Vehicle motion patterns for energy research: Comparison of annual mileage using vehicle- and person-based data

Michele Bolla, ETH Zurich
Maximilian Held, ETH Zurich
Lukas Küng, ETH Zurich
Gil Georges, ETH Zurich
Konstantinos Boulouchos, ETH Zurich

Conference paper STRC 2018

STRC | 18th Swiss Transport Research Conference
Monte Verità / Ascona, May 16-18 2018

Vehicle motion patterns for energy research: Comparison of annual mileage using vehicle- and person-based data

Michele Bolla
Aerothermochemistry and Combustion Systems Laboratory (LAV), ETH Zurich

Phone: +41 44 632 63 54
email: mbolla@lav.mavt.ethz.ch

Maximilian Held
Aerothermochemistry and Combustion Systems Laboratory (LAV), ETH Zurich

Phone: +41 44 632 03 53
email: held@lav.mavt.ethz.ch

Lukas Küng
Aerothermochemistry and Combustion Systems Laboratory (LAV), ETH Zurich

Phone: +41 44 633 85 66
email: lukas.kueng@lav.mavt.ethz.ch

Gil Georges
Aerothermochemistry and Combustion Systems Laboratory (LAV), ETH Zurich

Phone: +41 44 632 06 85
email: gil.georges@lav.mavt.ethz.ch

Konstantinos Boulouchos
Aerothermochemistry and Combustion Systems Laboratory (LAV), ETH Zurich

Phone: +41 44 632 36 68
email: boulouchos@lav.mavt.ethz.ch

May 2018

Abstract

Mobility patterns define the energy demand of passengers, respectively of the used vehicles. The daily and annual distribution of those patterns are of particular relevance to electric mobility; they define both the daily peak charging power demand and the overall need for additional electricity generation capacity. These are crucial inputs for life cycle assessments and energy research in general. A common approach to describe mobility patterns on a national level is to categorize a limited amount of personal mobility information (e.g. microcensus) and extrapolate it onto the whole population. The quality of the mobility behaviour distribution is highly dependent on the size and quality of the sample data. We present a comparison of two different approaches for describing the vehicle annual mileage. To that end, we analyse the distribution of mileage with data collected at different scales, i.e. daily resolved data from microcensus and yearly resolved data from maintenance logs. We conclude with a discussion on the potential complementarity of both approaches to gain further insights into individual mobility.

Keywords

Mobility – Mobility patterns – energy demand of vehicles – Household travel survey – annual mileage– Impact electrified mobility – Life-cycle assessment

1. Introduction

Mobility patterns define the energy demand of passengers. The daily and annual distribution of those patterns are of particular relevance to electric mobility; they define both the daily peak power demand and the overall need for additional electricity generation capacity. These are crucial inputs for life cycle assessments and energy research in general. The future of the Swiss energy system will be shaped by additional demand from electric mobility [1]. The ability to model realistic and temporally and geographically finely resolved energy demand patterns is crucial to energy research.

In this work we look at mobility using vehicle-based data to obtain a qualitative picture of the annual mileage, when classically people look at it via persons. For example in Switzerland with data from the Swiss household travel survey [2] (e.g. [3], [4]). It is noted here that the need for mobility is driven by persons, but the energy is demanded by the vehicles.

Annual mileage is a significant input to life-cycle assessment (LCA) consideration and underlies high uncertainties. Cox et al. [5] discuss the impact of different assumptions on this value. The standard approach of a LCA relies on European average mileages for different vehicle size classes. Our method to obtain more detailed mileage data would allow for higher resolution of this type of LCA input.

We aggregate the vehicle-based data from different calendar years into a temporally static representation. There is no intention to draw conclusions on the causalities governing annual mileage nor extrapolating the evolution of this data in time.

In this work we present a vehicle-based dataset denoted as to AMAG with the aim of an alternative description for the distribution of annual mileage in Switzerland.

The remainder of the paper is organized as follows. In the next section, the vehicle-based data is presented and its representativeness is evaluated by comparing it with the entire Swiss car fleet as collected in the motor vehicle information system (MOFIS) of the Federal Roads Office. In the results section, the aggregated annual vehicle mileage is presented as a function of vehicle age and population density. Subsequently, the average person-km in the various cantons is reconstructed and compared with data from the Household travel survey of the Swiss Federal Office of Statistics.

2. Vehicle-based data and methodology

2.1 Vehicle-based data

The vehicle-based data has been provided by AMAG AG, a Swiss car importer and dealer for Audi, Volkswagen, Skoda, Seat and until 2008 Porsche. The data consists of car maintenance-logs, where the vehicle mileage is recorded at every maintenance taking place in an AMAG AG subsidiary garage in Switzerland. The vehicle count for which multiple maintenances are available is approximately 300'000. This dataset is hereafter referred as to "AMAG". Various information about the car is available, such as: car brand and model, list price, date of first registration, fuel type, engine capacity and power, and curb weight. In addition, the Zip code and canton of the car's owner is available. As descriptive variables for our analysis we used vehicle age, canton and the population density in the zip codes. The latter was computed from population numbers and surface areas without lakes at the Zip code resolution.

The annual mileage of a single car is computed as the distance driven divided by the time difference between two maintenances. When computing the aggregated annual mileage, the single mileage between the maintenances is multiplied by the time between them. This annual mileage is attributed to the vehicle age in the middle of this period. The period between recorded maintenances represents the temporal resolution of the computed annual distance. The typical duration of this period is between 1 and 2 years, with a significant probability peak at 1 year as illustrated in Figure 1 in form of probability density function (PDF) (red line) and cumulative distribution (blue line). This means that for a large portion of vehicles the maintenance is determined by the calendar date rather than by the mileage. A minimum of 90 days between maintenances was considered and samples with a shorter interval have been omitted (less than 1% of total). This temporal resolution is sufficient in terms of input for introducing age conditioning using an age bin size of 1 years or higher. The scope here is not to explore seasonal mobility, but rather total driving distances in time during the aging of the vehicle, as potential input for LCA-oriented considerations.

