Contribution of aircraft types to noise levels across the NOMOS network of Amsterdam Airport Schiphol

Simons, Dick G.; Amiri-Simkooei, Alireza; Melkert, Joris A.; Snellen, Mirjam

DOI
10.1016/j.jairtraman.2025.102910

Publication date
2026

Document Version
Final published version

Published in
Journal of Air Transport Management

Citation (APA)
Simons, D. G., Amiri-Simkooei, A., Melkert, J. A., & Snellen, M. (2026). Contribution of aircraft types to noise levels across the NOMOS network of Amsterdam Airport Schiphol. *Journal of Air Transport Management*, *131*, Article 102910. https://doi.org/10.1016/j.jairtraman.2025.102910

Journal of Air Transport Management 131 (2026) 102910

Contents lists available at ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Contribution of aircraft types to noise levels across the NOMOS network of Amsterdam Airport Schiphol

Dick G. Simons [a], Alireza Amiri-Simkooei [a,*], Joris A. Melkert [b], Mirjam Snellen [a]

[a] Operations and Environment, Department of Control and Operations, Faculty of Aerospace Engineering, Delft University of Technology, The Netherlands
[b] Flight Performance and Propulsion, Department of Flow Physics and Technology, Faculty of Aerospace Engineering, Delft University of Technology, The Netherlands

ARTICLE INFO

Dataset link: https://www.eans.net/

Keywords:
Airport noise monitoring
Day–evening–night average level (Lden)
NOMOS measurement system
Fleet composition Schiphol
Bounded least squares
Certification-based weighted constraints

ABSTRACT

The increase in flight volumes in the aviation industry has significant socioeconomic implications that affect different aspects of our communities and economies. Although it has great economic benefits, it also causes annoyance and disturbance to communities living near airports. The latter requires understanding and prediction of the varying noise levels generated by various aircraft types. Noise assessment on a fleet level is traditionally achieved by using prediction models such as the DOC29. Such models need to be validated using real measurements. For Amsterdam Schiphol Airport, the so-called NOMOS (Noise Monitoring System) with 39 measurement stations is used for this purpose. We analyze the time series of these stations, collecting annual data for the period from 2006 to 2023. The main objective is to determine how the aircraft-generated noise at these stations can be assigned to 13 different aircraft types, taking into account the different noise levels produced by each aircraft type. This is performed by time series analysis of individual stations and the averaged time series over all stations. The results from two least-squares methods, namely unconstrained least squares (LS) and a proposed bounded least squares subject to weighted constraints (BLS + WC), are compared. The constraints are based on certification data as prior information in the least squares method, which is expected to enhance the model's performance. Based on the above two least squares methods, predictions are performed for 2022 and 2023. The results clearly demonstrate the superiority of the BLS + WC over the LS method. We further extend our analysis to predict noise levels for a hypothetical future year with more newer aircraft models. The results indicate a substantial reduction in the noise level compared to 2023. These findings can thus underscore the effectiveness of the proposed method in outperforming the LS and highlight the model's capability to forecast the impact of fleet modernization on noise reduction.

1. Introduction

Since its start, the impressive growth of the aviation industry has resulted in significant economic benefits. This advantage comes at the price of high and increasing levels of aviation-induced noise, resulting in annoyance and health problems (Franssen et al., 2004; Hansell et al., 2013; Basner and McGuire, 2018). Aircraft noise can adversely influence public health, although its impact is often less pronounced than that of individual lifestyle factors. Consequently, communities near airports now demand reductions in aviation activities.

The concerns are widely recognized, and measures to counteract the aviation-induced noise are taken by aircraft manufacturers, airliners and airports. There are continuous efforts to make aircraft more quieter, e.g. by increasing the bypass-ratio of the turbofan engines and applying acoustic lining (see e.g. Bertsch et al., 2015). In addition, flight procedures and operations producing less noise in populated areas have been established. The implementation of these measures is enforced through charges and regulations (see Morrell and Lu, 2000; Schiphol Airport, 2021).

Extensive research activities have been ongoing globally, particularly within the European Union through initiatives such as Aviation Noise Impact Management through Novel Approaches (ANIMA) (ANIMA Project Consortium, 2025) and Aircraft noise Reduction Technologies and related Environmental iMpact (ARTEM) (von Karman Institute for Fluid Dynamics, 2025) projects, as well as internationally by organizations such as the International Civil Aviation Organization (ICAO), the National Aeronautics and Space Administration (NASA) in the United States, and the Aviation Sustainability Center (ASCENT), a

* Corresponding author.
E-mail addresses: d.g.simons@tudelft.nl (D.G. Simons), a.amirisimkooei@tudelft.nl (A. Amiri-Simkooei), JAMelkert@tudelft.nl (J.A. Melkert), m.snellen@tudelft.nl (M. Snellen).

https://doi.org/10.1016/j.jairtraman.2025.102910
Received 14 January 2025; Received in revised form 16 August 2025; Accepted 22 September 2025
Available online 7 October 2025

research initiative supported by the Federal Aviation Administration (FAA). These efforts aim to further understand and hence mitigate aircraft noise and its associated health and environmental impacts.

These regulations are such that hard limits are imposed on the yearly cumulative noise levels at various locations around airports. These noise levels are traditionally calculated using so-called best-practice models or regulatory models, see e.g. Isermann and Vogelsang (2010). These models are based on legal compliance requirements, such as those described in Document 29 of the European Civil Aviation Conference (ECAC) (ECAC.CEAC, 2016), and are capable of calculating noise contours around airports with low computational cost and limited model inputs. The resulting contours, representing the noise impact of aircraft operations over large areas and e.g. for a full year, are typically employed to check compliance with noise limits and to estimate future aircraft impacts.

Various noise metrics can be used as a measure for annoyance (Vieira et al., 2020). For current regulatory purposes aircraft noise contours using the L_{den} metric (and the similar L_{night} metric) are commonly considered. L_{den} stands for the yearly-averaged day–evening–night average noise levels (Simons et al., 2022). However, L_{den} is not a direct measure of annoyance (Breugelmans et al., 2013; Gjestland, 2024). While it serves as a useful indicator of noise exposure, further processing and assumptions, such as population density and the correlation between L_{den} and reported annoyance, are needed to estimate the actual level of annoyance in a population.

We emphasize that the noise indicators used in this study are L_{den} and L_{night}, in accordance with the European Environmental Noise Directive (European Parliament and Council, 2002), where L_{den} incorporates specific time-weighted penalties for evening and night periods, unlike L_{day} which refers only to the daytime average. Noise monitoring using best-practice models, which inherently employ significant approximations, can result in deviations of the model predictions from the actual noise levels, which, in turn, has given rise to distrust in communities near airports (Bewoners Aanspreekpunt Schiphol, 2021). Obviously, validation of these models against real measurements can possibly reduce this distrust (Simons et al., 2022).

Previous research has investigated the relation between aircraft noise exposure and its impacts using noise metrics such as L_{den} and related measures derived from long-term monitoring data. For example, Zaporozhets (2016) highlights the shift in aircraft noise annoyance from traditional high-exposure zones to wider areas. They highlighted increased community reactions beyond standard noise contours. In another study, the frequency of noise-induced annoyance was examined in relation to the number of aircraft movements, leveraging noise metrics and operational data to quantify community exposure levels (Gjestland and Gelderblom, 2017). The U.S. aircraft noise policy was reevaluated, highlighting that changing flight patterns influence community annoyance (Cointin and Hileman, 2016). Although noise levels may have decreased due to quieter aircraft, the increased volume of flights and changes in arrival and departure patterns are significantly affecting public perception. Recent field data indicate rising aircraft noise annoyance at constant L_{pAeq} levels, likely due to increased flight movements, fleet changes, and shifting public attitudes (Guski, 2017). This suggests revisiting noise metrics and regulatory thresholds. The above-mentioned studies illustrate the important role of empirical noise metrics and time series data in modeling aircraft noise exposure and its social impacts. However, a direct link to fleet composition changes remains a less explored area, which our study aims to address.

Our study fills the above gap by introducing an empirical framework that interlinks year-to-year changes in measured L_{den} levels to the evolving fleet composition. To improve physical feasibility, we introduce a bounded least squares method with weighted constraints (BLS + WC) that considers aircraft certification levels as soft constraints. While regular regression is common in general modeling (Freedman, 2009), its application to aircraft noise monitoring has not yet been explored. This approach is novel in two aspects: (1) it uses a unique long-term and high-quality dataset of actual Lden measurements from 2006 to 2023 at one of the busiest airports in Europe; and (2) it integrates certification-related constraints into an empirical trend model, which balances model fidelity and its physical feasibility. These aspects not only allow for the explanation of past trends, but also simulate future scenarios based on hypothetical fleet changes. The results presented in this paper are part of a bigger research program where, among other things, the experimental data analyzed here is used for validation of the best-practice models. Such model-data comparisons are not presented here. However, in this contribution we describe the development of an empirical model that relates the measured L_{den} trend with the changing fleet composition at Schiphol Airport.

The paper is organized as follows. Section 2 briefly describes the NOMOS measurement system, and the data obtained from it. We also present the classification of different aircraft types along with their certification data and maximum take-off weights (MTOW). In Section 3, we present the methodology, which introduces two least-squares-based methods: (1) the (unconstrained) least squares (LS) method and (2) the bounded least squares subject to weighted constraints (BLS + WC) method. The latter is particularly of interest as it incorporates aircraft certification data as soft constraints in the least squares framework. This allows to derive an empirical model that correlates the measured L_{den} trend with the various aircraft types in the fleet at Schiphol over time. In Section 4, the results are presented and discussed. The section begins with a time series analysis of the measured yearly-averaged L_{den} and L_{night} data, making a distinction between aircraft noise and total measured noise. Linear least-squares curve fitting is applied to the data from NOMOS stations, averaged over all stations. Also, the precision of the obtained trends are determined. Additionally, maps using the data from individual NOMOS stations are presented, illustrating the obtained trends in L_{den} over the area. As one of the scenarios, the model is determined for the period 2006–2021 and is used to predict the situation for 2022 and 2023, comparing the predictions to actual data. We further explore the application of the predictive model to forecast values for a hypothetical future year where no empirical data is available but some old aircraft are replaced by newer ones. Section 5 summarizes the results and draws some conclusions.

2. Data and background

2.1. Noise monitoring system

This study uses data available at the website (European Aircraft Noise Services, European Aircraft Noise Services (EANS), 2021), which provides yearly-averaged L_{den} and L_{night} values obtained from the Noise Monitoring System (NOMOS) installed around Schiphol Airport. The NOMOS system consists of 43 measurement stations positioned in the Schiphol area, see Fig. 1.

Basically, each station is a calibrated microphone mounted on a 6–10 m high mast, which is either connected on a roof of a building or just on the ground. The microphones continuously measure the noise in the environment. We consider the yearly-averaged L_{den} and L_{night} data from the above mentioned website for the period 2006–2023 (see Crocker, 2008). A distinction is made between L_{den} and L_{night} data for aircraft noise events only and for the total noise measured.

