Estimating Traffic Contribution to Particulate Matter Concentration in Urban Areas Using a Multilevel Bayesian Meta-Regression Approach

Shahram Heydari (1) (Corresponding Author)
s.heydari@soton.ac.uk

Marko Tainio (2,3,4)
marko.tainio@ymparisto.fi

James Woodcock (4)
jw745@medschl.cam.ac.uk

Audrey de Nazelle (5)
anazelle@imperial.ac.uk

1. Department of Civil, Maritime, and Environmental Engineering
University of Southampton, UK

2. Sustainable Urban Programme
Finnish Environment Institute SYKE
Helsinki, Finland

3. University of Cambridge
MRC Epidemiology unit
Cambridge, United Kingdom

4. Polish Academy of Sciences
Systems Research Institute
Warsaw, Poland

5. Centre for Environmental Policy
Imperial College London
London, UK

ABSTRACT

Quantifying traffic contribution to air pollution in urban settings is required to inform traffic management strategies and environmental policies that aim at improving air quality. Assessments and comparative analyses across multiple urban areas are challenged by the lack of datasets and methods available for global applications. In this study, we quantify the traffic contribution to particulate matter concentration in multiple cities worldwide by synthesising 155 previous studies reported in the World Health Organization (WHO)'s air pollution source apportionment data for PM_{10} and $PM_{2.5}$. We employed a Bayesian multilevel meta-regression that accounts for uncertainties and captures both within- and between-study variations (in estimation methods, study protocols, etc.) through study-specific and location-specific explanatory variables. The final sample analysed in this paper covers 169 cities worldwide. Based on our analysis, traffic contribution to air pollution (particulate matter) varies from 5% to 61% in cities worldwide, with an average of 27%. We found that variability in the traffic contribution estimates reported worldwide can be explained by the region of study, publication year, PM size fraction, and population. Specifically, traffic contribution to air pollution in cities located in Europe, North America, and Oceania is on average 36% lower relative to the rest of the world. Traffic contribution is 28% lower among studies published after 2005 than those published on or before 2005. Traffic contribution is on average 24% lower among cities with less than 500,000 inhabitants and 19% higher when estimated based on PM_{10} relative to $PM_{2.5}$. This quantitative summary overcomes challenges in the data and provides useful information for health impact modellers and decision-makers to assess impacts of traffic reduction policies.

1. Introduction

Traffic is a significant source of air pollution in cities around the world (Pant and Harrison, 2013), and is also a major contributor to adverse health globally (Bhala et al., 2014; Anenberg et al., 2019). Due to adverse effect of traffic related air pollution, policies aimed at reducing vehicular sources are the most commonly evaluated interventions in air pollution accountability studies, and include, for example, requirements on vehicle emissions standards, traffic reductions strategies, diesel vehicle bans, fuel requirements, low emission zones, congestion charging, public transport restructuring, and even-odd number plate restrictions (Burns et al., 2019). However, the contribution from transport-related air pollution emissions to health burden, and the impacts and effectiveness of local transport-focused policies will vary from city to city depending on how much of the local air pollution is due to traffic itself. Both from a global perspective of comparison across cities, and from a local perspective of evaluating potential benefits of air pollution strategies, better understanding of how much traffic contributes to air pollution concentrations is needed.

Comparisons of effectiveness of urban strategies across multiple cities, and assessments of transferability of policies in different urban settings, are hindered by the lack of international emissions or source apportioned air quality datasets focused on urban areas. Air pollution policy scenario analysis would typically use dispersion modelling techniques making use of urban emissions inventory data as inputs. Such methods, however, are costly and complex to run (Jerrett et al 2005), in addition to lacking city-scale relevance for the existing and available global frameworks. For example, the global emissions inventory Emissions Database for Global Atmospheric Research (EDGAR) focusses on national air pollution emissions, thus limiting its usability in city level analyses (Crippa et al., 2018; Lelieveld et al., 2019). While proxies could be used to disaggregate such large scale databases to more spatially resolved local emissions inventories, this wouldn't remove the challenge of running air pollution models (accounting for dispersion and chemical and physical transformation) to estimate concentrations. Finally, while global air pollution mapping methods relying on satellite observations have seen improvements in spatial resolution (~ 11km x 11km at the equator), they do not provide source apportioned information needed for policy scenario analyses (Brauer et al., 2016).

Similarly, the WHO Global Ambient Air Quality Database provides comparable annual average air pollution concentration data from 4,300 cities in 108 countries. However, alone this is of limited use for predicting or assessing impacts of source-specific contributions and policies, such as transport policies. Source apportionment studies provide a break-down of source-specific contributions to air pollution concentrations (Belis et al., 2013; Thunis et al., 2019). Such data can be directly used to assess source-specific health burden, or used in a policy scenario context with simplifying assumptions on impacts of source-specific emission changes on concentrations. However, source apportionment studies are expensive to conduct, and performing such studies at a global scale would be beyond resources of typical research project.

The WHO and a group of researchers compiled a dataset of systematically reviewed source apportionment studies conducted between 1987 and 2014 (see Karagulian et al., 2015). The variation in quality and quantity of data, year of data collection and publication, the location of studies, and methods used to derive the source apportionment make the direct comparison of the studies presented in the WHO database challenging. To begin, ignoring heterogeneity between studies, caused by the variation in study level characteristics (e.g., the location of studies), can provide misleading results. In addition, the WHO database has a hierarchical (multilevel) structure since the reported estimates are sometimes nested within studies conducted by the same authors. This multilevel structure of the data may introduce dependencies between measurements reported by the same authors; for example, due to similarities in adopted estimation methods or investigation techniques. Such dependencies should be accounted for to ensure the reliability of the estimates.

Despite challenges and difficulties explained above, by setting proper eligibility criteria to select relevant studies and by employing a sound statistical method, we aim to provide a reliable quantitative synthesis of studies that have estimated traffic contribution to particulate matter ($PM_{2.5}$ and PM_{10}) in urban settings worldwide. To this end, we carry out a meta-analysis. As discussed by Thacker (1990), generalizability and statistical power increase when using meta-analysis. In fact, it is possible to generalize the results of a meta-analysis to a broader population. Also, the accuracy and precision of the estimates improve because more data is used to make statistical inferences. Specifically, we employ a Bayesian multilevel meta-regression model with the aim of synthesising previous studies available in the WHO database while capturing heterogeneity in the data. We account for within-study dependencies and uncertainty in all parameters and estimates. For a discussion on the advantages of the Bayesian approach in evidence synthesis, see, for example, Sutton and Abrams (2001). The proposed model allows us to predict uncertainties around the mean estimates, and these uncertainties can later be used, for example, in policy analysis, considering the most pessimistic or optimistic scenarios.