On average, a maintenance-log contains around 3 records (in addition to the first registration). This translates into 3 periods taken as distinct samples. The total dataset includes a total of 877k unique samples with the complete information of the descriptive variables.

We are aggregating all cars into the vehicle age and population density coordinates independent of the calendar year, i.e. we neglect that these events take place over different calendar years, most of them during the 2010-2017 period. The resulting aggregated annual mileage is therefore not assumed to be a function of the calendar year. From the dataset, the most representative year of the fleet is 2015.

Figure 1 Probability density function (red line) and cumulative probability (blue line) of time period (in years) between two maintenances for the AMAG dataset.

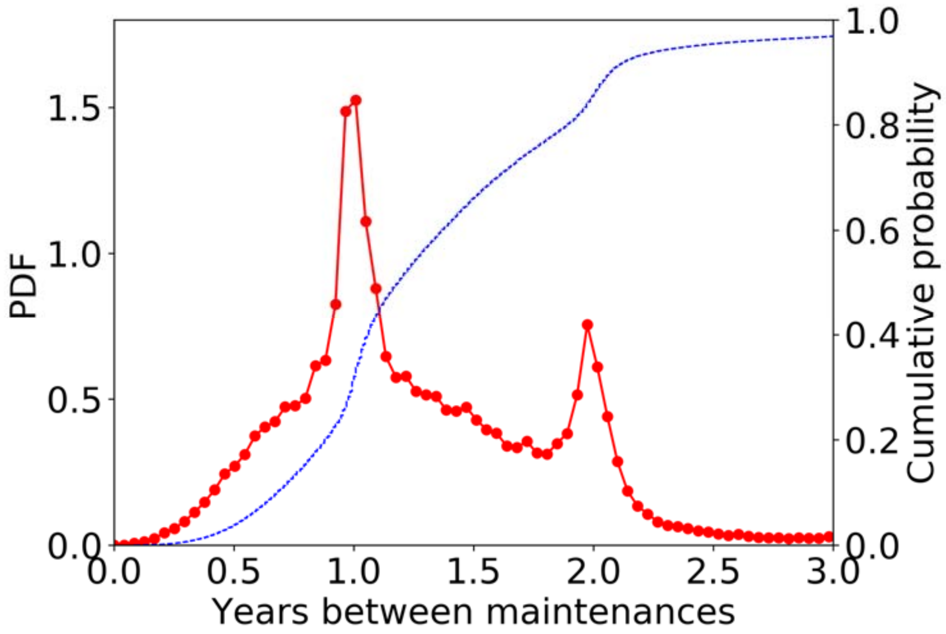
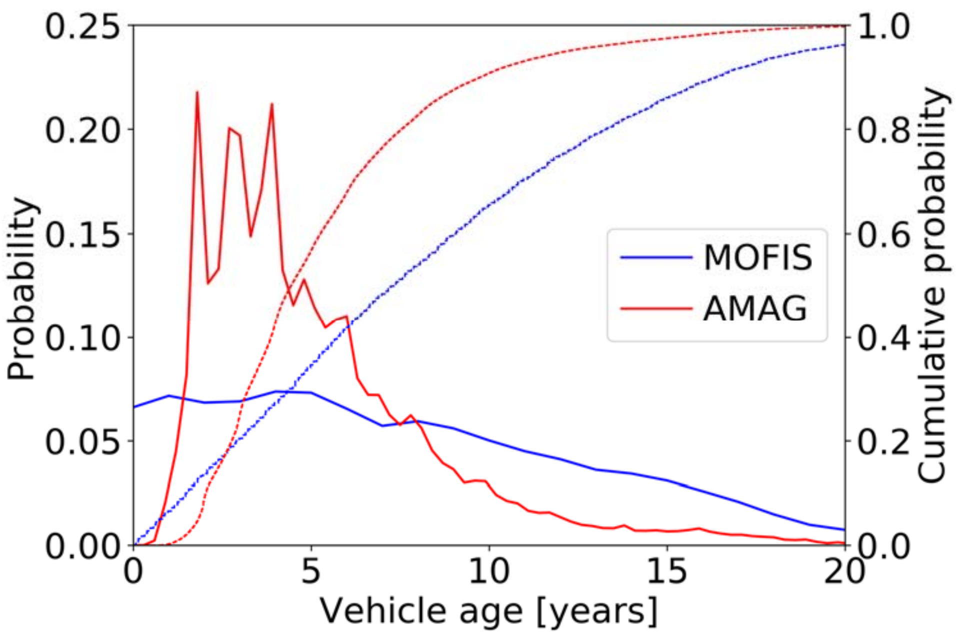


Figure 2 Probability density function (solid lines) and cumulative probability (dashed lines) of vehicle age for AMAG (red) and MOFIS (blue).