The main focus of the present contribution is on modeling L_{den} for aircraft noise only. The L_{den} metric has units dBA. As already mentioned, contours of L_{den} (and L_{night}) are calculated using best-practice models for the area around an airport. Typically, L_{den} values lie in the range 40–70 dBA.

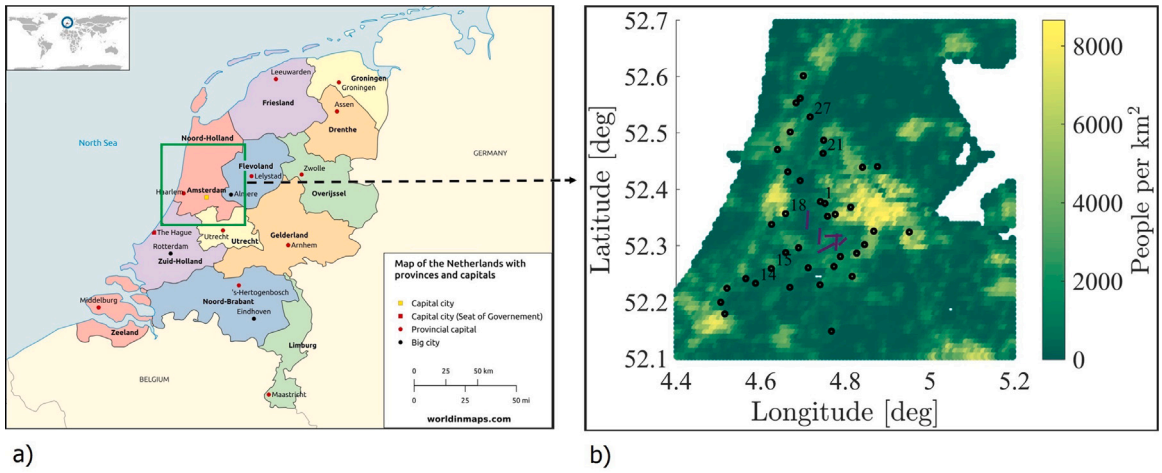

Fig. 1. Map of the Netherlands showing provinces, capital cities, and the location of NOMOS network around Schiphol Airport (a), WorldInMaps (2024); zoomed-in population density map of the Schiphol region with NOMOS stations, which highlights residential exposure (b), Simons et al. (2022).

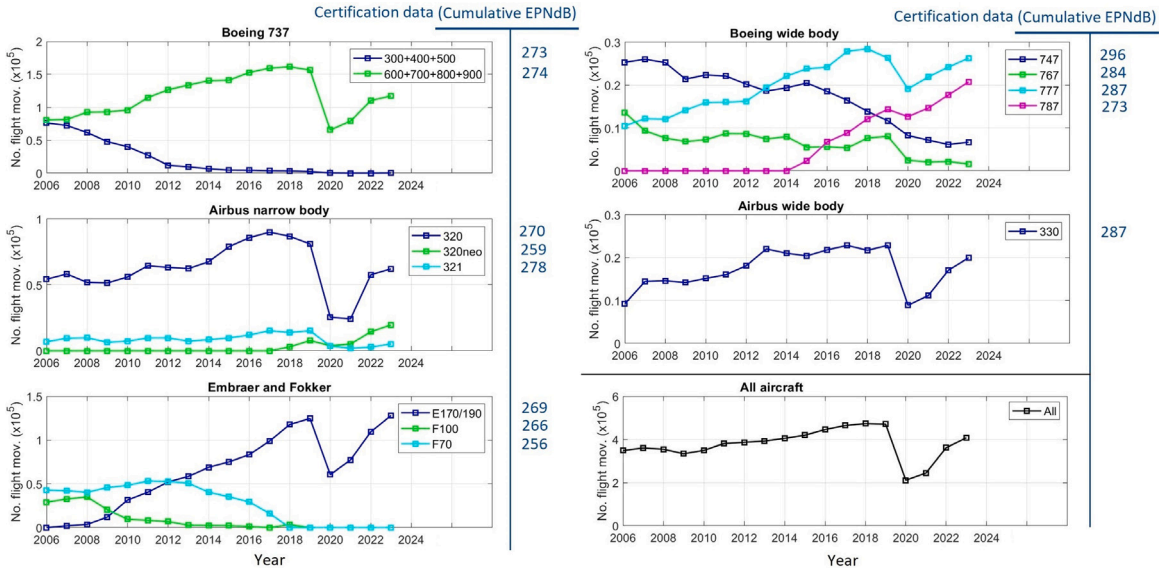

Fig. 2. Number of aircraft, classification and certification data around Schiphol Airport for the period of 2006–2023. The analysis focuses on 13 aircraft types, excluding recent additions due to insufficient data. The bottom-right panel presents the total number of operations across all 13 aircraft types.

2.2. Aircraft types and certification data

To link the above L_{den} data to aircraft types, we need details on the number of aircraft types over the years. The yearly statistics on aircraft movements by type are published by Amsterdam Airport Schiphol in their publicly available Traffic Review reports (Royal Schiphol Group, 2024). There are in total $p = 13$ aircraft types over $m = 18$ years from 2006 to 2023. These types are derived from larger fleet categories, namely Boeing 737 (2 types), Airbus narrow-body (3 types), Embraer and Fokker (3 types), Boeing wide-body (4 types), and Airbus wide-body (1 type), see Fig. 2. We excluded a few aircraft types, including the Embraer E295, Airbus A321neo, Airbus A350, and Boeing MAX 8, due to insufficient data. These types are recent additions to Schiphol airport's aviation sector, resulting in a total of $p = 13$ analyzed aircraft types.

Table 1 provides the certification data L_{EPN} for various aircraft types operating at Schiphol during take-off and landing, expressed in effective perceived noise level (EPNdB). Only the most commonly occurring aircraft types, as listed above, are included. For each aircraft, noise certification data are provided at three positions relative to the flight path: lateral, takeoff, and approach. Noise certification standards and corresponding data (L_{EPN} at lateral, takeoff, and approach) were taken from the official International Civil Aviation Organization (ICAO) (International Civil Aviation Organization, 2017), and from the EASA Type Certificate Data Sheets for Noise (TCDSN) (European Union Aviation Safety Agency, 2025). For each aircraft type, cumulative EPNdB values were derived by first plotting the L_{EPN} values against the certified maximum take-off weights (MTOW) available in the certification data, fitting a line using the least-squares method, and then taking the midpoint of the fitted line as the representative noise level for that type, see Simons et al. (2022). These EPNdB values are *cumulative*, meaning they are simply summed to represent total noise exposure.

Table 1 also includes each aircraft's MTOW and indicates whether the aircraft is classified as older or newer, with newer models typically being quieter at the same MTOW. As previously mentioned, the aircraft types are grouped into five main categories: Boeing 737, Airbus narrow-body, regional jets, Boeing wide-body, and Airbus wide-body. Fig. 3 shows the cumulative certification data plotted against the corresponding MTOW values. Blue markers represent the data for older

Table 1
Certification data L_{EPN} and maximum take-off weight (MTOW) for various operational aircraft types (older and newer) taking off and landing at Schiphol.

Aircraft category	Aircraft type	MTOW (Tonne)	Cumulative (EPNdB)	Old/new	Used here
Boeing 737	B737–300	61	273.5	old	✓
	B737–400	66	274.9	old	✓
	B737–500	58	271.2	old	✓
	B737–700	62	272.3	old	✓
	B737–800	74	275.3	old	✓
	B737 MAX	80	265.0	new	×
Airbus narrow-body	A319	66	269.0	old	×
	A320	75	272.3	old	✓
	A321	87	279.3	old	✓
	A320neo	75	258.0	new	✓
	A321neo	93	268.0	new	×
Embraer and Fokker (Regional jets)	E170/175	38	269.7	old	✓
	E190/195	48	269.2	old	✓
	F100	45	266.3	old	✓
	F70	37	255.7	new	✓
	E295	58	257.0	new	×
Boeing wide-body	B747–400	395	299.1	old	✓
	B767–300	182	283.7	old	✓
	B777–200	298	285.7	new	✓
	B777–300	349	289.1	new	✓
	B787–9	253	273.8	new	✓
Airbus wide-body	A330	235	287.0	old	✓
	A350	250	272.0	new	×

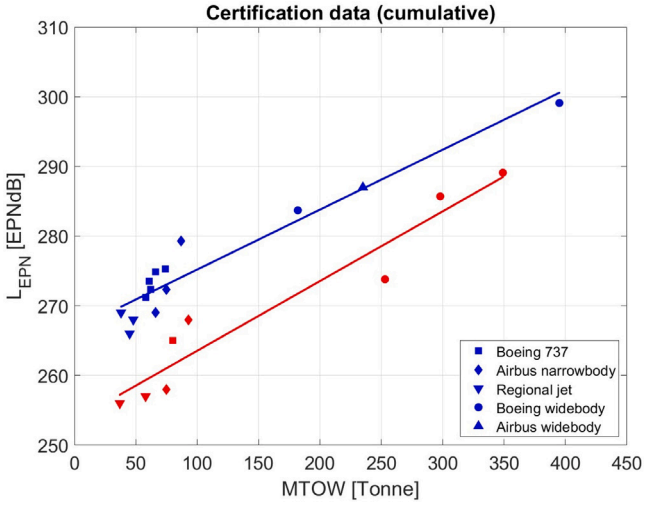

Fig. 3. Certification data L_{EPN} versus maximum take-off weight (MTOW) for various aircraft types taking off and landing at Schiphol. Blue and red data points indicate old(er) and new(er) aircraft types, respectively; lines indicate the least squares fit.

aircraft from Table 1, while red markers correspond to newer aircraft. As before, the aircraft are grouped into five categories using distinct symbols. Linear fits are applied separately to the data for old and new aircraft (see Section 3); the fitted lines are also plotted in the figure. It is evident that newer aircraft exhibit significantly lower cumulative noise levels, by approximately 10 EPNdB, compared to older models at similar MTOW values. This highlights the noise reduction achieved in newer designs.

3. Methodology

This section provides the methodologies used in this paper. It mainly includes the least squares (LS) method and the bounded least squares subject to weighted constraints (BLS + WC).

3.1. Establishing a linear observation model

The time series measurements are the yearly-averaged L_{den} and L_{night} in dBA of the NOMOS stations. Two linear models of observation equations will be used to conduct the L_{den} time series analysis. The first model is just the linear regression model as

$$L_{den}(t_i) = L_0 + r\ t_i \tag{1}$$

where L_0 is the intercept, r is the rate (two line parameters), and t_i is the time at year i. This leads to the ith row of the design matrix as $A_i = [1, \quad t_i]$, where $i = 1, \ldots, m = 18$. The results presented in Section 4.1 are based on the above linear model.

