Including congestion effects in urban road traffic CO₂ emissions modelling: Do Local Government Authorities have the right options?

Matt Grote *, Ian Williams, John Preston, Simon Kemp

Faculty of Engineering and the Environment, University of Southampton, Highfield Campus, Southampton SO17 1BJ, United Kingdom

Article info

Article history:

Keywords:
Road traffic model
Emissions model
Greenhouse gas
Air quality
Congestion
Local government

Abstract

Tailpipe emissions from vehicles on urban road networks have damaging impacts, with the problem exacerbated by the common occurrence of congestion. This article focuses on carbon dioxide because it is the largest constituent of road traffic greenhouse gas emissions. Local Government Authorities (LGAs) are typically responsible for facilitating mitigation of these emissions, and critical to this task is the ability to assess the impact of transport interventions on road traffic emissions for a whole network.

This article presents a contemporary review of literature concerning road traffic data and its use by LGAs in emissions models (EMs). Emphasis on the practicalities of using data readily available to LGAs to estimate network level emissions and inform effective policy is a relatively new research area, and this article summarises achievements so far. Results of the literature review indicate that readily available data are aggregated at traffic level rather than disaggregated at individual vehicle level. Hence, a hypothesis is put forward that optimal EM complexity is one using traffic variables as inputs, allowing LGAs to capture the influence of congestion whilst avoiding the complexity of detailed EMs that estimate emissions at vehicle level.

Existing methodologies for estimating network emissions based on traffic variables typically have limitations. Conclusions are that LGAs do not necessarily have the right options, and that more research in this domain is required, both to quantify accuracy and to further develop EMs that explicitly include congestion, whilst remaining within LGA resource constraints.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction

More than half the world’s population now live in urban areas (UNFPA, 2007), where concentrated travel requirements can often overwhelm transport systems during peak periods, leading to congestion. For urban road networks, carrying increasingly large numbers of vehicles results in increasingly large amounts of tailpipe emissions, including both greenhouse gases (GHGs) and pollutants detrimental to air quality (AQ), with the problem exacerbated by the stop-and-go nature of congestion increasing emissions yet further. This problem is particularly relevant in developing countries because towns and cities in the developing world are projected to constitute 80% of urban humanity by 2030 (UNFPA, 2007). Consequently,
urban areas produce a disproportionate amount of road traffic emissions compared to their geographic size, and should be a focus for efforts to mitigate such emissions.

In large part, responsibility for facilitating mitigation of these emissions falls to those Local Government Authorities (LGAs) which have urban road networks situated within their areas of administration. Hence, this research sets out to investigate whether (or not) LGAs have available the emissions modelling options they require to enable them to properly discharge this responsibility. The method employed for the investigation is an extensive and contemporary review of literature concerning road traffic data and its use by LGAs in emissions models (EMs). During the review, an absence was found of research specifically investigating the practicalities of LGAs engaging in the emissions modelling process. Strategies for mitigating climate change must tackle increasing urbanisation and its associated problems, and LGAs in all countries play a key role. Therefore, this is an important area for current research.

This article’s scope is limited to GHG emissions, in particular carbon dioxide (CO₂) because it is, by far, the largest constituent of transport’s GHG emissions, e.g. 99% on a CO₂-equivalent basis in the UK (DECC, 2014). Globally, the transport sector’s contribution to total CO₂ emitted from fuel combustion is 23%, of which road traffic is responsible for almost three-quarters (IEA, 2014). Although article scope is limited to CO₂, it is generally acknowledged that integrated strategies to reduce GHG emissions and AQ emissions will often result in significant co-benefits (DEFRA, 2009; EEA, 2009; King et al., 2010; Tiwary et al., 2013). The harmful nature of CO₂ manifests as a global phenomenon. Hence, it is the overall effect of a transport intervention (or combined effect of many small diffuse interventions) on emissions of CO₂ which is more important than any localised effects. Therefore, this article is concerned with estimation of emissions for an urban road network (or substantially large parts of a network) as a whole.

**Requirement for emissions models**

When instigating transport interventions, critical to the decision making process is the ability to assess environmental impacts, including the impact on road traffic emissions. To analyse this impact, it is necessary to quantify an intervention’s effect on emissions. However, it is impractical to measure real-world emissions at road network level due to the large number of vehicles and traffic conditions involved (Smit, 2006; Smit et al., 2010); and measurement is impossible when hypothetical scenarios are considered. Consequently, there is a requirement for EMs, which can offer a practical (and cheaper) alternative to real-world measurements.

Ultimately governments (typically LGAs in urban areas) are responsible for providing road infrastructure, and for maintaining local air quality to within established limits and achieving agreed GHG emission reduction targets. Therefore, LGAs must find the necessary resources for modelling the emissions impact of transport decisions. However, in most cases, public funds are limited. The global financial crisis of 2008 and subsequent austerity measures have increased the constraints on public funds, with many governments forced to make budget cuts (Lowndes and McCaughie, 2013). Hence, LGA resources are under pressure, meaning funds for modelling are scarce.

Models used by LGAs must strike a difficult balance. On the one hand, EMs must not be so simplistic that they fail to capture the majority of the emissions impact of potential interventions. On the other hand, there is a need to avoid model complexity. More complexity entails more time and expertise in building and running models and collecting the necessarily detailed input data. Such time and expertise is generally expensive, and beyond LGA budgets and decision making timescales.