2.2 Representativeness of vehicle-based data

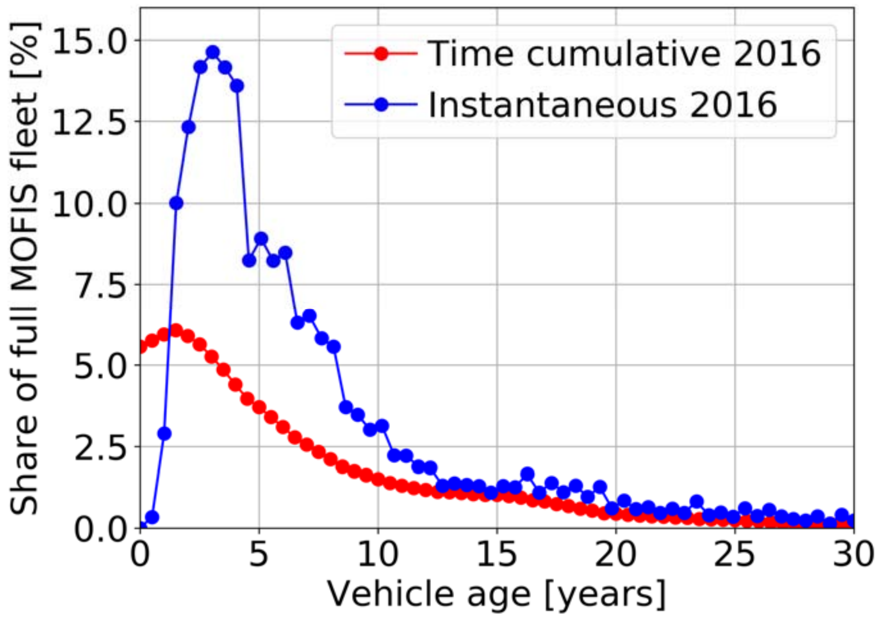
The representativeness of the available vehicle fleet with a maintenance-log is evaluated by comparing it with the entire Swiss car fleet as collected in the motor vehicle information system (MOFIS) of the Federal Roads Office ASTRA [6] accessed for the year 2016. The MOFIS fleet consists of 4.5 million matriculated cars. Figure 2 shows the PDF of vehicle age from MOFIS and AMAG. The AMAG dataset shows an over-representation of newer cars (age between 2 and 5 years). The median of vehicle age of the database is around 4 years as compared to 8 years for the full fleet. This discrepancy is expected, since AMAG AG as a car retailer and maintenance provider supplies its services to the first car owner and not necessarily to subsequent ones. Therefore the sample set is biased towards newer cars.

The distribution of the vehicle sample size as a function of vehicle age compared to the full MOFIS fleet is shown in Figure 3. It is important to note that the two datasets are of a different nature. AMAG data is aggregated for all calendar years and the temporal history of the vehicle during its life-time until the last maintenance is known. On the other hand, MOFIS provides the full list of the Swiss matriculated fleet on a given date (here we used 31.12.2016) without history.

For this reason, two different representations of the AMAG relative share of the full MOFIS fleet are provided, denoted as: (1) Instantaneous 2016 and (2) Time cumulative 2016. The first considers the instantaneous MOFIS fleet in 2016 and the AMAG vehicles at the date of their last maintenance (no temporal history). The second variant considers the same MOFIS and AMAG instantaneous fleets but the samples are collected by considering the entire life-time of the vehicles. For example, when counting the number of 10-year old cars, all cars with an age greater or equal to 10 years are considered. This procedure is conducted consistently for both datasets. For MOFIS it is inherently assumed that the shape of the vehicle age PDF is independent of time. From an LCA perspective, the time cumulative assessment of the relative share is more of interest because the characteristic life-time of the fleet is qualitatively included.

Results in Figure 3 show that the AMAG instantaneous share of cars between 2 and 4 years is between 10 and 15% and there is a decreasing relative share for older vehicles. Apart for cars up to 1 year (no maintenance conducted yet), the dataset contains at least 5, 2.5 and 1% of the full fleet for cars up to 8, 10 and 15 years. The time cumulative relative share of AMAG is lower because MOFIS exhibits a larger portion of cars older than 10 years compared to AMAG (approximately 40% against 10%). This contributes to a larger amount of time cumulative samples for cars younger than 10 years. The dataset contains at least 5, 2.5 and 1% of the full fleet for cars up to 4, 7 and 15 years.

Figure 3 Share of AMAG data compared to full MOFIS fleet as a function of vehicle age. Instantaneous 2016 (blue line) considers the AMAG vehicle age at the latest maintenance recorded and for MOFIS the vehicle age in 2016. Time cumulative 2016 (red line) considers the temporal history of every car from registration to the age in 2016 for both AMAG and MOFIS.



As a further step in the assessment of the representativeness of the AMAG dataset is its distribution of vehicles across the various regions (cantons) compared to the full fleet as displayed in Figure 4. The relative share of vehicles along cantons is similar between AMAG and MOFIS. This is an encouraging indication for the representative distribution of the vehicles along the different regions. The cantons with a small population (such as AI, AR and OW) don't have a sufficient sample size for a representative cantonal specific assessment.

Vehicle age is one of the determining factors that influences the annual mileage. Figure 5 illustrates the cantonal variability of the vehicle age distribution from the MOFIS database. Every thin line represents a canton and the red bold line corresponds to the overall Swiss vehicle age PDF. There is a cantonal variability, in particular for cars up to 5 years, but the overall decaying of presence of older cars is similar between cantons. This indicates that the vehicle age is a more appropriate descriptive variable compared to the canton.

Figure 4 Relative share of AMAG (red) and MOFIS (blue) fleets across cantons.