2. Materials and methods

2.1. Review framework

We centred our research on the World Health Organization (WHO) database on source apportionment studies for particulate matter in the air (PM_{10} and $PM_{2.5}$) (WHO, 2015). The database reports shares of various air pollution source categories, including industry, traffic, domestic fuel burning, natural sources (sea salt and dust), and unspecified sources of human origins. Note that as indicated by WHO (2015), "Traffic is a source category that includes primary PM emissions from exhaust, organic and inorganic gaseous PM precursors from the combustion of fuels and lubricants, particles through the wear of brake linings, clutch, and tires, including depositions onto the road which are re-suspended together with crustal/mineral dust particles and road wear material.” Traffic contribution measurements were based on either $PM_{2.5}$ or PM_{10}, which are considered sensible indicators of air quality (Burnett et al., 2014; Karagulian et al., 2015). The WHO database provides information relating to the receptor model source apportionment methods (e.g., principal component analysis, positive matrix factorization, and absolute principal component analysis) used by the different studies, population at the location of study, reference year, study year, site typology (urban, rural, etc.), the geographic coordinates of measurement locations, and pollutant concentrations for $PM_{2.5}$ or PM_{10}. For a detailed discussion of the WHO source apportionment database, see Karagulian et al. (2015).

2.2. Research questions

Our research questions were (i) “what is the magnitude of traffic contribution to air quality and its associated uncertainty in different cities worldwide on the basis of studies collected in the WHO database?”; (ii) “how can we explain the variation in traffic contribution estimates reported in previous studies?”, and (iii) does PM size fraction affect traffic contribution estimates reported in previous studies?

2.3. Selection of studies, eligibility criteria, and data extraction

In this paper, the focus was on traffic contribution to air pollution; therefore, studies with missing information relating to the traffic source were removed from the data. In addition, the source apportionment estimates were obtained from different site typologies such as urban and rural sites, thereby differing largely. This makes studies from dissimilar site typologies non-comparable. We therefore identified studies that were conducted in urban areas, creating a more homogeneous data, which allows for a more reliable quantitative summary. We then extracted data relating to the following items being available in the WHO database: authors, the year of publication, the year of study, the location of study (city, country, region, and continent), population, geographic coordinates, PM_{10} and $PM_{2.5}$ concentrations, site type (remote, urban, rural, etc.), and method of estimation. In this research, we grouped countries into 13 regions mostly according to geographic proximity and the study conducted by Karagulian et al. (2015): North America, Central Europe, East Asia, East/West Africa, India, Middle East, Northwestern Europe, Oceania, South/Central America, Southeastern Asia, Southern Asia, Southwestern Europe, and Western Europe. The list of countries in each region is provided in Table A of the supplementary material.

2.4. Meta-regression

An important challenge in carrying out any quantitative summary is often related to the fact that factors such as study protocols, measurement conditions, and site characteristics may vary from one study to another. One should thus account for such variations; otherwise, the results will be only a rough representation of the reality (Joseph, 2000). In this paper, we deal with this issue using a rigorous statistical approach: a Bayesian multilevel meta-regression, as described below. We employed a multilevel model since some studies are a collection of more than one study campaigns, reporting two or multiple estimates of traffic contribution obtained from different locations. The multilevel model ensures that within-study dependencies are accounted for.

Traffic contribution (the outcome of interest) to particulate matter concentration is in the form of a rate being in the interval [0, 1]. Therefore, we first logit transformed the outcome and assumed a normal density for the logit transformed estimates. Let y_i denote the logit transformed traffic contribution reported in study campaigns ($i = 1, 2,\ldots, N$). Let $\boldsymbol{Z} = (Z_1, Z_2,\ldots, Z_s)$ be the vector of study-level explanatory variables (e.g., year of publication) with the corresponding coefficients $\boldsymbol{\beta} = (\beta_1, \beta_2,\ldots, \beta_s)$, where $s = (1,2,.., S)$ denote studies. Let η_s denote the random intercepts (here, study effects), which follows a normal density with parameters μ_η (mean) and v_η (variance). Let $\boldsymbol{X} = (X_1, X_2,\ldots, X_l)$ be the vector of observation-level explanatory variables (e.g., city population) with the corresponding coefficients $\boldsymbol{\gamma} = (\gamma_1, \gamma_2,\ldots, \gamma_l)$. We can then write

$$
\begin{gathered}
y_i \sim Normal\ (\mu_i, v) \\
\mu_i = \eta_s + \boldsymbol{\beta Z}_s + \boldsymbol{\gamma X_i} \\
\eta_s | \mu_\eta, v_\eta \sim normal(\mu_\eta, v_\eta)
\end{gathered}
\tag{1}
$$

where v is the observation-level variance, and μ_η and v_η are, respectively, the mean and the variance of the study-specific random effects. Non-informative priors normal(0,1000) and inverse-gamma(0.001,0.001) were, respectively, placed on the mean μ_η and the variances v and v_η. A sensitivity analysis was conducted for the variances using uniform priors for the standard deviations associated with the above variances and did not result in any significant variation in the results. Note that in the above model, we do not allow the variance v to vary across observations. This is because the study-specific random intercepts η_s vary from one study to another and most studies include only one reported measurement, which would not allow a varying variance v to be estimated properly.

In this paper, the varying intercepts recognise the fact that between-study differences, which cannot be captured entirely through study-specific characteristics, may exist. The variance of the study-specific random effects (v_η) indicates the between-study variation. Our model explains part of the variations of the reported traffic contribution estimates using study characteristics as indicated above. However, the remaining variation, which is unexplained, is captured through the between-study variance. The observation-level variance corresponds to within-study variations, and together with the between-study variation constitute the overall variation in the data that is random and that cannot be explained by the explanatory variables used in our study. Based on the abovementioned variances, intra-study correlation can be obtained from Equation (2).

$$
\gamma = \frac{v_\eta}{v_\eta + v} \tag{2}
$$

The value of γ varies from 0 to 1; a value close to 1 indicates a high within-study correlation. This estimated value indicates the level of similarity among measurements reported in the same study (or by the same set of authors). To draw posterior inferences, we used WinBUGS (Lunn et al., 2000) for our Markov chain Monte Carlo simulations running two chains each containing 20,000 iterations. The posterior inferences are based on the final 15,000 total iterations since the first 5,000 iterations were discarded for convergence requirements. This number of iterations was sufficient based on the Gelman-Rubin statistic (Gelman and Rubin, 1992), history plots, and Monte Carlo errors.

