

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## Journal of Transport &amp; Health

journal homepage: [www.elsevier.com/locate/jth](http://www.elsevier.com/locate/jth)

# Transport-related noise exposure in a representative sample of Île-de-France residents: A data-enrichment approach



Nina Kreuzberger<sup>a,b,\*</sup>, Basile Chaix<sup>a</sup>, Ruben Brondeel<sup>a</sup>, Julie Méline<sup>a</sup>,  
Tarik El Aarbaoui<sup>a,b</sup>

<sup>a</sup> INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health (IPLESP), Nemesis team, F75012 Paris, France

<sup>b</sup> EHESP School of Public Health, 15 avenue du Professeur Léon-Bernard, 35043 Rennes, France

## 1. Introduction

Noise is an environmental stressor which has been linked to various adverse health outcomes (Babisch, 2011) such as hearing loss (Lie et al., 2016), stress hormone imbalance (Selander et al., 2009), hypertension (Van Kempen and Babisch, 2012), sleep disturbance (Hume et al., 2012), long-term elevated risk of stroke (Halonen et al., 2015) and diabetes (Dzhambov, 2015). To minimize the negative impact of noise on health, daily average levels below 50–55 dB(A) are recommended for the general population (European Environment Agency, 2010).

Many of the associations between noise and health were either assessed in occupational settings (Skogstad et al., 2016) or were based on residential assessments of outdoor noise from modeled noise exposure of selected sources, typically road traffic, railway systems, and aircrafts (Sørensen et al., 2012; Floud et al., 2013). However, noise exposure is not limited to a unique source or a single environment, as people move between places for activities. Therefore, stationary assessment might not accurately reflect personal noise exposure, in particular during the day.

Recent studies took the movement between different microenvironments into account and described personal exposure by following individuals continuously with a noise dosimeter (Taimisto et al., 2013; Kraus et al., 2015; Neitzel et al., 2009). A study done by Kraus et al. (2015) showed that one of the environments with high noise levels was traffic (mean equivalent sound levels ( $L_{Aeq}$ ) between 74.0 and 78.3 dB(A) depending on the transport mode. The average values during transport exceeded the noise level of 70 dB(A) which was established for prevention of hearing loss, indicating room for improvements in the protection of the general population during transport in metropolises. More extensive information on the intensity, duration, and frequency of exposure to noise in a variety of settings would be needed to estimate average daily exposures of a population and evaluate the impact that different scenarios of intervention may have. Unfortunately, studies that deploy noise dosimeters and thus offer very precise information on noise exposure over multiple environments usually result in small sample sizes due to high costs and work-load in data collection and analysis.

To circumvent the limitations of both large-sample residential noise studies based on modelled maps and small studies based on noise dosimeters, the current study proposes a data-enrichment approach to estimate noise levels for a large representative sample from a smaller study sample of the same background population with available noise dosimeter data. A similar approach has been successfully applied to physical activity during transportation by Brondeel et al. (2016). The objectives of the present study were a) to describe personal noise exposure during daily trips with different transport modes in the Île-de-France region, b) to specify a model that predicts personal noise exposure in the dosimeter-based RECORD MultiSensor study, c) to cross-validate this model with a train-test set procedure and d) to apply the developed and validated model to a larger, representative dataset of the same background population (Enquête Globale Transport – EGT) in order to estimate transport-related noise exposure of Île-de-France residents.

\* Correspondence to: UMR-S 1136 Inserm – UPMC-Paris 6, Faculté de Médecine Saint-Antoine, 27 rue Chaligny, 75012 Paris, France.

E-mail addresses: [nina.keuzberger@gmail.com](mailto:nina.keuzberger@gmail.com) (N. Kreuzberger), [basile.chaix@iplesp.upmc.fr](mailto:basile.chaix@iplesp.upmc.fr) (B. Chaix), [ruben.brondeel@umontreal.ca](mailto:ruben.brondeel@umontreal.ca) (R. Brondeel), [julie.meline@irsn.fr](mailto:julie.meline@irsn.fr) (J. Méline), [tarik.el-aarbaoui@iplesp.upmc.fr](mailto:tarik.el-aarbaoui@iplesp.upmc.fr) (T. El Aarbaoui).

<https://doi.org/10.1016/j.jth.2019.02.002>

Received 25 November 2018; Received in revised form 2 February 2019; Accepted 3 February 2019

Available online 12 February 2019

2214-1405/ © 2019 Elsevier Ltd. All rights reserved.

## 2. Materials and method

### 2.1. RECORD MultiSensor study

#### 2.1.1. Sample

As described elsewhere (Chaix et al., 2012), the participants of the RECORD study were recruited in 2007–2008 during free preventative health checkups offered by the French National Health Insurance System for Salaried Workers (CNAMTS) at four different sites in the Île-de-France region affiliated with the IPC medical Center. The Île-de-France region is the most populated of the 18 regions in France and contains the city of Paris. Inclusion criteria were a) being born between 1928 and 1979, b) residing in one of 112 a priori selected municipalities, and c) being able to complete the questionnaires of the study in French with minimal help. The same participants were followed up and new participants were invited in a second wave of the RECORD study (2011–2012). During this second wave, 286 participants were invited to participate in the RECORD MultiSensor study from September 2013 to June 2015. Written informed consent was obtained from all participants. The RECORD MultiSensor study was approved by the French Data Protection Authority.