Another factor in the accuracy/complexity trade-off is that, although more complex models are generally more accurate than less complex models, they also require more detailed input data (Smit et al., 2006). Finely detailed input data are more susceptible to errors in estimation, measurement, or assumptions. Therefore, a lack of quality input data may offset any accuracy gained through increased model complexity (Smit et al., 2010; Ramos et al., 2011; Zhu and Ferreira, 2013). This is illustrated by Alonso (1968) who distinguished between two error sources. Firstly, specification error which arises due to models being simplified representations of real-world phenomena; and secondly, measurement errors in input data. Total model error is the sum of these two sources. Fig. 1 shows there is an optimal model complexity where total prediction error is minimised.

**Major factors influencing emissions**

*Distance, speed and vehicle category*

Tailpipe emissions (i.e. from fuel combusted in-vehicle) are typically estimated through multiplying activity data by emission factors (EFs) (Smit et al., 2010). Hence, distance travelled by a vehicle (vehicle-kilometres, VKMs) has a large influence on emissions (i.e. greater activity gives greater emissions). Vehicle speed is another important influence on emissions, because road traffic EFs are strongly dependent on speed (Smit et al., 2008b; Abou-Senna and Radwan, 2013). Vehicle CO₂-equivalent: amount of CO₂ emitted that would cause the same time-integrated radiative forcing, over a given time horizon, as an emitted amount of another GHG.
category also has a considerable influence on emissions. Different vehicle categories have different EFs due to factors such as vehicle mass, fuel specification, engine size, aerodynamics, and emissions control technology.

Importance of congestion

Congestion is the deterioration of smooth, free-flowing traffic conditions due to increased travel demand and/or reduced traffic movement capacity (Smit et al., 2008a). It is commonly accepted that under the stop-and-go traffic conditions associated with congestion there is an increase in the number of acceleration and deceleration events experienced by vehicles, which results in increased emissions (Chen and Yu, 2007; Barth and Boriboonsomsin, 2008; Smit et al., 2008a; Madireddy et al., 2011). Congestion has been repeatedly identified as a major factor when estimating road traffic emissions (Smit et al., 2008a), and ranks alongside VKMs, vehicle speed and vehicle category, as one of the most important influences (De Haan and Keller, 2000; Int Panis et al., 2006; Smit et al., 2008a,b).

Barth and Boriboonsomsin (2008) compared CO₂ emissions from cars during steady-state activity (i.e. constant speed) to emissions during real-world activity (i.e. including dynamics due to congestion) having the same average speed. According to results, the increase in emissions at 45 km/h (a typical average speed on urban roads⁴) due to congestion was approximately 40%. This study assumes that all the dynamics of real-world driving patterns³ are due to congestion. This is a reasonable assumption because the important issue is to capture as much as possible (within resource constraints) of the influence of vehicle dynamics, regardless of whether or not they are labelled as congestion. In effect, the term congestion is used as a proxy for vehicle dynamics, regardless of source.

Congestion can be considered at multiple scales, for example around a single intersection, along a certain corridor (series of links⁴ and intersections), or for a network as a whole. Assessment of the emissions impact of transport interventions at these different scales may require LGAs to use different types of road traffic data and EMs. However, the scope of this article is predicting the impact of interventions on CO₂ emissions at the network level and, consequently, it is LGA options for inclusion of congestion at this scale that formed the focus of research.

Other factors

There are many other factors which influence emissions including: driver behaviour and gear-shift strategies; road gradient; payload; cold starts; ambient temperature; increasing vehicle age or lack of maintenance; use of auxiliaries; and vehicles using alternative drivetrains or fuels. The justification for this article’s focus on congestion, rather than other factors, is

---

² According to the UK Government’s Department for Transport, the average speed on major urban roads in the South-East region of the UK during the peak hour is approximately 47 km/h.

³ Fine grained time series (e.g. 1 Hz) of speed points for an individual vehicle.

⁴ Uni-directional road section between two intersections.
that congestion (arguably) ranks alongside VKMs, vehicle category, and vehicle speed as one of the most important influences on CO₂ emissions; and, therefore, should also be explicitly⁵ included in EMs. In general, the influences of other factors are smaller than that of congestion. However, it is acknowledged that all factors should be included in EMs where possible. Indeed, the influences of most other factors are usually included to some extent, typically through assumptions, either of a constant value (e.g. zero road gradient), or that emissions tests, the data from which are used as the basis for EM construction, cover the real-world distribution of values these factors can take (e.g. vehicle age).

This article is concerned with estimation of emissions at network level (or substantially large parts of a network, e.g. >1 km²). Therefore, the validity of assumptions used to include other factors is strengthened because random errors introduced by not fully accounting for vehicle-specific values should (to a certain extent) average out (Smit et al., 2008b). For example, negative gradients will offset positive gradients, and lighter than average vehicle loads will offset heavier than average vehicle loads. A final point is that, in general, LGAs have more influence over congestion than for example gradient, vehicle loading, ambient weather, use of auxiliaries, etc., and it is interventions affecting congestion in which LGAs are interested.