2.4.1. Selection of explanatory variables

Different subsets of available variables were considered to capture the variability in the traffic contribution estimates reported by previous studies. We considered geographic coordinates (latitude and longitude), population, region where a study was conducted, year of publication, method of analysis, and PM size fraction, which were provided for each study. Based on an exploratory data analysis phase, we identified the most important variables. We tested different sets of variables to identify the most relevant ones. For example, we used publication year (a common variable often used in any meta-analyses); however, it was not found to be statistically important in the model. We then created different categorical variables based on the year of publication and tested these in our model. A categorical variable based on 2005, created indicating whether the publication date was after 2005, was found to be relevant in explaining traffic contribution estimates reported in the previous research; and therefore, this was included in the final model. Similarly, population was dichotomised: less or more than 500,000 inhabitants. A categorical variable was created indicating whether traffic contribution estimates were obtained based on source contribution analyses of PM_{10} or $PM_{2.5}$. Also, we created a categorical variable to examine whether the source apportionment method of each study affected reported estimates. Lastly, we created two groups of regions: predominantly developed countries (Europe, North America, and Oceania) and predominantly low/middle income countries (Africa, Latin America, and Asia). The latter variable was also selected based on testing different regions in the model.

3. Results

3.1. Study selection

After identifying relevant studies according to the criteria explained in Section 2.3, our final dataset contains 293 reported estimates obtained from 155 studies (Table B of the supplementary material). These studies were conducted in 6 continents, covering 169 cities located in 48 countries, from 1987 to 2014. Note that some studies report two or more estimates depending on the number of locations considered in their study design.

3.2. Study characteristics

Sample characteristics are summarised in Table 1. Close to three quarters of the observations obtained for the analysis stemmed from studies in Europe, North America and Oceania. Africa was highly under-represented, contributing only 7 of the 293 data points (Table 1) due to the fact that only few studies were available in the extant literature. Most of the data points stemmed from studies conducted after 2005 (75%), in cities with more than 500,000 inhabitants (64%), and based on $PM_{2.5}$ (62%) rather than PM_{10} measurements (Table 1). In our final data, the reported traffic contribution to particulate matter concentrations varied across cities from around 2% to 77.1%, with a mean (and standard deviation) of 26.62% (15.79%) at a global level.

3.3. Meta-regression: traffic contribution and explanatory variables

The results of our meta-regression are reported in Table 2 from which the magnitudes of association between a series of explanatory variables and traffic contribution estimates can be inferred. We found that the region of study, publication year, PM size fraction (whether PM_{10} or $PM_{2.5}$), and population can explain traffic contribution estimates reported in the previous literature collected in the WHO database. However, the variable representing the source apportionment method was not found to have an effect in our study. Note that in Bayesian statistics, instead of point estimates obtained in classical statistics, posterior densities of parameters of interest are obtained. The posterior mean densities and their 95% Bayesian intervals are reported in Table 2. Since the logit transformed traffic contribution was modelled, as explained in Section 2.4, odds ratios (the exponentiated values of the regression coefficients) needed to be calculated in order to interpret the regression coefficients. The average effect of explanatory variables (over the entire sample) on the outcome of interest can be inferred from odds ratios reported in Table 2 (see Section 4.1 for further details).

The mean (μ_η) and the variance (v_η) of the study-specific random effects and their respective uncertainties (Bayesian intervals) are reported in Table 2, as well as the estimated observation-level variance (v). Based on the abovementioned variances, an intra-study correlation of 0.67 is obtained from Equation (2) (see Section 2.4), suggesting that 67% of the total variance in the outcome of interest is accounted for by the apparent grouping (observations nested within studies) in the data. Note that the need for standardization has been highlighted and efforts made to develop common source apportionment protocols to reduce between-study variabilities due to the use of different methods (Solomon et al., 2014; Xin e al., 2015; Zhang et al., 2017). Our results imply that between-study differences in the reported traffic contribution measurements are relatively large. This indicates that study protocols, measurement conditions, and site characteristics varied between most studies, highlighting the importance of accounting for such variations in order to draw reliable statistical inferences.

3.4. Estimated traffic contribution and its associated uncertainty at city and region levels

Our approach allows us to estimate expected traffic contribution for different cities and regions within our model estimation procedure. Doing so, we can obtain uncertainties for city and region level estimates. Specifically, we calculated expected traffic contribution for 169 cities in the data and for 13 regions (North America, Central Europe, East Asia, East/West Africa, India, Middle East, Northwestern Europe, Oceania, South/Central America, Southeastern Asia, Southern Asia, Southwestern Europe, and Western Europe). Figure 1 displays the posterior expected traffic contribution (and associated uncertainty in terms of 95% Bayesian intervals) for these cities while being grouped into different regions shown in different colours. Based on our estimates, Pullman (USA) with a mean (standard deviation) of 5% (3%) has the lowest traffic contribution to particulate matter, while Colombo (Sri Lanka) with a mean (standard deviation) of 61% (14.7%) has the highest traffic contribution to particulate matter. Expected traffic contribution estimates and their respective uncertainties are displayed in Figure 2 for 13 regions.

4. Discussion

4.1. Explaining sources of heterogeneity in the WHO database

For convenience, we interpret the effects of the explanatory variables (sources of heterogeneity) on traffic contribution estimates reported in the previous studies in terms of percentage change. Specifically, traffic contribution to air pollution in European, North American, and Oceanian cities is on average 36% (i.e., 100*(0.642-1)) lower relative to the rest of the world. It is 28% lower among studies published after 2005 than those published on or before 2005. Traffic contribution is on average 24% lower among cities with less than 500,000 inhabitants and is 19% higher when estimated based on PM_{10} compared to $PM_{2.5}$.