The results of Section 4.1 will indicate a decreasing trend in the observed L_{den} data. We therefore aim to investigate a possible correlation between the observed trend in the measured L_{den} over the years with changes in fleet composition at the airport. A question arises regarding the extent to which the observed noise reduction can be attributed to the quantity and quality (indicated by certification data) of the most commonly used aircraft types taking off and landing at Schiphol. To establish a linear model of observation equations for L_{den}, we may now write (a similar formulation can be developed for L_{night})

$$L_{den}(t_i) = \alpha_0 + \alpha_1 N_1(t_i) + \cdots + \alpha_p N_p(t_i) \tag{2}$$

where $N_j(t_i)$ is the number of aircraft type j, $j = 1, \ldots, p = 13$ for the given period from $t_1 = 2006$ to $t_m = 2023$ (see Fig. 2). The above equation leads to the ith row of the design matrix A in Eq. (4) as

$$A_i = [1, \quad N_1(t_i), \ \ldots, \ N_p(t_i)] \tag{3}$$

where i runs from 1 to m. In this context, the design matrix, also referred to as the coefficient matrix in estimation theory, includes a column of ones to represent the offset term α_0 introduced in Eq. (2), in addition to the columns corresponding to the explanatory variables N_i's.

The above linear relationships (i.e. Eqs. (1) and (2)) can be rewritten in a compact matrix form using the following linear model of observation equations:

$$y = Ax + e, \quad \mathrm{D}(y) = Q_y \tag{4}$$

where y is a vector of m observations (L_{den} values), e is a vector of m residuals, x is a vector of n unknown parameters, and A is an $m \times n$

design matrix. Q_y is the $m \times m$ covariance matrix of observations y and D(.) denotes the dispersion operator. In the model of $y = Ax + e$, $\mathrm{D}(y) = Q_y$, all the lowercase italic variables like y, e, and x are vectors, and all uppercase italic variables like A and Q_y are matrices. The n-vector x (with $n = p+1$) is assumed to be unknown and will be estimated using the least squares method (Teunissen, 2000a). In this paper, we provide the results of two least-squares-based methods tailored for our application, which include the unconstrained Least Squares (LS) and Bounded Least Squares subject to Weighted Constraints (BLS + WC).

3.2. Least squares method

To derive the unconstrained least squares solution for estimating x, the least squares principle states that the squared norm of the residuals e should be minimized. This leads to the following objective function for our minimization problem:

$$\phi = \frac{1}{2}\|e\|^2 = \frac{1}{2}(y - Ax)^T Q_y^{-1}(y - Ax) \tag{5}$$

where $\|.\| = (.)^T Q_y^{-1}(.)$ represents the norm of a vector. Noting that $y^T Q_y^{-1} y$ in the above equation is constant (independent of x), we aim to minimize the following objective function:

$$\min_x \phi = \min_x \left(\frac{1}{2}x^T N x - u^T x\right) \tag{6}$$

where $N = A^T Q_y^{-1} A$ is the $n \times n$ normal matrix and $u = A^T Q_y^{-1} y$ is an n-vector. Differentiation of the above objective function with respect to the unknowns x results in the following estimate $\hat{x}$ (as a global minimizer of Eq. (6))

$$\hat{x} = \left(A^T Q_y^{-1} A\right)^{-1} A^T Q_y^{-1} y = N^{-1} u \tag{7}$$

The least squares estimate of the observations y and the residuals e are

$$\hat{y} = A\hat{x}, \quad \hat{e} = y - \hat{y} \tag{8}$$

The covariance matrices of the above estimators are given as

$$\begin{cases} Q_{\hat{x}} &= \left(A^T Q_y^{-1} A\right)^{-1} \\ Q_{\hat{y}} &= A Q_{\hat{x}} A^T \\ Q_{\hat{e}} &= Q_y - Q_{\hat{y}} \end{cases} \tag{9}$$

which express the precision of the estimates $\hat{x}$, $\hat{y}$, and $\hat{e}$, respectively.

Validating the fit between observed data and the linear model is essential, with hypothesis testing playing a key role in identifying model mis-specifications such as the detection and removal of outliers. The data snooping procedure based on hypothesis testing in linear model (including the w-test for outlier detection) has been elaborated in Appendix.

3.3. Bounded least squares (BLS)

Because the above least squares method is unconstrained, the estimated x is not guaranteed to be realistic for some applications. Many engineering data-processing techniques benefit from introducing different types of constraints on the parameters of interest. These constraints may include weighted constraints as well as hard equality and/or inequality constraints. An example of an inequality-constrained adjustment model is the non-negative least squares (NNLS) problem, which addresses a linear least squares problem while enforcing the estimated parameters to be non-negative. This model is encountered in many engineering applications such as aviation industry, control systems, and artificial intelligence. For example, in aviation acoustics, the non-negativity constraints can be used when estimating the contributions of different aircraft types to the overall noise levels around airports. It is widely recognized that such contributions must be calculated as non-negative values ($x \geq 0$ indicating $\alpha_i \geq 0$ in Eq. (2)).

In the NNLS problem, the objective function in Eq. (6) is modified to include the non-negativity constraints (Amiri-Simkooei, 2016):

$$\min_x \phi = \min_x \left(\frac{1}{2}x^T N x - u^T x\right), \quad \text{subject to} \quad x \geq 0 \tag{10}$$

The above minimization problem subject to non-negativity constraints is a convex quadratic programming problem (Bemporad, 2015). Convex functions play crucial roles in many optimization problems due to their advantageous properties. For example, they guarantee a unique solution for their minimum.

In convex optimization problems, the Karush–Kuhn–Tucker (KKT) conditions are the necessary and sufficient conditions to obtain the optimal solution. The Lagrange function for the NNLS problem is then given as

$$L(x, \mu) = \frac{1}{2}x^T N x - u^T x - \mu^T x \tag{11}$$

where μ represents the Lagrange multipliers (or dual variables). The following conditions must be satisfied to obtain the optimal solution:

$$\frac{\partial L(x,\mu)}{\partial x} = Nx - u - \mu = 0, \quad \text{with} \quad x \geq 0, \quad \mu \geq 0, \quad \mu^T x = 0 \tag{12}$$

There are standard methods to solve the above quadratic programming problem (see e.g. Coleman and Li, 1996; Bro and De Jong, 1997; Franc et al., 2005).

In a generalized form, the non-negativity constraints can be replaced by the bounded constraints, where the unknown vector x is assumed to have a lower and upper bounds as $\ell^B \leq x \leq u^B$. The bounded least squares (BLS) is an extension of non-negative least squares, which has been introduced by Amiri-Simkooei (2016), Franc et al. (2005). BLS has been developed to replace the non-negative constraints with bounded constraints. The algorithm is presented in Fig. 4. The above KKT conditions are particularly relevant in our BLS method as they formally characterize the optimality of the solution to our constrained formulation, see Boyd and Vandenberghe (2004).

3.4. BLS with soft certification constraints

In addition to the above bounded constraints, the linear model in Eq. (4) can incorporate additional types of constraints known as weighted constraints (WC), or soft constraints, hereinafter referred to as BLS + WC. The equality constraints serve to stabilize the linear systems and enhance their accuracy and reliability. Such soft constraints can always be considered as an additional linear model (Amiri-Simkooei, 2019):

$$c = C^T x + e_c, \quad \mathrm{D}(c) = Q_c \tag{13}$$

which can be added, as prior information, to the linear model $y = Ax + e$. The covariance matrix Q_c determines the weighting assigned to these weighted constraints. This will consequently modify the objective function to

$$\phi = \frac{1}{2}\|e\|^2 + \frac{1}{2}\|e_c\|^2 \tag{14}$$

leading to

$$\min_x \phi = \min_x \left(\frac{1}{2}x^T N x - u^T x\right) \tag{15}$$

where the soft constraints can also be incorporated as

$$\begin{cases} N &= A^T Q_y^{-1} A + C Q_c^{-1} C^T \\ u &= A^T Q_y^{-1} y + C Q_c^{-1} \mathrm{c} \end{cases} \tag{16}$$

The unconstrained least squares estimate of x is then

$$\hat{x} = \left(A^T Q_y^{-1} A + C Q_c^{-1} C^T\right)^{-1} \left(A^T Q_y^{-1} y + C Q_c^{-1} c\right) \tag{17}$$

which is worth noting again that the solution is not guaranteed to be bounded, e.g. it is possible for it to become negative (cf. Eq. (7)). For this application we also aim to obtain the BLS solution, so minimizing

Implementation of BLS

Input:
1. design matrix A $(m \times n)$ of observation equations
2. observation vector y $(m \times 1)$
3. lower and upper bounds vectors ℓ^B and u^B $(n \times 1)$
4. convergence threshold using small value for ε

BEGIN
compute $N = A^T Q_y^{-1} A$ and $u = A^T Q_y^{-1} y$
solve for a bounded $\ell^b \leq \hat{x} \leq u^b$ from N and u as:

initialize vectors $x^{[0]} \leftarrow 0$ and $\mu^{[0]} \leftarrow -u$
set iteration counter $j = 0$
begin
  do for $k = 1$ to n
    compute $x_k^{[j+1]} = \max(\ell_k^B, x_k^{[j]} - \mu_k^{[j]}/N_{k,k})$
    calculate $x_k^{[j+1]} = \min(u_k^B, x_k^{[j+1]})$
    compute $\mu^{[j+1]} = \mu^{[j]} + (x_k^{[j+1]} - x_k^{[j]}) N(:,k)$
    update vector $\mu^{[j]} \leftarrow \mu^{[j+1]}$
  end do
  update vector $x^{[j]} \leftarrow x^{[j+1]}$
  increase counter $j \leftarrow j + 1$
  while $\|x^{[j]} - x^{[j-1]}\|_{Q_{\hat{x}}^{-1}} > \varepsilon$ **repeat** j
end

obtain $\hat{x} = x^{[j]}$, $\hat{y} = A\hat{x}$, and $\hat{e} = y - \hat{y}$
END

Fig. 4. Symbolic algorithm for implementation of bounded least-squares (BLS) estimation in the linear model of observation equations $y = Ax + e$; $x^{[j]}$ is the vector of unknown parameters x estimated in iteration j.

Eq. (16) subject to $\ell^B \leq x \leq u^B$. This can easily be implemented by applying the algorithm outlined in the previous section, with the matrix N and vector u as those provided in Eq. (16).

