverbetering.

Net als bij de analyse op vluchtsoort-niveau is het ook voor de analyse op vliegtuigtype-niveau van belang om te weten welke vliegtuigtypen per meetpost het vaakst worden gemeten. Figuur 14 geeft een overzicht van het aandeel vliegtuigtype per meetpost, zoals geregistreerd in de ANOMS-database, opgesplitst voor starts en landingen. In Figuur 14 is te zien dat vooral vliegtuigen van Airbus, Boeing en Embraer de vloot op Schiphol domineren. Voor de gewichtsklasse 'Medium' (maximaal startgewicht tot 136000 kg) zijn dit de Boeing 737-varianten, de toestellen uit de Airbus A320-serie en de Embraer 190 en 170/175. Voor de gewichtsklasse 'Heavy' (maximaal startgewicht hoger dan 136000 kg) vliegen de Airbus A330-varianten, de Boeing 777-, 787- en 747-varianten het vaakst op Schiphol. Hiernaast is er een aantal andere vliegtuigtypen dat ook op diverse meetposten rondom Schiphol wordt gemeten, zoals van de fabrikanten Bombardier, ATR en wat oudere vliegtuigtypen, zoals de A310 die nog steeds als vrachtvliegtuigen worden ingezet. Deze typen vallen grotendeels onder de groepering 'overig'. Alle vliegtuigtypen die op een gegeven meetpost een bijdrage leveren van minder dan 3 procent van het totale vliegverkeer (gemeten op die meetpost) worden in deze categorie gegroepeerd. Figuur 14 geeft dan ook aan dat de verhouding van gemeten vliegtuigtypen per meetpost kan verschillen.



Figuur 14 Aandeel vluchtsoort per meetpost locatie, opgesplitst op basis van vliegtuigtype,

In Figuur 15 zijn de verschillen op SEL-niveau per vliegtuigtype na filtering te zien, opgesplitst naar starts en landingen. Het toepassen van de validatiefilters in hun huidige (strengere) vorm kan een significant effect hebben op het aantal beschikbare metingen, zeker als er naar kleine groepen van data zoals op vliegtuigtype-niveau wordt gekeken. Voor veel vliegtuigtypen valt het aantal beschikbare metingen na toepassing van validatiefilters onder de ondergrens van 50 vliegbewegingen. De vliegtuigtypen waarvoor dit gebeurt, staan niet in Figuur 15.



Figuur 15 Verschil in SEL-waarden per meetpost, opgesplitst op basis van vluchtsoort en vliegtuigtype na toepassing van validatiefilters. De mediaan-waarde en het interkwartiel-bereik zijn aangetoond. Categorie 'M' staat voor medium gewichtsklasse en 'H' voor heavy gewichtsklasse.

In het algemeen kan worden vastgesteld dat de verschillen voor de meeste vliegtuigtypen na het toepassen van de validatiefilters binnen een marge van -2 tot +4 dBA SEL vallen. Door het toepassen van de validatiefilters worden grotere verschillen op SEL-niveau grotendeels verwijderd. Figuur 15 laat verder zien dat er niet één vliegtuigtype aan te merken is waarvoor het verschil structureel op alle meetposten boven de rest uitsteekt. Het blijkt dat de meetlocatie waar de vergelijking wordt gemaakt een grotere invloed heeft op het waargenomen verschil. Wel kunnen de vliegtuigtypen B737-800, A319/A321 en de E175/E190 op een aantal meetposten worden gezien als de vliegtuigtypen met een iets hoger verschil tussen meten en rekenen op SEL-niveau.

Het kan waardevol zijn om in een vervolgtraject nader te onderzoeken waarom de verschillen voor deze vliegtuigtypen op deze meetposten boven gemiddeld uitkomen, zeker als ze ook een significante bijdrage leveren aan de L_{den} -verschillen op totaal-niveau. De exacte oorzaken van

de verschillen per vliegtuigtype kunnen aan diverse factoren liggen, zowel aan de rekenkant als de meetkant. Aan de rekenkant kan het bijvoorbeeld zijn dat de onderliggende geluidstabellen niet-representatief zijn voor de operationele werkelijkheid of dat de modellering van bepaalde vliegprocedures verbeterd kan worden. Hiernaast kan het verschil ook liggen aan aannames in de invoergegevens (gekozen gewichtsklasse, derating, versnelling boven 10.000 ft et cetera) zoals onder andere genoemd in het rapport over het toepassingsbereik van metingen en berekeningen (Sahai 2024). Een verdiepend onderzoek in de mogelijke factoren die tot verschillen leiden vanuit de modelleringskant valt buiten scope van de huidige opdracht voor de PAMV en is voor een vervolgtraject beoogd.

Wat betreft de meetkant is voor het validatieonderzoek al rekening gehouden met de validatie-eisen op de meetlocatie en aanvullende meeteisen aan wind, neerslag en elevatiehoek. Hierdoor zijn de factoren die tot verstoorde metingen kunnen leiden grotendeels geminimaliseerd. Wel blijft het gebruik van de drempelwaarde een factor aan de meetkant die de vergelijking van metingen met berekeningen bij lage geluidbelasting bemoeilijkt. In Figuur 16 wordt het effect van de drempelwaarde op de vergelijking op vliegtuigtype-niveau nader onderzocht door de berekende en gemeten distributies te vergelijken voor drie representatieve vliegtuigtypen op twee meetposten. De distributies worden vergeleken op meetpost 1 (Zwanenburg, gelegen in een hoog geluidbelasting gebied) en op meetpost 32 (Castricum, gelegen in een gebied met lagere geluidbelasting). Dit wordt gedaan voor de typen Embraer E190, Boeing 737-800 (B738) en de Boeing 747-400 (B744).