A lower traffic contribution to air quality in Europe, North America, and Oceania could be partly explained by more stringent environmental and exhaust regulations/policies (e.g., EURO-emission standards) and a newer vehicle fleet in these regions, when compared to other emissions sources. Also, more effective traffic operations and traffic management strategies such as low emission zones and road pricing (for a discussion in this regard see Bigazzi and Rouleau (2017)), which are more common in Europe, North America and Oceania compared to the rest of the world, provide further support for this finding. A higher penetration rate of low emission vehicles in predominantly high-income regions can be another explanation for a lower traffic contribution to air quality in these regions.

Similar factors discussed above may explain a lower traffic contribution observed at a global level after 2005 as well as an outgrowth of increased awareness among transportation agencies, local authorities, and general public with respect to the impact of traffic on air pollution and its harmful health effects. Lower traffic congestion could explain our finding related to relatively low-populated cities (i.e., those with less than 500,000 inhabitants). For example, a more stabilized operating speed leads to less acceleration and breaking resulting in fuel savings; and consequently, a reduction in traffic related air pollution. Also, in general, people travel longer distances in larger cities compared to small cities. Finally, one possible explanation for a higher traffic contribution estimates based on PM_{10} compared to $PM_{2.5}$ could be that part of re-suspended dust due to natural sources is attributed to traffic (see Karagulian et al., 2015 for a similar discussion). Finally, high levels of competing sources (e.g., crop burning, cooking and heating, coal-fired power stations, and desert dust) in some developing countries, such as those in Middle East, diminish the relative contribution of traffic in these countries (Kumar et al., 2015, Karagulian et al., 2015; Gautam et al., 2019).

4.1.1. Investigating variations in absolute concentrations attributable to traffic

With respect to the above discussions, caution should be taken in interpreting relative (percentage) contributions since changes in the competing sources of particulate matter may have had an impact on the relative contributions from traffic. This said, a more conclusive inference would be possible by taking into account absolute concentrations attributable to traffic (in $\mu g/m^3$). Based on the WHO database, we therefore conducted further investigation in this regard, using statistical software Stata (StataCorp, 2015); the results are described below.

Average absolute $PM_{2.5}$ and PM_{10} concentrations due to traffic were lower in Europe, North America, and Oceania compared to the rest of the world (respectively, 5.43 $\mu g/m^3$ and 12.61 $\mu g/m^3$ vs. 17.689$\mu g/m^3$ and 38.63 $\mu g/m^3$). Similarly, $PM_{2.5}$ and PM_{10} concentrations were lower in cities with less than 500,000 inhabitants (respectively, 5.49 $\mu g/m^3$ and 12.35 $\mu g/m^3$ vs. 10.75 $\mu g/m^3$ and 24.54 $\mu g/m^3$). However, while percentage traffic contribution estimates reported in studies published after 2005 are smaller, in general, than those published on or before 2005, average absolute traffic-related concentrations increase globally: $PM_{2.5}$ increases from 5.82 $\mu g/m^3$ to 10.29 $\mu g/m^3$ and PM_{10} increases from 14.13 $\mu g/m^3$ to 20.41 $\mu g/m^3$. This could be partly explained by rapid motorisation of transport world-wide, particularly in low- and middle-income countries (Kitamura and Mohamad, 2009; Chalya et al., 2012).

A more detailed investigation indicates that average $PM_{2.5}$ concentrations slightly decreases from 5.59 $\mu g/m^3$ in the before 2005 period to 5.34 $\mu g/m^3$ in the after 2005 period in prevalently high-income countries. In developing countries, however, we observed almost two folds increase in $PM_{2.5}$, from 9.07 $\mu g/m^3$ to 18.18 $\mu g/m^3$ during the same period. For PM_{10}, in Europe, North America, and Oceania, average traffic-related PM_{10} concentrations remain largely stable (slight decrease from 12.58 $\mu g/m^3$ to 12.63 $\mu g/m^3$ between the two periods), however, in contrast to $PM_{2.5}$, a relatively significant decrease from 52.99 $\mu g/m^3$ to 38.07 $\mu g/m^3$ in the rest of the world during this period is observed. The latter would partly explain why percent traffic contributions to PM show a decreasing trend in estimates provided by studies published after 2005.

4.2. City- and region-level estimates

With respect to traffic contribution estimates at city level, our results imply that, in general, there is a relatively large variability even within each region (see Figure 1). Less variability in traffic contribution estimates can be seen in Northwestern Europe, where all mean estimates of traffic contribution to particulate matter are smaller than the global average value. However, in a few cities the upper bound of the uncertainties exceeds the global average considerably. For cities situated in Middle East, North America, and Oceania regions, most mean estimates of traffic contribution are less than the global average while in many occasions the upper bounds are bigger than the global average. Southwestern Europe have mixed results with most cities exceeding the global average when considering uncertainties. The mean traffic contribution estimates in Western and Central Europe are mostly lower than the global average; however, the upper bounds of the 95% Bayesian interval is above the global average. Traffic contribution in most cities in Latin America, Africa, Eastern, Southern, and Southeastern Asia, and India (prevalently low- and middle-income continents) is larger than the global average.

The East/Western Africa regions is highly underrepresented by only three cities; in fact, there are large uncertainties around East/West Africa estimates, which has the largest traffic contribution to particulate matter, followed by India and then Southeastern Asia. Traffic contribution in Northwestern Europe is the lowest, followed by Central Europe, North America and Middles East. In terms of uncertainty estimates, regions with multiple records such as Southwestern Europe and North American regions have the narrowest 95% Bayesian intervals. In order to better understand the extent of traffic contribution to particulate matter, more monitoring stations and source apportionment studies are thus needed in cities or regions where the levels of uncertainty are relatively high. In this regard, our study provides valuable insights for 169 cities worldwide, allowing researchers and local authorities to target cities where further research is most warranted to reduce the amount of uncertainty in traffic contribution estimates. Note also that a low traffic contribution estimate in a city does not mean that traffic is of a minor concern since pollutant concentrations produced by traffic in that city might be relatively high. This an important concern especially when it comes to investigating the effect of traffic-related air pollution on human health.