A final remark on the construction of the constrained model $c = C^T x + e_c$ using the certification data is in order. The certification data of the $p = 13$ aircraft classes are given in Fig. 2. Let us assume that the certification values for two aircraft classes i and j are given as d_i and d_j, converted to intensity from the EPNdB values. Consequently, we introduce a constraint such that

$$\frac{\alpha_i}{\alpha_j} = \frac{d_i}{d_j} \tag{18}$$

indicating the unknown coefficients in Eq. (2) should be proportional to their corresponding certification data. This is however a soft constraint, which could be represented as

$$\alpha_i d_j - \alpha_j d_i + e_{ij} = 0 \tag{19}$$

where the error term e_{ij} indicates the softness of equation. The above formulation can simply be used to form the constraints equations $c = C^T x + e_c$, with $c = 0$. For example, this equation can lead to one constraint equation as follows:

$$0 = c_{ij}^T x + e_{ij} = \begin{bmatrix} 0 & \cdots & d_j & 0 & \cdots & d_i & 0 & \cdots \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} + e_{ij} \tag{20}$$

The aforementioned constrained certification equation has been formulated for all pairs of α_i and α_j where $i, j = 1, \ldots, p$, making in total $p(p-1)/2$ constraint equations. For example, we consider the first constraint between α_1 and α_2 and the last constraint between α_{p-1} and α_p. For this case, the constrained matrix C^T is then of size $p(p-1)/2 \times n$.

3.5. Prediction using the established empirical model

The observation model $y = Ax + e$ subject to possible equality and inequality constraints can be used to establish a predictive model $y_p = A_p x$. Once we have obtained the least squares solution $\hat{x}$ for either the unconstrained least squares (LS) or the BLS with weighted constraints (BLS + WC), we can use the predictive model to predict $\hat{y}_p$. This can be performed simply if the design matrix A_p of the predictive model is available, which results in $\hat{y}_p = A_p \hat{x}$. For example, if the training observations L_{den} data are available for time instants t_1 to t_m, we may use $y(t_1)$ to $y(t_m)$ to establish A and estimate $\hat{x}$. To predict L_{den} for the time instant t_{m+1}, we need to know the number of aircraft types at that time instant and therefore establish the A_p as

$$A_p(t_{m+1}) = [1, \quad N_1(t_{m+1}), \ \ldots, \ N_p(t_{m+1})] \tag{21}$$

which subsequently provides the prediction as $\hat{y}_p = A_p \hat{x}$ at t_{m+1}.

In the next section, we implement predictions for one and two years ahead and compare the results with real data. Additionally, we simulate a hypothetical year where all old aircraft are replaced with the newer generations, assumed to be quieter, to predict the noise level for that year.

4. Results and discussion

4.1. Observed trends in measured L_{den} and L_{night}

In this section, we present the results for the individual NOMOS stations and for the data averaged over all NOMOS stations. The NOMOS network consists of 43 monitoring stations; however, only 39 stations were used in our analysis, as three stations had data coverage below 50% over the 18-year period (2006–2023) and were therefore excluded. Fig. 5 shows the percentage of available data for these 39 stations, ranging from 50% to 100%. The yearly-averaged L_{den} for aircraft noise only, for the period 2006–2023, is shown in Fig. 6 for NOMOS stations 2 and 21, as two representative examples. A noticeable dip is evident in 2020, which is attributed to the COVID-19 pandemic period. This drop is not an anomaly or data error, but rather reflects the substantial reduction in air traffic during the global lockdowns. While this effect is prominent in the examples shown, similar behavior is observed across some other NOMOS stations (not shown). The impact of COVID-19 is a well-known phenomenon, and its influence on aircraft noise levels, during this period, has been taken into account in the subsequent subsections when assigning noise levels to aircraft types (see also Eq. (2)).

A linear least-squares fit (Eq. (1)) applied to the data shows that in these cases a significant decreasing trend (i.e., negative slope) is obtained. As observed from the figures, a few data points deviate notably from the local trend. These deviations may stem from atypical but real events, such as altered flight procedures, temporary shifts in runway usage, or specific operational anomalies. For example, the relatively high values observed in 2017 at Station 2 might reflect such factors. As already mentioned, the decrease in 2020 can be attributed to the well-documented reduction in air traffic during the COVID-19 pandemic. To test the robustness of the trend, we applied an outlier removal procedure based on the test statistic w_i, as described in Appendix, using a two-sided 5% significance level. The recalculated slopes after removing detected outliers (shown as red squares) are also indicated in the figures (green dashed line). While there are differences, the presence of negative slopes remains consistent across most stations, both before and after outlier removal.

The (updated) slopes obtained for all NOMOS stations are mapped in Fig. 7a. It is observed that nearly all stations exhibit a negative trend in aircraft noise levels, with only 2 out of 39 stations (yellow circles) showing a slightly positive trend. For comparison, a map of the obtained slopes for the total noise (i.e. background noise plus aircraft noise) measured is shown in Fig. 7b. In contrast to aircraft noise, a

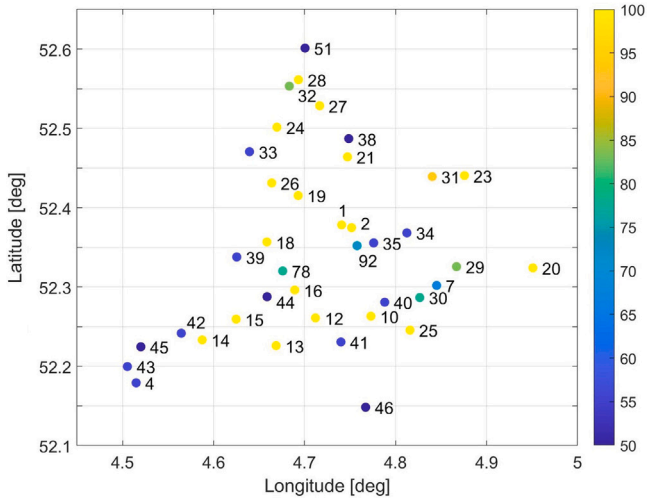

Fig. 5. Percentage of available data for each of 39 NOMOS stations used in our analysis (2006–2023), with station names indicated alongside each point. Data availability ranges from 50% to 100%.

considerably larger number of stations display a positive trend in total noise, while negative trends are less prominent (see Fig. 7a). This divergence suggests that, while aircraft noise has generally decreased over time, other sources of environmental noise have either remained stable or increased at several locations.

To illustrate this, Fig. 7b provides three examples (indicated by orange circles) of stations with a positive trend in total noise. Station 42, for example, is located in a residential area with significant community and background noise, and is situated adjacent to a water canal where sound reflections may further amplify the noise levels (Fig. 8). Similar conditions apply to stations 28 and 46. Station 28, in particular, is not only close to a road and residential surroundings but is also densely surrounded by trees, which can influence the propagation and reflection of ambient noise. These environmental and situational factors likely contribute to the observed positive trends in total noise at these locations, and may also reflect increases in road traffic or other local activities not directly related to aircraft operations.

It should be also noted that the separation between aircraft noise and total noise is based on event detection algorithms that classify individual noise events as aircraft-related when they coincide with aircraft movements tracked by radar or transponder data (e.g. using ADS-B data). This methodology is standard in airport noise monitoring systems (see e.g. European Commission, 2020), which allows for a reliable distinction between aircraft and background noise sources. While misclassification may occur at the event level, the long-term trends remain still robust for our analysis.

This result is also observed when the aircraft noise data are presented in a different way. In Fig. 9 (a and c), we show all measured yearly-averaged L_{den} and L_{night} data in one graph, together with the L_{den} and L_{night} data per year, averaged over all NOMOS stations. The averaged data are linearly least-squares fitted (Eq. (1)). For L_{den}, the estimated slope of the averaged data and its precision is $\hat{r} \pm \sigma_{\hat{r}} = -0.52 \pm 0.04$ dBA/year, i.e. a decrease in L_{den} of almost 9 dBA for the period 2006–2023. Similar conclusions can be made for the aircraft noise L_{night}. The estimated trends are not only statistically significant but also substantial. For example, a 3 dBA decrease in L_{den} for aircraft noise corresponds to a reduction of a factor of two in number of flight movements when we assume the aircraft in the fleet are the same, i.e. having the same loudness.

For comparison, similar results for the total noise are also shown in Fig. 9 (b and d). Now, for L_{den}, the slope and its precision turns out to be $\hat{r} \pm \sigma_{\hat{r}} = -0.17 \pm 0.04$ dBA/year, i.e. a decrease in L_{den} of 3 dBA for the period 2006–2023. Note that, on average, the L_{den} for total noise is more than 10 dBA higher than the L_{den} for aircraft noise alone. This means that the contribution of background noise, i.e. all noise other than that due to aircraft, to the L_{den} for total noise is dominant (Simons et al., 2022). The origin of the observed L_{den} negative trend for aircraft noise is believed to be due to changes in the fleet composition at Schiphol, at least for a substantial part (see Section 4.2). The origin of the observed L_{den} trend for total noise is less obvious and might be due to an increase in background noise, the decrease in aircraft noise (as it is still present in the total noise) or due to be a combination of both. Also, a decrease in microphone sensitivity due to overdue maintenance can play a role (see e.g. Simons and Snellen, 2023).

The results for L_{night}, both for aircraft noise and the total noise, are also shown in Fig. 9. Similar trends and conclusions as those described above for L_{den} can be drawn for L_{night}.

4.2. Establishing an empirical predictive model

To further support our previous statement regarding the underlying cause of the estimated negative trend in L_{den} attributed to aircraft noise, which is presumed to originate from shifts in the fleet composition at Schiphol, we establish an empirical model between measured L_{den} trend and fleet composition over time. The model described here indicates that if we know the number and composition of the aircraft types operating around Schiphol airport, we can predict the L_{den} associated with those aircraft. This is referred to as the empirical predictive model, which can, in principle, be applied to both L_{den} and L_{night}; however, for the sake of brevity, we focus here on L_{den} only.

We note that our analysis is based on airport-level averages of L_{den} and does not incorporate detailed operational data such as runway usage or the distribution of takeoffs and landings per aircraft type, as such data are not available for the full study period (2006–2023). Therefore, in this model, the implemented noise abatement flight procedures and operations at Schiphol have not been taken into consideration (Simons et al., 2022); the contribution of these less noisy operations is subject to further research. Therefore, the model focuses on the relationship between fleet composition and average noise levels, rather than operational factors.

Our motivation for averaging was to enable modeling and prediction of overall trends in average total aircraft noise exposure. While individual NOMOS stations may be differently affected by arriving or departing aircraft, plotting the typical flight paths over the station map in 2018 (Fig. 7a) shows no consistent directional bias in average exposure across the network. The distribution of arrivals and departures appears random with respect to station locations and their long-term average noise levels. This supports the assumption that, when averaged across a sufficiently large and spatially distributed set of stations, such directional effects tend to cancel out. Furthermore, the sensitivity analysis in Section 4.3 confirms that the estimated coefficients, linking noise levels to fleet mix composition, are robust across different station selection strategies, which reinforces the validity of our analysis.