#### 2.1.2. Data collection and processing

For the RECORD study, sociodemographic variables were obtained during the health checkup where participants filled in the IPC medical questionnaire and the RECORD questionnaire. Household income needed to be recoded from a categorical to a continuous variable to be comparable with the EGT income variable. Educational level was summarized in three categories (No diploma of secondary education, diploma of secondary or lower tertiary education, diploma of higher tertiary education). The location of residence was indicated as inner city (Paris), inner suburbs or outer suburbs.

During the MultiSensor study, the participants carried several sensors (wearable class II dosimeter Wed007, ACOEM; GT3X+ accelerometer, ActiGraph; and BT-Q1000XT GPS receiver, QStarz) from day of inclusion up to day eight. Measurements from day one, 03:00 a.m. up to day eight, 03:00 a.m. were used for analysis. The GPS data of each participant was automatically pre-processed in order to identify visited places and start and end time of each trip stage (Wolf et al., 2004), which is the portion of a trip undertaken with a single transportation mode. These pre-processed GPS tracks were consolidated during a phone mobility survey with the participants, resulting in a detailed timetable of visits at activity places and trips segmented into trip stages for the 7 days of observation, as described in more detail in Brondeel et al. (2016) and Chaix et al. (2014).

Participants were instructed to wear the dosimeter at their belt and the microphone close to the ear, and to charge the device during night. Noise was measured continuously as A-weighted equivalent sound level ( $L_{Aeq,1s}$ ) per second in a range of 40 to 120 dB (A). All  $L_{Aeq,1s}$  values of a trip stage were extracted and summarized average  $L_{Aeq}$  which is defined as follows:

$$L_{Aeq,T} = 10 \log \left( \frac{\sum_{i=1}^n 10^{\frac{L_{Aeq,i}}{10}}}{n} \right)$$

Where:

$L_{Aeq,T}$  = Averaged equivalent noise level

$L_{Aeq}$  = Equivalent noise level for each second

$n$  = The total number of one second intervals from  $i$  to  $n$  per trip stage

In total, 78 participants of the MultiSensor study carried all devices necessary for inclusion in the prediction model (i.e., the noise dosimeter). However, observations of one participant were excluded due to dosimeter failure.

### 2.2. Global transport survey

The global transport survey (EGT; STIF - OMNIL - DRIEA, 2012) is a household survey investigating travel behavior in the Île-de-France region every ten years to assess transport usage in the region, with the latest available data from 2010. Each member of randomly selected households was interviewed face-to-face with regard to all trips performed on the day preceding the interview, resulting in 42,529 surveyed persons. Trips with multiple transport modes were separated into trip stages. The locations of the start and end points of each trip were defined within a grid with a spatial resolution of 100 m × 100 m squares, hereafter called tiles. For the current analysis, the age range was limited to the age range of the participants of the RECORD study who carried noise dosimeters (34–74 years old at the date of investigation), resulting in a sample of 18,929 out of 35,511 participants with trips during the past day. A comparison of demographics of the two samples can be found in Table 1. The mean age in the EGT sample retained for the analysis was 51.6 years (sd = 10.9). The RECORD MultiSensor study included relatively more men, more people with higher education and more people living in intramural Paris and the inner suburbs.

### 2.3. Noise map

To comply with the Environmental Noise Directive, the non-governmental organization Bruitparif is responsible for integrating

**Table 1**

Overview of the demographic characteristics of the RECORD MultiSensor and EGT study samples.

Variable	RECORD MultiSensor	EGT
<b>Sample size</b>		
Number of participants	77	18,929
Number of trips	2183	82,046
Number of trip stages	3729	207,923
<b>Demographic information</b>		
Age		
34-43 years (%)	28.6	34.0
44-53 years (%)	32.5	27.7
54-63 years (%)	22.1	23.0
64-74 years (%)	16.9	15.3
Gender		
Female (%)	36.4	52.5
Male (%)	63.6	47.5
Educational level		
No diploma of secondary education (%)	18.2	38.1
Diploma of secondary or lower tertiary education (%)	31.2	27.0
Diploma of higher tertiary education (%)	50.6	34.9
Household Income (€, mean)	3940	3169
<b>Location of residence</b>		
Inner city (Paris, %)	24.7	13.9
Inner suburbs (Petite Couronne, %)	41.6	35.5
Outer suburbs (Grande Couronne, %)	33.7	50.5

RECORD = Residential Environment and CORonary heart Disease study cohort, MultiSensor branch; EGT = Global Transport Survey.

modelled noise databases of Île-de-France municipalities. A noise map for the years 2007 to 2011 was constructed to represent the yearly averaged indicator for day, evening and night sound levels ( $L_{den}$ ) per transport-related noise source (road, railway and air traffic), starting at 55 dB(A) in classes of 5 dB(A). Additionally, the original noise layers were gathered and processed to obtain a map that covers the noise range between 30–80 dB(A) (Méline et al., 2015; Méline et al., 2013).

The three traffic source layers were combined to produce one integrative noise layer which was used to compute yearly averages of noise levels at the start and end location of each trip stage at a resolution of 100 m  $\times$  100 m tiles in both the RECORD MultiSensor and the EGT studies. ArcMap 10.5 Geographic Information System (ESRI, 2016) and ArcGIS python scripts (Python version 3.4) were used to intersect the required tiles with the polygons of the final noise layer. The process is illustrated in *Supplemental material A*. For each tile, the mean and quartiles of noise level and the percentage of surface covered by noise levels equal to or higher than 75 dB(A) were calculated using the R software, Version 3.3.2 (R Core Team, 2016). These indicators are hereafter referred to as noise map predictors.