Road traffic data

Road traffic data required as inputs to EMs are available from numerous sources. Traffic counts (manual or automatic) record the number of vehicles passing a location, and can also include vehicle category data. Automatic Number Plate Recognition (ANPR) cameras read vehicle licence plates allowing vehicle category to be determined, and journey time data to be produced from the time taken for vehicles to travel between different camera locations. The long-established Moving Car Observer (MCO) method uses observers in a test vehicle to derive average journey time and traffic flow data. Queue length surveys manually record the number of queueing vehicles or queue length in metres, and often include associated delay times. The utility of these sources for predicting emissions can be limited, either because their availability is restricted to only a few locations (i.e. they lack link-by-link resolution), or because they involve resource-intensive data acquisition, or both. However, they can be a useful addition to the emissions modelling process, either in combination with other data or through providing data for the calibration and validation of models.

Vehicle category data for a given nation (or sub-national region) are usually available from fleet models, which are typically provided by a country’s government (or other delegated authority). Measurements of the characteristics of the road network itself (e.g. link length, number of lanes, link curvature, intersection layouts, number of intersections per kilometre, speed limits, signal timings or roadside land use) provide another useful source of road traffic data.

Intelligent transport systems technologies

Intelligent Transport Systems (ITS) are defined as any application of information and communication technology to transport, which includes several technologies that can serve as sources of road traffic data. Floating car data (probe vehicles) can be provided by a number of different in-vehicle devices, such as Bluetooth, GPS, mobile telephony and Wi-Fi. These devices can provide information on traffic flow, average speeds, delays, travel times, and driving patterns. A drawback of floating car data is privacy. Gathering identifiable data requires driver permission, which may not be forthcoming from private citizens. Installing devices on captive fleets could be more practical, but resistance may still be encountered from reluctant workers or business owners. Another problem is penetration, i.e. the number of vehicles from which data can be gathered compared to the total number of vehicles. A small sample size decreases the likelihood that the samples will be representative of the traffic conditions on all parts of a network (De Kievit et al., 2014b). An example of the use of floating car data can be found in the compilation of the 2010 London Atmospheric Emissions Inventory⁶ where traffic average speeds for approximately 62% of major road links were available from GPS data provided by Trafficmaster⁷ (GLA, 2014).

Automatic Vehicle Identification using Radio Frequency Identification Devices (RFID) provides similar data to ANPR cameras. Vehicles are fitted with RFID tags (sometimes called transponders), typically in the form of labels attached to the windscreen or licence plate, which pass the vehicle’s details to a roadside tag reading unit. However, vehicle identification is limited to those carrying transponders (De Kievit et al., 2014b).

Based on the vehicle telematics data available from ITS, Traffic Congestion Indices (TCIs) (also known as Traffic Performance Indices, TPIs) can be produced using methods such as comparison of measured travel times with free-flow measured travel times or, less commonly, comparison of the marginal cost of congestion with the average cost of congestion. For example, TomTom produce a Traffic Index for 218 cities worldwide based on GPS data. INRIX also produce global congestion data for urban areas (e.g. the Urban Mobility Scorecard Annual Report), and the Texas A&M Transportation Institute produce similar data for the USA. In the UK, Mott MacDonald produce Strat-e-gis Congestion which provides historic congestion data based on GPS.

---

⁵ Explicit: within model-user control, with values specified for use as EM inputs. Implicit: outside model-user control, but may be included through underlying EM assumptions.

⁶ Inventory of all atmospheric emissions sources in Greater London.

⁷ Trafficmaster is a division of Teletrac, one of the largest fleet companies in the UK and USA.
Road traffic models

Road Traffic Models (RTMs) are often used as a source of EM input data. RTMs represent how travel demand is satisfied by the road network, with demand generally expressed in the form of an origin–destination (O–D) demand matrix, populated by the number of trips between each pair of defined origin and destination zones. RTMs are usually classified according to scale. Macro-RTMs typically consider road traffic as an aggregated flow, with the flow of traffic through links and intersections described by relationships between variables such as: traffic density (vehicles/km), traffic average speed (km/h) and traffic flow (vehicles/h) (Lighthill and Whitham, 1955; Kotsialos et al., 2002). Demand is assigned to the network iteratively, aimed at finding an equilibrium solution that replicates route choice through the network. Demand, and the resulting values for each link’s density, speed and flow, are assumed to be constant for the entire modelled period which is not a very realistic assumption for congested networks (Ortúzar and Willumsen, 2011), particularly over longer time periods (e.g. a peak hour).

Micro-RTMs consider the motions and interactions of individual vehicles based on combining detailed network characteristics with detailed driver behaviour sub-models (car following, lane choice, and gap acceptance) (Papacostas and Prevedouros, 2005; Ortúzar and Willumsen, 2011; Ramos et al., 2011). Hence, driving patterns for each vehicle are available as outputs. However, micro-RTMs are typically calibrated and validated for aggregate traffic measures (e.g. traffic flow, average speed, average delay and queue length), rather than for driving patterns of individual vehicles (Hirschmann et al., 2010; Song et al., 2012, 2013; Toffolo et al., 2013). Therefore, driving pattern outputs from micro-RTMs are rarely validated properly, and do not necessarily accurately represent real-world driving patterns (Song et al., 2012, 2013).

