Big Data Research 6 (2016) 11–25

Contents lists available at ScienceDirect

Big Data Research

www.elsevier.com/locate/bdr

Big Data for Supporting Low-Carbon Road Transport Policies in Europe: Applications, Challenges and Opportunities

Michele De Gennaro [a,*], Elena Paffumi [b], Giorgio Martini [b]

[a] AIT, Austrian Institute of Technology GmbH, Mobility Department, Giefinggasse 2, Vienna, 1210, Austria
[b] European Commission – Directorate General Joint Research Centre, Institute for Energy and Transport, Sustainable Transport Unit, via Enrico Fermi 2749, Ispra 21027, Italy

ARTICLE INFO

Article history:
Received 21 December 2015
Received in revised form 5 April 2016
Accepted 29 April 2016
Available online 2 June 2016

Keywords:
Big data
Data mining
Low-carbon road transport
European transport policies
Environmental policies and climate change mitigation
Smart cities

ABSTRACT

Big data is among the most promising research trends of the decade, drawing attention from every segment of the market and society. This paper provides the scientific community with a comprehensive overview of the applications of a data processing platform designed to harness the potential of big data in the field of road transport policies in Europe. This platform relies on datasets of driving and mobility patterns collected by means of navigation systems. Two datasets from conventional fuel vehicles collected with on-board GPS systems have been used to perform an initial pilot study and develop its core algorithms. They consist of 4.5 million trips and parking events recorded by monitoring 28,000 vehicles over one month. The presented analyses address: (1) large-scale mobility statistics, (2) potential of electric vehicles in replacing conventional fuel vehicles and related modal shift, (3) energy demand coming from electric vehicles, (4) smart design of the recharge infrastructure and Vehicle-to-Grid, and (5) real-world driving and evaporative emissions assessment and mapping. The developed methodology and the presented outcomes demonstrate the potential of big data for policy assessment and better governance, focusing on the challenges and on the huge opportunities offered for future developments. This paper ultimately aims to show how big data can inspire smart policies together with public and private investments to enable the large scale deployment of the next generation of green vehicles, offering an unprecedented opportunity to shape policies for future mobility and smart cities.

1. Introduction

1.1. The European road transport policy framework and the use of big data for supporting low-carbon road transport policies in EU

In order to enable transport emissions reduction in Europe and to meet the Kyoto protocol objectives, the EC White Paper 2011 sets the de-carbonisation of transport as a priority, defining ten goals to be achieved over the next twenty to forty years, [1]. As far as road transport is concerned, these include:

- Halve the use of conventional fuel cars in urban areas by 2030, phasing them out in the cities by 2050;
- Establishing an European framework for multi-modal information management systems;
- Move towards full application of "user-pays" and "polluter-pays" principles and private sector engagement to eliminate distortions in the taxation, including harmful subsidies.

Moreover, in February 2015, the European Commission has unveiled a Strategy and Action Plan for creating an Energy Union [2], endorsed by the European Council on 19 March 2015, that includes several actions and initiatives in three key areas which can directly or indirectly reduce greenhouse gas emissions from the transport sector: (1) switching towards carbon free or less carbon intensive fuels, (2) improving vehicle efficiency and (3) managing transport demand. All these goals call for the widespread of low-carbon vehicle technologies (i.e. Hybrid and Battery Electric Vehicles, HEVs and BEVs), together with smart systems capable to harness the potential of digital mobile technologies in storing and processing data, in order to provide the user with smart transportation solutions in real-time as well as enable the implementation of a smart taxation system.

Transportation is a complex world. It is a mix of technologies, social behaviours, choices of single users and stochastic events, nested within a geographical, environmental and economic sce-

* Corresponding author. Tel.: +39 3396917924.
E-mail addresses: michele.degennaro.phd@gmail.com, michele.degennaro@ait.ac.at (M. De Gennaro).

http://dx.doi.org/10.1016/j.bdr.2016.04.003

Nomenclature

Acronyms

AC	Alternating Current
BEV	Battery Electric Vehicle
CATI	Computer-Assisted Telephone Interviews
CO_2	Carbon dioxide
CS2-JU	Clean Sky 2 Joint Undertaking
DC	Direct Current
EEA	European Economic Area
EU	European Union
EU-ETS	EU Emission Trading Scheme
HEV	Hybrid Electric Vehicle
GHG	Greenhouse Gas
GIS	Geographic Information Systems
GPS	Global Positioning System
ICT	Information and Communication Technology
LDV	Light Duty Vehicle
POI	Point of Interest
SUV	Sport Utility Vehicle
TEMA	Transport tEchnology and Mobility Assessment platform
USA	United States of America
VOC	Volatile Organic Compound
V2G	Vehicle to Grid

nario. For this reason the design of the transportation network and its regional regulatory framework involves know-how from engineering, geography, environmental sciences, economy and social sciences. This process must be supported by harmonised datasets from different sources together with data processing methodologies developed across different scientific fields in order to handle real world complexity. Recent studies prove the potential of big data in this respect, being used to measure commuting efficiency in megalopolis [3], to explore public transport users' behaviours [4], to simulate individual mobility choices in carpooling [5] and to classify activity patterns [6,7], with applications in the fields of mobility networks design and infrastructures [8–10] and multi-modal transportation systems [11]. On one hand these studies constitute interesting advances of big data in transport, but, on the other hand, they are limited to single case studies and applications, mostly grounded on data averaging and data aggregation approaches.

More in general, the potential of ICTs in support of mobility needs is a well-known research topic, funded by a number of European projects under the FP7-ICT call for projects 2011. Among the most relevant projects, we must cite the REDUCTION [12], the PEACOX [13] and the eCOMPASS [14] projects, which addressed novel ICT solutions for optimising driving behaviour, routing and multi-modality for passenger and freight transport fleet in cities, together with the ICT-EMISSIONS project [15], that addressed CO_2 emission estimation at a regional scale via monitoring of vehicle fleets. All these parent applications explore interesting uses of data in transport, and might be interesting data source for future big data studies.

Nowadays big data is applied to a number of different disciplines and in 2013 Forbes.com indicates big data among the "Top 10 Strategic Technology Trends of the Year" and "Top 10 Critical Tech Trends For The Next Five Years", [16] and [17]. The term "data science" dates back to the 1960s, when the newly born information technologies posed the problem of how to store, manage, process and retrieve growing amount of data, in a way which was never experienced before. In 1962, John W. Tukey wrote "The Future of Data Analysis" [18], where he presented a visionary perspective of how mathematics, statistics, data analysis and informatics can be merged in a novel discipline with unprecedented potential. After nearly 50 years, at the beginning of the 21st century, data is become "big". Of course this is related to the "size", but it would be quite reductive to condense the meaning of "big data" solely to its bytes count. The big data means "bits and pieces of the real world puzzle", which, if adequately processed, is capable to offer unprecedented insights on a number of real world phenomena. This is potentially capable to impact social dynamics, choices and behaviours, public response to events, market trends, services and goods' demand, enabling to open the doors to a number of applications, re-inventing and re-structuring existing companies and evolving novel concepts in almost every business segment. This work aims to present how big data can be used for supporting low-carbon transport policies in Europe.