The less prominent "Covid dip" (approximately 1 dBA reduction) in the L_{den} time series in 2020 and 2021, as shown in Fig. 9, despite the significant reduction in air traffic in 2020, is likely due to several contributing factors rather than a single cause. First, while the total aircraft movements dropped to less than half of 2019 levels, there was a notable modal shift in traffic composition: cargo aircraft, typically louder than passenger aircraft, increased their share to 10.5% of all flights, nearly four times their proportion in 2019. This shift likely offset a portion of the noise reduction expected from reduced movements alone. Second, in theory, halving the number of flight operations would yield an approximate 3 dBA decrease in L_{den}. However, the precision of our noise data, as illustrated in Fig. 9, shows a standard deviation of around 0.8 dBA, corresponding to a 99% confidence interval of ±2.0 dBA. This variability may have masked part of the expected drop. Third, while some individual stations (e.g., stations 2 and 21 in Fig. 6)

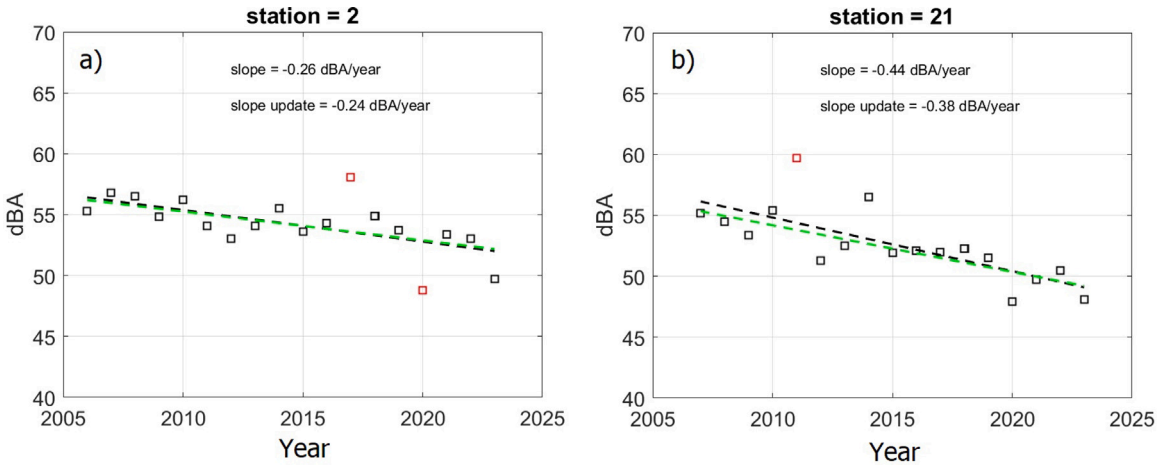

Fig. 6. The yearly-averaged L_{den} for aircraft noise only, its linear least-squares fit and outlier removal for two typical NOMOS stations (2 and 21).

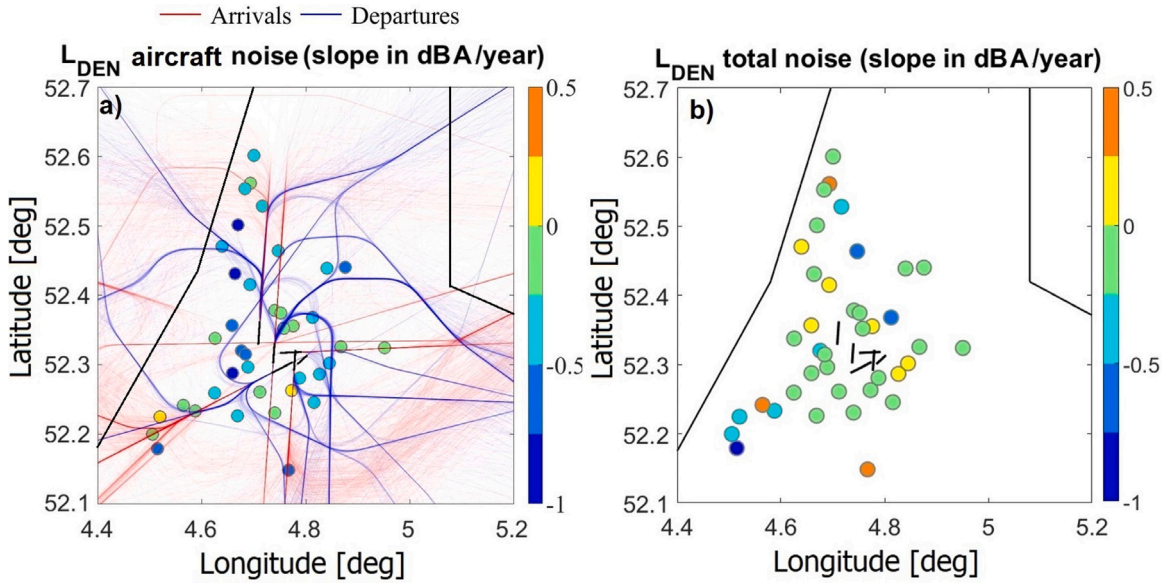

Fig. 7. Scatter plot of estimated slopes for all NOMOS stations based on least-squares fitting after outlier removal, showing (a) aircraft noise and (b) total noise. In panel (a), typical arrival (red lines) and departure (blue lines) flight paths at Schiphol Airport for 2018 are also shown for reference.

Fig. 8. Location of station 42 near a water canal and residential area, where reflections and community noise may elevate ambient levels.

show a dip during the Covid period, this was not consistent across all stations, which further diminishes the visibility of the overall effect.

We establish a linear model $y = Ax + e$ using the model in Eqs. (2) and (3). To link the above L_{den} data to aircraft types, we require information on the number and distribution of aircraft types over time. In total, we include $p = 13$ aircraft types spanning $m = 18$ years from 2006 to 2023. These aircraft types are grouped into broader fleet categories: Boeing 737 (2 types), Airbus narrow-body (3 types), Embraer and Fokker (3 types), Boeing wide-body (4 types), and Airbus wide-body (1 type), see Fig. 2. A few aircraft types like the Embraer E295, Airbus A321neo, Airbus A350, and Boeing MAX 8 were excluded due to insufficient data. This refers to either a relatively low number of movements (i.e., takeoffs and landings), a limited number of operational years at Schiphol, or both. In addition, including more aircraft types would increase the number of unknowns in the linear model $y = Ax+e$, which reduces data redundancy and can affect the stability of least squares estimation. Therefore, we focus on a subset of 13 aircraft types with sufficient temporal and operational representation for robust modeling. The final design matrix A is thus of size $m \times n = 18 \times 14$.

We aim to train and implement a predictive model using two scenarios (Table 2). For each scenario we split the 18 data points into

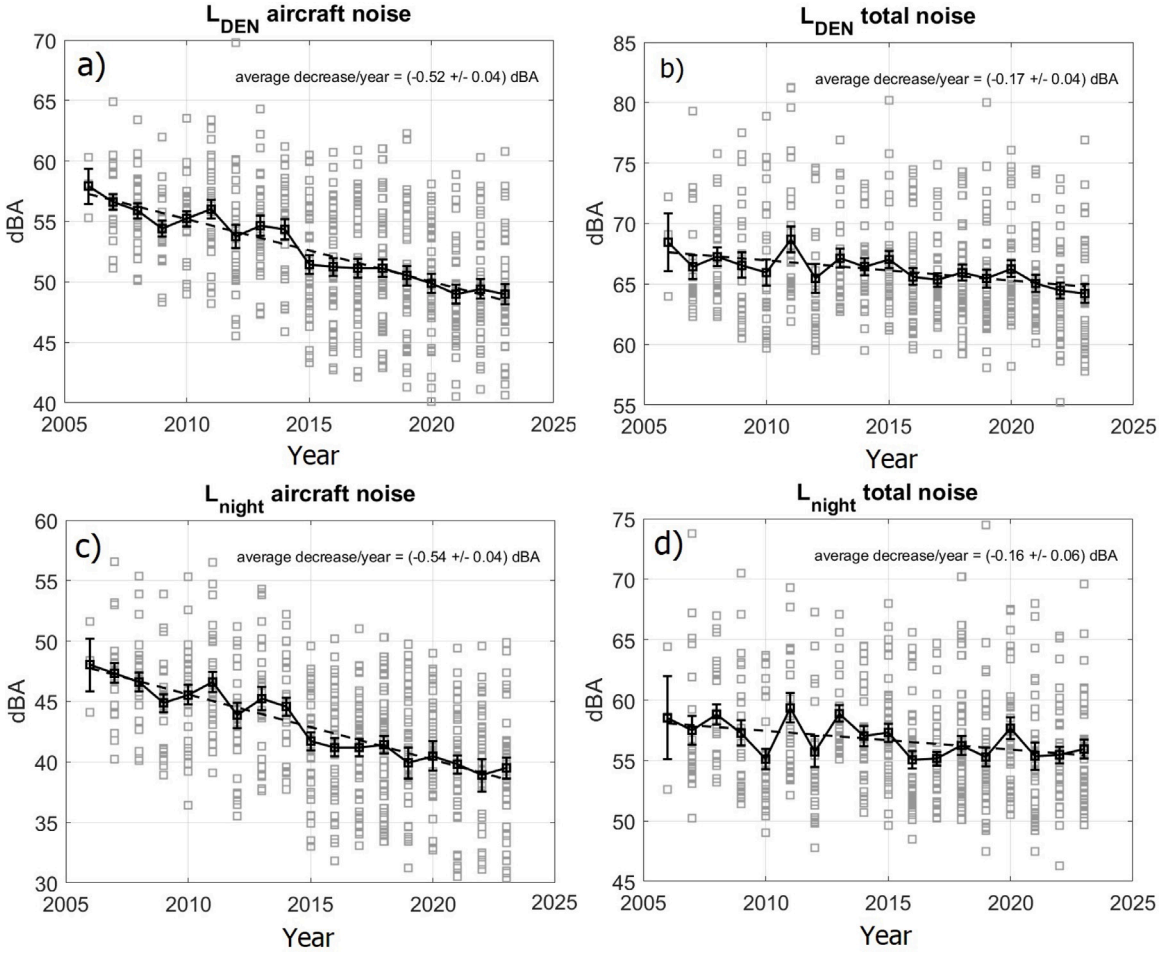

Fig. 9. Comparison between trends in aircraft (a and c) and total (b and d) noise levels over the period 2006–2023: L_{den} (a and b) and L_{night} (c and d). In each panel, gray squares represent individual stations, black squares denote mean values over all stations, the dashed black line indicates the least-squares fit, and error bars show the corresponding standard deviations.

Table 2
Details of two scenarios for training and testing data splitting used to establish a predictive model. $m-n$ is the redundancy of the training model $y = Ax + e$ (Scenario 2 has less redundancy ($m-n=2$ vs. $m-n=3$ of Scenario 1) because it uses one fewer training data point, 16 vs. 17.

Scenario	Training	Prediction	Size of A $(m \times n)$	Size of A_p	$m-n$
1	2006–2022	2023	17 × 14	1 × 14	3
2	2006–2021	2022, 2023	16 × 14	2 × 14	2

'estimation' and 'prediction' as explained in Section 3. The estimation process is the training step (based on historical data), and the prediction is the testing step (to test the performance of the prediction for future events). For example, for scenario 1, we use the first 17 data points (from 2006 to 2022) for training and the last data of 2023 for prediction (testing).