Door de vergelijking van berekende en gemeten SEL-distributies in Figuur 16 kunnen twee belangrijke constatering worden gedaan:

1. Het effect van de drempelwaarde verschilt per meetlocatie (conform de analyse in het onderzoek naar het toepassingsbereik): De berekende en gemeten distributies komen op meetpost 1 heel goed met elkaar overeen. Op meetpost 32 is de vergelijking minder goed, vooral voor SEL-waarden lager dan 70 dB.
2. Het effect van de drempelwaarde op de linkerkant van de gemeten distributie verschilt per vliegtuigtype: Voor een relatief licht vliegtuigtype zoals de E190 ontbreekt een significant deel van de metingen, omdat dit type verder weg van de luchthaven en in gebieden met een lage geluidbelasting lage geluidniveaus produceert die vaker onder de drempelwaarde vallen. Voor een zwaarder type zoals de B738 treedt dit effect ook op, maar in mindere mate. Ook voor dit type wordt het linkerdeel van de distributie niet gemeten, maar worden er meer hogere geluidniveaus gemeten. Een zwaar vliegtuigtype zoals de B744 veroorzaakt ook op grote afstanden en in gebieden met een lage geluidbelasting hogere geluidniveaus, waardoor een groter deel van de metingen voor dit type in het meetsysteem wordt geregistreerd.

Figuur 16 bevestigt het beeld en de constatering die in het onderzoek naar het toepassingsbereik van metingen en berekeningen is gedaan

(Sahai 2024) dat de vergelijking van metingen en berekeningen van vliegtuiggeluid bij lage geluidbelasting minder goed mogelijk is dan bij hoge geluidbelasting. Om het effect van ontbrekende gemeten vliegtuigbewegingen beter te analyseren, zou het waardevol kunnen zijn om nieuwe, meer gerichte metingen met meetposten in gebieden met een lage geluidbelasting uit te voeren. Een optie hierbij zou kunnen zijn om de meetdrempelwaarde op de gekozen locaties te verlagen of voor een korte periode bemande metingen uit te voeren. Een herhaling van de validatie met metingen die met het nieuwe detectie algoritme zijn uitgevoerd waarbij de drempelwaarde achteraf op basis van het geluid wordt bijgesteld, zou de vergelijking bij lage geluidbelasting ook kunnen verbeteren.



Figuur 16 Vergelijking van de distributies van berekende en gemeten geluidniveaus (SEL) voor de Embraer 190 (boven), Boeing 737-800 (midden) en Boeing 747-400 (onder) voor meetpost 1 (links) en meetpost 32 (rechts).

2.3.3

Discussie: in hoeverre is er sprake van een systematisch verschil?

In de analyse uit de voorafgaande paragrafen is te zien dat, alhoewel de rekenresultaten in het algemeen goed met metingen correleren, er een verschil zichtbaar is tussen beiden. Gemeten waarden van de geluidbelasting vallen hierbij doorgaans hoger uit dan berekende waarden. Het verschil (voor de gekoppelde dataset) is relatief klein bij hoge geluidbelasting en neemt met afnemende geluidbelasting geleidelijk toe. Op het eerste gezicht kan deze constatering wijzen op een systematische afwijking in het model, dat wil zeggen dat het model systematisch de geluidbelasting lager berekent dan de metingen.

Een conceptversie van dit rapport is door de PAMV-expertgroep bekeken, waarbij de expertgroep het consortium over de kwaliteit en juistheid van het uitgevoerde onderzoek heeft geadviseerd. Naar de mening van de expertgroep is het feit dat de berekeningen systematisch lager uitvallen dan de metingen een belangrijke reden dat omwonenden minder vertrouwen hebben in berekeningen dan in metingen. Zij verzoekt daarom om het verminderen en/of opheffen van het systematische verschil tussen rekenen en meten als doel op te nemen in het eventuele vervolg op de onderzoeksfase van PAMV. Hiernaast adviseert de expertgroep om een duidelijker onderscheid aan te brengen tussen het optreden van het systematische verschil en de overige spreiding tussen rekenen en meten. Verder adviseert de expertgroep om aan te geven in welke mate de onderzochte onderliggende factoren het systematische verschil en de overige spreiding kunnen verklaren.