#### 2.4. Statistical analysis

The 77 RECORD MultiSensor study participants considered in the present study undertook 2537 trips consisting of 4365 trip stages. After exclusion of trip stages that used rare transport modes (e.g., skateboard), crossed or were out of the Île-de-France department border, missed more than 50% of one-second noise  $L_{Aeq}$  values or lasted for less than one minute, 3729 trip stages remained in the dataset for analysis.

To ensure consistency with the literature and the Environmental Noise Directive,  $L_{Aeq,T}$  values were used as the outcome for the prediction models per transport mode. A multilevel model (MLM) was specified to predict trip stage level personal noise exposure from trip stage variables. A random intercept was specified at the individual level to account for the hierarchical structure of the data with various trip stages per participant, while variables at the trip stage level were added as fixed effects. Person level variables such as sex, age and educational level were excluded due to the small sample size.

For the cross-validation of MLM, the dataset was randomly split in half 10,000 times at the person level to generate pairs of independent datasets, hereafter called train and test sets. The random intercepts that were estimated for each participant in a train set could not be used for the corresponding test set, because participants of one dataset were excluded from the other. Instead, the estimated overall population intercept was used (which is called marginal prediction and used when the final aim is to conduct inference towards a wider population; Welham et al., 2004).

Based on the success of a previous data-enrichment approach by Brondeel et al. (2016), the non-parametric machine learning algorithm 'Random Forests' (RF) with 1000 trees per forest was used in addition to the MLM. RF consist of an ensemble of decision trees that are created by introducing two sources of randomness to the CART algorithm, bagging of observations and subsampling of variables at each tree node (Breiman, 2001). Bagging refers to the construction of each tree based on a bootstrapped subset of approximately two thirds of the unique observations, while at the same time only one-third of model predictors are used per node in the decision trees. An example and more information on RF can be found in *Supplemental material B* and publications by e.g. Tan et al. (2005), Strobl et al. (2009), and James et al. (2014). The algorithm is more robust against overfitting than single decision trees and has an internal validation mechanism, the out-of-bag  $R^2$ , i.e., an  $R^2$  that is estimated by predictions for each trip stage using every tree

for which this trip stage was not used in developing the model. It has been found that the out-of-bag  $R^2$  can be overestimated when the data is clustered, because trip stages of each given individual are split between those used to grow the forest and those in the out-of-bag sample (Karpievitch et al., 2009). Therefore, as a more robust validation procedure, the RF models were then cross-validated by splitting the dataset 100 times into a train and test set, containing 70% and 30% of the randomly sampled participants, respectively. The number of splits was lower for RF than for MLM due to the complex algorithm which requires longer computation times.

As a performance indicator for both model types, the proportion of variance explained ( $R^2$ ) by the full models in the train sets and test sets were calculated with the following formula to allow comparability:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

The denominator represents the total variation of the response variable  $y$  from the mean, while the numerator represents the variation of  $y$  from the predicted outcome.  $R^2$  can take negative values for MLM and RF which indicate that the prediction of the fitted model is worse than the horizontal line drawn through the overall mean noise exposure across all trip stages. Our a priori strategy was to apply the prediction model calibrated from the RECORD MultiSensor Study to the EGT dataset only if the model fit could be considered as good and the predictions accurate. The multilevel model was built with the R package ‘nlme’ (Pinheiro et al., 2016). Random forests were built with ‘randomForest’ R package (Liaw and Wiener, 2015) which implements Breiman’s RF (Breiman, 2001).

### 3. Results

#### 3.1. Personal dosimeter and map-based assessments of noise

A summary of personal noise exposures according to demographic and trip stage variables can be found in Table 2. The mean duration of a trip stage was 9.7 min. The overall average personal noise exposure across all trip stages for  $L_{Aeq,T}$  was 71.6 dB(A) (sd = 8.6 dB(A)).

Noise exposure was not very variable between age groups (range: 69.9–72.4 dB(A)) and was on average 2 dB(A) higher for men than for women. Personal noise values were on average higher in the second half of the week as compared to the beginning of the week and highest during summer compared with other seasons. Expected differences in exposure between transport modes were

**Table 2**

Descriptive statistics for average equivalent sound levels ( $L_{Aeq,T}$  in dB(A)) of trip stages in RECORD, overall and separated by participant and trip stage characteristics.

		$N_{Trip}$	Mean (sd) $L_{Aeq,T}$	P25 – p75 $L_{Aeq,T}$
Overall		3729	71.6 (8.6)	68.0 – 76.8
Age	34-43 years	1097	72.0 (9.7)	68.1 – 77.5
	44-53 years	1205	72.4 (6.6)	68.4 – 76.6
	54-63 years	806	69.9 (8.9)	66.8 – 75.2
	64-74 years	605	71.5 (9.4)	68.4 – 77.4
Sex	Female	1466	70.4 (10.0)	67.0 – 76.5
	Male	2247	72.4 (7.5)	68.4 – 76.9
Transport Mode	Walking	2061	70.5 (8.6)	66.9 – 75.6
	Bike	143	73.5 (7.2)	70.7 – 77.6
	Motorbike	155	79.5 (10.3)	74.7 – 84.8
	Car	720	71.2 (8.3)	67.4 – 76.1
	Bus	143	70.2 (6.7)	67.9 – 73.3
	Tram	46	73.8 (4.6)	72.1 – 75.8
	Metro	332	75.1 (7.6)	72.9 – 79.3
	RER, Train <sup>a</sup>	113	71.1 (8.0)	68.6 – 75.1
Day of the week	Monday	583	70.5 (9.1)	67.6 – 75.4
	Tuesday	618	70.5 (8.8)	67.1 – 75.7
	Wednesday	584	71.6 (8.7)	68.2 – 76.8
	Thursday	594	73.1 (6.6)	68.6 – 77.4
	Friday	511	72.4 (8.0)	68.1 – 77.3
	Saturday	490	72.0 (9.5)	68.5 – 77.7
	Sunday	333	71.1 (9.3)	67.1 – 76.3
Season	Spring	1359	69.9 (8.2)	66.9 – 74.5
	Summer	1126	73.7 (8.8)	69.4 – 78.4
	Autumn	492	71.8 (9.5)	68.2 – 77.9
	Winter	736	72.4 (6.8)	68.2 – 76.8