The results of the above two scenarios, and for the two methods of LS and BLS + WC, are presented in Fig. 10. A few observations are highlighted. (i) For the L_{den} prediction, it is evident that while there is consistency among the BLS + WC method to forecast the first or second year, the LS method can overestimate (frames a and c) the L_{den}. Notably, BLS + WC exhibits more stable and consistent performance than the unconstrained LS in both scenarios 1 and 2, particularly in terms of generalization to testing data. In the next paragraph, we interpret these results through the lens of the bias–variance trade-off, which highlights why BLS + WC generalizes better despite not always outperforming LS in every year. (ii) As another scenario (the results not presented here), we have also tested the prediction results using the NNLS method, so without considering the weighted certification constraints. The prediction results are not much better than the unconstrained LS, and still worse than BLS + WC. These findings suggest the importance of considering the soft constraints related to the certification data when forecasting L_{den} values using BLS. This offers new opportunities and potential challenges for future studies on L_{den} noise monitoring around airports. (iii) We also compared the performance of the methods on training and testing data separately, as shown in Fig. 11. While the LS approach fits the training data closely (i.e., small residuals), it produces large root mean square errors (RMSE) on the test set, which is indicative of overfitting. In contrast, BLS + WC yields consistent results across both estimation and prediction phases for both scenarios. We note that the RMSE in scenario 1 is based on a one-year prediction and in scenario 2 on a two-year prediction. To further evaluate robustness, we also tested a three-year ahead prediction scenario, training the model on data from 2006 to 2020 and testing on 2021, 2022, and 2023. Although these results are not shown in the paper, they continue to demonstrate stable predictive performance for BLS + WC. The RMSE for this extended prediction is 1.32 dBA under this scenario.

A remark on the last observation mentioned above, regarding the application of LS and BLS + WC, is in order. The observed differences in performance between LS and BLS + WC highlight a classic issue in statistical modeling: the bias–variance trade-off (Geman et al., 1992). The LS method, which fits the training data very well, is likely to have a low bias but a high variance. This means that while the model approximates the training data closely (called over-fitting), it becomes

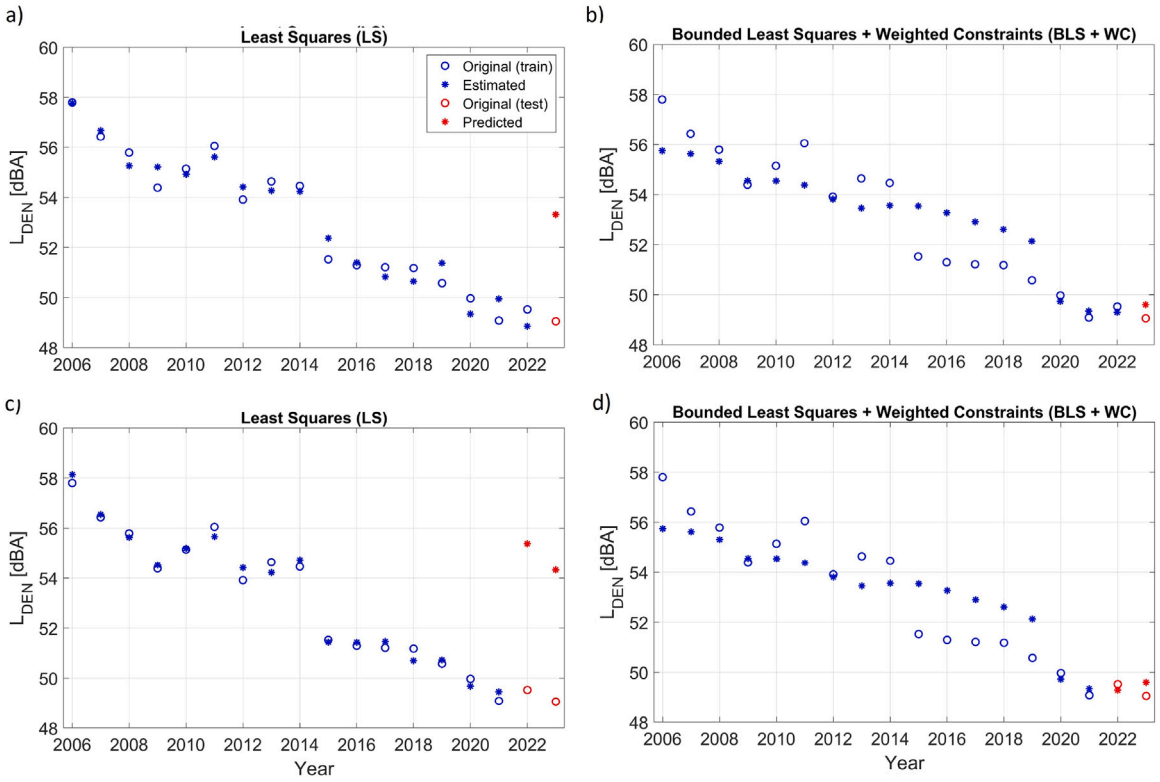

Fig. 10. Estimation and prediction performance of L_{den} data using LS (a and c) and BLS + WC (b and d) methods in Scenario 1 (a and b) and Scenario 2 (c and d).

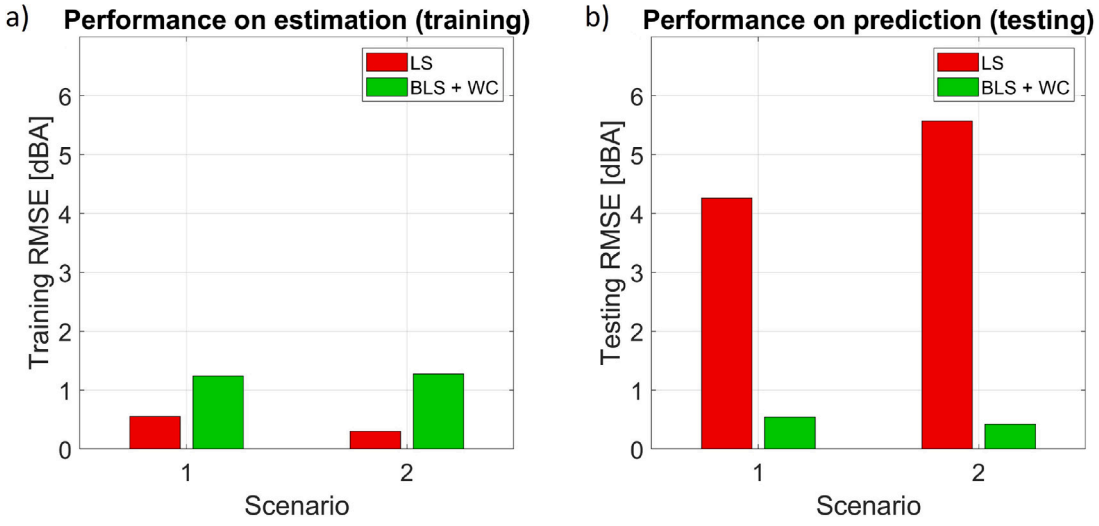

Fig. 11. Performance of LS and BLS + WC methods on training (estimation) and testing (prediction) datasets, measured by root mean squared error (RMSE).

very sensitive to small fluctuations and noise, which leads to poor performance on unseen data. On the other hand, the BLS + WC method, which fits the training data less perfectly but performs better on testing data, likely has higher bias but lower variance. This implies that BLS + WC, being a simpler model, does not capture all the complexities of the training data, but its simplicity makes it more reliable and less sensitive to noisy data, leading to better performance in the testing data set.

Before running the predictive model for a simulated year using the estimated coefficients (next section), we plot the estimated α's of the BLS + WC model against the certification data shown in Fig. 2. The results are depicted in Fig. 12. For the LS model (not shown here), many α values are estimated as negative, which is not physically interpretable.

Fig. 12 shows that the estimated coefficients are significantly correlated with the certification data, as also shown by the fit. For example, the loudest aircraft (i.e. B747) has the largest estimated coefficient α. This correlation results from incorporating certification data as soft constraints in the BLS + WC model, which regularizes the α estimates and helps ensure physically meaningful values. In contrast, the LS method can produce negative coefficients, which are not physically feasible. A closer match to the certification data can be achieved by increasing the weight of the regularization, but this would reduce the fit quality. The current balance reflects a trade-off between fidelity to

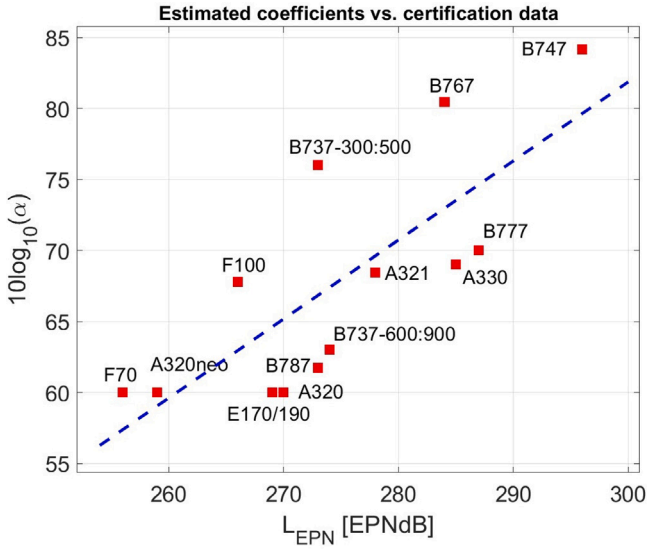

Fig. 12. Estimated coefficients α's (converted to $10\log(\alpha)$) versus the certification data in the BLS + WC model; the blue dashed line shows the least squares fit.

certification data and predictive performance. This regularization effect is absent in the LS and NNLS methods.

4.3. Sensitivity analysis

To investigate the robustness of the estimated coefficients and assess the potential influence of spatial and temporal variations in the measurement network, we performed a sensitivity analysis using three different station selection strategies. The analysis aimed to determine whether the overall trends are sensitive to the subset of NOMOS stations included in the averaging process.

1. Test Case I: This case included 14 stations located within 10 km of the center of the NOMOS network (representing the central area near Schiphol runways). These stations are generally influenced by both approach and departure operations and offer a compact spatial sample.
2. Test Case II: This case included 25 stations located beyond 10 km from the network center. These peripheral stations are often more strongly influenced by specific operational modes (e.g. mainly approach or departure).
3. Test Case III: This case included 18 stations with complete data availability (100% coverage over the entire 18-year period), thereby eliminating potential biases introduced by data gaps.

The estimated rates are $\sigma_r^I = -0.47 \pm 0.04$, $\sigma_r^{II} = -0.52 \pm 0.04$, and $\sigma_r^{III} = -0.40 \pm 0.04$ for Test Cases I, II, and III, respectively. The estimated coefficients linking noise levels to fleet mix composition in Eq. (2), derived from each of the three test cases, closely match those obtained from the full set of stations (i.e., averages over all available stations).