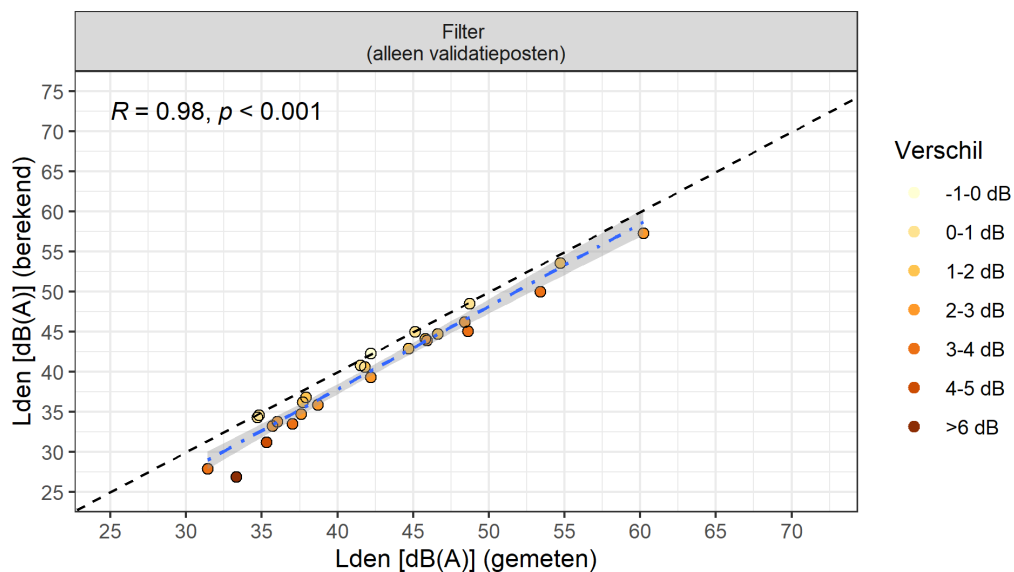
Het PAMV-consortium heeft naar aanleiding van de uitkomsten van de PAMV-uitwerkfase geadviseerd om een aantal vervolgonderzoeken uit te voeren, naast een aantal taken dat structureel kan worden uitgevoerd. Hierbij heeft het consortium geadviseerd om onderzoek naar de oorzaken van het verschil tussen meten en rekenen als vervolgonderzoek op te nemen. Hierbij zal inderdaad de focus worden gelegd op het verminderen of opheffen (zover mogelijk) van het systematische verschil, afkomstig van het model. Met het systematische verschil wordt hierbij een verschil bedoeld dat vrijwel volledig aan het model toe te schrijven is. Of en in hoeverre het model daadwerkelijk systematisch de geluidbelasting onderschat, is op basis van de uitgevoerde analyse in dit onderzoek nog niet met zekerheid te constateren. Daarvoor zou in principe nader onderzoek naar de oorzaken van het verschil nodig zijn, wat als onderdeel van vervolgonderzoek wordt beschouwd. Het aanbrengen van een duidelijker onderscheid tussen het mogelijke systematische verschil en de overige spreiding in het verschil wordt in deze paragraaf geadresseerd. Dit wordt met een (compacte) verdiepende regressieanalyse van het verschil tussen meten en rekenen gedaan. De analyse wordt opgesplitst in een analyse van het systematische verschil en een analyse van de optredende spreiding in het verschil per meetpost.

Analyse systematisch verschil

Bij het interpreteren van de metingen moeten we rekening houden met de drempelwaarde en de externe (weers)omstandigheden. Door het toepassen van de drempelwaarde op de meetwaarden is er statistisch gezien sprake van een bias in de gemeten data. In de meetdata is verder een significante bijdrage te zien van metingen die potentieel door

externe omstandigheden beïnvloed zijn en die als geluidsversterkend werken, zoals wind en lage elevatiehoekwaarden. Dit leidt voor een deel tot een oneerlijke vergelijking tussen meten en rekenen met meer meetdata beschikbaar bij hoge geluidbelasting en minder meetdata bij lage geluidbelasting. Hiernaast geldt dat het model uitgaat van ideale omstandigheden, waarbij de effecten van wind, reflecties, afscherming en dergelijke niet gemodelleerd worden. Om potentieel verstoorde metingen niet voor de vergelijking te gebruiken, zijn in dit onderzoek de validatiefilters toegepast. De toepassing van deze filters leidt tot nog minder data die voor de vergelijking beschikbaar blijven, waarbij veel metingen bij lage geluidbelasting door de filters worden uitgefilterd. Door het toepassen van de filters komen de gemeten en berekende resultaten dichterbij elkaar (zie Figuur 7).

Figuur 17 zoomt in op de rechterplot van Figuur 7 en de vergelijking tussen meten en rekenen voor de gefilterde situatie wordt afzonderlijk gepresenteerd. Hiernaast wordt in Figuur 17 aanvullende informatie gepresenteerd, zoals het verschil tussen meten en rekenen per meetpost en het betrouwbaarheidsinterval van de trendlijn over het hele meetbereik.




Figuur 17 Ingezoomde plot voor de gefilterde situatie voor het verschil tussen gemeten en berekende geluidbelasting met visuele weergave van het verschil per meetpost en het betrouwbaarheidsinterval.