RECORD: Residential Environment and CORonary heart Disease study cohort, MultiSensor branch;  $N_{Trip}$  = Number of trip stages per factor level;  $L_{Aeq,T}$  = Time-averaged equivalent sound level; sd = Standard deviation; p25 = 25th percentile; p75 = 75th percentile; <sup>a</sup> Regional TER and SNCF trains.

**Table 3**

Fixed effects of two-level regression model with trip stage variables as fixed effects and person identifier as random effect on the full dataset to predict average equivalent sound levels ( $L_{Aeq,T}$ ) in dB(A) from personal dosimetry.

	Coefficient	95% CI
Intercept	67.02	[65.22; 68.83]
Mode: Walking	<i>Reference</i>	
Mode: Bike	1.42	[-0.08; 2.92]
Mode: Motorbike	5.95	[4.24; 7.65]
Mode: Car	1.23	[0.44; 2.02]
Mode: Bus	-0.42	[-1.76; 0.93]
Mode: Tram	1.83	[-0.49; 4.15]
Mode: Metro	3.68	[2.71; 4.64]
Mode: Train	-0.82	[-2.48; 0.83]
% noise > 75, start <sup>a</sup>	2.60	[0.61; 4.58]
% noise > 75, end <sup>a</sup>	3.13	[1.15; 5.11]
Season: Spring	<i>Reference</i>	
Season: Autumn	0.95	[-0.95; 2.85]
Season: Summer	-0.09	[-2.02; 1.84]
Season: Winter	0.80	[-1.00; 2.59]
Day: Monday	-0.72	[-1.77; 0.34]
Day: Tuesday	-0.61	[-1.67; 0.45]
Day: Wednesday	0.60	[-0.46; 1.67]
Day: Thursday	1.64	[0.57; 2.71]
Day: Friday	1.24	[0.15; 2.33]
Day: Saturday	0.93	[-0.15; 2.02]
Day: Sunday	<i>Reference</i>	
Rush hour <sup>b</sup>	0.41	[-0.16; 0.98]
Time: Morning	<i>Reference</i>	
Time: Afternoon	1.19	[0.63; 1.76]
Time: Evening	1.38	[0.48; 2.28]
Time: Night	1.10	[-1.84; 4.04]
Air Distance	0.09	[0.02; 0.15]
Paris, start	<i>Reference</i>	
Inner suburbs, start	0.78	[-0.59; 2.16]
Outer suburbs, start	1.89	[0.48; 3.30]
Paris, end	<i>Reference</i>	
Inner suburbs, end	0.00	[-1.37; 1.37]
Outer suburbs, end	0.64	[-0.76; 2.04]

95% CI: 95% confidence interval, dB(A): A-weighted decibels; <sup>a</sup> percentage of surface within start and end tile of departure covered by polygons representing noise exposure equal to or higher than 75 dB(A); <sup>b</sup> 7:00 h–9:00 h and 16:00 h–19:00 h.

recorded, with motorbike showing the highest exposure followed by subway trip stages, and then by tram and bike trip stages. Lower personal exposures were documented for car, train, walking and bus trip stages.

### 3.2. Multilevel model

#### 3.2.1. Model on the full dataset

A comparison of the empty model to predict  $L_{Aeq,T}$  with and without participant-level random effects justified the choice of a multilevel model ( $\chi^2 = 568.73$ ,  $df = 1$ ,  $p < 0.0001$ ;  $\Delta BIC = 560.74$ ). The variance partition coefficient of the multilevel model without fixed effects was 0.20, indicating that around 20% of the variance in noise exposure can be attributed to between-person variation.

The coefficients for the fixed effects and corresponding 95% confidence intervals (95% CI) of the multilevel model can be found in Table 3. The full model with random intercepts could explain 25.4% of the overall variance in  $L_{Aeq,T}$ , while the model that ignored random intercepts resulted in an  $R^2$  of 0.10. Several transport modes, i.e. motorbike, car and metro, showed higher noise exposures than walking. The predicted personal noise exposure  $L_{Aeq,T}$  increased by 2.6 dB(A) (95% CI: 0.61 to 4.58 dB(A)) and 3.13 dB(A) (95% CI: 1.15 to 5.11 dB(A)), respectively for a 0 to 100% change of the noise map predictor. It was found that trip stages starting in the outer suburbs compared to intramural Paris were associated with higher personal noise exposure. Trip stages on Thursdays and Fridays exposed participants to more personal noise compared with Sundays. Afternoons and evenings showed higher exposure than mornings ( $\beta = 1.19$  dB(A), 95% CI: 0.63 to 1.76 dB(A) and  $\beta = 1.38$ , 95% CI: 0.48 to 2.28 dB(A), respectively). Each kilometer increase in straight-line distance from start to end of a trip stage resulted in an increase of a bit less than 0.1 dB(A). All other variables were not associated with personal noise exposure.