1.2. Transport emissions in Europe

In the European Union (EU), transport contributes to nearly one-third of the carbon dioxide (CO_2) emissions and is the only major sector where emissions increased over the last decade despite the economic downturn [19]. The EU needs to reduce the Greenhouse Gases (GHGs) emissions by 20% below 1990 levels by 2020, and by 80-to-95% by 2050, under the Kyoto Protocol, [20,21]. Transport, i.e. road, rail waterborne and air transport modes together, will contribute to this goal by reducing its GHGs emissions below 1990 levels by 60% by 2050, [1]. EU accounts of 35.3 billion tonnes CO_2-equivalent emissions in 2013 [22], with approximately 11 billion (i.e. one-third) tonnes coming from transport. Road transport accounts for approximately 72%, rail and waterborne transport account together for 15% while air transport accounts for the remaining 13% of the total transport GHGs emission in EU, as per [23]. Nearly two-thirds of road transport emissions originate from light duty vehicles (LDV), while the remaining one-third originates from heavy duty vehicles (HDV) [24], representing respectively 83.3% of total inland surface passenger transport (LDV) and 9.2% of total inland surface passenger transport (HDV, i.e. buses) plus 74.9% of total inland surface freight transport (HDV). Rail transport only accounts for 7.5% of total inland surface passenger transport and 18.2% of total inland surface freight transport, while waterborne accounts for the remaining 6.9% of the inland surface freight transport, with a passenger share which is negligible, as per [25,26]. By considering carbon-intensity per mode (i.e. CO_2 grams per kilogram of payload per kilometre), rail and waterborne are definitively greener solutions, accounting for nearly one-third of the specific emissions compared to road or air transport [27]. However, despite their high carbon-reduction potential, rail and waterborne are still under-exploited solutions and this is mainly related to the poor inter modality of the networks between the Member States and to their low cost-efficiency compared to road transport, [1]. On top of this, automotive and oil industry still have a very important economic weight in most of the countries and this make a change even more difficult. Recent statistics estimates that the number of worldwide circulating vehicles is approximately one billion units (2013), exhibiting an annual growth rate of 4% till 2020, [28]. Such growth is mainly due to the increasing wealth in emerging countries (i.e. 76% of the 2020 market is forecasted not to be in EU or US, [29]), with a long term growth up to 2.4 billion circulating vehicles in 2050 (i.e. +140% with respect to the 2013 figure, traded off by a worldwide population increase of only 30%, [30]). Regarding air transport mode, this is typically not included in road, waterborne and rail transport compared statistics, being it natively cross-borders and not "inland surface". As per road mode, air transport is also a carbon-intensive solution (i.e. their specific emissions are comparable) and fast-growing. In fact air passengers in, from and to the EU went from 810 million in 2009 to 920 million in 2013 (+13.5% over 5 years, i.e. 86% of air-carried payload), while air cargo went from 11.8 million tonnes in 2009 to 15.0 million tonnes in 2013 (+27.1% over 5 years, i.e. 14% of air-carried payload), as per [31]. The air cargo share on freight transport is negligible when compared to the inland surface value, while, for passenger transport, the air carriers account for a significant share of kilometres travelled per person, and therefore for a considerable overall carbon footprint. However, despite the carbon intensity drawback of air transport, its peculiar short travel time allows this mode to play a role which can be hardly replaced by road and rail transport, bridging European countries and people and playing a key societal role for the cultural exchange and the integration of the European citizens.

Apart from the impact of GHG emissions on climate change, transport is also a significant source of noxious air pollutants which are proven to have serious implications on human health. The World Health Organisation highlights that tens of thousands of deaths per year can be attributed to road transport-related air pollution, similar to the death toll from traffic accidents [32]. Moreover recent epidemiology and toxicology literature reviews have shown that there is a causal relationship between human exposure to traffic-related gaseous emissions and exacerbation of respiratory and cardiovascular diseases for people living within 500 meters from major roads [33]. These findings are confirmed by a large number of similar studies, from different regions of the world, such as [34,35] and [36], even at levels far below those causing the severe public health issues which have been arising in far-east megalopolis [37]. Similarly to road transport, pollutants from civil aviation are proven to have significant health impact and turbofan engines are considered primary sources of $PM_{2.5}$ and gaseous aerosol precursors, causing health damages that may be distributed over regional-to-continental scales. A recent estimation attributes nearly 210 deaths per year in US which can be directly related to pollutants from aviation [38], and similar figures are presumably likely to be valid in EU too. Based on this motivation, the emissions from aviation, including landing, take-off and ground activities, are receiving increasing attention from the regulatory bodies. In 2012 the emissions from all flights from, to and within the European Economic Area (EEA), have been included in the EU Emission Trading Scheme (EU-ETS) according to the entry into force of the Directive 2008/101/EC, [39], while US EPA announced to move the first step towards regulating GHGs emissions and NO_x emissions from aircraft in response to section 231 of the Clean Air Act, [40]. Additionally the EU Clean Sky 2 Joint Undertaking (CS2-JU) Programme [41], addresses the reduction by 50% of CO_2 emissions through reduction of fuel consumption and the reduction of 80% of the NO_x emissions as main priorities for the European aviation industry. There is hence the necessity to investigate the potential of innovative vehicle technologies in road, rail and aerial transport under real world constraints, in order to minimize transport environmental impact and maximize the overall system efficiency and services, to meet the climate change mitigation and the sustainability challenges of the upcoming decades.

1.3. The TEMA platform

The objective of this paper is to provide the scientific community with an overview of a data processing platform TEMA (Transport tEchnology and Mobility Assessment platform) developed by the authors over nearly three years, designed for harnessing the potential of big data in the field of transportation policies in Europe. The platform is conceived as a flexible and modular platform and is explicitly designed for multi-purpose applications. The topics addressed so far are:

- quantifying the real world potential of deploying electrified vehicles within urban areas accounting for different electric vehicles penetration shares under different technological and infrastructural constraints;
- quantifying and geo-referencing the shift from oil to electric energy and the impact on the electricity distribution grid of the deployment of EVs;
- evaluating driving and evaporative real world emission from the current fleet of conventional vehicles, and the gaseous emissions reduction potential from the introduction of new vehicle technologies;
- estimate future market competition and new business opportunities offered by the diverse considered scenarios.

Together with the developed applications and challenges to overcome, such as handling and analysis large datasets, this paper aims to present the challenges and the opportunities offered by this approach and, more in general, by big data for transport policies. The innovativeness of this work consists in its multi-disciplinary characteristic and applicability to different fields, showing how the approach and the algorithms developed for TEMA are capable to support the development of effective transport regulation in the areas of real world driving and gaseous evaporative emissions, new vehicle technologies deployment and alternative fuel infrastructures design, in term of environmental impact, energy efficiency, climate change and sustainability. Section 2 of this article presents the methodological steps undertaken to develop TEMA, providing the reader with a quick overview of the adopted and implemented models applied to two datasets as pilot, while section 3 provides the key results of the several applications developed so far, summarising the contents of a number of scientific papers from the authors, i.e. [42–48] and [49].