Het is te zien in Figuur 17 (en ook in Figuur 7) dat er een hoge mate van voorspelbaarheid is van de berekende waarde als functie van de gemeten waarde. Dit kan met een simpele, lineaire regressie $Y = a \cdot X + b$ worden uitgedrukt. Wel bevat de voorspelling met deze trendlijn ook een bepaald niveau van onzekerheid, waarmee rekening dient te worden gehouden bij het bepalen van een mogelijk systematisch verschil in het model. Door deze onzekerheid nader te analyseren, zowel bij hoge als lage geluidbelasting, kan meer inzicht worden verkregen in het aandeel van het verschil dat systematisch door het model wordt veroorzaakt en het resterende verschil dat niet direct verklaarbaar is. In Tabel 2 is meer informatie over de trendlijn

weergegeven. Het gaat daarbij om de richtingscoëfficiënt, de interceptie, de standaardfout en het (statistisch) significantieniveau van beide componenten van de trendlijn. Hiernaast wordt ook de overall correlatie tussen meten en rekenen gegeven die voor deze trendlijn geldt.

Tabel 2 Lineaire regressie coëfficiënten voor de trendlijn voor het voorspellen van een berekende waarde met een gemeten waarde.

	Waarde	Standaardfout	Significantieniveau	
Richtingscoëfficiënt (a)	1.033	0.041	<0.0001*	
Intercept (b)	-3.51 dB	1.75 dB	0.0565	
$LLLLLLL \nearrow \text{mmmmmm} = a. \quad LLLLLLL \text{mmmmmm}$ + b		1.49 dB	<0.0001*	0.9607

De richtingscoëfficiënt van de trendlijn heeft een waarde van 1.033 ± 0.041 . De interceptie heeft een negatieve waarde van $-3.51 \text{ dB} \pm 1.75 \text{ dB}$, maar met een significantieniveau iets boven 0.05. Verder geldt er een zeer hoge mate van correspondentie tussen de gemeten en berekende geluidbelasting (zoals ook al visueel gezien in Figuur 7 en Figuur 17). Belangrijk om hierbij op te merken, is dat de voorspelling die met deze trendlijn wordt gemaakt ook een onzekerheidsmarge kent.

Volgens deze vergelijking is het verschil tussen meten en rekenen afhankelijk van de geluidbelasting. In Tabel 3 wordt het verschil voor twee representatieve waarden van de geluidbelasting getoond, bij lage en hoge geluidbelasting. Dit wordt gedaan zowel voor de gefilterde waarde van de geluidbelasting als voor de ongefilterde waarde. Zoals te zien in Tabel 3 ligt het bereik van dit verschil tussen 1.50 en 2.45 dB voor respectievelijk de hoogst gemeten L_{den} en de laagst gemeten L_{den} .

Tabel 3 Verschil tussen gemeten en berekende geluidbelasting voor twee representatieve waarden van de geluidbelasting met indicatie van de standaardfout in de voorspelling en het 95% betrouwbaarheidsinterval.

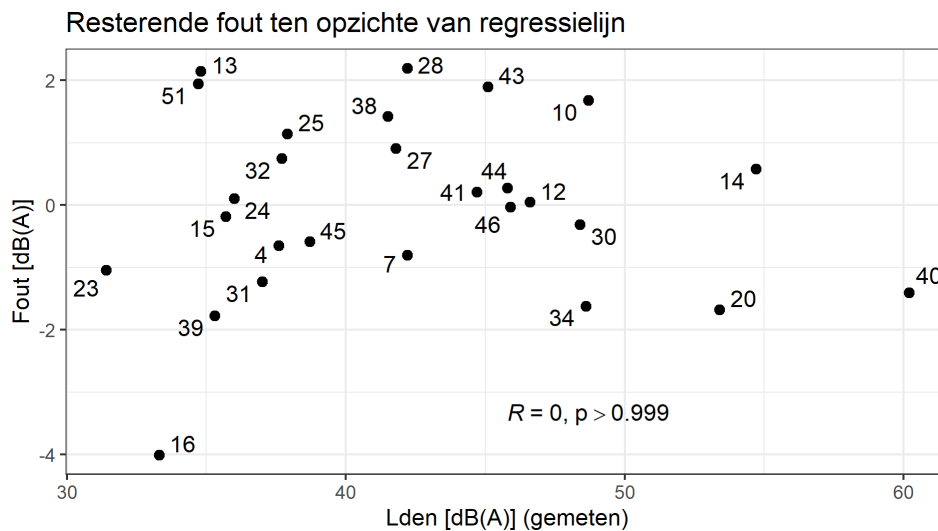
Gemeten L_{den} [dB]	Gemeten L_{den} (geen filters) [dB]	Verschil met voorspelde $LLLLLLL \nearrow \text{mmmmmm}$	Standaardfout voorspelling	95% betrouwbaarheidsinterval
31.4	42.1	+2.45 dB	0.53 dB	1.08 dB
60.2	67.9	+1.50 dB	0.79 dB	1.62 dB