### 3.2.2. Cross-validation

With a division into train and test set at the person level, the average  $R^2$  of the 10,000 models on the train sets when using random intercepts was 0.27 (range: 0.14 to 0.40). When using the population intercept instead for a prediction in the train set, the average  $R^2$  was 0.12 (range: 0.02 to 0.24). In contrast, the average  $R^2$  that was obtained by applying each train set model to its corresponding test set was 0.01 (range: -0.77 to 0.14), indicating a very low predictive ability for other participants than those based on which the model was estimated. For further post-hoc examination, multilevel models were also estimated in each test set, and the regression coefficients for the transport modes and noise map predictors were compared between each train and test set pair of models.

### 3.3. Random forests

#### 3.3.1. Model on the full dataset

Two versions of RF were built, one including and one excluding person level variables (sex, age, occupation, educational level, household income, and number of people living in a household). Several noise map predictors were entered into the models (mean, median, 25th and 75th percentile of the proportion of the 100 m tiles at the start and end of trip stages covered by noise polygons), as random forests are able to deal with multicollinearity of the predictors (Hayes et al., 2015; Strobl et al., 2008). As RF result in slightly different predictions for each run due to the randomness introduced by bagging and random selection of variables at each tree node, RF were checked for consistency of prediction accuracy by iterating each version (one with and one without person level variables) 50 times. The model including person level variables obtained an average out-of-bag  $R^2$  of 0.37 (range: 0.37–0.37) and an average train set fit of 0.84 (range: 0.84–0.84). Excluding the person level variables, the model was still able to result in a prediction accuracy of 0.27 (range: 0.26–0.27) for the out-of-bag  $R^2$  and 0.81 (range: 0.81–0.81) of variance explained for the dataset itself. However, it should be noted that the out-of-bag observations originated from the same participants as the observations used to build the trees.

#### 3.3.2. Cross-validation

To test if RF can predict noise exposure for new participants, the dataset was split 100 times randomly into a train set (70% of the sample) and a test set at the person level. For the train set, the RF procedure automatically estimated an  $R^2$  based on out-of-bag observations. In addition, we recalculated  $R^2$  for the whole train set and for the test set. The average and range of the out-of-bag  $R^2$ , the whole train set  $R^2$ , and test set  $R^2$  per version of the model are shown in Table 4. The model including the person level variables achieved an out-of-bag  $R^2$  of 0.39 (range: 0.26–0.48) and a good average fit on the train dataset. However, the average  $R^2$  for predictions on the test set (for participants that were not considered to grow the model) was 0.01, ranging from -0.12 to 0.10. Similarly, the predictions for the test sets of the model with only trip stage variables resulted in an  $R^2$  of -0.02 (range: -0.15 to 0.10). The predictive ability of the model was thus close to zero and its range was very broad. Based on these results, the estimated models were not applied to obtain noise level estimates in the EGT dataset.

## 4. Discussion

### 4.1. Summary of the results

The current study aimed to describe noise exposure during daily trips, to build a model that can predict this noise exposure, and to apply the model to a larger sample that comes from the same source population. In line with previous findings in different locations (Kraus et al., 2015; Taimisto et al., 2013; Neitzel et al., 2009), noise exposure summarized as the  $L_{Aeq,T}$  on average exceeded the recommended level of 70 dB(A) for all transport modes. With an average transport time of 93.2 min per participant per day, personal transport is a considerable source of noise exposure.

The multilevel model showed that approximately 20% of the variance in noise exposure in our sample was at the person level. Around 9.6% of the variance could be explained by the fixed effects, where transport mode, day of the week and time of the day, the noise map predictor, the straight-line distance between start and end tiles were found to be important predictors. The  $R^2$ , and thus the predictive ability of the models, was relatively low even when estimated for the participants who were used to derive the models. However, our cross-validation procedure splitting the trip stages at the participant level showed that the prediction accuracy was even lower and included negative values when models were applied to the participants who were not considered to derive the model. These negative values indicate that the mean across all predictions can therefore explain as much of the variance in the data as our model. Although the non-parametric RF can describe the train data itself very well, the out-of-bag  $R^2$  (estimated from trip stages from the same participants as those used to derive the model) indicated that RF performed better than MLM but worse than expected. RF

**Table 4**

Average and range of the out-of-bag  $R^2$  estimated in the train set (%), of the  $R^2$  estimated in the whole train set  $R^2$  (%) and of the test set  $R^2$  (%) per version of the model predicting the noise outcome  $L_{Aeq,T}$ .

Model	Out-of-bag $R^2$		$R^2$ train set		$R^2$ test set	
	Mean	Range	Mean	Range	Mean	Range
Model with person level variables	0.39	0.26; 0.48	0.84	0.81; 0.87	0.01	-0.12; 0.10
Model without person level variables	0.29	0.18; 0.38	0.81	0.78; 0.85	-0.02	-0.15; 0.10

estimates were not better than MLM estimates when applied to other participants than those used to derive the models. These inaccurate predictions did not allow us to apply the model to the EGT dataset, therefore objective d could not be realized.