The aim of the present work is to present in a homogeneous format the structure of the developed platform, and the results obtained considering two pilot datasets. This allows to focus the attention on the challenges and the nearly unlimited opportunities offered by big data for transport applications, potentially inspiring smart policies and public and private investments, paving the way towards low-carbon mobility and ultimately towards low-carbon society.

2. Methodology

2.1. Step 1: pre-processing the data

Every big data application relies on a set of data, and the correct pre-processing of the data is a key for harnessing the data potential and application. In order to develop an effective pre-processing methodology two large datasets of driving and mobility patterns collected by means of GPS have been analysed as pilot study. The driving pattern databases refer to the Italian provinces of Modena and Firenze and are representative of a large sample of the LDVs fleet registered in these provinces. The databases have been acquired from a private company (i.e. Octo Telematics [50]) which equipped a large number of vehicles with GPS black boxes on behalf of a major insurance company. The fleet sample contains a mix of light duty vehicles registered either to the name of private citizens or to commercial activities (more details reported in [42]). These vehicles belong to different segments, and they have been selected in order to match the owners' distribution with the Italian average data, with the objective of guaranteeing the highest

Table 1
Overview of the analysed data for the provinces of Modena and Firenze.

	Surveyed vehicles (% of the total)	Analysed sample (% of the surveyed)	Number of records [$\cdot 10^6$]	Analysed km [$\cdot 10^6$]	Analysed trips [$\cdot 10^6$]
Province of Modena	52,834 (12.0%)	16,263 (30.7%)	16.920	14.98	2.642
Province of Firenze	40,459 (5.9%)	12,478 (30.8%)	33.361	20.66	1.870

representativeness of the data sample. They allow reconstructing the complete activity pattern (i.e. the sequence of trips and parking events) which characterises each vehicle over an analysis period of one month (i.e. May 2011). The data acquisition campaign has originally involved 52,834 conventional fuel vehicles registered in the province of Modena and 40,459 vehicles registered in the province of Firenze (i.e. respectively 12.0% and 5.9% of the fleet in these provinces). The on-board data-logging devices recorded time, GPS position coordinates, engine status, instantaneous speed and cumulative distance. The data is collected to be continuous in time, and the acquisition frequency autonomously adapts to trade-off the size of data and accuracy of the information. A double filtering was preliminary applied to the raw data in order to delete the vehicles driven for more than 50% of the trips outside the province borders (in order to focus only on those vehicles which show a predominantly local usage, i.e. urban vehicles) and all the trips with a length less than 30 meters and/or duration less than 30 seconds (not representative of a real mobility demand). A trip is defined as an event which starts with the switch-on of the internal combustion engine and stops with the switch-off of the internal combustion engine from the user. Start-and-stop feature of modern vehicles is not considered in this analysis, being no representative of real trips. As consequence of these filters the databases are reduced to 16,263 vehicles for Modena (30.7% of the original size, 3.6% of the province's fleet) and to 12,478 vehicles for Firenze (30.8% of the original size, 1.7% of the province's fleet). All the analyses are carried out on these sub-sets only, thus targeting the urban mobility in the developed applications. For both provinces about the 91% of the analysed vehicles were registered to the name of physical persons (later labelled as "private") while the remaining part were vehicles registered to the name of a commercial activity (later labelled as "commercial") and the age distribution of the owners of the private vehicles resembles the distribution of the vehicles' owner of the fleet in the province. These samples account for 2.642 and 1.870 million trips, representing approximately 15.0 and 20.7 million driven kilometres and exhibiting a mean value of 162 trips and 921 kilometres per vehicle per month in Modena, and 150 trips and 1,656 kilometres per vehicle per month in Firenze. A summary of the datasets is given in Table 1.

Data time-continuity and the ability to fully track each vehicle during the analysed period are key aspects for ensuring the representativeness of the analyses presented in this work. TEMA platform has been designed for handling similar datasets, allowing adapting the developed methodologies to different fleet sample, geographical areas and infrastructural constraints.

2.2. Step 2: exploiting the data potential via a customised data processing platform

The data processing platform is the key tool to exploit the data potential, and its effective customisation is a key conceptual step to extract value from the data. This process entails several technical challenges related to the definition of the data input format, the implementation and customisation of a data processing algorithms, the generalisation and profiling of these algorithms and their interface with visualisation and post-processing toolsets. Aside the complexity of the algorithms themselves, the main challenges to overcome in implementing TEMA have been:

- the structuring of the input data and the definition of the data exchange format among the different modules of the platform;
- the profiling of the data processing algorithms, making them capable to manage and process millions of data lines (i.e. trip/parking events in this specific case) with an acceptable computational burden;
- the post-processing of the results, and, more specifically, the techniques to adopt to render the results of the analyses in a format that does not lose the complexity of the information contained in the data and that is, at the same time, easy to understand;
- providing the link between the data and the potential policy assessment.

This section aims to provide the reader with a simplified overview of the developed modules of TEMA: the pre-processors (i.e. later referred as "Module 0") and five data processing modules (i.e. respectively labelled from "Module 1" to "Module 5"). A schematic overview of the platform is provided in Fig. 1, where the modules are embedded within the dashed boxes. The main conceptual steps are reported, highlighting how each module contributes to deliver one or more results relevant to address specific aspects, applications and policy support.

Module 0 (Pre-processor): in order to prepare the input to be processed by the data processing modules, the raw data must be cleansed, checked and organised in a lean and usable format. For this reason the raw data has been submitted to a cleansing and consistency check procedure, targeted to identify and restore trips which eventually show non-consistent data series and generic errors of acquisition (e.g. trips not starting with an "engine switch-on" status and/or not ending with an "engine switch-off" status). After this step, the cleansed data is submitted to a data aggregation procedure, divided in three sub-aggregation steps: sub-aggregation by day, sub-aggregation by week and sub-aggregation by month. This procedure allows deriving three sub-datasets from the original cleansed data, merging the information at different level of detail, and deriving leaner and smaller databases to be used for the different applications. Table 2 reports the sequence of these aggregations steps, starting from the number of records of the raw data on the left up to the final aggregation level on the right. As far as the province of Modena is concerned the data accounts for 16.920 million records, equivalent to 2.642 million trips. This reduces to approximately 397,000 records aggregated by day, 87,000 records aggregated by week and 16,263 records aggregated by month (i.e. total no. of vehicles in the database), decreasing by approximately 80% per each aggregation steps (compression rate, i.e. the data amount reduction between the step i-th and the step $(i-1)$-th). A similar consideration can be made for the database of the province of Firenze, which reduces from 33.361 million records and 1.870 million trips to approximately 292,000 records aggregated by day, 65,700 records aggregated by week and 12,478 records aggregated by month, exhibiting a similar compression rate. The output of this module consists of three cleansed and aggregated datasets that are saved and stored in order to be used by Modules 1-to-5.