Op basis van de eerder uitgevoerde analyse en onder andere Figuur 9 van dit rapport is te concluderen dat de meetdataset het meest compleet is bij geluidbelastingwaarden van hoger dan circa 50-55 dB L_{den} ongefilterd (ongeveer 40-45 dB gefilterd). De vergelijking is dan ook het meest betrouwbaar vanaf dit bereik van de geluidbelasting. De mate van onderschatting van het model is volgens de regressieanalyse dan circa 1.50 dB + een variatie als functie van L_{den} . Voor hoge geluidbelasting geldt een afwijking tussen meten en rekenen van $1.50 \pm 0.79 \text{ dB}$ en bij lage geluidbelasting van $2.45 \pm 0.53 \text{ dB}$. Rekening houdend met deze onzekerheden in de voorspelling kan het op dit moment niet met zekerheid worden gezegd of het model de geluidbelasting minimaal met 1.5 dB systematisch onderschat. Verder is het voorlopig onbekend welk aandeel van dit verschil in de voorspelde

waarde volledig aan het model kan worden toegeschreven. Bij lage geluidbelasting is te vermoeden dat een deel van het waargenomen verschil van 2.45 dB op basis van de bias in de meetdata worden verklaard. Welk aandeel bij lage geluidbelasting alleen door de modellering te verklaren is, verdient nader onderzoek. Indien bij een toekomstige validatie meer meetdata beschikbaar komt, zou de onzekerheid in de vergelijking kunnen worden gereduceerd en het systematische verschil met meer zekerheid kunnen worden bepaald.

Analyse spreiding in het verschil

De spreiding in het verschil tussen meten en rekenen per meetpost kan worden bepaald als de standaarddeviatie van de residuele fout, wat het verschil is tussen de voorspelde waarde en de gemeten waarde. Deze heeft een waarde van 1.49 dB voor het hele meetbereik van de trendlijn (zie Tabel 2) en kan worden beschouwd als de mate van variatie onverklaarbaar door het lineaire regressiemodel. Deze spreiding kan het resultaat zijn van spreiding in gemeten data die tot een zekere mate vrijwel altijd zal plaatsvinden (door bijvoorbeeld de grote afstanden tussen het vliegtuig en de meetpost en de variatie in de weersomstandigheden). De waarde van 1.49 dB geldt trouwens alleen voor de geanalyseerde data voor 2019. Voor andere jaren zou deze spreiding een andere waarde kunnen hebben. Zoals te zien is in Figuur 18 laat deze spreiding geen trend zien als functie van de geluidbelasting.



Figuur 18 Spreiding in de residuele fout tussen meten en rekenen per meetpost.

Samenvattend:

Op basis van de in deze paragraaf gepresenteerde compacte regressieanalyse is geprobeerd om nader inzicht te verkrijgen in het mogelijke systematische verschil tussen metingen en berekeningen van vliegtuiggeluid en de overige spreiding in dit verschil. Uit de analyse blijkt dat het niet eenduidig te bepalen is of en in hoeverre het model systematisch de geluidbelasting onderschat. Het verschil tussen meten en rekenen bedraagt op basis van de regressieanalyse 1.5 dB plus een variatie als functie van de L_{den} . Dit verschil neemt met afnemende geluidbelasting geleidelijk toe tot een waarde van 2.45 dB bij de laagst

gemeten L_{den} . Door de onzekerheid in de met de trendlijn bepaalde voorspelde waarde is het echter op dit moment niet mogelijk om te bepalen in hoeverre dit verschil aan het model is toe te schrijven. Om dat te bepalen, is nader onderzoek nodig naar de mogelijke oorzaken van het verschil. Daarmee kan voor verschillende geluidbelastingwaarden worden vastgesteld welk aandeel van het waargenomen verschil door de modellering wordt veroorzaakt. Wat betreft de spreiding in het verschil is een waarde van 1.49 dB bepaald, op basis van de variatie in de residuele fout per meetpost voor het gebruiksjaar 2019.

3 Conclusies en aanbevelingen

Dit onderzoek is uitgevoerd binnen de kaders van de PAMV met het doel om het Doc.29-model, specifiek de implementatie van dit model voor Schiphol, te valideren. Hiervoor zijn berekeningen die zijn uitgevoerd met het Doc.29-model vergeleken met metingen op meetposten van het NOMOS-meetsysteem rondom Schiphol.

Het onderzoek is uitgevoerd in opvolging van twee aanbevelingen gedaan tijdens de PAMV-verkenningsfase:

1. Signaalfunctie: Het ontwikkelen van een methodiek om de kwaliteit van geluidberekeningen te beoordelen.
2. Modelontwikkeling: Het starten van een structureel en langjarig modelvalidatieprogramma met de focus op het Doc.29-rekenmodel.

De overeenkomst tussen metingen en berekeningen van vliegtuiggeluid is in dit rapport op basis van een statistische analyse bepaald. Hiermee is getracht om statistisch te duiden hoe goed de samenhang tussen metingen en berekeningen is, hoe hoog het absolute verschil tussen metingen en berekeningen is en hoe groot de spreiding in dit verschil is. Daarnaast hebben de analyses in dit rapport ook het doel om nader inzicht te verkrijgen in de verschillen op totaal niveau, op vluchtsoort-niveau (starts/landingen) en op vliegtuigtype-niveau. Deze informatie kan in een eventueel vervolgtraject voor modelverbetering worden gebruikt.