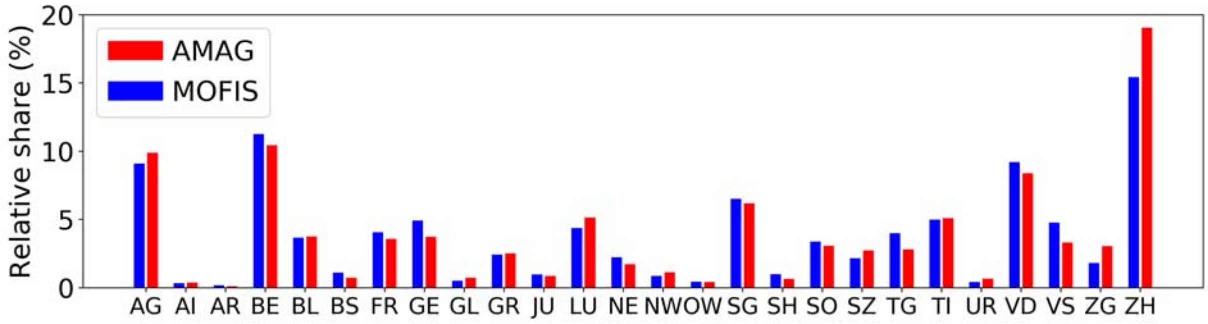
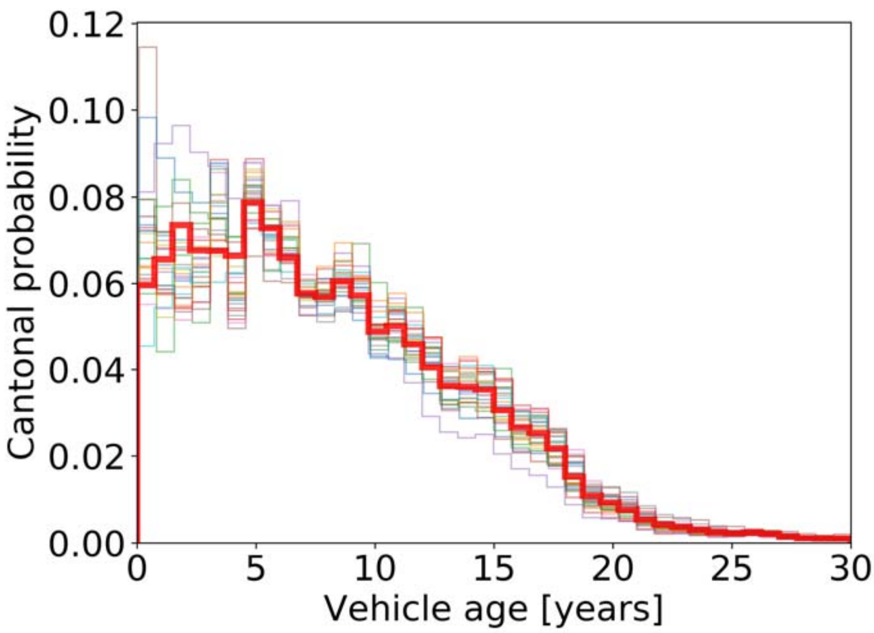


Figure 5 Probability density function of vehicle age for full Swiss car fleet MOFIS in 2016 in every canton (thin lines) and overall in Switzerland (bold red line).



3. Results and discussion

3.1 Mean annual mileage

Figure 6 shows the aggregated mean and median of annual mileage conditioned on vehicle age. The x-axis has been clipped at 30 years because the amount of cars is very limited for higher ages. The binning size in age when computing the conditional mean and median has been chosen to increase linearly in age to guarantee a sufficient sample size for older cars. The half-width age bin size $dt/_2$ was defined as:

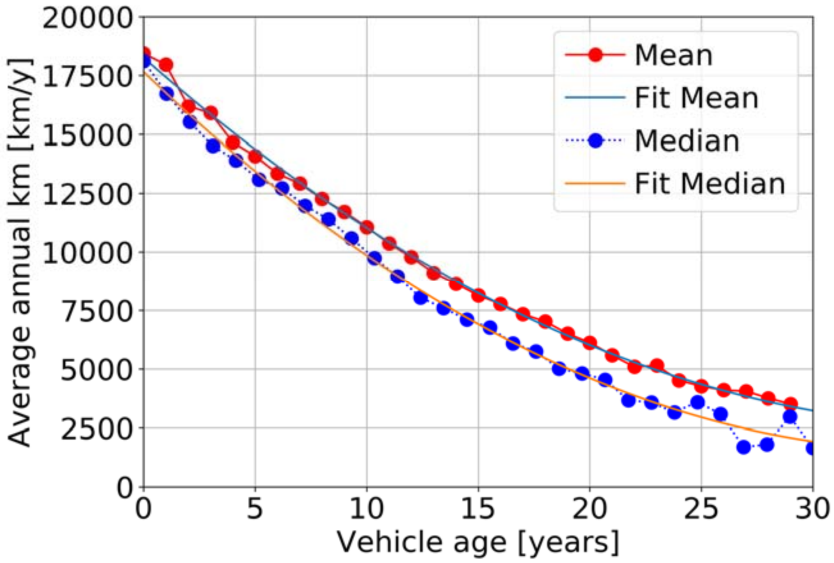
$$\frac{dt}{2} = 0.5y + 0.05 \cdot age$$

As expected, there is a clear decrease of the annual driving for older cars. A quadratic regression line is added for illustration purpose. The mean annual drive is approximated with the parabolic function:

$$\frac{km}{y}(age) = 18{,}250\frac{km}{y} - 833\frac{km}{y^2} \cdot age + 11\frac{km}{y^3} \cdot age^2$$

Starting from an average of 18,250 km/y for new cars there is initially a nearly linear decrease of around 833 km/y every year for the first 10 years. Later the rate of decrease slows down.

Figure 6 AMAG mean (red solid line with circles) and median (blue dotted line with circles) of average vehicle annual mileage as a function of vehicle age. Quadratic regression fits are added for reference.



3.2 Distribution of annual mileage

The median is smaller than the mean during the entire vehicle life-time and indicates a right-skewed distribution of the annual distance. This is confirmed by the PDF of annual mileage for different ages as shown in Figure 7 (upper half). The increasing binning size in age guarantees a comparable sample size (around 100k) for the first 8 years. Later the bin size is further increased but the sample size is still reduced. At 18 years there are 15k samples and this is considered to be acceptable for a qualitative assessment of the PDF shape. The result of the AMAG full sample is drawn as a reference with red circular markers.