Meso-RTMs are a third classification often distinguished between micro-RTMs and macro-RTMs. Vehicle motions and interactions are considered, but in less detail than in micro-RTMs. For example, SATURN® groups vehicles into platoons, and uses a platoon-dispersion module to simulate the movement of vehicles over the network accounting for interaction of individual vehicles (using less detailed driver behaviour sub-models than at micro-scale), but only calculates entry and exit times for each link, rather than calculating full driving patterns.

The resources (costs, labour, and computer runtimes) demanded by RTMs increase with network size. Hence, for a given resource constraint, as the size of network to be modelled increases, the detail of traffic data generated is effectively reduced. For large urban road networks, the outputs of RTMs are usually limited to data for the traffic on each link, rather than for individual vehicles (Smit et al., 2008b). In other words, micro-RTMs are limited to micro-scale geographic areas (e.g. several links and intersections). As network scale increases, meso-RTMs and macro-RTMs are used instead, providing less detailed output data.

Urban traffic control data

Urban Traffic Control (UTC) systems coordinate traffic signals to achieve good progression for vehicles through urban networks. An example of a widely used system is SCOOT, which operates in over 250 cities and towns worldwide (Bretherton et al., 2011). The data used to control the signals are generated by inductive loop detectors (ILDs) installed under the road surface, which send updates of vehicle presence every 250 ms in the form of 1s and 0s (denoting occupied or unoccupied respectively).

An advantage of UTC data is that it can be considered a by-product of the traffic signal control system (Marsden et al., 2001), allowing EM input data to be collected without using additional resources (Reynolds and Broderick, 2000). Another advantage is that UTC data provide an ‘on the ground’ picture of the real-world situation, in contrast to RTM outputs which are a modelled representation of the real-world. However, a disadvantage is that data availability is limited to certain point locations on the network, i.e. data are only available where ILDs are installed. Another disadvantage is that, where traffic average speed is required as an EM input, this is typically space-mean-speed (i.e. based on average travel time over longer distances, e.g. link length or longer), whereas ILDs provide estimates of time-mean-speed (i.e. average of vehicle spot speeds). Also, because UTC data are real-world measurements, the data are not available for assessment of hypothetical scenarios.

Alongside their primary function of detecting vehicle presence, some enhanced ILDs also provide vehicle category data. Enhanced ILDs rely on the resulting change in inductance due to the passage of a vehicle over the detector being distinct for different vehicle categories. To further assist with vehicle category classification, enhanced ILD data can be supplemented with axle count data (from pneumatic tubes or piezoelectric sensors). However, whilst enhanced ILDs can typically distinguish between two-wheel vehicles, LDVs, rigid HGVs, articulated HGVs and buses/coaches, they cannot further disaggregate vehicle categories (e.g. by fuel type, mass, compliance with different emissions standards, or vehicle age). Additionally, enhanced ILDs are not universally installed, and upgrading existing, conventional ILDs would involve expense and time.

Emissions models

In this review, types of EM are broadly dealt with in order of complexity, moving from simpler EMs to the more complex, as depicted in Fig. 2. The system of EM classification selected to provide a framework for the review is that published in Smit
et al. (2010). However, it is acknowledged that there is no definitive, universally agreed classification system; and that under any given system some examples of EMs may defy easy classification.

Average speed emissions models

Average Speed EMs calculate EFs for each vehicle category as a function of traffic average speed (space-mean-speed). Most road traffic EMs are currently based on average speed (Boulter et al., 2012). A suggested reason for this prevalence is that, particularly for larger urban networks, readily available data are often restricted to estimates of traffic average speed for each link (Smit et al., 2008b). A limitation of Average Speed EMs is that they cannot account for the fact that trips with differing vehicle operation characteristics will all have differing emissions, but could all result in the same average speed (Int Panis et al., 2006; Toffolo et al., 2013). This is a particular problem at low average speeds, such as those in congested urban areas where the possible range of operational characteristics for a given average speed is large (Boulter et al., 2012, 2009; Ramos et al., 2011). However, when applied to a whole network (or substantially large parts of a network) this inaccuracy should be subject to a certain amount of averaging out, i.e. the increase in emissions due to stop-and-go conditions at particular network locations (model under-estimating) is offset by the decrease in emissions from free-flowing conditions at other locations (model over-estimating).

Average Speed EMs do implicitly account for some congestion influence because the driving cycles11 used in vehicle emissions tests that generate the data from which EMs are developed will have a dynamic speed–time profile. However, the driving cycles used during model development cannot be varied by the model user so as to reflect the particular congested situation of interest (Smit et al., 2008a). Common examples of Average Speed EMs include COPERT12 in Europe, MOBILE13 and EMFAC14 in the USA, TRL EFs 200915 in the UK, and the EM built-in to the SCOOT UTC system. Although all these EMs perform emissions calculations using average speed, specific methods for calculations can vary between EMs. For example, calculations based on an average speed for each link, or a single average speed for an entire network, or weighting average speeds by VKMs travelled at each speed. In addition to variation with average speed, EFs in MOBILE and EMFAC are also road type-specific, and so could arguably be classified as Traffic Variable EMs.