Algemene conclusies

De analyse in dit onderzoek is uitgevoerd voor een dataset waarin iedere gemeten vliegtuigbeweging aan een berekende vliegtuigbeweging is gekoppeld (de zogenoemde gekoppelde dataset). Op basis van de analyse kunnen de volgende algemene conclusies omtrent de vergelijking tussen metingen en berekeningen worden getrokken:

- Er is een sterke correlatie tussen metingen en berekeningen en allebei volgen dezelfde trends. De correlatie wordt sterk verbeterd door gebruik te maken van validatiemeetposten.
- Het toepassen van de aanvullende validatie-eisen voor wind, neerslag en elevatiehoek verkleint het verschil tussen metingen en berekeningen en elimineert grotendeels de spreiding in het verschil.
- De waarde van de geluidbelasting op basis van metingen komt in het algemeen hoger uit dan de berekende geluidbelasting.
- Door het gebruik van een drempelwaarde op de NOMOS-meetposten worden geluidniveaus van lager dan ongeveer 70 dBA SEL minder vaak gemeten, terwijl deze geluidniveaus veel vaker worden berekend. Dit heeft met name op meetposten die verder van de luchthaven en in gebieden met een lage geluidbelasting liggen een effect op de vergelijking tussen meten en rekenen.

- De (medianen van de) verschillen op SEL-niveau na het toepassen van validatiefilters liggen voor de meerderheid van meetposten binnen een bereik van -2 tot +4 dBA SEL.
- Het verschil tussen metingen en berekeningen is relatief klein voor hoge geluidbelasting en neemt met afnemende geluidbelasting toe. Een regressieanalyse van het verschil laat zien dat het verschil bij hoge geluidbelasting 1.5 dB L_{den} bedraagt en neemt tot circa 2.5 dB toe bij de laagste gemeten geluidbelasting.
- De vergelijking bij lage geluidbelasting is echter minder goed mogelijk door het lage aantal metingen dat bij lagere waarden van de geluidbelasting beschikbaar blijft, primair door het toepassen van de drempelwaarde.

Conclusies op basis van vergelijking op vluchtsoort- en vliegtuigtype-niveau

Bij de analyse op vluchtsoort-niveau zijn losse analyses uitgevoerd voor starts en landingen. De analyses laten zien dat het verschil tussen metingen en berekeningen bij hoge geluidbelasting iets groter is voor landingen dan voor starts. Bij lage geluidbelasting blijft het verschil voor starts redelijk constant, maar neemt voor landingen licht toe.

De analyses op vliegtuigtype-niveau laten zien dat er geen specifiek vliegtuigtype aan te merken is waarvoor het verschil tussen metingen en berekeningen structureel hoger uitkomt. De locatie waar de vergelijking wordt uitgevoerd, lijkt bepalender te zijn voor de mate van de verschillen.

De vergelijking van berekende en gemeten SEL-distributies op vliegtuigtype-niveau leidt tot twee belangrijke constatering:

1. Het effect van de drempelwaarde verschilt per meetlocatie, waarbij de distributies goed met elkaar overeenkomen bij meetposten met een hoge geluidbelasting, maar minder goed voor meetposten met een lage geluidbelasting (vooral de linkerkant van de gemeten distributies, dat wil zeggen waarden die lager zijn dan 70 dB SEL, ontbreekt dan).
2. Het effect van de drempelwaarde op de linkerkant van de gemeten distributie verschilt per vliegtuigtype. Voor een relatief licht vliegtuigtype zoals de E190 ontbreekt een groot deel van de metingen, omdat dit type verder weg van de luchthaven en in gebieden met een lage geluidbelasting lage geluidniveaus produceert die vaker onder de drempelwaarde vallen. Voor een iets zwaarder type zoals de B737-800 treedt dit effect ook op maar in mindere mate. Een zwaar vliegtuigtype zoals de B747-400 veroorzaakt ook op grote afstanden en in gebieden met een lage geluidbelasting hogere geluidniveaus, waardoor een groter deel van de passages van dit type in het meetsysteem wordt geregistreerd.

De analyse op vliegtuigtype-niveau bevestigt de constatering uit het onderzoek naar het toepassingsbereik van metingen en berekeningen dat de vergelijking van metingen en berekeningen van vliegtuiggeluid bij lage geluidbelasting minder goed mogelijk is dan bij hoge

geluidbelasting. Dit is grotendeels toe te wijzen aan het gebruik van de drempelwaarde.

Effect van ontbrekende vliegtuigbewegingen op de vergelijking tussen meten en rekenen

Niet iedere vliegtuigbeweging kan gemeten worden op iedere meetlocatie. Indien een vliegtuig niet in de buurt van een bepaalde meetpost komt, is het geluid daar niet hoorbaar. En als het geluid wel hoorbaar is, maar niet boven de drempelwaarde komt, zal de meting niet meegenomen worden. Om het effect van ontbrekende vliegtuigbewegingen nader te analyseren, is een vergelijking gemaakt voor een dataset waarin geen koppeling is gemaakt tussen de gemeten en berekende vliegtuigbewegingen. De berekeningen in deze dataset zijn uitgevoerd voor alle vliegtuigbewegingen van en naar Schiphol in het gebruiksjaar 2019. Een grote hoeveelheid van deze vliegtuigbewegingen komt niet in de gekoppelde dataset terug, omdat ze niet boven de drempelwaarde kwamen. De vergelijking van de geluidbelasting voor alle berekende vliegtuigbewegingen en alle gemeten vliegtuigbewegingen (zonder koppeling van beiden) laat zien dat vooral de berekende waarden bij lage geluidbelasting omhoog gaan indien alle vliegtuigbewegingen worden berekend. De gemeten waarden blijven dan onveranderd. De correlatie voor deze dataset is in het algemeen minder goed dan voor de gekoppelde dataset. De op deze wijze uitgevoerde vergelijking kent twee primaire nadelen:

1. Relatief weinig metingen worden met een groot aantal berekeningen vergeleken (zeker bij meetposten met een lage geluidbelasting).
2. Het blijft onduidelijk wat de bijdrage van de ontbrekende vliegtuigbewegingen op de *gemeten* geluidbelasting is.