The age-specific distributions remain right-skewed during the entire period as confirmed by the lower median value compared to the mean as shown in Figure 6. For older cars the probability peak experiences a progressive increase of its value and a shift towards smaller distances.

Using the AMAG age-conditioned annual mileage PDF, the mileage distribution of the effective fleet (drawn as blue line with circles) can be extrapolated by employing an integration over the MOFIS age-PDF. This has the effect of a correction for the biased age PDF of the AMAG data compared to MOFIS. This operation results into a shift of the distribution towards lower distances because the MOFIS fleet contains a larger portion of older cars (lower mileages).

This relatively smooth and apparently well-behaved increase and shift towards smaller distances of the probability peak motivates the hypothesis of a self-affine relation between the distributions of different ages. This has been verified by introducing a normalized annual mileage $\frac{\mathrm{km}_{NORM}}{\mathrm{y}}$ as follows:

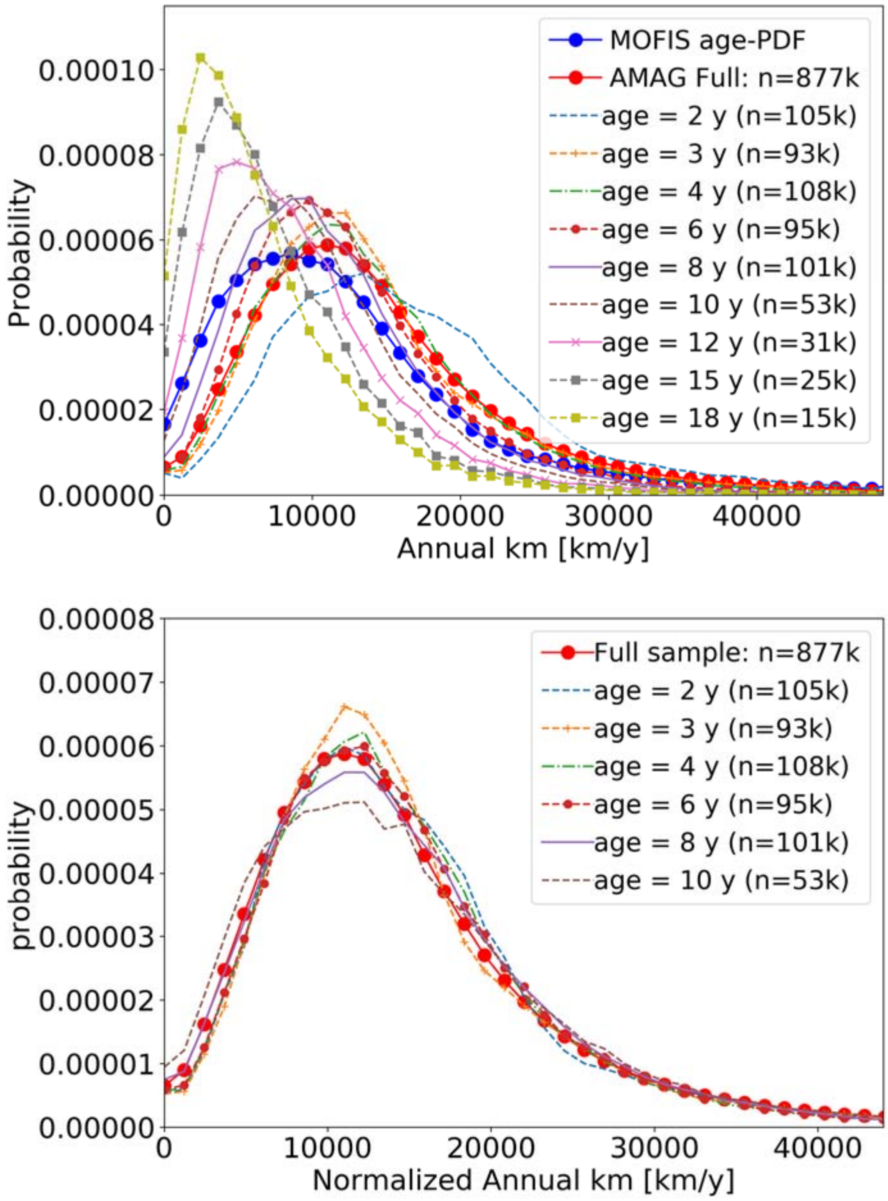
$$\frac{\mathrm{km}_{NORM}}{\mathrm{y}}(\mathrm{t}) = \frac{\mathrm{km}}{\mathrm{y}}(\mathrm{t}) \cdot \frac{\overline{\frac{\mathrm{km}}{\mathrm{y}}}(\mathrm{t})}{\overline{\frac{\mathrm{km}}{\mathrm{y}}}(\mathrm{t}_{REF})}$$

where $\frac{\mathrm{km}}{\mathrm{y}}(\mathrm{t})$, $\overline{\frac{\mathrm{km}}{\mathrm{y}}}(\mathrm{t})$ and $\overline{\frac{\mathrm{km}}{\mathrm{y}}}(\mathrm{t}_{REF})$ are the annual mileage of a single car at age t, average mileage at age t and t_{REF}. The latter is a reference age, here taken to be 4 years due to the high probability of this age in the dataset.

Results are illustrated in the lower half of Figure 7. The PDF of the normalized annual mileage during the first 10 years collapse close to the reference line, in particular the heavy tails merge practically into a single line. It is observed that for cars older than 10 years there is an overproportional increase of the close-to-zero distances that cannot be captured by an affine transformation.