Traffic situation emissions models

In Traffic Situation EMs the parameters of emissions tests, and their associated average EFs, are correlated to specific traffic situations. This results in each traffic situation being referenced to an average EF. Different traffic situations are characterised by road type (e.g. motorway with 120 km/h limit, or main road outside built-up area) and a qualitative description of congestion (e.g. free flowing, or stop-and-go). The user specifies a traffic situation, and then appropriate average EFs for different vehicle categories are weighted according to traffic composition (Smit et al., 2010; Boulter et al., 2012). Explicit account for congestion influence is achieved through the user-defined qualitative description of traffic conditions. An example of a Traffic Situation EM is HBEFA16 which is widely used within Europe. However, a disadvantage of HBEFA is that it is designed specifically for use in Germany, Austria, Switzerland, Sweden, Norway and France, with traffic situations representative of conditions in those countries, which means its application elsewhere is questionable (Boulter et al., 2012; De Kievit et al., 2014a).

Traffic variable emissions models

Traffic Variable EMs predict EFs based on variables aggregated for the traffic as a whole (Smit et al., 2010). The inclusion of other traffic variables (in addition to traffic average speed) as indicators of congestion allows congestion influence to be accounted for explicitly and quantitatively. The definition of ‘traffic variable’ is extended here to encompass network characteristics because of their influence on traffic movement capacity. Whilst this EM category is somewhat lacking in examples of fully developed, commercially available EMs, the following paragraph details examples of research in this domain.

An early example of a Traffic Variable EM is provided by the Traffic Energy and Emissions-Kinematic Correction Factor (TEE-KCF), which attempted to overcome the limited ability of Average Speed EMs to account for congestion through use of a KCF. Its application initially involves calculating emissions using traffic average speed and COPERT. Then link-specific KCFs are calculated based on values for traffic average speed, traffic density, effective green-time ratio,17 and link length. Finally, COPERT emissions are multiplied by the KCF (Smit, 2006; Boulter et al., 2012). Research by Reynolds (1996) investigated the use of SCOOT UTC data to predict roadside concentrations of AQ pollutants. An attempt was made to produce a Traffic Variable Concentration
Model that predicted roadside concentration of carbon monoxide (CO) based on SCOOT values for traffic delay and flow. A study by Smit et al. (2008b) investigated improving the accuracy of emissions predictions through application of a link-specific average speed distribution, rather than a single traffic average speed. In other words, supplementing traffic average speed with data on how the average speeds of individual vehicles are distributed around the average for the traffic as a whole. This method was seen as being a closer approximation to reality, and so was expected to improve on the accuracy of emissions predictions based solely on traffic average speed. Jeng et al. (2013) investigated enhanced ILDs that went beyond vehicle category classification, and worked on the principle that the resulting change in inductance due to vehicle passage is unique to that vehicle. The study used ‘raw’ ILD data (rather than UTC system outputs) to provide improved estimates of traffic average speed and vehicle category mix, which could then be used as inputs to existing EMs. Song et al. (2015) developed the Delay Correction Model (DCM) to predict emissions from buses traversing intersections based on traffic variables commonly used to describe intersection performance. The DCM is applied using intersection delay (seconds), number of stops (during intersection crossing), and intersection type (arterial/arterial or arterial/collector) as inputs, from which a Delay Correction Factor (DCF) is calculated. A baseline EF, which represents the case where a bus encounters no delay during intersection crossing, is then multiplied by the DCF to find the EF particular to the situation of interest.

**Cycle variable emissions models**

Cycle Variable EMs calculate EFs for individual vehicles as a function of variables derived from a vehicle’s driving cycle, for example: number of stops per km, average speed, maximum acceleration, idle time, etc. A vehicle’s driving pattern is typically required as input, which means congestion influence is explicitly included. However, the necessary driving patterns for each vehicle can only be acquired from a micro-RTM or vehicles equipped with a GPS device (Smit et al., 2010).
VERSIT+ LD\textsuperscript{18} (for Light Duty Vehicles) was originally developed as a Cycle Variable EM. It consisted of statistical models that were constructed using multiple linear regression analysis of emissions test data to find an empirical relationship between EF and driving cycle variables for each vehicle category (Smit et al., 2007). However, major changes were made to VERSIT+ LD in 2009, and it is now better described as a Modal EM (Ligterink and De Lange, 2009). The Network Emissions Model (NEMO) was developed during ARTEMIS.\textsuperscript{19} Emissions for individual vehicles are calculated based on values of cycle variables (Smit, 2006; Palmer, 2007). However, rather than using driving patterns for each vehicle as inputs, NEMO is designed to calculate cycle variables from a single driving pattern representative of all traffic. This simplification was introduced to avoid difficulties associated with collecting driving patterns for every vehicle and to reduce computing time. In this respect, NEMO could arguably be classified as a Traffic Variable EM because a single driving pattern is used to represent all traffic on a link.

Modal emissions models

Modal EMs calculate EFs for individual vehicles as a function of vehicle or engine operating modes (Smit et al., 2010). A vehicle’s driving pattern is typically required as input, which means congestion influence is explicitly included. The latest generation of Modal EMs predict EFs for operating modes at high temporal resolutions (e.g. 1 Hz), i.e. a vehicle’s mode, and associated EF, is calculated on a second-by-second basis. At this temporal resolution, Modal EMs are typically termed Instantaneous EMs (IEMs).