De vergelijking van de berekende geluidbelasting voor alle vliegtuigbewegingen en alleen gekoppelde vliegtuigbewegingen laat zien dat het verschil tussen beiden tot een geluidbelasting van ongeveer 55 dB L_{den} verwaarloosbaar is, maar neemt met afnemende geluidbelasting duidelijk toe. Bij een geluidbelasting van 45 dB L_{den} bedraagt het verschil tussen beide datasets gemiddeld 4 dB L_{den} . Deze analyse bevestigt dat het NOMOS-meetsysteem een aanzienlijk aantal relevante vliegtuigbewegingen bij lage geluidbelasting niet kan meten.

3.1 Aanbevelingen

Toekomstige vergelijking meten en rekenen

Voor toekomstige vergelijkingen tussen berekeningen en metingen van vliegtuiggeluid kan op basis van het validatieonderzoek een aantal aanbevelingen worden gedaan:

- Bij vergelijkingen tussen meten en rekenen dient rekening gehouden te worden met het effect van de drempelwaarde op de gemeten geluidbelasting. Als blijkt dat op bepaalde locaties geen goede vergelijking gemaakt kan worden door de drempelwaarde, kan onderzocht worden of de drempelwaarde op specifieke locaties kan worden verlaagd. Hierbij is het van belang om vast te stellen in hoeverre de lagere drempelwaarde niet leidt tot minder nauwkeurige meetresultaten doordat metingen van lage geluidniveaus gevoeliger zijn voor stoorgeluid.

- Het wordt aanbevolen om de vergelijking in ieder geval voor het jaar 2022 en later te herhalen om het effect van een nieuw detectie algoritme voor metingen op de NOMOS-meetposten te analyseren. Indien door toepassing van dit algoritme meer (betrouwbare) data voor de vergelijking bij lage geluidbelasting beschikbaar komen, zal dat de statistische vergelijking tussen meten en rekenen versterken. Tevens zou meer data bij lage geluidbelasting helpen om het toepassingsbereik van het model concreter te formuleren.
- Net als genoemd in het onderzoek naar het toepassingsbereik van metingen en berekeningen is het op dit moment niet bekend of de uitkomsten van de vergelijking voor Schiphol ook voor de resterende regionale luchthavens gelden. Het wordt aanbevolen om de nut en noodzaak van modelvalidatie voor regionale luchthavens vast te stellen.

Ontwikkeling signaalfunctie voor structurele validatie

Het validatieonderzoek heeft de basis gelegd voor de ontwikkeling van het validatiesysteem (de zogenoemde signaalfunctie) waarmee de verschillen tussen meten en rekenen structureel gemonitord kunnen worden. Ook kan met dit systeem een eventuele toekomstige toename van de verschillen worden gesignaleerd. Het wordt aanbevolen om het werk dat voor het validatieonderzoek is uitgevoerd te gebruiken voor de ontwikkeling van de signaalfunctie. De signaalfunctie zal dan monitoren in hoeverre de verschillen tussen meten en rekenen voldoende klein blijven, of dat juist onderzocht moet worden waarom de verschillen groter zijn geworden dan gewenst en wat daarvoor de mogelijke oplossingen zijn. Het PAMV-consortium heeft over de signaalfunctie en structurele monitoring van het verschil tussen meten en rekenen een apart advies voor het ministerie van IenW opgesteld, waarin diverse mogelijkheden voor een structurele validatie worden toegelicht.

Uitvoering modelverbetering

Alhoewel het validatieonderzoek heeft gedemonstreerd dat rekenresultaten van het Doc.29-model een goede overeenkomst met metingen laten zien ligt het verschil voor ongeveer een derde van de validatiemeetposten hoger dan 2 dB L_{den} . Conform de PAMV-validatieaanpak dient bij verschillen tussen gemeten en berekende geluidbelasting van groter dan 2 dB L_{den} gezocht te worden naar mogelijkheden voor modelverbetering of naar een verklaring voor het verschil in de meetwaarden. Deze stap is in het huidige onderzoek niet uitgevoerd. Het wordt aanbevolen om deze verschillen in een vervolgtrajec voor modelverbetering nader te onderzoeken. Tijdens de modelverbetering kan het ook waardevol zijn om de vergelijking van het model voor een ander gebruiksjaar dan 2019 aanvullend uit te voeren om de robuustheid van het model te toetsen.

Referenties

BAC-Australia (2005). Volume D: Airspace - Aircraft Noise Modeling Methodology. Brisbane, Australia.