Figure 7 Upper diagram: Probability density function of vehicle annual mileage (in km/y) for different vehicle age classes. AMAG full (red line with circles) contains all AMAG samples and MOFIS age-PDF (blue line with circles) considers AMAG age-conditioned distributions integrated over the MOFIS age-PDF.

Lower diagram: Probability density function of normalized vehicle annual mileage (in km/y) for different vehicle age classes and AMAG full sample.



From a modelling perspective the self-affinity between years is a desirable property of which one can take advantage to describe the data with a smaller number of parameters. The description of the annual mileage distribution becomes a function of age only. The shape of the normalized annual mileage distribution can be parametrized with a Gamma or Weibull function as common practice (e.g. [7]). A parametrization is outside the scope of this work.

3.3 Influence of population density

In the following, the dependence of annual distance on the population density is explored. The database includes the Zip code of the car's owner and the population density is computed at the spatial resolution of the Zip code for the year 2016. The population density is often used as a simple indicator for the urbanization degree. Small values (e.g. up to 200 abitants/km^2) are typical for rural areas and larger values (e.g. larger than 2000 ab/km2) for cities. Intermediate values are more likely to represent peripheral agglomeration. For example, Loder et al. [4] have subdivided the population density into tercils - denoted as high, medium and low - based on the cumulative distribution.

As a first step, the population density PDF is compared for AMAG and MOFIS vehicles and for the population as shown in Figure 8. A similar distribution is obtained for all three datasets, with a heavily right-skewed distribution. It is noted that the Swiss mean and median population density is around 200 and 700 ab/km^2, respectively. The main difference between AMAG and MOFIS is the distribution for low densities (i.e. in the country-side). In these rural regions, the MOFIS fleet exhibits a higher probability than the population. This indicates that the private motorization share is higher in rural than in the urban regions, where public transportation represents a valid alternative. The trend is reversed at population densities higher than 2000 ab/km^2, where the car density is lower than the population.

As compared to the MOFIS fleet, the AMAG dataset shows a lower probability in the rural regions and, on the other hand, a larger representation in the intermediate densities (between 500 and 2000 ab/km2). This is mainly motivated by the observation that the AMAG dataset under-represents cars older than 7 years compared to MOFIS (see Figure 2). For high population density regions (larger than 2000 ab/km^2) the two vehicle fleet datasets provide nearly the same distribution.

Figure 8 Probability density function (left axis) and cumulative probability (right axis) of population density based on AMAG fleet (red), MOFIS fleet (blue) and population (green dotted).

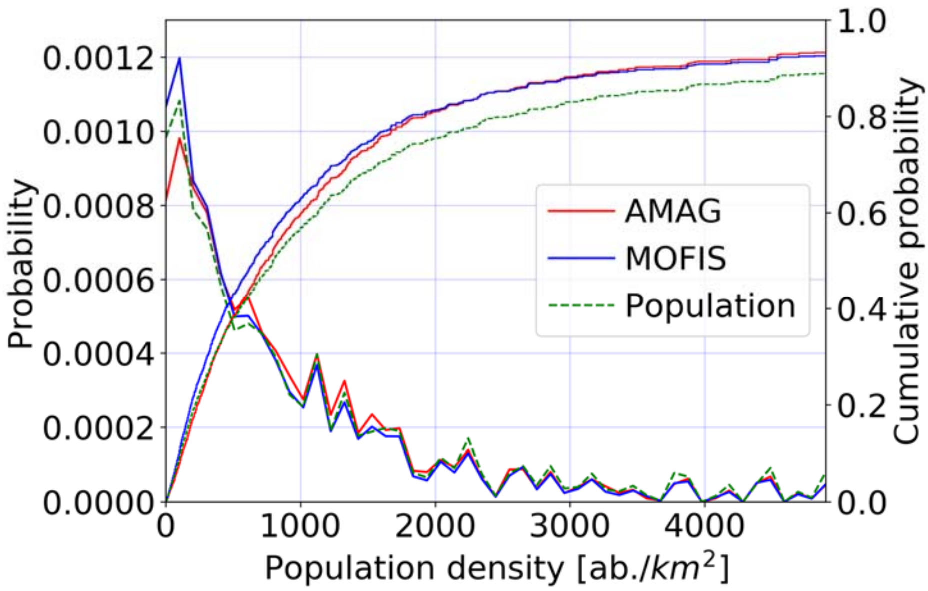
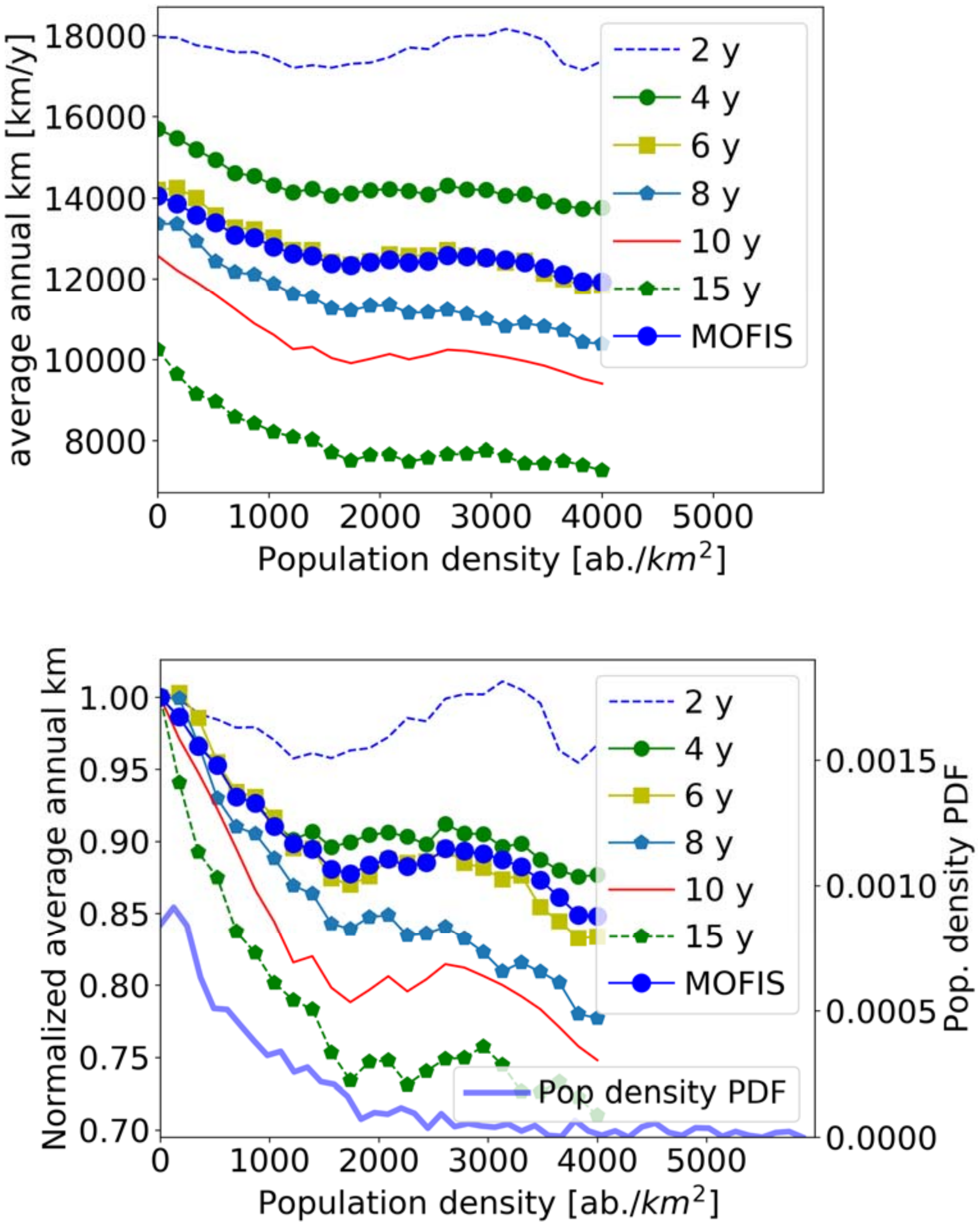