Motor Vehicle Emission Simulator (MOVES) is the EPA’s official EM. Operating modes in MOVES are characterised by a combination of vehicle speed and Vehicle Specific Power\textsuperscript{20} (VSP), and are divided into bins, with each speed–VSP bin having an associated EF. From a driving pattern, a link’s operating mode distribution (amount of time vehicles spend in each speed–VSP bin) is calculated (Zhao and Sadek, 2013). Hence, similar to NEMO, MOVES could arguably be classified as a Traffic Variable EM because a single operating mode distribution is used to represent all traffic on a link. The Passenger car and Heavy duty Emissions Model\textsuperscript{21} (PHEM) uses a vehicle’s driving pattern, vehicle characteristics data, and a model of gear shift behaviour, to compute instantaneous values of engine power and engine speed. These values are then used to determine the associated EF (Boulter et al., 2012; Hausberger et al., 2009).

Emissions models to assess ITS

This section does not constitute a separate class of EM. Instead, because ITS assessment has generated substantial research into predicting urban network emissions (using various EM types), this work is grouped together here for convenient review. AMITRAN\textsuperscript{22} and ECOSTAND were concerned with developing standard methodologies for evaluating the effects of ITS on CO2 emissions. Both projects concluded that network geographic scale has a major bearing on the appropriate type of EM for assessment. For localised interventions, a detailed EM is required that predicts emissions for individual vehicles. However, when assessing ITS interventions on larger scales, a less detailed EM is acceptable and may be the only practical option (De Kievit et al., 2014a; Jonkers et al., 2014).

CARBOTRAF produced a Traffic Variable EM that predicts real-time network CO2 emissions based on ILD data supplemented by data from specialist roadside sensors called smart eye Traffic Data Sensors\textsuperscript{23} (TDS) which determine the proportion of accelerating vehicles in the traffic. Alternative traffic management options that satisfy traffic demand with reduced network CO2 emissions are then offered to the traffic control centre operator (Litzenberger et al., 2012; North and Hu, 2012).

The ICT-Emissions approach was to model in detail the local effect of an ITS intervention, and then extrapolate the effect across the network. Two methods were proposed to integrate between detailed, local-scale modelling and less detailed, larger-scale modelling. The first used a local-scale micro-RTM to determine local changes to traffic’s speed–flow–density relationships, which were then used to adjust a network-scale macro-RTM, the average speed outputs of which were used as inputs to COPERT. The second method involved constructing an extended Average Speed EM. Driving cycles were selected to represent conditions with and without the ITS intervention in place, and an IEM used to calculate associated emissions. In this way, two average speed emission functions are generated for each vehicle category, one with the ITS intervention and one without. A macro-RTM is then used to generate average speed inputs to the extended Average Speed EM (Toffolo et al., 2013).

---

\textsuperscript{18} VERkeers SIUtatie – distributed by the Netherlands Organisation for Applied Scientific Research (TNO).

\textsuperscript{19} Assessment and Reliability of Transport Emission Models and Inventory Systems – a European Commission 5th Framework project.

\textsuperscript{20} VSP (kW/tonne) is calculated from a vehicle’s instantaneous speed and acceleration, and represents the power demand placed on a vehicle (by rolling resistance, aerodynamic resistance, acceleration resistance, and road gradient resistance) divided by vehicle mass.

\textsuperscript{21} Developed during the ARTEMIS project.

\textsuperscript{22} AMITRAN, ECOSTAND, CARBOTRAF & ICT-Emissions are all European Commission 7th Framework projects.

\textsuperscript{23} A proprietary sensor produced by the Austrian Institute of Technology.
Discussion

Readily available road traffic data

The results of the literature review indicate that the detail level of road traffic data readily available for collection by LGAs within their resource constraints is data aggregated at traffic level rather than data disaggregated at individual vehicle level; i.e. traffic variables rather than driving patterns. In general, driving patterns lack availability because they are difficult to collect, rarely used in traffic engineering to describe road network performance (Song et al., 2015), and their simulation in micro-RTMs is of questionable accuracy.

Two sources of traffic variables readily available on a link-by-link basis are UTC data and RTM output data. UTC data are particularly appealing because they are a by-product of the traffic signal control system. RTM data are required for instances when UTC data are not available, such as for assessment of hypothetical situations or for areas of the road network where ILDs are not installed. Reliance on other road traffic data sources to provide EM input data is likely to incur additional costs for collection and processing on a link-by-link basis, and/or involve sources that are not widely used by LGAs.

UTC data are more appropriate for the assessment of smaller, tactical interventions because their real-world nature means assessment can only be achieved post-intervention through comparing before and after emissions estimates. Examples of smaller interventions include altering traffic signal timings, re-routing particular vehicles (e.g. HGVs or buses), restricting vehicle loading/unloading to certain times, or prohibiting on-street parking at certain locations (Reynolds, 1996; Reynolds and Broderick, 2000). When assessing these interventions, it should be emissions for the whole network (or substantially large parts of the network) that are analysed, rather than local scale emissions, because assumptions about random errors averaging out become less valid for localised assessments (Smit et al., 2008b).

The impact of large, strategic interventions typically requires assessment prior to implementation, when there will be no real-world UTC data available for hypothetical post-intervention scenarios. Therefore, an RTM is required to simulate the effect of an intervention on traffic, and can be used as a source of traffic data for emissions modelling. Examples of large interventions include substantial alterations to the road infrastructure, variation of area-wide speed limits, or provision of large car parking facilities or park-and-ride schemes.