The results, shown in Fig. 13, demonstrate a high degree of consistency in the estimated coefficients across all configurations. This confirms that the long-term trends identified in the analysis are not significantly driven by a particular spatial subset or temporal sampling bias.

4.4. Running the predictive model for a simulated year

In this section, we explore the application of the predictive model to forecast values for a hypothetical future year, where no empirical data is available. Therefore the exercise aims to demonstrate the model's ability to provide predictions based on existing trends, patterns and expectations of future fleet mix scenarios. The input variables for the hypothetical year are similar to those of 2023 for the number of aircraft types. However, the following assumptions were made to include more newer aircraft in the prediction:

(1) The Boeing 737 fleet will remain unchanged, meaning that no Boeing 737 (300, 400 and 500) will be operational, but Boeing 737 (600, 700, 800 and 900) will continue to operate as they do currently.
(2) All Airbus narrow body aircraft (320 and 321) will be replaced by the 320neo.
(3) All Embraer aircraft (E170/190) will continue to operate as they do now. Note that F70 and F100 are no longer operational.
(4) All Boeing wide body aircraft will be replaced by the 787.
(5) The Airbus aircraft (i.e. 330) will continue to operate as it does now.

The rationale behind the assumption that all Boeing wide-body aircraft will be replaced by the B787 is based on the fact that aircraft with similar MTOW, whether passenger or cargo, can be substituted by newer and quieter models of equivalent capacity. For example, older types like the B747 and B777 may, in certain contexts, be replaced with the B787, which is available in both passenger and cargo configurations. The share of B747 operations has already declined significantly in recent years. This *optimistic* fleet modernization scenario is consistent with long-term fleet renewal projections in the Boeing Commercial Market Outlook (2024–2044) and the EUROCONTROL Aviation Outlook 2050, both of which foresee replacement of older aircraft with quieter and more efficient types (Boeing, 2024; Eurocontrol, 2022).

The prediction results for the simulated year are presented for both scenarios in Fig. 14, and they appear to be identical. For the simulated year, we obtain 43.5 dBA, i.e. a reduction of 5.5 dBA in the L_{den} data is observed compared to 2023 (49.0 dBA). This reduction is substantial (see Section 4.1).

A decrease in L_{den} of nearly 9 dBA over the period 2006–2023 has already been shown in Fig. 9. However, achieving such a significant reduction in the near future is unlikely. There are several reasons why the above-mentioned L_{den} reduction of 5.5 dBA remains challenging. One key assumption is the replacement of the B747 (the loudest aircraft) and the B777 (the second loudest) with the B787. This replacement accounts for approximately 4 dBA of the estimated 5.5 dBA reduction. However, it is important to note that the B747 is still in operation as a cargo aircraft, and a substantial number of B777s remain in service. Consequently, their contribution to the noise levels will persist in the coming years. Furthermore, the B767, the third loudest aircraft, has already experienced a significant decline in operations at Schiphol Airport, leading to a notable reduction in its noise contribution (the remaining B767 fleet contributes less than 0.5 dBA to overall noise levels). Our analysis also assumes that the Boeing 737 fleet and all Embraer and Fokker aircraft (E170/190 and F100) will remain unchanged from their 2023 configurations. The replacement of Airbus narrow-body aircraft (A320 and A321) with the newer A320neo models contributes a reduction of approximately 1 dBA.

We can therefore conclude that the 5.5 dBA reduction is likely to occur over the long term. In this timeframe, additional sources of noise reduction may also become significant. For example, quieter aircraft models such as the Embraer E295, Airbus A321neo, Airbus A350, and Boeing MAX 8 were excluded from our current analysis due to insufficient data, as these aircraft types are relatively new to operations at Schiphol airport. Once these aircraft become fully integrated into the airport's fleet in the long term, they are expected to contribute further to the noise reduction, and therefore to the L_{den} values.

For comparison, another scenario was evaluated in which, instead of replacing all Boeing wide-body aircraft with the B787, only the older

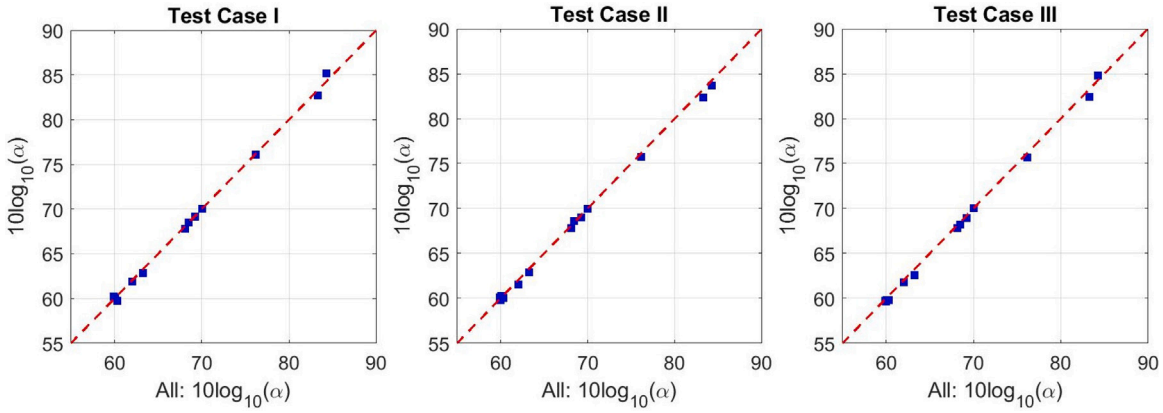

Fig. 13. Comparison of estimated coefficients based on different subsets of NOMOS stations. Three test cases are compared against the baseline case using all available stations: Test Case I includes 14 stations within 10 km of the network center; Test Case II includes 25 stations located beyond 10 km from the center; and Test Case III includes 18 stations with complete data availability over the full 18-year period.

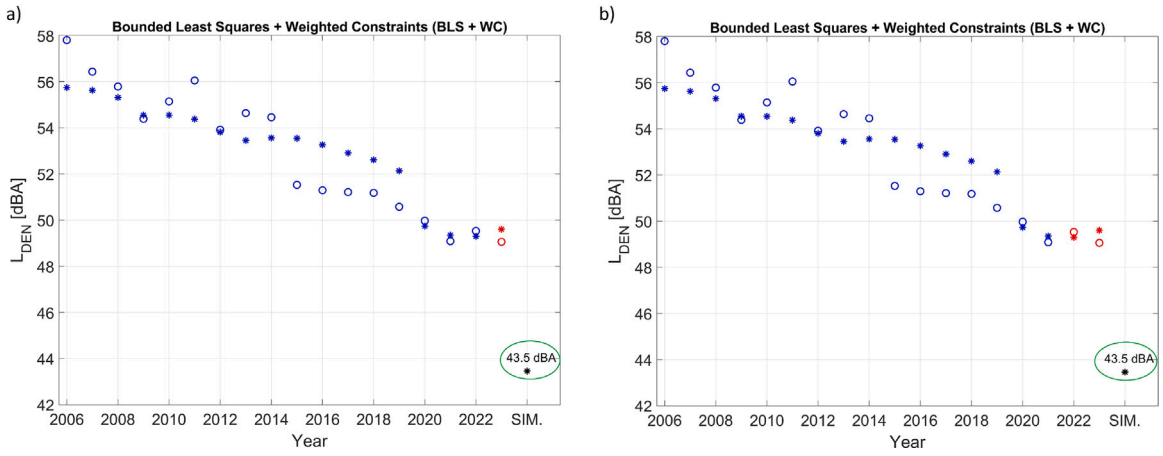

Fig. 14. Prediction of L_{den} for a hypothetical future year using the BLS + WC model in Scenario 1 (a) and Scenario 2 (b).

B747 (the loudest aircraft in the fleet) was assumed to be replaced by the newer B777. In this case, the projected noise reduction was limited to approximately 2.5 dBA, still underscoring the significant influence of wide-body fleet renewal on the overall noise mitigation potential.

5. Summary and conclusions

Traditional noise models, while cost-effective, often differ from actual measurements, which lead to community distrust and the need for validation with real monitoring data like noise monitoring systems (NOMOS). In this study, we analyzed NOMOS data from 2006 to 2023, observing a significant decreasing trend of −0.5 dBA/year in yearly averaged L_{den}. This trend indicates a considerable reduction in noise levels over this period, primarily attributed to changes in the fleet composition at Schiphol Airport. The established empirical model demonstrates that by knowing the number and types of aircraft operating at the airport, we can predict the associated L_{den} value.

Using the least squares methods, including the unconstrained least squares (LS) and bounded least squares subject to weighted constraints of certification data (BLS + WC), we assigned measured L_{den} to 13 different aircraft types. The BLS + WC method showed superior performance, with significantly smaller prediction errors compared to the unconstrained LS method. This highlights the importance of incorporating certification data as soft constraints, which improves the accuracy and reliability of noise predictions.

We further extended our analysis to predict noise levels for a hypothetical future year with newer aircraft models. The results indicate a substantial reduction of ~ 5.5 dBA in noise levels compared to 2023, primarily due to the replacement of the B747 and B777 with the quieter B787 and the introduction of the Airbus A320neo model. These findings underscore the potential for fleet modernization to significantly reduce noise pollution around airports.

This study highlighted the high potential of using empirical data to model average annual aircraft noise levels based on fleet composition. Our findings showed that changes in aircraft types over time can explain a large portion of the long-term variation in average L_{den} values at Schiphol Airport. These insights can support airport management to evaluate the effectiveness of fleet modernization strategies and estimating the scale of fleet changes needed to meet regulatory noise thresholds. While the current analysis is constrained by the lack of detailed operational data (e.g., runway usage and flight paths), future research could benefit from integrating such variables to improve model accuracy. In addition, the use of more relevant noise metrics and advanced analytical methods like machine learning could further enhance predictive capabilities and community trust in noise management efforts.

CRediT authorship contribution statement

Dick G. Simons: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Visualization, Validation, Conceptualization. **Alireza Amiri-Simkooei:** Writing – original draft, Software, Methodology, Investigation, Formal analysis, Visualization, Validation, Conceptualization. **Joris A. Melkert:** Review and editing,

Investigation, Data curation. **Mirjam Snellen:** Review and editing, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors acknowledge the use of aircraft noise data provided by the European Aircraft Noise Services (EANS) and aircraft movement statistics published by Amsterdam Airport Schiphol in their publicly available Traffic Review reports.

Appendix. Hypothesis testing in linear models

In many engineering applications, validating the fit between observed data and the linear model is crucial. In this context, hypothesis testing plays a significant role to identify mis-specification in the functional model. Among such mis-specifications, outliers identification and removal is an essential step. This will guarantee that the final LS estimates are unbiased (Teunissen, 2000b). For simplicity, we assume that the observations are independent and have identical variance, leading to the covariance matrix as $Q_y = \sigma^2 I_m$, where I_m is an identity matrix of size m. Two cases can usually be considered: (1) 'σ known', and (2) 'σ unknown'.