4.3. Comparison to previous research

Our work built up on the review work reported by Karagulian et al. (2015). Karagulian et al.'s population-weighted estimates are obtained for $PM_{2.5}$ and PM_{10} separately. However, our estimations are obtained while combining data from both types of PM by using a categorical explanatory variable in our meta-regression. This is because one aim of the study was to understand whether reported estimates of traffic contribution by previous studies are affected by PM size fraction considered in those studies. That is answering our third research question (see Section 2.2). The only way, not only to answer this question but also to quantify the magnitude of difference was to include both types of PM in the model, considering the entire data. Also, when using more data in the analysis almost always more reliable statistical inferences are obtained. We only focus on traffic contribution while Karagulian et al. (2015) considers the five sources mentioned in Section 2.1. Finally, we considered only urban sites while Karagulian et al.'s study has considered urban areas and a limited number industrial sites.

Therefore, caution must be taken in comparing the results of the two studies. Also, while Karagulian et al. (2015) do not provide estimation uncertainties, we estimated uncertainties. Karagulian et al. (2015) estimated that 25% of the measured $PM_{2.5}$ and PM_{10} was due to traffic. In our study, the estimated average global traffic contribution was 25.7% (with a 95% Bayesian interval of [24% - 27%]). With respect to region-level estimates, for example, Karagulian et al. (2015) estimated that traffic contribution to particulate matter was 21% and 12% based on $PM_{2.5}$ and PM_{10}, respectively, in Northwestern Europe. We found that the mean value of traffic contribution to particulate matter in Northwestern Europe was 15.8% with a 95% Bayesian interval ranging from 12.2% to 20.2%. Recall that in estimating these values we used both $PM_{2.5}$ and PM_{10} at the same time. This is while the average value for this region, as reported by previous research (Karagulian et al., 2015), was 14.4% (see Figure 2).

Note that uncertainties provided in our research are based on Bayesian (credible) intervals that have an intuitive interpretation in contrast to classical confidence intervals (Gelman et al., 2004). A 95% Bayesian interval indicates that traffic contribution to particulate matter is in that interval with 95% probability. However, this cannot be inferred from confidence intervals; the probability of an estimate being in a confidence interval is either zero or one (Sim and Reid, 1999). An important implication of the region-level uncertainty estimates obtained in our study is that they allow for estimating a range of traffic contribution to particulate matter in cities where monitoring stations do not exist or where no previous source apportionment study has been conducted. The certainty regarding this probable range dependents on the number of studies available in a specific region. For example, as displayed in Figure 2, any conclusions regarding the range of traffic contribution is more certain for cities located in regions such as Europe and North America compared to those in India or Africa.

Another practical implication of the uncertainty estimates at both city- and region-levels is that decision makers would be able to consider the most optimistic and pessimistic scenarios, respectively, based on the lower and upper bounds of the uncertainty intervals estimated here. This could be useful, for example, in prioritising policies that aim at improving air quality in urban areas. That is, to determine whether traffic contribution is significant enough compared to other sources to warrant action targeting traffic management strategies. The uncertainty estimates also enable local authorities to estimate the range of benefits, in terms of air quality, expected from implemented or planned traffic management strategies. Lastly, the estimated uncertainties would allow researchers to investigate the reliability of health impact assessment – with respect to the adverse effect of traffic-related air pollution on human health – conducted in the past or to be conducted in the future. A similar discussion in this regard is provided by Burnett et al. (2018).

4.4. Limitations

In our meta-analysis, we considered studies available in the WHO database from 1987 to 2014 as the aim was to summarise this database. However, several other relevant studies that estimated traffic contribution to air quality after 2014 do exist. In particular, in areas with very few estimates, especially Africa, recent literature could possibly have provided more results to analyse. While we recognise this important limitation, we stress that our meta-analytical approach was to make use of an existing homogenised database. As our analysis is based on a large sample of studies (155 studies containing 293 records), including results for most areas of the world, the pool and coverage are large enough to provide valuable insights, especially thanks to our proposed meta-analytical framework. This paper suggests a valuable and novel method for synthesizing existing knowledge and deriving uncertainty estimates of traffic contributions to particulate matter. Once further studies are made available and homogenised in a common dataset, these can be readily included in the analysis.

There are further limitations related to the data available within the WHO database. We could have improved our estimations of uncertainties, for example, had the variances for traffic contribution estimates been available in the WHO database. Finally, most of the source apportionment methods used in the original studies included in the database are only able to account for contributions from the primary particles emitted from traffic (Karagulian et al. 2015). Secondary particles, which are formed in the air from chemical reactions between gasses, can be formed from local sources and are also generated from traffic or other sources beyond the local city boundaries impacting regional air quality. While source apportionment methods can be used to account for non-linear sources of particulate matter (Clappier e al., 2017), for traffic sources specifically, this study could only estimate source contribution from primary emissions, thus we are likely to under-estimate traffic contributions to city-wide particulate matter. Although contributions from traffic to secondary aerosols found in cities will typically be relatively small (Querol et al., 2004), it will vary by region, thus how much is under-estimated will vary by region as well.

5. Summary

To obtain reliable estimates of traffic contribution and their associated uncertainties at a global level, we conducted a quantitative synthesis of previous studies collected in the WHO database (WHO, 2015). Due to significant differences between studies carried out in sites of different types (urban, rural, remote, industrial, etc.) and in line with our research question, we only considered studies conducted in urban areas. This was necessary to make sure studies were similar enough in order to be included in our meta-analysis. Consequently, 155 studies, conducted in 169 cities worldwide, met our inclusion criteria. We employed a Bayesian random effects multilevel meta-regression model that allowed us to account for both between- and within-study variations and uncertainties while accommodating the hierarchical structure of the data. Note that accounting for the hierarchical structure of the WHO database is needed as measurements reported by the same study are often more alike than those reported by different studies. To our knowledge, this is the first study attempting to draw conclusive inferences regarding the contribution of traffic to $PM_{2.5}$ and PM_{10} using such an approach.

According to our results, the expected mean (standard deviation) of traffic contribution to air quality varies from 5% (3%) to 61% (14.7%) globally. We found that the region of study, publication year, PM size fraction (whether PM_{10} or $PM_{2.5}$), and population can explain variation in traffic contribution estimates reported in previous literature, collected in the WHO database. Specifically, reported traffic contribution estimates in Europe, North America, and Oceania are lower than other continents. These estimates are lower among studies carried out after 2005 than those conducted on or before 2005. A lower traffic contribution estimates reported in Europe, North America, Oceania, and post-2005 could be related to more stringent policies and newer vehicle fleets. Traffic contribution is lower among cities with less than 500,000 inhabitants, perhaps because of lower traffic congestion, and is higher when estimated according to PM_{10} since re-suspended dust is partly attributed to traffic. Using a large number of previous studies conducted worldwide, our study identifies some of the most important sources of heterogeneity in the estimated traffic contribution by a relatively large sample of previous studies.