#### 4.2. Interpretation and limitations

The idea for the current study was based on the successful application of RF for the prediction of the number of minutes spent in moderate-to-vigorous physical activity during trips (Brondeel et al., 2015, 2016). Even though comparable datasets and partly the same methodology were used, there are substantial differences in the predicted outcomes which require a different set of predictors. Unfortunately, a limitation inherent to the data-enrichment approach, here from a detailed dosimeter study to a large representative household study, is the necessity to use only predictors available in both datasets. Some potentially relevant factors for noise exposure may have been missed due to the unavailability in either one or both datasets. For example, if we were interested only in the RECORD MultiSensor study participants, additional variables such as the travel speed, the exact distance covered, and information based on more precise coordinates such as GPS tracks in combination with the noise map could have been entered as predictors to allow for more efficient use of data. Although the predictors used in the current MLM were associated with the personal exposure, the coefficient of 3.1 was small for an increase from 0% to 100% of the surface of the tile exceeding 75 dB(A). We expected the RF to account for the unequal importance of the noise map predictors for the different transport modes (e.g., outdoor road, air, and railway traffic noise may not be meaningful to predict the personal exposure for the underground subway trip stages). In addition, the specific map we used represented the annual average noise exposure without any daily, weekly or seasonal variation.

Due to the relatively small participant number, some variables contained numerous factor levels. One of these variables was the specific transport line covered in each of the public transport trips; however, many of these lines occurred with low frequencies in our sample and could thus not be entered into the model, but may have implicitly incorporated information on speed, crowding, environmental factors, and type and age of vehicles typically. For instance, train and subway noise levels are influenced by the presence of squealing noises and public audio system, train type and external (Tabacchi et al., 2011; Soeta and Shimokura, 2011), while car noise differs by car type, opening of windows, type of road, number of intersections, etc. For the same reason, person level demographic variables were excluded, which could contain valuable information on behavioral aspects linked to noise exposure. The person-level clustering indicates that there may be some factors at the individual level that influence noise exposure and may explain the non-generalizability of estimated associations to participants who were unknown to the model. One major influence potentially arose from speech because some participants may preferably perform their trips in the company of friends or relatives, and contribution of speech to overall noise levels is dependent on the background levels (Ryherd et al., 2012), increasing unpredictability of the measurements. Other sources of person level variation, dependent on the transport mode, may be route and speed preferences, population density and other characteristics of areas commonly visited, vehicle type, sound system usage, and driving style (Mikulec et al., 2011). Future prediction models based on a larger dataset may show if the transport line and person level predictors can capture some of the variability in noise exposure.

Due to this person level variation, RF were tested for prediction because they are known to capture complex interactions without explicit specification and are commonly used for clustered data (Karpievitch et al., 2009). Therefore, the huge difference between the out-of-bag estimate and the test set prediction (on independent participants) was unexpected. Martin (2015) showed that random forests based on the CART algorithm performed poorly in his simulated condition with a high intra-class correlation of 0.30 and a small sample size (20 clusters with 15 observation each) (Martin, 2015), resulting in negative values for  $R^2$ . The explanation of his results adapted to our findings may sound as follows: the trip stages undertaken by a given participant resemble each other more than when compared with the trip stages by other participants. Therefore, when a tree in a forest is exposed to trip stages of a participant, its knowledge is higher for other trip stages of this participant than for trip stages of other participants. As a consequence, the RF majority vote is more adapted to the participants used for forest construction than for new participants (Martin, 2015).

Predictions for the known participants was not outstanding, possibly due to the lack of specificity of the predictors as outlined before, but potentially also due to the way noise exposure was summarized. We chose the  $L_{Aeq,T}$  due to its comparability with current regulations. However, noise exposure expressed as a single value may not describe the variation of soundscapes between micro-environments sufficiently, as the  $L_{Aeq,T}$  can take the same value for different patterns of exposure. It is possible that different transport modes have very diverse patterns of noise exposure in regard to the frequency bands, rhythm, and noise level peaks. Noise summary measures with different meanings are for example the number of noise events or percentiles of one-second interval  $L_{Aeq}$  values such as  $L_{A95}$  and  $L_{A5}$  (Can et al., 2016). More detailed noise data will be available from a currently ongoing study (MobiliSense) in our lab, which may allow for more discriminable summary measures in the future. Data and scripts can be provided upon request.

#### 4.3. Conclusion

In line with previous studies, the current study found that noise exposure during daily travels is in general high for all transport modes. While models with moderate prediction ability for the participants used to develop the models could be built based on both MLM and RF, the prediction accuracy for other participants was low. Accordingly, we decided to not conduct a prediction of noise levels for the trip stages of the large representative EGT sample. The results may be explained by limitations in the predictors that were used, by the small sample size of the model development set, and by the way noise was summarized.

## Acknowledgments

We would like to thank Claire Merrien, Audrey Pelletier and Valérie Thériau for their efforts in the data collection of the RECORD MultiSensor Study.

## Funding

The RECORD MultiSensor Study was supported by the Ministry of Ecology (DGITM); Cerema (Centre for the Study of and Expertise on Risks, the Environment, Mobility, and Planning); INPES (National Institute for Prevention and Health Education); Ile-de-France Mobilités (Ile-de-France Transportation Authority); and DRIEA (Regional and Interdepartmental Direction of Equipment and Planning) of Île-de-France.

## Conflict of interest

None.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <https://doi.org/10.1016/j.jth.2019.02.002>.