ECAC.CEAC (2016a). Report on Standard Method of Computing Noise Contours around Civil Airports, ECAC.CEAC Doc 29. Neuilly-sur-Seine, European Civil Aviation Conference. **Volume I: Applications Guide.**

ECAC.CEAC (2016b). Report on Standard Method of Computing Noise Contours around Civil Airports, ECAC.CEAC Doc 29. Neuilly-sur-Seine, European Civil Aviation Conference. **Volume II: Technical Guide.**

EU (2015). Commission Directive (EU) 2015/996 of 19 May 2015 establishing common noise assessment methods according to Directive 2002/49/EC of the European Parliament and of the Council, European Union.

Eurocontrol (2019). 'The Aircraft Noise and Performance (ANP) Database : An international data resource for aircraft noise modellers.' 2.2. from <https://www.aircraftnoisemodel.org/>.

FAA (2008). Integrated Noise Model (INM) Version 7.0 Technical Manual. Washington DC, USA, Office of Environment and Energy, FAA.

Giladi, R., Menachi, E. (2020). "Validating Aircraft Noise Models." MDPI proceedings.

Heblij, S. J., J. Derei and R. H. Hogenhuis (2017). Toepassing ECAC Doc29 voor het bepalen van de geluidbelasting van het vliegverkeer van Schiphol - Methode zoals gevolgd bij de MER NNHS Schiphol. Amsterdam, Nationaal Lucht- en Ruimtevaartlaboratorium.

Hogenhuis, R. H. and S. J. Heblij (2018). Trendvalidatie van Doc.29-berekeningen. Amsterdam, Nationaal Lucht- en Ruimtevaartlaboratorium.

ICAO (2018). Recommended Method for Computing Noise Contours Around Airports. Montreal, Canada, ICAO.

ISO-20906/TC (2009). Acoustics — Unattended monitoring of aircraft sound in the vicinity of airports. Geneva, International Organization for Standardization.

Joosten, E., Mabjaia, N., van Loon R., den Hollander, H. (2020). Geluidmonitor 2019 : Meting en validatie van geluidproductie door rijkswegen en spoorwegen, RIVM.

KNMI (2023). Uurgegevens van het weer in Nederland. De Bilt, KNMI.

Merino-Martínez, R., et al. (2019). 'Improving Aircraft Noise Predictions Considering Fan Rotational Speed.' Journal of Aircraft **56**(1): 284-294.

Ministerie van Infrastructuur en Waterstaat, I. (2018). Kamerstuk - Over de rol die metingen van vliegtuiggeluid kunnen spelen in relatie tot berekeningen ervan. Den Haag, Tweede Kamer der Staten Generaal 2018–2019, 31936, nr. 518.

Ministerie van Infrastructuur en Waterstaat, I. (2019). Kamerstuk - Rapport verkenningfase 'programmatische aanpak meten vliegtuiggeluid' en monitoring- en evaluatieprogramma Lelystad d.d. 18 december 2019. I. e. Waterstaat. Den Haag, Tweede Kamer der Staten Generaal 31936-711.

Noise, I. T. S. (1978). Acoustics — Procedure for describing aircraft noise heard on the ground, ISO.

Plotkin, K. J., Page, J.A., Gurovich, Y., Hobbs, C.M. (2013). Detailed Weather and Terrain Analysis for Aircraft Noise Modeling. Cambridge MA, Wyle Environmental and Energy Research & Consulting (EERC).

Rhodes D. P., e. a. (2018). Guidance on comparing calculated aircraft noise levels with measurements, UK Civil Aviation Authority.

Rhodes, D. P., White, S., Havelock, P. (2002). VALIDATING THE CAA AIRCRAFT NOISE MODEL WITH NOISE MEASUREMENTS, CAA, UK.

Sahai, A., Hogenhuis, R., Heblj S.J., Vinkx K. (2022). Validatieopstartfase - Voorstel voor een validatiemethodiek Programmatische Aanpak Meten Vliegtuiggeluid (PAMV), RIVM, NLR, To70.

Sahai, A., Hogenhuis, R.H., Heblj S.J., Smetsers R., Assink, J. (2021). Programmatische Aanpak Meten Vliegtuiggeluid: Nationale meetstrategie. Kaders en inventarisatie meetsystemen, RIVM.

Sahai, A., Mabjaia, N., Wartenberg, T. Hogenhuis, R., Heblj S.J., Vinkx, K. (2024). Programmatische Aanpak Meten Vliegtuiggeluid (PAMV): Toepassingsbereik metingen en berekeningen van vliegtuiggeluid, RIVM.

Schiphol (2022). Verbeterde geluidsmetingen in omgeving Schiphol. Amsterdam, Schiphol.

Simons, D. G., et al. (2015). 'Assessment of noise level variations of aircraft fly-overs using acoustic arrays.' *Journal of Aircraft* **52**(5): 1625-1633.

Smetsers, R., et al. (2019). Vliegtuiggeluid: meten, berekenen en beleven - Een verkenning van wensen en ontwikkelopties. Bilthoven, RIVM: 164.

Trow, J., Allmark, C. (2018). The benefits of validating your aircraft noise model. *Euronoise 2018*. Crete, EAA.

Dit is een uitgave van:

**Rijksinstituut voor Volksgezondheid
en Milieu**

Postbus 1 | 3720 BA Bilthoven

Nederland

www.rivm.nl

maart 2024

**De zorg voor morgen
begint vandaag**