The strength of our approach is that that our estimations of traffic contribution are statistically more reliable compared to observed values (measurements reported in previous research) since we are employing a meta-regression. This allows us to pool strength over multiple studies while adjusting for between- and within-study differences based on a set of explanatory variables. Also, we are employing a Bayesian hierarchical model that improves estimations statistically by borrowing strength from other similar estimates when the data are sparse (Richardson and Best, 2003). A valuable insight provided by our research relates to the estimation of uncertainties associated with traffic contribution estimates at both city and region levels. This is in particular appealing when it comes to defining the range of traffic contribution to air quality in cities where source apportionment studies are missing. Our quantitative summary overcomes challenges in data gaps and provides useful information for health impact modellers and policy-makers to assess impacts of traffic reduction policies.

Acknowledgments

This study was supported by the project “Towards an Integrated Global Transport and Health Assessment Tool (TIGTHAT)”, funded by Medical Research Council (MRC) Global Challenges Research Fund, UK (number: RG87632-SJ). The work of the first author was partly funded by the Natural Sciences and Engineering Research Council of Canada.

References

1. Anenberg, S., Miller, J., Henze, D., Minjares, R., 2019. A global snapshot of the air pollution-related health impacts of transportation sector emissions in 2010 and 2015. Retrieved from the International Council on Clean Transportation, https://www.theicct.org/publications/health-impacts-transport-emissions-2010-2015

2. Belis, C., Karagulian, F., Larsen, B., Hopke, P., 2013. Critical review and meta-analysis of ambient particulate matter source apportionment using receptor models in Europe. Atmospheric Environment 69, 94-108.

3. Bhalla, K., Shotten, M., Cohen, A., Brauer, M., Shahraz, S., Burnett, R., Leach-Kemon, K., Freedman, G., Murray, C., 2014. Transport for Health: The Global Burden of Disease from Motorized Road Transport. World Bank Group: Washington, DC.

4. Bigazzi, A., Rouleau, M., 2017. Can traffic management strategies improve urban air quality? A review of the evidence. Journal of Transport & Health 7, Part B, 111-124.

5. Brauer, M., Freedman, G., Frostad, J., Van Donkelaar, A., Martin, R., Dentener, F., Dingenen, R., Estep, K., Amini, H., Apte, J., Balakrishnan, K., Barregard, L., Broday, D., Feigin, V., Ghosh, S., Hopke, P., Knibbs, L., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J., Shaddick, G., Anderson, H., Vos, T., Forouzanfar, M., Burnett, R., Cohen, A., 2016. Ambient Air Pollution Exposure Estimation for the Global Burden of Disease 2013. Environmental Science & Technology 50, 79-88.

6. Burnett, R., Chen, H., Szyszkowicz, M., et al., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proceedings of the National Academy of Science of the USA 115, 9592-7.

7. Burnett, R., Pope, C., Ezzati, M., et al., 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environmental Health Perspect 122, 397–403.

8. Burns, J., Boogaard, H., Turley, R., Pfadenhauer, L., van Erp, A., Rohwer, A., Rehfuess, E., 2014. Interventions to reduce ambient particulate matter air pollution and their effect on health (Protocol). Cochrane Database of Systematic Review 1. doi:10.1002/14651858.CD010919

9. Chalya, P., Mabula, J., Dass, R., Mbelenge, N., Ngayomela, I., Chandika, A., 2012: Injury characteristics and outcome of road traffic crash victims at Bugando Medical Centre in Northwestern Tanzania. Journal of Trauma Management Outcomes 6:12012.

10. Clappier, A., Belis, C., Pernigotti, D., Thunis, P., 2017. Source apportionment and sensitivity analysis: two methodologies with two different purposes. Geoscientific Model Development 10, 4245–4256.

11. Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., van Aardenne, J.A., Monni, S., Doering, U., Olivier, J.G.J., Pagliari, V., et al., 2018. Gridded Emissions of Air Pollutants for the period 1970–2012 within EDGAR v4.3.2. Earth Syst. Sci. Data Discuss.

12. Gautam, S., Yadav, A., Tsai, C., Kumar, P., 2016. A review on recent progress in observations, sources, classification and regulations of PM2.5 in Asian environments. Environmental Science and Pollution Research 23 (21), 21165–21175.

13. Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis. Chapman & Hall/CRC.

14. Gelman, A., Rubin, D., 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7 (4), 457–472.

15. Jerrett, M., Bennett, R., Ma, C., Pope, C., Krewksi, D., Newbold, K., Thurston, G., Shi, Y., Finklestein, N., Calle, E., and Thun, M., 2005. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 16, 727-736.

16. Joseph, 2000. A brief introduction to meta-analysis. Division of Clinical Epidemiology, Department of Medicine, Montreal General Hospital, 1650 Cedar Avenue, Montreal, Quebec, H3G 1A4, Canada, and Department of Epidemiology and Biostatistics, McGill University, 1020 Pine Avenue West, Montreal, Quebec, H3A 1A2, Canada. http://www.medicine.mcgill.ca/epidemiology/Joseph/courses/EPIB-669/joseph2000.pdf

17. Karagulian, F., Belis, C., Dora, C., Prüss-Ustün, A., Bonjour, S., Adair-Rohani, H., Amann, M., 2015. Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level. Atmospheric Environment 120, 475–483.

18. Kitamura, R., Mohamad, J., 2009. Rapid motorization in Asian cities: urban transport infrastructure, spatial development and travel behaviour. Transportation 36 (3), 269e274.

19. Kumar, P., Khare, M., Harrison, R.., Bloss, W.J., Lewis, A.C., Coe, H., Morawska, L., 2015. New Directions: air pollution challenges for developing megacities like Delhi. Atmospheric Environment 122, 657–661.

20. Lelieveld, J., Evans, J., Fnais, M., Giannadaki, D., Pozzer, A. 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525: 367-371.