Figure 9 (upper half) shows the average annual mileage doubly conditioned on population density and vehicle age. Despite the reduced mileage for older cars as observed in Figure 6, there is a significant decrease in the average mileage when moving from low to high population densities. The relative dependence on the population density is increasingly more pronounced for older cars. This observation is further emphasized in Figure 9 (lower half) by drawing the normalized average annual mileage where every age-specific curve is normalized by its own value in the first population density bin. The population density PDF of the AMAG dataset is illustrated on the secondary axis to provide a feeling of the potential impact in the different density regions.

It is interesting to note that the usage of newer cars is similar between urban and rural regions, whereas older cars are over-proportionally more employed on the country-side. For example, 10 years old vehicles in the cities drive on average around 20% less km than in the rural region. Therefore, it appears that older cars maintain a more important relative functionality in the rural regions. Another consideration is that during the life time of the vehicle there seems to be a net flux of cars moving towards the country-side.

As a first-order correction of the biased vehicle age PDF of the AMAG dataset towards newer cars, the average annual distance conditioned on population density and age is integrated over the MOFIS age PDF to obtain the estimated overall dependence on population density. This consideration presumes that the population density distributions are not significantly different between the datasets. The result becomes, as expected, more similar to the 6 years old cars and are more representative of the effective fleet.

Figure 9 Average annual mileage (in km/y) (upper diagram) and normalized average annual mileage (no units) (lower diagram) as a function of population density for different vehicle age classes. The MOFIS line (blue line with circles) considers AMAG age-conditioned distributions integrated over the MOFIS age-PDF. The population density PDF is drawn on the right axis for reference.



3.4 Comparison with person-based data

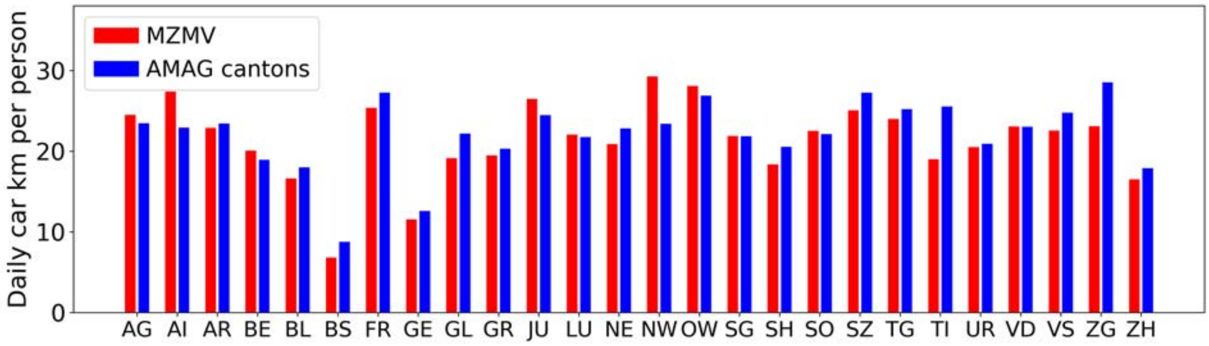
In the previous section, the results of vehicle annual mileage obtained from the vehicle-based data have been presented. In the following, the overall vehicle-km from Swiss matriculated cars in 2016 are compared with the data from Federal Office of Statistics [8], which are based on the household travel survey conducted in 2015 [2]. In total, 49,595 million vehicle-km and 4.524 million cars are reported. This results into an average annual mileage of 10,963 vehicle-km/y. From the AMAG dataset the reconstructed fleet average annual mileage is computed by integrating the AMAG mean mileage conditioned on age (Figure 6) over the MOFIS age PDF (Figure 2). This results into 11,800 km/y. It should be noted that the age binning size can influence the results, leading to a variability of up to approximately 150 km/y (~ 1%). The AMAG dataset is 7.6% higher than the one estimated from [8]. Given the qualitative nature of this comparison, this discrepancy is considered to be reasonable. For example, the person-based data [8] considers only distances driven in Switzerland, and the survey is based on daily mobility.