Module 1 (Statistical mobility): this module is built to load the aggregated datasets and perform a large number of statistical analyses to investigate diverse aspects of the mobility patterns of the

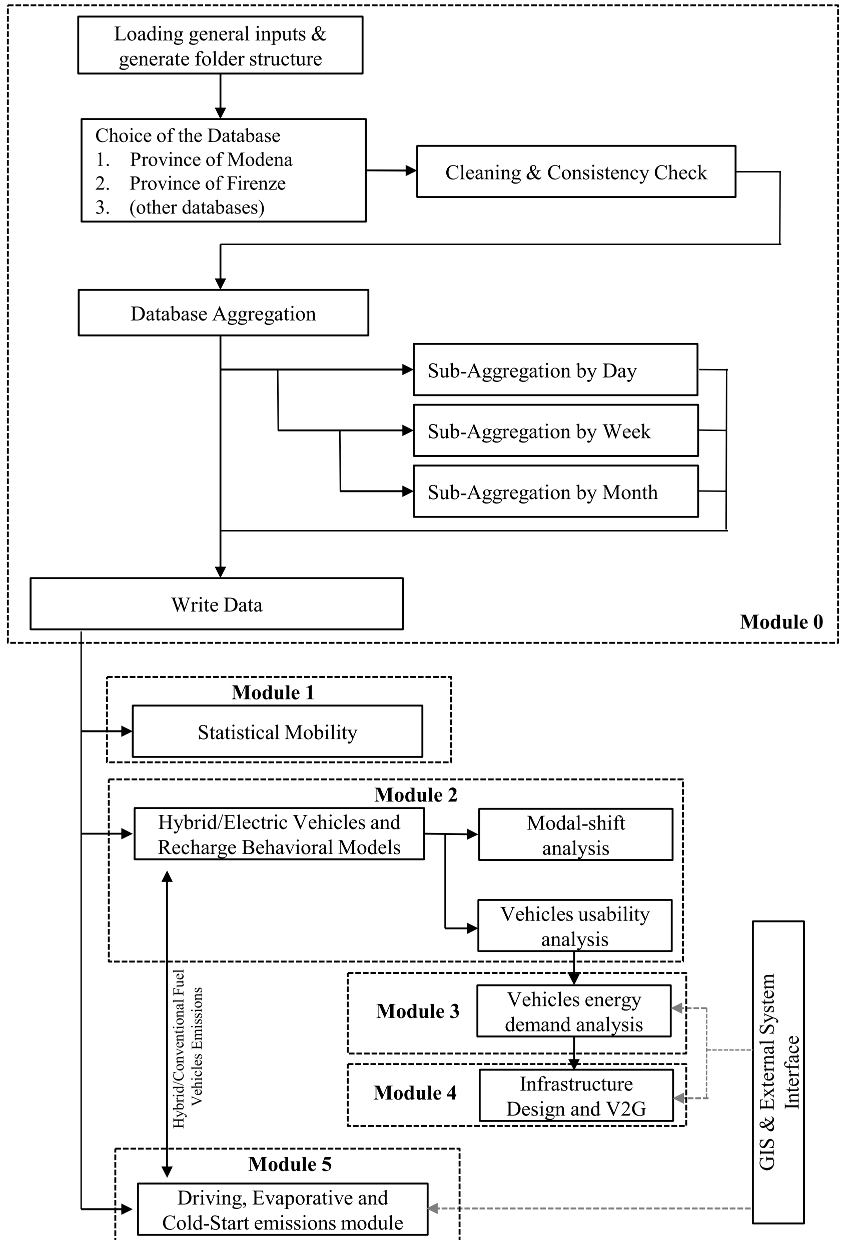


Fig. 1. Overall architecture of the TEMA platform.

Table 2
Aggregation steps and compression rates.

	No. of records [$\cdot 10^6$]	No. of trips [$\cdot 10^6$]	Aggregation by day no. of records [$\cdot 10^6$] (Compression rate [%])	Aggregation by week no. of records [$\cdot 10^6$] (Compression rate [%])	Aggregation by month no. of records [$\cdot 10^6$] (Compression rate [%])
Province of Modena	16.920	2.642	0.397–(−85.0%)	0.0870–(−78.1%)	0.0163–(−81.3%)
Province of Firenze	33.361	1.870	0.292–(−84.4%)	0.0657–(−77.5%)	0.0125–(−81.0%)

monitored fleet sample. The output of Module 1 consists in approximately 250 different statistical analyses and results, including the minima, maxima, mean and median values plus cumulative and probability distribution functions of a number of variables such as trip distance, trip duration, trip speed, parking duration, fleet shares in motion and parked in time, averaged per day and per week. A comprehensive overview of the statistical outputs of this module is provided in [42], showing how large statistics are needed to derive and characterise the activity patterns of the vehicles. The results of this module can contribute to generate additional statistics, complementary to the official statistics based on surveys. Big data is, in fact, a potential source of additional statistics, as indicated by the several working groups which are focusing on this aspect [51,52].

The output of this module can feed the other modules (2 to 5), when specific analysis based on aggregated data is performed, as per [45].

Module 2 (Modal shift and vehicle usability): this module is built to simulate the energy efficiency and environmental performances of different conventional fuel vehicles, HEVs, and BEVs in relation to the activity patterns derived from the data. The recorded sequence of trips and parking events during the analysed period is applied to each of these vehicle technologies to understand to which extend they can replicate the recorded mobility behaviour. The vehicle is characterised by a number of performance parameters (i.e. vehicle mass, averaged real world driving energy consumption and battery size for the BEVs, emission factors for HEVs and conventional fuel vehicles) derived either from literature or from experimental studies from the authors, as per [53–55] and [56]. In case of HEVs/BEVs the vehicle's model assumes to initiate the activity pattern with a fully charged battery, associating each trip to an energy consumption event and each parking to a recharge opportunity, which happens if the recharging constraints are met. In this case, the parking is labelled as a recharge event, with an associated energy demand in a specific time window and in a specific space location. If a trip or part of it cannot be made in electric mode by a HEVs because the energy demand is higher than the energy stored in the battery, the trip is switched to the internal combustion engine mode of the HEVs, thus accounting for driving gaseous emissions according to given emission factors. If this occurs in the case of BEVs the trip is deleted from the sequence of the electric trips for the BEVs, and the vehicle is labelled to fail that specific trip. In the case of conventional fuel vehicles, each trip is associated to a fuel consumption event and the driven segments are associated to driving gaseous emissions according to emission factors related to different emission categories and weighted on the route length. Several BEVs/HEVs vehicles models are considered, ranging from a small hatchback to a large Sport Utility Vehicle (SUV). In particular, the models implemented account for:

- 10 BEVs ranging from a mass of 450 [kg] (i.e. small quadricycle) to 2,100 [kg] (large SUV), with a battery size from 13.0 to 85.0 [kWh] and an energy consumption from 70 to 265 [Wh/km];
- 5 HEVs in plug-in configuration, ranging from a mass of 1,540 to 1,770 [kg], with a battery size from 4.4 to 16.0 [kWh];
- 1 generic HEV/BEV model, for parametric analyses with different battery energy and fuel consumption inputs.