21. Lunn, D., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10 (4), 325–337.

22. Pant P., Harrison, R., 2013. Estimation of the contribution of road traffic emissions to particulate matter concentrations from field measurements: a review. Atmospheric Environment 77, 78–97.

23. Querol, X., Alastuey, A., Ruiz, C., Artiñano, B., Hansson, H., Harrison, R.,Buringh, E., ten Brink, H., Lutz, M., Bruckmann, P., Straehl, P., Schneider, J., 2004. Speciation and origin of PM10 and PM2.5 in selected European cities. Atmospheric Environment 38 (38), 6547-6555.

24. Richardson, S., Best, N., 2003. Bayesian hierarchical models in ecological studies of health environment effects. Environmetrics 14, 129–147.

25. Sim, J., Reid, N., 1999. Statistical inferences by confidence intervals: issues of interpretation and utilization. Physical Therapy 79 (2), 186-195.

26. StataCorp (2015). Statistical Software: Release 14. College Station, TX: StataCorp LP.

27. Solomon, P., Crumpler, D., Flanagan, J., Jayanty, R., Rickman, E., McDade, C., 2014. US National PM2.5 Chemical Speciation Monitoring Networks-CSN and IMPROVE: description of networks. Journal of the Air and Waste Management Association 64, 1410–1438.

28. Sutton, A., Abrams, K., 2001. Bayesian methods in meta-analysis and evidence synthesis. Statistical Methods in Medical Research 10 (4), 277-303.

29. Thacker, S., 1988. Meta-analysis. A quantitative approach to research integration. JAMA 259 (11), 1685-1689. doi:10.1001/jama.1988.03720110047033

30. Thunis, P, Clappier, A, Tarrason, L, Cuvelier, C, Monteiro, A, Pisoni, E, Wesseling, J, Belis, CA, Pirovano, G, Janssen, S, Guerreiro, C, and Peduzzi, E. 2019. 'Source apportionment to support air quality planning: Strengths and weaknesses of existing approaches', Environment International 130, 104825.

31. WHO, 2015. World Health Organisation's Source Apportionment Database for PM_{10} and $PM_{2.5}$ Updated to August 2014. World Health Organization, Geneva, Switzerland. http://www.who.int/quantifying_ehimpacts/global/source_apport/.

32. Xin, J., Wang, Y., Pan, Y. et al. The campaign on atmospheric aerosol research network of China: CARE-China. Bull American Meteorological Society 96: 1137– 55.

33. Zhang, Y., Cai, J., Wang, S., He, K., and Zheng, M.: Review of receptor-based source apportionment research of fine particulate matter and its challenges in China. Science of the Total Environment 586, 917– 929.

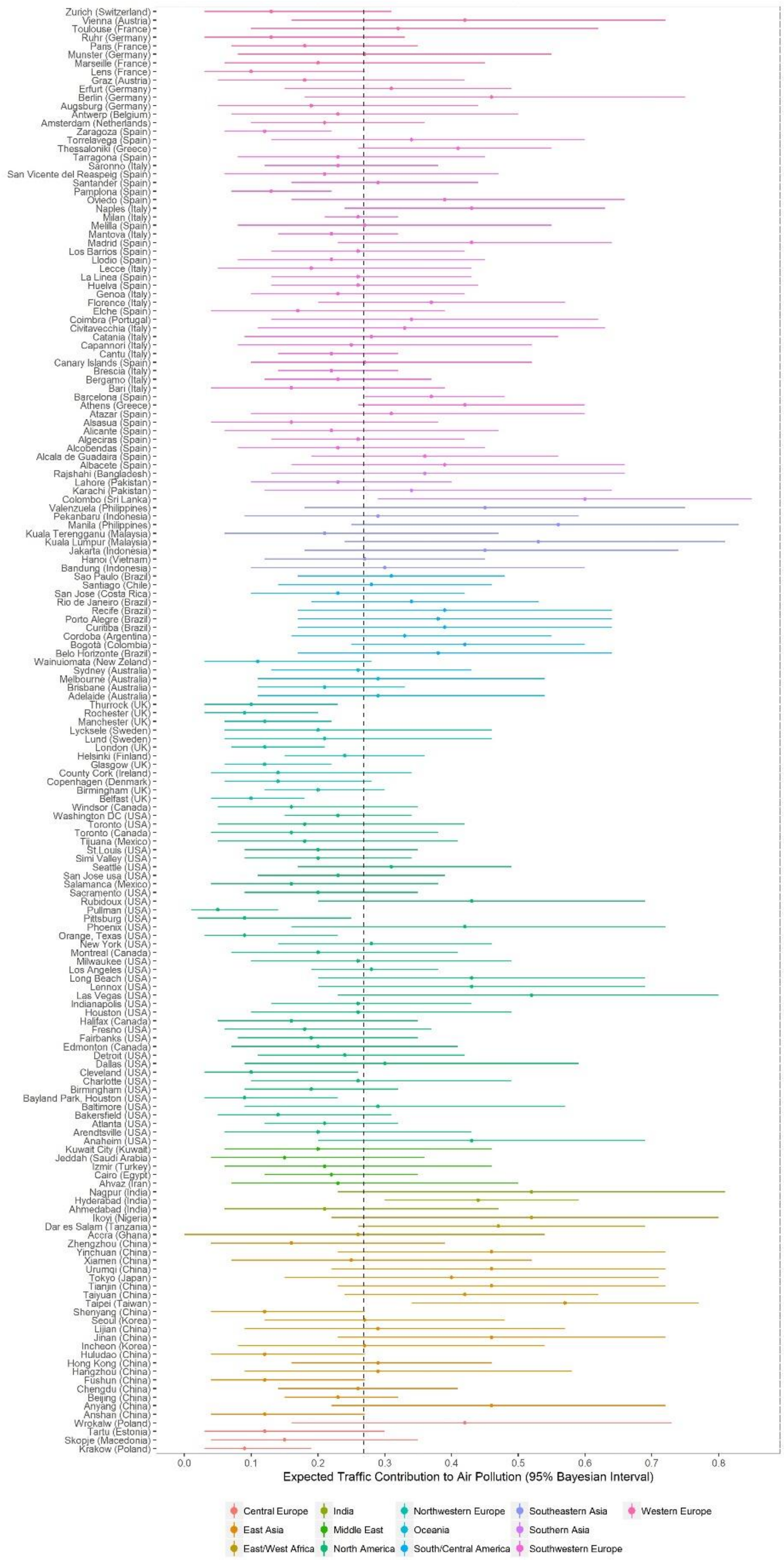

Figure 1. Expected traffic contribution to air quality (and associated uncertainty) for 169 cities worldwide. Dashed line indicates the overall (global) mean value of traffic contribution to particulate matter.