In the following, the average mileage in every canton is considered to verify whether the AMAG data provide a similar trend in the various regions as the person-based data [2]. Figure 10 shows the average daily distance per person in every canton from microcensus [2] (red) and AMAG (blue). Data from microsensus has been directly sampled for every canton separately. The AMAG annual mileages have been related to daily km per person by dividing by 365 days and the canton-specific private motorization degree (expressed in # vehicles by 1,000 abitants over 18 years old). The latter information is extracted from MOFIS [6] and STATPOP (Population and household statistics) [9] and aggregated at the cantonal level. The AMAG data has been normalized to match the overall Swiss average daily travel distance per person. The focus here is to look at the relative variability between cantons using two different sources.

The average daily distance driven by car per person is 20 km. The variability between cantons is considerable, where distances range from 7 km/day in Basel Stadt (BS) um to 26-28 km/day in various cantons. The AMAG data reproduces qualitatively the general trend. It is interesting to note that in BS and Geneva (GE) the low distance driven per person are captured by sampling AMAG data. The main contribution to these low distances is mainly due to the lower private motorization degree. Analysis of the AMAG data indicate that the differences between cantons in daily distance per vehicle are less than half as compared to the ones per persons. Differences in private motorization degree plays also an important role.

The overall consistent behaviour across cantons is in line with the representative distribution of the AMAG data across cantons as illustrated in Figure 4. This observation provide further support for the usefulness of this data in describing annual mileage also on a regional level.

Figure 10 Average daily km per person in every Swiss canton from person-based data (MZMV [2]) and vehicle-based data (AMAG). Data are obtained by sampling separately in every canton. AMAG mileage is normalized to match overall Swiss daily mileage per person.



4. Conclusions

A vehicle-based dataset denoted as to AMAG has been introduced and presented with the aim of an alternative description for the distribution of annual mileage in Switzerland. The representativeness of the dataset with respect to the full fleet (MOFIS data) has been assessed in terms of vehicle age distribution, relative share of vehicle across cantons and relative percentage of samples in vehicle age. The dataset is well distributed along regions but it over-represents newer cars.

The vehicle age was found to be the main influential variable, more than the population density. During the vehicle life-time the mean annual driving is clearly reducing. A quadratic regression curve was found to fit the data. A new vehicle was observed to drive on average 18,250 km and after 5 and 10 years this annual distance is decreased by 25 ad 40%, respectively.

The age-conditioned annual mileage PDF was found to exhibit a high degree of self-affinity between the different ages. This information is beneficial from the modelling descriptivism, where the vehicle age is the main degree of freedom.

The influence on the population density is assessed. Despite the higher mileage in the rural regions as compared to urban areas, there is a preferential increase of the relative difference between rural and urban areas for older vehicles. For example, 10 years old vehicles in the cities drive on average around 20% less km than in the rural region.

The Swiss total annual vehicle-km in 2016 from AMAG was found to be 7.6% higher than the one obtained from microsensus. The variability of daily distance across cantons was found qualitatively to be consistent with microsensus.

Overall, the presented data represent a valid alternative to person-based data for the description of annual mobility patterns in the context of LCA and energy research in general.

5. Acknowledgment

The authors gratefully acknowledge the financial support of the Swiss Federal Office of Energy (BFE) [SI/501311-01] and the Swiss Competence Center for Energy Research in Efficient Technologies and Systems for Mobility (SCCER mobility), funded by Innosuisse. We thank the AMAG AG, the Federal Statistical Office (BFS) and the Federal Roads Office (ASTRA) for providing the datasets.

6. References

[1] Prognos, "Die Energieperspektiven für die Schweiz bis 2050," UVEK, Basel, 2012.

[2] BFS, Mikrozensus Mobilitaet und Verkehr (MZMV), https://www.bfs.admin.ch/bfs/de/home/statistiken/ mobilitaet-verkehr/erhebungen/mzmv.html.

[3] Weber, S., & Farsi, M. (2014, September). Travel distance and fuel efficiency: An estimation of the rebound effect using micro-data in Switzerland. In 7th International Workshop on Empirical Methods in Energy Economics, Zurich.

[4] Loder, A., Tanner, R., & Axhausen, K. W. (2017). The impact of local work and residential balance on vehicle miles traveled: A new direct approach. Journal of Transport Geography, 64, 139-149.

[5] Cox, B. L., & Mutel, C. L. (2018). The environmental and cost performance of current and future motorcycles. Applied Energy, 212, 1013-1024.

[6] ASTRA, Motor vehicle information system (MOFIS), http://www.astra.admin.ch.

[7] Plötz, P., Jakobsson, N., & Sprei, F. (2017). On the distribution of individual daily driving distances. Transportation research part B: methodological, 101, 213-227.

[8] BFS, Leistungen des privaten Personenverkehrs auf der Strasse: Methodenbericht 2017.

[9] BFS, Population and household statistics (STATPOP), https://www.bfs.admin.ch/bfs/en/home/statistics/population/surveys/statpop.html.