All the considered models are built to simulate the real world performances and energy consumption of HEVs and BEVs technologies already available on the market, including the average consumption of auxiliary systems, the energy losses during the recharges and the battery self-discharge which occurs when the vehicle is parked for long time, thus allowing as realistic predictions as possible. As far as the recharging constraints concern, 15 different recharging strategies are defined according to different individual behaviours of the drivers combined with the power of the different recharging infrastructures which might be available (e.g. non-aggressive/aggressive recharge, on-peak/off-peak recharge, indirect price-based/direct smart-grid recharge, with the mono/tri-phases AC/DC recharge stations at a rated power from 3.3 to 50 [kW]).

This module allows hence to:

- Evaluate the share of trips and kilometres of the databases which might be driven electric, given a HEVs/BEVs model, thus calculating the usability of different vehicle technologies and electrified drivetrain architectures under real world driving conditions;
- Evaluate the fleet share affected by trip failure events and their rate of occurrence, thus calculating the technology-shift and the modal-shift induced by adopting different HEVs/BEVs technologies in urban driving;
- Evaluate the real world road transport emissions of a large-scale fleet of conventional fuel vehicles and HEVs, by considering different fleet shares and fleet mix, as explained in Module 5 section.

Further details on the vehicles and developed recharge models are provided in [42] and [43].

Module 3 (Energy demand): this model is built to process and aggregate in space and time the distributed electric energy demand derived from the Module 2 during the recharges when some specific HEVs/BEVs models are adopted. In practice, by assuming a specific vehicle model, or, eventually, a specific fleet mix characterised by different shares of the considered models, and assuming one or more recharging strategies among the 15 available, it is possible to calculate the sequence of recharging events and the associated energy demand which occurs by replicating the activity patterns of the fleet. These results can be integrated in time and space to visualise the maps of the electric energy demand over the analysed area and the electric power curve over the analysed time period. Further details on this module are provided in [44], while the implementation of its GIS interface, that is a key element to exploit the potential of this module, is described in section 2.3.

Module 4 (Infrastructural design and V2G applications): this model is built to design a recharging infrastructure based on the driving and parking patterns recorded by the data, the electric energy demand calculated by Module 3 and networks of potential locations for installing recharging spots, such as Points of Interests (POIs). This approach allows deriving a detailed map of the recharging infrastructure, or, more in general, of the alternative fuel infrastructure needed to serve the calculated mobility demand. The association of the energy demand events to specific locations is based on a minimum distance criterion, by assuming that the offer-demand match is mainly driven by geographical constraints and that the mobility demand is not influenced by shifting from conventional fuel vehicles to HEVs/BEVs. The model derives a set of performance parameters for each point of recharge and the time-dependent energy demand per POI, thus optimising the infrastructure needed in each location to meet the maximum, mean or median value of the electricity demand, and predicting if, and to which extent, the specific location might be a candidate for installing recharging stations, accounting for profitability of the site and eventual market competitors. Additionally, by matching the offer-demand curve, the module can simulate the implementation of Vehicle-to-Grid (V2G) applications, investigating, for example, how the effect of the recharge/discharge of the batteries of the parked HEVs/BEVs can be controlled to ensure delivering the requested energy to the fleet with the minimum load on the electric energy distribution grid.

This module is also natively interfaced with the GIS, opening up to a number of different applications, such as market analysis for electric energy distribution, optimisation of the alternative fuel infrastructure, simulation of the smart control algorithms for smart-grid and smart-cities under real world constraints. Additionally a simple econometric model is natively embedded within the infrastructure design algorithm to address recharge location profitability and future competition scenarios. At least four of the six pillars of the smart city concept (i.e. smart mobility, smart environment, smart economy and smart governance [57,58]) are addressed with this approach, linking big data and data mining technique to transport, energy and air quality policies. Further details on the module are provided in [46].

Module 5 (Gaseous emissions): this module is built to estimate the driving gaseous emissions (i.e. carbon mono and di-oxides, nitrogen oxides and hydrocarbons) related to the usage patterns of the conventional fuel vehicles or to the trips shares driven with the internal combustion engine of the HEVs fleet. As a matter of example the driving CO_2 emission can be calculated with a simplified procedure based on the fuel consumption, as reported by the

EMEP/CORINAIR Emission Inventory Guidebook – 2007, chapter 7 (road transport) of the European Environment Agency [59,60]. This procedure allows estimating the CO_2 emission of the fleet on the basis of the overall fuel consumption, assuming that the carbon content of the fuel is fully oxidized into CO_2 [61,62], the emission factors and the number of vehicles and kilometres driven per each vehicle category and class. Some corrective coefficients need to be applied to the baseline emission factors because of the different vehicle age, improved fuels and technologies, road gradient and vehicle load, thus accounting for real world driving emissions [63]. The result of this methodology consists in deriving an average fleet-weighted value of emission in [g/km], and calculating time and space distribution of CO_2 sources from road transport over a given geographical area based on real world activity data.

This module implements two additional emission models: the evaporative emission model and a cold-start emissions model. The first model is related to emissions from parked vehicles and is based on the implementation of the hot-soak, permeation and breathing emissions mechanisms for the production and release into the atmosphere of evaporative Volatile Organic Compounds (VOCs), combined with an activated carbon canister semi-empirical model as per [47]. The second model is based on the fact that uncontrolled hydrocarbons emission happens when the catalyst is operated at low temperature, i.e. cold-start of the internal combustion engine, that is, when the vehicle is started after a parking event whose duration sets above a threshold, e.g. above six hours.

2.3. Step 3: the interface with GIS and external systems

In order to enhance the potential of TEMA, the setup of an interface with GIS is of fundamental importance. This allows exploiting the full potential of the implemented models for a more comprehensive interpretation of the calculated geographical distributions and rendering the space dependency of the results. More in detail the GIS interface directly applies to the Modules 3, 4 and 5, being those related to space-dependent results (i.e. space distribution of the energy demand, geo-localisation of the alternative fuel infrastructure and space distribution of driving, evaporative and cold-start emissions, as per Fig. 1). This interface is natively built in the data processing platform, and based on a dynamic link of the results mapping routine with digital mapping systems retrieved from the web [64]. In order to handle the geo-referenced results, each analysed area is embedded in an analysis window, defined by the minima and maxima values of latitude and longitude, appropriately set to include the targeted area. For example the province of Modena is embedded by an analysis window extending from 44.10 to 45.00 [deg] of latitude north and from 10.40 to 11.40 [deg] of longitude east resulting in an area of approximately 7,391 [km^2], whereas the province of Firenze is embedded by an analysis window extending from 43.40 to 44.30 [deg] of latitude north and from 10.60 to 11.80 [deg] of longitude east resulting in an area of approximately 9,631 [km^2]. These windows can be divided in squared terrain tiles of variable size, concentrating the calculated variables in the centroid of each tile for rendering purposes. The smaller the size of the tile, the higher the resolution of the depicted results; however to trade-off the computational burden of the GIS interface with the accuracy of the results' rendering, a terrain tile size of 0.25 [km^2], i.e. 500 [m] per edge, has been considered, resulting in approximately 29,500 tiles for the province of Modena and 38,500 tiles for the province of Firenze. Additionally, the modules can be interfaced with external systems and datasets for additional specific analyses and applications. An example is provided for Module 4: in this case the energy demand calculated from different scenarios of fleet penetration of BEVs has been linked to the networks of POIs in order to derive the recharge infrastructure layout. For this specific application a POI database for commercial GPS applications has been downloaded from the web [65], and the POIs have been organized in classes according to five categories of infrastructures. In total 551 airports and air fields, 28,144 petrol stations, 1,507 shopping malls, 4,688 car parking lots and 700 bus parking lots have been included, accounting in total for 35,590 POIs distributed all over Italy. The databases has been then filtered to derive POIs sub-sets related to the analysed provinces, resulting in 423 POIs in the province of Modena and 632 POIs in the province of Firenze, to be then interfaced with the energy demand results and with GIS for rendering purposes. The POIs datasets is only an example of how other systems and datasets can be embedded within the developed platform, to enhance its computational capabilities and applications.