Each individual observation is screened for the presence of an outlier (Teunissen, 2000b; Baarda, 1968). This is usually performed by computing a so-called w-test statistic. An important application of the w-test is thus *blunder* detection. A blunder (called also outlier) affects just a single observation. To screen the observations and identify outliers, we usually formulate m alternative hypotheses, corresponding to m observations. If the covariance matrix of observables is diagonal ($Q_y = \sigma^2 I_m$), with σ^2 known, the expression for the w-test statistic reduces to a simple form as

$$w_i = \frac{\hat{e}_i}{\sigma_{\hat{e}_i}} \tag{A.1}$$

with $\sigma_{\hat{e}_i} = (Q_{\hat{e}})_{ii}^{1/2}$ the standard deviation of the least-squares residual i, for $i = 1, \ldots, m = 18$ (see Eq. (9) for $Q_{\hat{e}}$). This test statistic is referred to as the *normalized* residual. This test statistic can be tested within a given confidence level $1 - \alpha$. When σ is known, the test statistic will follow a standard normal distribution under the null hypothesis H_0, expressing that the data is not an outlier.

When σ is not known, it can also be estimated using the least squares residuals. The least squares estimate of the variance σ^2 is then (Amiri-Simkooei, 2007)

$$\hat{\sigma}^2 = \frac{\hat{e}^T \hat{e}}{m - n - 1} \tag{A.2}$$

where $m - n - 1$ is the redundancy of the functional model under the alternative hypothesis. This hypothesis assumes that one observation contains a blunder and is therefore removed from the analysis. This estimate will also affect the covariance matrix $Q_y = \sigma^2 I_m$ as $\hat{Q}_y = \hat{\sigma}^2 I_m$. This consequently leads to the modification of $Q_{\hat{x}}$, $Q_{\hat{y}}$, and $Q_{\hat{e}}$ as $\hat{Q}_{\hat{x}}$, $\hat{Q}_{\hat{y}}$, and $\hat{Q}_{\hat{e}}$, where the estimated $\hat{Q}_y$ is used instead of Q_y. The w-test statistic is accordingly modified as

$$w_i = \frac{\hat{e}_i}{\hat{\sigma}_{\hat{e}_i}} \tag{A.3}$$

with $\hat{\sigma}_{\hat{e}_i} = (\hat{Q}_{\hat{e}})_{ii}^{1/2}$ the standard deviation of the least-squares residual i, for $i = 1, \ldots, m$. This test statistic can also be tested within a given confidence level $1 - \alpha$. Under H_0, the above test statistic will have a Student t distribution with a $m - n - 1$ degrees of freedom.

Data availability

The yearly-averaged L_{den} and L_{night} data used in this contribution can directly be obtained from the European Aircraft Noise Services (EANS) website, Schiphol, https://www.eans.net/.

References

Amiri-Simkooei, A., 2007. Least-squares variance component estimation: theory and GPS applications (Ph.D. thesis). Delft University of Technology, Publication on Geodesy, 64, Netherlands Geodetic Commission.

Amiri-Simkooei, A., 2016. Non-negative least-squares variance component estimation with application to GPS time series. J. Geod. 90 (5), 451–466.

Amiri-Simkooei, A., 2019. Unified least-squares formulation of a linear model with hard constraints. J. Surv. Eng. 145 (4), 04019012.

ANIMA Project Consortium, 2025. ANIMA – Aviation Noise Impact Management through Novel Approaches. https://anima-project.eu/. (Accessed 09 April 2025).

Baarda, W., 1968. A Testing Procedure for Use in Geodetic Networks. Tech. rep., 2, (5), Netherlands Geodetic Commission, Publ. on Geodesy, New Series, Delft.

Basner, M., McGuire, S., 2018. WHO environmental noise guidelines for the European region: a systematic review on environmental noise and effects on sleep. Int. J. Environ. Res. Public Heal. 15 (3), 519.

Bemporad, A., 2015. A quadratic programming algorithm based on non-negative least squares with applications to embedded model predictive control. IEEE Trans. Autom. Control 61 (4), 1111–1116.

Bertsch, L., Simons, D.G., Snellen, M., 2015. Aircraft noise: The major sources, modelling capabilities, and reduction possibilities.

Bewoners Aanspreekpunt Schiphol, 2021. 2020 Jaarrapportage. Technical Report, Bewoners Aanspreekpunt Schiphol, Amsterdam, pp. 1–52, www.minderhinderschiphol.nl. (Accessed 26 February 2024).

Boeing, 2024. Commercial market outlook 2024–2043. https://www.boeing.com/commercial/market/commercial-market-outlook. (Accessed 12 August 2025).

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.

Breugelmans, O., Houthuijs, D., Swart, W., Hajema, K., Hogenhuis, R., van Poll, R., 2013. Aircraft noise and health: is Lden the proper indicator to describe the relationship? In: ISEE Conference Abstracts 25, vol. 2013, (1), p. 3712.

Bro, R., De Jong, S., 1997. A fast non-negativity-constrained least squares algorithm. J. Chemom.: A J. Chemom. Soc. 11 (5), 393–401.

Cointin, B., Hileman, J., 2016. US civil aircraft noise annoyance survey design. In: Inter-Noise and Noise-Con Congress and Conference Proceedings, vol. 253, (1), Institute of Noise Control Engineering, pp. 6972–6977.

Coleman, T.F., Li, Y., 1996. A reflective Newton method for minimizing a quadratic function subject to bounds on some of the variables. SIAM J. Optim. 6 (4), 1040–1058.

Crocker, M.J., 2008. Fundamentals of acoustics, noise, and vibration. pp. 1–16, Handbook of noise and vibration control.

ECAC.CEAC, 2016. Doc. 29. Report on Standard Method of Computing Noise Contours around Civil Airports. Volume 2: Technical Guide, Tech. Rep., fourth ed. European Civil Aviation Conference (ECAC).

Eurocontrol, 2022. Eurocontrol aviation outlook 2050. https://www.eurocontrol.int/publication/aviation-outlook-2050. (Accessed 12 August 2025).

European Aircraft Noise Services (EANS), 2021. European aircraft noise services (EANS) – schiphol. https://www.eans.net/. (Accessed 26 February 2024).

European Commission, 2020. Environmental Noise Directive 2002/49/EC: Common Noise Assessment Methods in Europe (CNOSSOS-EU). Tech. rep., European Commission, URL https://ec.europa.eu/environment/noise/cnossos-eu_en.htm.

European Parliament and Council, 2002. Directive 2002/49/EC of the European parliament and of the council of 25 June 2002 relating to the assessment and management of environmental noise. Official Journal of the European Communities, L189, pp. 12–25. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32002L0049.

European Union Aviation Safety Agency, 2025. Type certificate data sheets for noise (TCDSN). URL https://www.easa.europa.eu/en/document-library/type-certificates/tcdsn. (Accessed 12 August 2025). Available online via the EASA document library.

Franc, V., Hlavac, V., Navara, M., 2005. Sequential coordinate-wise algorithm for the non-negative least squares problem. In: International Conference on Computer Analysis of Images and Patterns. Springer, pp. 407–414.

Franssen, E., Van Wiechen, C., Nagelkerke, N., Lebret, E., 2004. Aircraft noise around a large international airport and its impact on general health and medication use. Occup. Environ. Med. 61 (5), 405–413.

Freedman, D.A., 2009. Statistical Models: Theory and Practice, second ed. Cambridge University Press.

Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. Neural Comput. 4 (1), 1–58.

Gjestland, T., 2024. Measuring community response to noise—Factors affecting the results of annoyance surveys. Int. J. Environ. Res. Public Heal. 21 (4), 420.

Gjestland, T., Gelderblom, F.B., 2017. Prevalence of noise induced annoyance and its dependency on number of aircraft movements. Acta Acust. United Acust. 103 (1), 28–33.

Guski, R., 2017. The increase of aircraft noise annoyance in communities. Causes and consequences. In: 12th ICBEN Congress on Noise As a Public Health Problem. Zurich.

Hansell, A.L., Blangiardo, M., Fortunato, L., Floud, S., De Hoogh, K., Fecht, D., Ghosh, R.E., Laszlo, H.E., Pearson, C., Beale, L., et al., 2013. Aircraft noise and cardiovascular disease near Heathrow airport in London: small area study. BMJ 347.

International Civil Aviation Organization, 2017. Annex 16 to the Convention on International Civil Aviation: Environmental Protection, Volume I – Aircraft Noise, eighth ed. International Civil Aviation Organization, Montréal, Canada, Document AN/896.

Isermann, U., Vogelsang, B., 2010. AzB and ECAC Doc. 29 — Two best-practice European aircraft noise prediction models. Noise Control Eng. J. 58 (4), 455–461.

Morrell, P., Lu, C.H.-Y., 2000. Aircraft noise social cost and charge mechanisms–a case study of Amsterdam Airport Schiphol. Transp. Res. Part D: Transp. Environ. 5 (4), 305–320.

Royal Schiphol Group, 2024. Annual traffic reviews. URL https://www.schiphol.nl/en/schiphol-group/traffic-review/. (Accessed June 2025).

Schiphol Airport, 2021. AMS airport charges, levies, slots and conditions. https://www.schiphol.nl/nl/route-development/pagina/ams-airport-charges-levies-slots-and-conditions/. (Accessed 26 February 2024).

Simons, D.G., Besnea, I., Mohammadloo, T.H., Melkert, J.A., Snellen, M., 2022. Comparative assessment of measured and modelled aircraft noise around Amsterdam Airport Schiphol. Transp. Res. Part D: Transp. Environ. 105, 103216.

Simons, D.G., Snellen, M., 2023. An Introduction to General Acoustics and Aircraft Noise. Tech. Rep. AE4431-23, Control and Operations Department, Delft University of Technology, Aircraft Noise.

Teunissen, P.J.G., 2000a. Adjustment theory: an introduction. Delft University Press, Website: http://www.vssd.nl. Series on Mathematical Geodesy and Positioning.

Teunissen, P.J.G., 2000b. Testing theory: an introduction. Delft University Press, https://pure.tudelft.nl/ws/portalfiles/portal/214132690/Testing_Theory.pdf. Series on Mathematical Geodesy and Positioning.

Vieira, A., Snellen, M., Simons, D.G., 2020. Experimental assessment of sound quality metrics for takeoff and landing aircraft. AIAA J. 59 (1), 240–249.

von Karman Institute for Fluid Dynamics, 2025. ARTEM – Aircraft noise Reduction Technologies and related Environmental iMpact. https://www.vki.ac.be/index.php/316-h2020/h2020-artem/797-h2020-artem. (Accessed 09 April 2025).

WorldInMaps, 2024. World in maps: Interactive world maps and data. URL https://worldinmaps.com/. (Accessed 26 July 2025).

Zaporozhets, O., 2016. Aircraft noise models for assessment of noise around airports—improvements and limitations. Environ. Rep.: Aviat. Clim. Chang. 50–56.