## References

- Babisch, W., 2011. Cardiovascular effects of noise. *Noise Health* 13, 201–204. <https://doi.org/10.4103/1463-1741.80148>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45. <https://doi.org/10.1023/a:1010933404324>.
- Brondeel, R., Pannier, B., Chaix, B., 2015. Using GPS, GIS, and accelerometer data to predict transportation modes. *Med. Sci. Sport. Exerc.* 47, 2669–2675. <https://doi.org/10.1249/mss.0000000000000704>.
- Brondeel, R., Pannier, B., Chaix, B., 2016. Associations of socioeconomic status with transport-related physical activity: combining a household travel survey and accelerometer data using random forests. *J. Transp. Health* 3, 287–296. <https://doi.org/10.1016/j.jth.2016.06.002>.
- Can, A., Aumond, P., Michel, S., De Counsel, B., Ribeiro, C., Botteldooren, D., Lavandier, C., 2016. Comparison of noise indicators in an urban context. In: Proceedings of 45th International Congress and Exposition on Noise Control Engineering (Inter-noise 2016) Hamburg, Germany. pp. 5678–5686. (Retrieved from: <https://hal.archives-ouvertes.fr/hal-01373857/file/doc00025834.pdf>); Last accessed 1 February 2019).
- Chaix, B., Kestens, Y., Bean, K., Leal, C., Karusisi, N., Meghrief, K., Burban, J., Fon Sing, M., Perchoux, C., Thomas, F., Merlo, J., Pannier, B., 2012. Cohort profile: residential and non-residential environments, individual activity spaces and cardiovascular risk factors and diseases – the RECORD cohort study. *Int. J. Epidemiol.* 41, 1283–1292. <https://doi.org/10.1093/ije/dyr107>.
- Chaix, B., Kestens, Y., Duncan, S., Merrien, C., Thierry, B., Pannier, B., Brondeel, R., Lewin, A., Karusisi, N., Perchoux, C., Thomas, F., Méline, J., 2014. Active transportation and public transportation use to achieve physical activity recommendations? A combined GPS, accelerometer, and mobility survey study. *Int. J. Behav. Nutr. Phys. Act.* 11, 124. <https://doi.org/10.1186/s12966-014-0124-x>.
- Dzhambov, A., 2015. Long-term noise exposure and the risk for type 2 diabetes: a meta-analysis. *Noise Health* 17, 23–33. <https://doi.org/10.4103/1463-1741.149571>.
- Esri, 2016. ArcGIS Desktop 10.5. Environmental Systems Research Institute, Redlands, CA.
- European Environment Agency, 2010. Good Practice Guide on Noise Exposure and Potential Health Effects. European Environment Agency, Copenhagen (Retrieved from: <https://www.eea.europa.eu/publications/good-practice-guide-on-noise>); Last accessed 02 February 2019).
- Floud, S., Blangiardo, M., Clark, C., De Hoogh, K., Babisch, W., Houthuijs, D., Swart, W., Pershagen, G., Katsouyanni, K., Velonakis, M., Vigna-Taglianti, F., Cadum, E., Hansell, A.L., 2013. Exposure to aircraft and road traffic noise and associations with heart disease and stroke in six European countries: a cross-sectional study. *Environ. Health* 12, 89. <https://doi.org/10.1186/1476-069X-12-89>.
- Halonen, J.I., Hansell, A.L., Gulliver, J., Morley, D., Blangiardo, M., Fecht, D., Toledano, M.B., Beevers, S.D., Anderson, H.R., Kelly, F.J., Tonne, C., 2015. Road traffic noise is associated with increased cardiovascular morbidity and mortality and all-cause mortality in London. *Eur. Heart J.* 36, 2653–2661. <https://doi.org/10.1093/eurheartj/ehv216>.
- Hayes, T., Usami, S., Jacobucci, R., Mcardle, J.J., 2015. Using Classification and Regression Trees (CART) and random forests to analyze attrition: results from two simulations. *Psychol. Aging* 30, 911–929. <https://doi.org/10.1037/pag0000046>.
- Hume, K., Brink, M., Basner, M., 2012. Effects of environmental noise on sleep. *Noise Health* 14, 297–302. <https://doi.org/10.4103/1463-1741.104897>.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2014. *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated.
- Karpievitch, Y.V., Hill, E.G., Leclerc, A.P., Dabney, A.R., Almeida, J.S., 2009. An introspective comparison of random forest-based classifiers for the analysis of cluster-correlated data by way of RF + +. *PLoS One* 4, e7087. <https://doi.org/10.1371/journal.pone.0007087>.
- Kraus, U., Breitner, S., Hampel, R., Wolf, K., Cyrus, J., Geruschkat, U., Gu, J., Radon, K., Peters, A., Schneider, A., 2015. Individual daytime noise exposure in different microenvironments. *Environ. Res.* 140, 479–487. <https://doi.org/10.1016/j.envres.2015.05.006>.
- Liaw, A. & Wiener, M., 2015. 'randomForest' - Breiman and Cutler's Random Forests for Classification and Regression. CRAN, R package version 4.6-12. (Retrieved from: <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf>); Last accessed: 2 February 2019).
- Lie, A., Skogstad, M., Johannessen, H.A., Tynes, T., Mehlum, I.S., Nordby, K.-C., Engdahl, B., Tambs, K., 2016. Occupational noise exposure and hearing: a systematic review. *Int. Arch. Occup. Environ. Health* 89, 351–372. <https://doi.org/10.1007/s00420-015-1083-5>.
- Martin, P.D., 2015. Efficiently exploring multilevel data with recursive partitioning. Dissertation, University of Virginia. (Retrieved from: <https://dpmartin42.github.io/extras/dissertation.pdf>); Last accessed: 2 February 2019).
- Méline, J., Van Hulst, A., Thomas, F., Chaix, B., 2015. Road, rail, and air transportation noise in residential and workplace neighborhoods and blood pressure (RECORD Study). *Noise Health* 17, 308–319. <https://doi.org/10.4103/1463-1741.165054>.
- Méline, J., Van Hulst, A., Thomas, F., Karusisi, N., Chaix, B., 2013. Transportation noise and annoyance related to road traffic in the French RECORD study. *Int. J. Health Geogr.* 12, 44. <https://doi.org/10.1186/1476-072x-12-44>.
- Mikulec, A.A., Lukens, S.B., Jackson, L.E., Deyoung, M.N., 2011. Noise exposure in convertible automobiles. *J. Laryngol. Otol.* 125, 121–125. <https://doi.org/10.1017/s0022215110002355>.
- Neitzel, R., Gershon, R.R.M., Zeltser, M., Canton, A., Akram, M., 2009. Noise levels associated with New York City's mass transit systems. *Am. J. Public Health* 99, 1393–1399. <https://doi.org/10.2105/ajph.2008.138297>.
- Pinheiro J., Bates D., Debroy S., Sarkar D. & Core Team, R., 2016. nlme: Linear and Nonlinear Mixed Effects Models. CRAN, R package version 3.1-128. (Retrieved from: <https://cran.r-project.org/web/packages/nlme/nlme.pdf>); Last accessed: 2 February 2019).
- R Core Team, 2016. *R: A Language and Environment for Statistical Computing*. Foundation for Statistical Computing, Vienna, Austria.