3. Results

To illustrate the capability of TEMA, the results obtained applying its various modules to the two datasets referring to the province of Modena and Firenze are here below reported.

3.1. Urban mobility: key-results from large-scale statistics

The urban mobility is investigated through large scale statistics via the Module 1 described in section 2.2. The results for the two provinces show that the share of the fleet that is in motion at the same time (averaged over the four weeks of the analysed period, i.e. May 2011) never exceeds 11.72% in Modena and 10.36% in Firenze, with a mean value of 4.29% in Modena and 4.47% in Firenze. By looking at the weekly averaged time dependant values, the share of the driving fleet exhibits three peaks during the working days, i.e. from Monday to Friday, approximately at 7.30, at 12.00 and at 18.30, and two peaks during the weekend, i.e. Saturday and Sunday, approximately at 12.00 and at 19.00. By averaging this data on a daily-basis rather than on a weekly-basis, the averaged parked fleet share is most of the time above 90%, reaching a value above 99% from 1.00 and 5.00 in the morning, when almost all the vehicles are parked. These results suggests that, by shifting the conventional fuel vehicles to HEVs/BEVs, there is a large possibility to recharge, being the vehicles parked for most of the time, supporting the assumption of a single charging event per day, probably overnight when the maximum number of vehicles are parked. By distinguishing between private vehicles and commercial vehicles substantial differences in the driving patterns arise. In particular the commercial fleet share in motion increases up to 15.23% in Modena and 16.57% in Firenze, suggesting a more intense activity of commercial vehicles compared to that of the private vehicles during the working days. On the other hand the activity of the commercial vehicles is slightly below that of the private vehicles during the weekend.

General statistics on the urban fleet show that the time-averaged trip has a length between 5 and 20 [km], a trip duration between 10 and 20 minutes, a trip speed between 25 and 40 [km/h] and parking duration between 2 and 12 hours, daily and nightly values respectively. These results highlight how the urban mobility demand is more fragmented during the day than during the night. Additionally urban driving speed reduces of approximately 60% during the day with respect to the night value due to the increase of the road congestion. By looking at the cumulative and probability distribution results approximately 50% of the trips have a driving length below 3.5 [km], a duration below 8 minutes and an average speed below 25 [km/h], while 90% of the trips are below 20 [km], lasting less than 30 minutes, with an average speed below 50 [km/h], with very similar results for both provinces. As far as parking duration is concerned, 50% of the parking events lasts less than 50 minutes (i.e. eventually suitable for quick recharges), and 90% of the events take less than 700 minutes

($\sim$11.5 hours). Additionally half of the vehicles in the sample make less than 6 trips and 20 [km/day] and 30 trips and 200 [km/week], being parked for more than 90% of the time. Approximately 78% of the vehicles in the sample travel up to 50 [km/day] and approximately 9% of the vehicles in the sample exceed 100 [km/day], reducing to 3% exceeding 150 [km/day]. The practical implication of these results is that approximately 7 out of 10 among the urban vehicles exhibiting a predominant urban usage pattern never show a trip length above 100 [km], a value compatible with the driving range of most of the BEVs already available on the market, implying that they can be targeted for early adoption of BEVs. These results provide a detailed picture of the urban mobility in the analysed areas, giving back a scenario characterised by short and frequent trips alternated to numerous parking events lasting for 2–4 hours in the day and for 8–12 hours in the night. This suggests the possibility to have a large widespread of the current HEVs/BEVs technologies as well as possible synergies between fast recharge during the day and slow recharge during the night. A detailed description of the statistical results of the urban mobility, with all the relevant graphs and tables, can be found in [42]. As mentioned already above, this approach represents a new insight towards official statistic from big data analysis.

3.2. Potential of electric vehicles in replacing conventional fuel vehicles and modal-shift

The results of this analysis has the objective to derive the performances of different conventional fuel vehicles, HEVs, and BEVs following the activity patterns calculated from the data, to quantify the potential of new low-carbon vehicle technologies in reducing overall driving emissions and replacing conventional fuel vehicles in urban areas. The results of this analysis strongly depends on the assumptions made, in terms of vehicle model adopted, i.e. energy consumption and battery capacity coupled with recharging constraints and user's behaviour, i.e. non-aggressive/aggressive recharge, on-peak/off-peak recharge, indirect price-based/direct smart-grid recharge. By considering a small-to-medium sized BEVs (battery size between 16.0 and 24.0 [kWh] and energy consumption between 180 and 210 [Wh/km]), regardless of the recharging strategy, the vast majority of the trips of the two databases, i.e. more than 80%, can be driven electric. This is a result derived at a fleet level, that is by considering the trips of the fleet as events not correlated one to each other, whereas by looking at the results at a vehicle level, approximately 10% to 25% of the fleet is capable to drive only electric, depending on the recharge strategy, meaning that the trip sequence of the vehicle is never interrupted by trips failure events. These results lead to two conclusions:

- most of the real world urban mobility demand can be served by BEVs, thus enabling a shift from conventional fuel vehicles to BEVs of about four-fifths of the urban mobility;
- a non-negligible share of the fleet, i.e. approximately one-fifth, would not suffer any range limitation as consequence of this shift.

The remaining four-fifths of the fleet are instead subject to the modal-shift to cover a mobility demand which is not compatible with BEVs, although this modal-shift results to be rather limited. In fact it is estimated to be below one trip per week by increasing the BEVs fleet share from 25% to 50%, between one and two trips per week by increasing the BEVs fleet share to approximately 65% and between two and five trips per week by increasing the BEVs fleet share to approximately 80%. These results clearly highlight how, given the real world mobility demand, the target value of the EC White Paper for electric urban mobility, i.e. 50% of urban mobility shifted to HEVs/BEVs by 2030, can be already reached with the current generation of BEVs with a minimum modal-shift, thus implementing better synergies between urban on-road transport and other modes of transport, e.g. public transport. In other words this implies that ambitious goals for electric vehicles in urban areas can be realistically achieved only by better exploiting the potential of the current BEVs technologies, and that the shift of urban mobility towards low-carbon systems mostly depend on overcoming cost-scalability barriers and customers' acceptance of BEVs. All the details of this analysis are extensively presented in [42] and [43].