Table 1. Summary statistics of the compiled data

	Frequency	Percent
Explanatory variables		
Study was conducted in Europe, North America or Oceania		
No (0)	82	28.0
Yes (1)	211	72.0
Study published after 2005		
No (0)	73	24.9
Yes (1)	220	75.1
City population is less than 500,000 inhabitants		
No (0)	186	63.5
Yes (1)	107	36.5
Traffic contribution estimates were based on PM_{10}		
No (0)	182	62.1
Yes (1)	111	37.9
Number of records in each region		
Africa	4	1.4
Central and Eastern Europe	5	1.7
East Asia	34	11.6
India	6	2.0
Middle East	7	2.4
North America	62	21.2
Northwestern Europe	24	8.2
Oceania	8	2.7
South and Central America	17	5.8
Southeastern Asia	9	3.1
Southern Asia	5	1.7
Southwestern Europe	95	32.4
Western Europe	17	5.8

Note: Total number of observations is 293.

Table 2. Meta-regression estimation results explaining traffic contribution to particulate matter

Variables	Mean	Std. Dev.	95% Bayesian interval		Odds ratios	Std. Dev.	95% Bayesian interval	
Europe, North America or Oceania	-0.443	0.162	-0.768	-0.129	0.642	0.105	0.464	0.878
Published after 2005	-0.326	0.179	-0.670	0.029	0.722	0.133	0.511	1.030
Population < 500,000 people	-0.269	0.097	-0.459	-0.079	0.764	0.074	0.631	0.924
PM_{10}	0.171	0.091	-.008	0.350	1.191	0.108	0.991	1.419
Study-specific random effect mean	-0.577	0.212	-0.986	-0.154	-	-	-	-
Study-specific random effect variance	0.556	0.091	0.393	0.748	-	-	-	-
Observation-level variance	0.265	0.033	0.208	0.336	-	-	-	-

Note: For description of the explanatory variables, see Table 1.

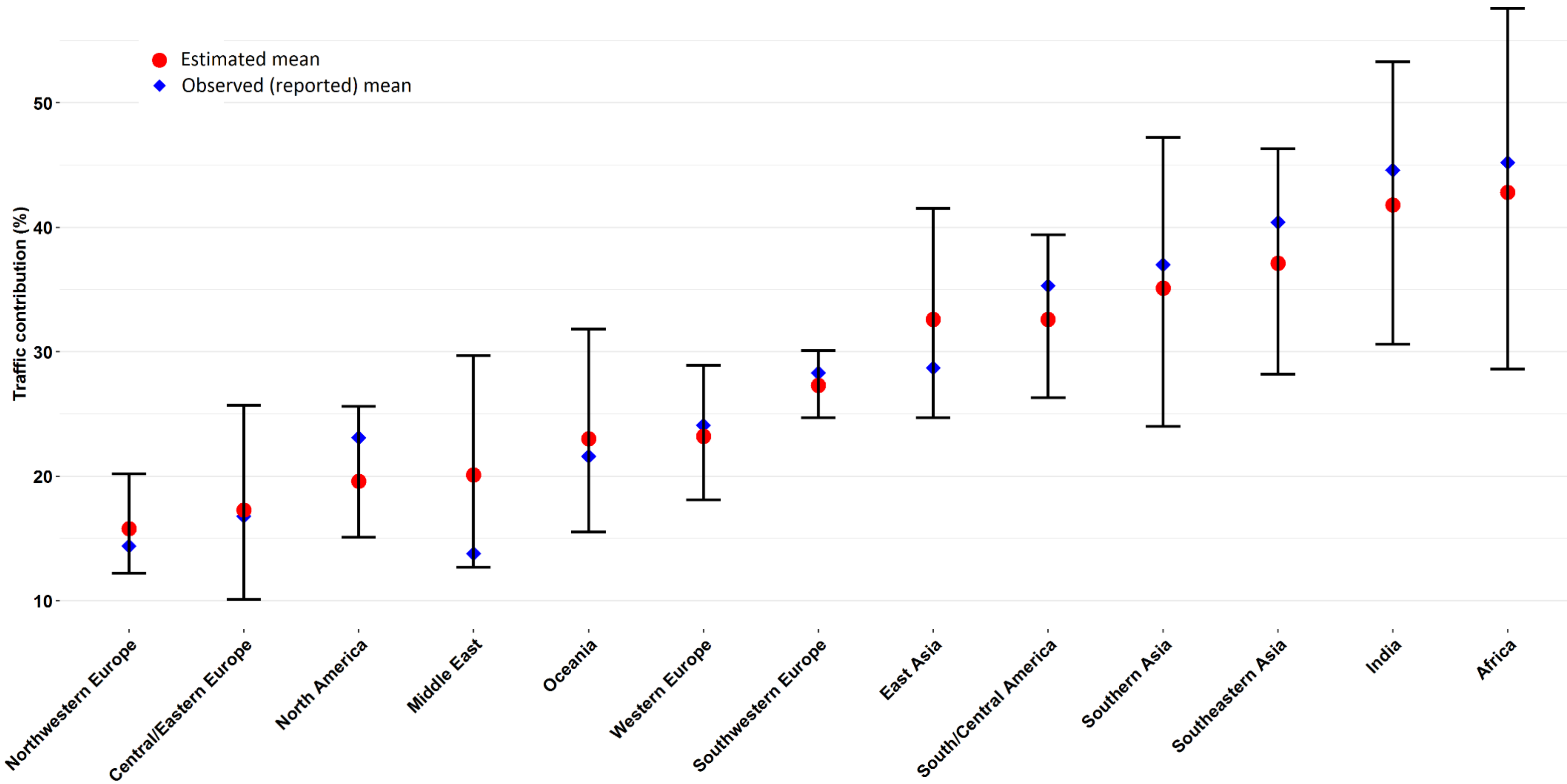

Figure 2. Estimated average traffic contribution (and its 95% Bayesian interval) for different regions.