- Ryherd, S., Kleiner, M., Wayne, K.P., Ryherd, E.E., 2012. Influence of a wearer's voice on noise dosimeter measurements. *J. Acoust. Soc. Am.* 131, 1183–1193. <https://doi.org/10.1121/1.3675941>.
- Selander, J., Bluhm, G., Theorell, T., Pershagen, G., Babisch, W., Seiffert, I., Houthuijs, D., Breugelmans, O., Vigna-Taglianti, F., Antoniotti, M.C., Velonakis, E., Davou, E., Dudley, M.-L., Järup, L., 2009. Saliva cortisol and exposure to aircraft noise in six European countries. *Environ. Health Perspect.* 117, 1713–1717. <https://doi.org/10.1289/ehp.0900933>.
- Skogstad, M., Johannessen, H.A., Tynes, T., Mehlum, I.S., Nordby, K.C., Lie, A., 2016. Systematic review of the cardiovascular effects of occupational noise. *Occup. Med.* 66, 10–16. <https://doi.org/10.1093/occmed/kqv148>.
- Soeta, Y., Shimokura, R., 2011. The impact of external environments on noise inside a train car. *Noise Control Eng. J.* 59, 581–590. <https://doi.org/10.3397/1.3628526>.
- Sørensen, M., Andersen, Z.J., Nordsborg, R.B., Jensen, S.S., Lillelund, K.G., Beelen, R., Schmidt, E.B., Tjønneland, A., Overvad, K., Raaschou-Nielsen, O., 2012. Road traffic noise and incident myocardial infarction: a prospective cohort study. *PLoS One* 7, e39283. <https://doi.org/10.1371/journal.pone.0039283>.
- Stif - Omnif - Driea, 2012. Enquête globale transport (EGT): La mobilité en Île-de-France. (Retrieved from: [http://www.omnil.fr/IMG/pdf/egt2010\\_enquete\\_globale\\_transports\\_-\\_2010.pdf](http://www.omnil.fr/IMG/pdf/egt2010_enquete_globale_transports_-_2010.pdf)); Last accessed: 2 February 2019).
- Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. *BMC Bioinforma* 9. <https://doi.org/10.1186/1471-2105-9-307>.
- Strobl, C., Malley, J., Tutz, G., 2009. An introduction to recursive partitioning: rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychol. Methods* 14, 323–348. <https://doi.org/10.1037/a0016973>.
- Tabacchi, M., Pavon, I., Ausejo, M., Asensio, C., Recuero, M., 2011. Assessment of noise exposure during commuting in the Madrid subway. *J. Occup. Environ. Hyg.* 8, 533–539. <https://doi.org/10.1080/15459624.2011.600237>.
- Taimisto, P., Yli-Tuomi, T., Pennanen, A., Vuorisalo, I., Samaras, Z., Keuken, M.P., Lanki, T., 2013. Noise exposure during commuting in three European cities. *Internoise Innsbruck, Austria*. 1–5. (Retrieved from: <https://repository.tudelft.nl/view/tno/uuid%3A9b2682ff-5a95-4e80-a8f6-fc60a2df1523>); Last accessed: 2 February 2019).
- Tan, P.-N., Steinbach, M., Kumar, V., 2005. *Introduction to Data Mining, First Edition*. Addison-Wesley Longman Publishing Co., Inc.
- Van Kempen, E., Babisch, W., 2012. The quantitative relationship between road traffic noise and hypertension: a meta-analysis. *J. Hypertens.* 30, 1075–1086. <https://doi.org/10.1097/HJH.0b013e328352ac54>.
- Welham, S., Cullis, B., Gogel, B., Gilmour, A., Thompson, R., 2004. Prediction in linear mixed models. *Aust. New Zealand J. Stat.* 46, 325–347. <https://doi.org/10.1111/j.1467-842X.2004.00334.x>.
- Wolf, J., Schönfelder, S., Samaga, U., Oliveira, M., Axhausen, K., 2004. Eighty weeks of Global Positioning System traces: approaches to enriching trip information. *Transp. Res. Rec.: J. Transp. Res. Board* 1870, 46–54. <https://doi.org/10.3141/1870-06>.