3.3. Energy demand from electric vehicles in urban areas

Based on the results presented in the paragraph above, the electric energy demand in time and space from different shares of BEVs in the real world mobility can be derived. As per the results described in section 3.2, also these results strongly depend on the assumptions made on the vehicle model and on the recharge strategy. However, by assuming the small-to-medium sized BEVs as above, shifting a fleet share going from 10-to-25% from conventional fuel vehicles to BEVs implies an electric energy demand increase on the province from 0.4% to 5.1% (value scaled up to the province fleet size and referred to the total electric energy consumption of the province, as per recorded data [66]). By considering that the domestic sector accounts for approximately 18% of the total electric energy demand in the province of Modena and 26% of the total electric energy demand of the province of Firenze, the increase of the domestic demand, thus considering the BEVs as additional domestic devices, ranges from 2.2% to 19.6%. Time dependent energy demand is strongly depend on the recharge strategy adopted, and therefore, on the time constraints imposed by the strategy itself. The strategies that allow the recharge all over the 24 hours, imposing constraints only on the parking duration, provide time dependent energy demand curves that practically resemble the driving and parking patterns of the fleet, reproducing the three activity peaks in the working days and the two peaks in the weekend, as highlighted in section 3.1. On the other hand, the strategies that impose specific time constraints, i.e. overnight recharge or smart recharge controlled by a time window synchronised with the minimum energy demand of the province, provide time-bounded energy demand capable to fill the valley of the electric energy demand profile, and suggesting potential for synergies between domestic/industrial energy demand and BEVs fleet. The power level also varies, depending on the recharging strategies and on the recharge power set by the infrastructure (assumed to be 3.3 [kW] of mono-phase AC infrastructure, 10 [kW] for the tri-phases AC infrastructure and 50 [kW] for the DC infrastructure, and scaled down to 2, 9 and 40 [kW] to account for the effect of the recharge profile), ranging from 0.4 [MW] to more than 4 [MW] for the different analysed cases [43]. By analysing the spatial distributions of the recharging events, it can be derived that, being the fleet sample characterised by vehicles with a predominant local driving patterns, the recharge demand is clustered around urbanised areas within the provinces, with one-fourth of the events taking place in a circle of 5 [km] of radius around the main city (i.e. Modena and Firenze, respectively), to increase to approximately two-thirds by considering a circle of 15 [km] of radius. Some differences between the provinces can be observed, depending on the diversities of their urban planning, as well as on a clustering of the events around minor urbanised areas, involving Carpi, Mirandola, Sassuolo and Vignola for the province of Modena, Empoli and Prato for the province of Firenze. By considering the space dependant results, the spatial energy demand can be either depicted with geo-referenced contours plots, thus deriving values ranging from approximately 200 [kWh/km^2/day] to 1,000 [kWh/km^2/day], or based on the demand per terrain tile representation, given the

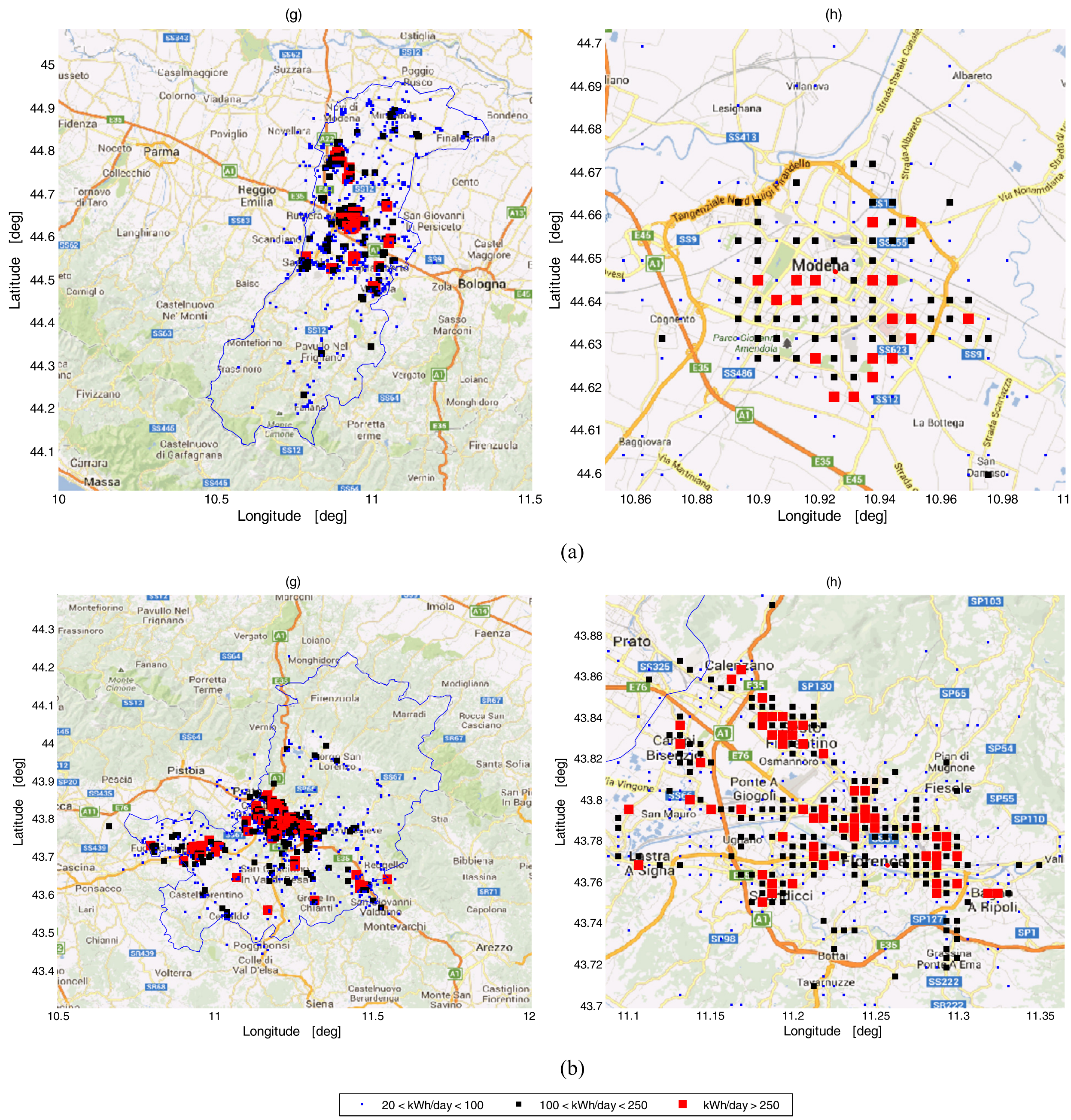
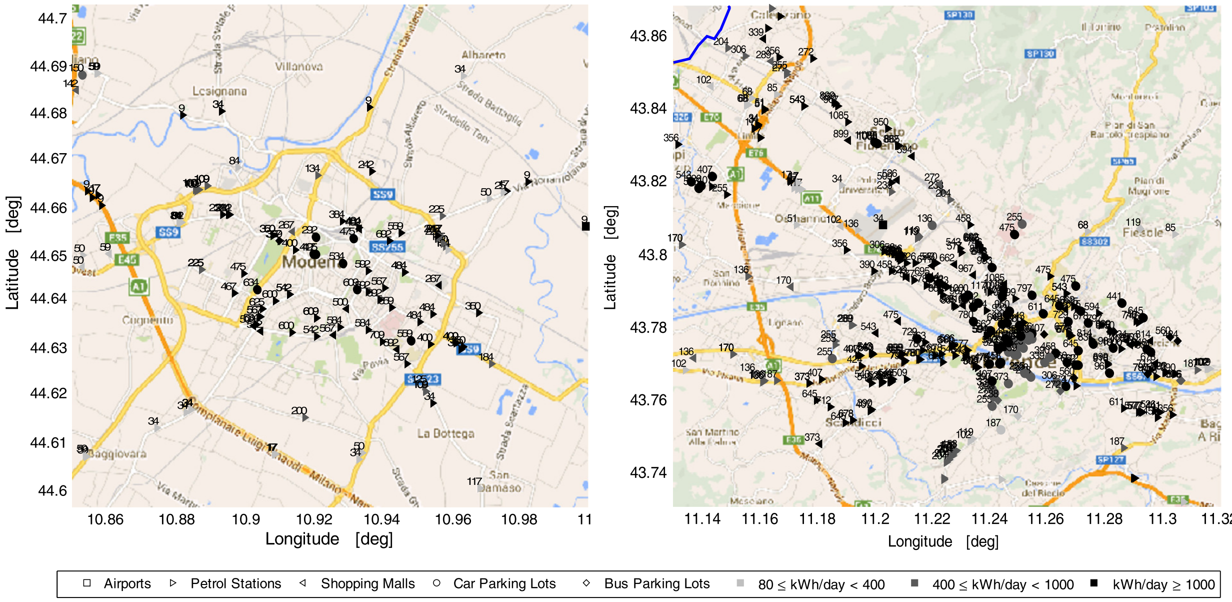