Network characteristics are an additional source of traffic variables readily available to LGAs. Data on these characteristics have the advantage of being fairly easily measured by LGAs (or on their behalf); and having been measured once, are not subject to change very often. The rise in vehicle telematics is also a potential source of traffic variables that could come to satisfy LGA traffic data requirements, and investigation of relationships between TCIs and CO₂ emissions is an area for future work.

Optimal EM complexity

Resulting from the literature review and based on the ready availability to LGAs of traffic variables, a hypothesis has been formed that optimal model complexity for LGAs is represented by Traffic Variable EMs. This is because using less complex models (i.e. Average Speed or Traffic Situation EMs) does not fully utilise all traffic variables readily available to LGAs, which offer potential to improve accuracy through explicitly including congestion influence. It has been suggested that including only the influence of vehicle category and speed, and ignoring other influences, could be a simplification that distorts the information supporting policy making (Ligterink et al., 2012).

Using more complex models (i.e. Cycle Variable or Modal EMs) requires LGAs to collect and process accurate driving patterns for each vehicle on the network, which is impractical within existing and likely future resource constraints. Even when LGAs invest in micro-RTMs, because they are typically calibrated for aggregate traffic measures rather than for individual vehicle driving patterns, there is uncertainty about whether simulated driving pattern outputs are accurately representative of the real-world. A case study by Vieira da Rocha et al. (2015) found some evidence that the effect of inaccuracies in simulated driving patterns causing inaccuracies in predicted emissions tends to average out when lots of driving patterns are used to predict emissions from groups of vehicles. However, even if these findings apply universally, the significant data processing requirements of collecting (accurate or inaccurate) driving patterns for every vehicle in a network, and then inputting them to Cycle Variable or Modal EMs are likely to be off-putting for LGAs.

EM options for LGAs

A reason for the prevalence of Average Speed EMs in network emissions modelling is the common availability of traffic average speed data for use as inputs (which should be space-mean-speed, but does not mean attempts aren’t made in practice to use time-mean-speed). However, other traffic variables that can be used as indicators of congestion are also readily available to LGAs. Hence, a Traffic Variable EM that includes the influence of congestion is a practical option for LGAs. A further point worthy of reiteration is the advantage of Traffic Variable EMs over Average Speed EMs whereby a Traffic Variable EM can distinguish between different traffic conditions that happen to result in the same average speed. It is also worth noting the possibility that a Traffic Situation EM, which includes the influence of congestion qualitatively, may be able to compete with a Traffic Variable EM in terms of accuracy and resource consumption.
If LGAs want to use Traffic Variable EMs (or Traffic Situation EMs), what options do they have? The examples detailed here are typical of the options available, and serve to illustrate some of the obstacles that LGAs face. TEE-KCF was a fully developed example of a Traffic Variable EM, but the emissions data on which it is based are now over 10 years out-of-date. The application of average speed distributions by Smit et al. (2008b) was a case study investigation, which has not been developed into a format usable by LGAs as an EM. Development of the DCM by Song et al. (2015) was accomplished for buses traversing intersections in Beijing, China. Therefore, it would need extension to cover other vehicle categories, traversing links as well as intersections, in other global regions, before it can have wider application. NEMO has not achieved traction as a widely used EM, and has not replaced the wide-spread use across Europe of COPERT and HBEFA.

Both HBEFA (in Europe) and MOVES (in the USA) are fully developed, up-to-date EMs that are currently used by LGAs. However, both would require extension to be usable in other global regions. In general, transferability is an obstacle to be overcome when considering an EM developed in one global region for application in another. This is because vehicle category classifications can be markedly different between regions and there is often no easy way to map between them. Even within a given global region (e.g. Europe) where the system of vehicle category classification is standardised, transferability between sub-regional areas (e.g. European countries) may not be straightforward. The reason for this being differences in factors such as network characteristics or vehicle fleet compositions.

Reynolds (1996) attempted to predict roadside CO concentrations based on traffic variables output from SCOOT. Unfortunately, no relationship could be established, and no model was produced. Whilst this work concerned a different pollutant, and concentrations rather than emissions, the results suggest it may prove difficult to develop a Traffic Variable EM for CO₂ based solely on outputs from a UTC system. Rather than UTC system outputs, Jeng et al. (2013) investigated the use of ‘raw’ enhanced ILD data. However, this study was not aimed at developing an EM itself, but was a method to provide more accurate road traffic input data to existing EMs, and additionally requires LGAs to invest in upgrading conventional ILDs to enhanced versions.

CARBOTRAF is a Traffic Variable EM predicting network CO₂ emissions. However, CARBOTRAF is pre-loaded with a database of network CO₂ emissions for each alternative intervention (in CARBOTRAF these are alternative traffic management options), which have been pre-calculated using driving patterns from a micro-RTM as inputs to an IEM. CARBOTRAF works because the range of alternative traffic management options is limited, allowing their emissions to be pre-calculated. Extending CARBOTRAF to assess an open-ended range of interventions would entail assessing each alternative with a micro-RTM and an IEM, which is just the kind of resource-intensive process likely to be beyond LGA budgets. Also, CARBOTRAF is reliant on the specialist TDS, which would require LGAs to meet associated installation costs. ICT-emissions proposes two methods for including changes to vehicle dynamics in network level emissions modelling, both based on Average Speed EMs. However, as both methods extrapolate a micro-scale effect across a network as a whole, they are more appropriate for blanket interventions that affect all links at once in a similar fashion (e.g. area-wide promotion of eco-driving) rather than for interventions that affect different links in different ways (e.g. altering signal timings on certain links). Additionally, the first method relies on localised changes to speed–flow–density relationships due to an ITS intervention being replicated across an entire network, which may not be valid; and the second method requires the resource-intensive process of constructing an extended Average Speed EM to be repeated for each alternative intervention.

It could be argued that if an LGA has invested in building an RTM, then an efficient way to calculate emissions would be using the RTM’s built-in EM (if it has one). However, built-in EMs come with problems. For example, the distributors of SATURN suggest that the built-in Traffic Variable EM is extremely crude, and that emissions predictions would be best handled using a stand-alone EM (Atkins Limited, 2014). The AIMSUN built-in IEM highlights another problem, which is the mismatch in vehicle category disaggregation between RTMs and EMs (Tofolo et al., 2013). RTMs are typically only concerned with highly aggregate vehicle categories; whereas, to accurately predict emissions, EMs are typically concerned with highly disaggregate vehicle categories. The AIMSUN IEM provides emission functions for only five aggregate vehicle categories, which is consistent with the vehicle category aggregation one might expect in RTMs (TSS, 2013). It is possible to define many more vehicle categories in AIMSUN, each with different emission functions. However, acquiring the data to define all the emissions-related characteristics of each new vehicle category would be a substantial task. Additionally, a large increase in the number of vehicle categories would entail a large increase in complexity of O–D demand matrices in order to accurately represent the proportion of vehicles in each category travelling between each origin–destination pair. A further problem that affects micro-RTMs with built-in Modal EMs is the aforementioned issue of micro-RTM calibration for aggregate traffic variables rather than for accurate driving patterns, leading to uncertainty about the simulated driving patterns used as inputs to the built-in Modal EM.

The examples discussed in this section demonstrate that extant methodologies in the domain of predicting network emissions based on traffic variables typically have (to a greater or lesser extent) limitations. These must be addressed if an EM that explicitly includes congestion, whilst remaining within LGA resource constraints, is to be widely established.

Requirement for further research

The hypothesis that a Traffic Variable EM is the optimal model complexity for LGAs is based on qualitative literature review results. It has not been possible to quantitatively validate this hypothesis because there is only limited literature concerning EM validation and quantification of prediction errors (Smit et al., 2010; Kousoulidou et al., 2013). In general, it is not possible to validate an EM at network level in the strict scientific sense because it is not possible to measure true emissions
for a network due to the large number of vehicles and traffic conditions involved. Instead, only partial validation is possible for specific localised situations over relatively short time periods (Smit, 2006). A meta-analysis of EM validation studies by Smit et al. (2010) concluded that “there is inadequate understanding of the uncertainties in traffic emission models and the main factors affecting prediction errors”. Another, interesting finding of this meta-analysis was that there was no conclusive evidence for more complex EMs systematically being more accurate than less complex EMs. Part of the explanation for this could be that more complex EMs are being used with input data at less than ideal accuracy, i.e., accuracy gains from decreasing specification error ($e_s$) are offset by increasing measurement error ($e_m$).

Additional research is required comparing the predictions of Average Speed, Traffic Situation and Traffic Variable EMs (EMs likely to be useable within LGA resource constraints) with real-world emissions before the hypothesis can be assessed quantitatively. Maybe Average Speed or Traffic Situation EMs will prove equivalent to Traffic Variable EMs in terms of accuracy; although, interestingly, the EPA felt that MOBILE (Average Speed EM) needed replacing with MOVES. If results show that Traffic Variable EMs do out-perform their rivals, then further research would be required to develop existing or new methodologies into a widely available EM for LGAs.

Conclusions

Returning to the question posed in this article’s title, the conclusion is that the qualitative evidence indicates that LGAs do not necessarily have the right options to accurately include the effects of congestion on emissions of CO2 (or other pollutants) from traffic on urban road networks. Based on readily available road traffic data, the hypothesis is put forward that the optimal model complexity for LGAs is a Traffic Variable EM. More complex models, that calculate emissions for individual vehicles, are discounted as an option because they require driving patterns as inputs, which are difficult to collect, rarely used in traffic engineering to describe road network performance, and simulated with questionable accuracy in micro-RTMs.

In contrast, traffic variables are used to describe network performance and are more readily available to LGAs. Average Speed and Traffic Situation EMs are both viable options because they use traffic variables as inputs, are well established, and are less complex than Traffic Variable EMs. However, they may not fully realise the explanatory power of all traffic variables available to LGAs, which could be used to explicitly include congestion influence and improve emissions prediction accuracy.

Whilst the superiority of Traffic Variable EMs over Average Speed and Traffic Situation EMs in terms of accuracy has not been established quantitatively, there is enough qualitative evidence in the literature to conclude that this subject warrants further research. Co-benefits of such research would be in addressing any limitations of existing methodologies in the domain of predicting network emissions based on traffic variables and further developing EMs that explicitly include congestion whilst remaining within LGA resource constraints.

Acknowledgements

This research was jointly funded by the Engineering and Physical Sciences Research Council (Grant number EP/G036896/1) and Southampton City Council.

References