Fig. 2. Extract from the geo-referenced energy demand results, province of Modena (a) and province of Firenze (b), whole province area (left, blue line indicates the province border) and zoom on the city area (right). These results refer to a medium-sized BEVs (i.e. battery size of 24.0 [kWh] and averaged energy consumption of 210 [Wh/km]) coupled with a smart recharge strategy (i.e. recharge allowed only in a time window of 4 hours, i.e. ±2 hours, around the minimum of the electric energy demand, at the power of 2 [kW]), referring to a conventional fuel o BEV fleet shift of nearly 20%. The results are reported in integral form in [44]. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

geographic representation described in section 2.3. Knowing the energy demand per load categories, it is possible to identifying demand hot-spots to enable the infrastructure design presented in section 3.4. By assuming three ranges of energy demand, i.e. averaged energy demand between 20 and 100 [kWh] per day, between 100 and 250 [kWh] and above 250 [kWh], and associating to these ranges small-sized black squares, medium-sized blue squares and big-sized red squares, it is possible to derive a visual energy demand map as reported in Fig. 2, where an extract of the results of the geo-referenced energy demand per tile for the provinces of Modena and Firenze is reported, for a medium-sized vehicle and a smart recharge strategy (conventional fuel to BEV fleet shift of nearly 20%). Note that this result is extracted among the several scenarios considered and reported in integral form in [44]. This representation provides a quick overview of where the demand spots are located, quantifying averaged daily values of energy requested by the BEVs fleet as well as defining time dependency per different locations. It also enables to draft a detailed picture of the potential energy supply market which might develop as result of shifting from conventional fuel vehicles to BEVs, as well as estimating the potential for business opportunities, profitability of public and private investments and shift from oil to energy market, with consequent reduction of driving emissions and GHGs.

3.4. Recharge infrastructure design and V2G results

Based on the energy demand maps presented in section 3.3, the following step is to derive a customer-driven recharge infrastruc-



(a) (b)

Fig. 3. Extract from the geo-referenced customer driven infrastructure results, province of Modena (a) and province of Firenze (b), zoom on the city area. The results are derived under the same assumptions of Fig. 2, in order to enable a visual comparison of the electric energy demand-offer match. The results are reported in integral form in [46].

ture which relies on optimal energy demand-offer match, calculating how many recharging spots are needed and where they must be installed in order to serve the BEVs with the requested energy delivery service. As already mentioned in the description of the Module 4, the recharge infrastructure design relies on networks of potential locations where to install recharging spots, i.e. POIs networks built for commercial GPS applications. Different classes of POIs have been considered, and the electric energy demand-offer match algorithm is driven by geographical constraints, assuming that the energy demand calculated by the recharging model can be served by all neighbourhood locations within a specified area, within a distance of 1 [km] from the parking place. The potential demand can be either not met by any POI, thus constituting a match failure, or transferred to two or more POIs, if the demand event is happening in an area dense of potential suppliers, suggesting market competition between different suppliers. As result of such algorithm, the model predicts a Geographic Key Performance Indicator ($\mathrm{Geo_{KPI}}$), representative of the capability of a specific POI to be as close as possible to the demand, and a Repetitiveness index (R), representative of the rate of occurrence of the recharges of the same vehicle at the same POIs, characterising the potential customers as recurring or occasional customers. The set of parameters calculated by the model is finally complemented by the daily averaged energy demand profile per POI, and, therefore by the sizing of the infrastructure based on the average number of recharging spots required. The results show that the $\mathrm{Geo_{KPI}}$ assumes mid values in the densely populated areas, increasing to higher values in some isolated spots in rural areas outside the main districts, representing the fact that more recharging locations might serve the same demand in the city areas, hence offering more choices to the customers than in isolated areas. Similarly the R index assumes mid values in the city areas, whereas it tends to increase to higher values around isolated spots, resembling the fact that a large customers pool might be served in densely populated areas rather than in rural areas. An example of the recharge infrastructure layout is reported in Fig. 3, based on the same assumptions of medium-sized vehicle and a smart recharge strategy as per Fig. 2. By comparing these pictures it is possible to figure out how the demand-offer matching algorithm works, and how it allows deriving the geo-referenced layout of the customer driven recharge infrastructure. Note that also this result has been extracted among the several scenarios considered and reported in detail in [46]. In terms of global number of installed recharging spots, the results depends on the assumptions made; however it is noticed that for all the considered cases approximately 95% of the POIs in both provinces are suitable locations for installing recharging spots, with a number of plugs which is typically two-to-four times higher the circulating BEVs and an energy offer which is four-to-eight times higher than the demand. It is important to highlight that the model predicts the layout of the infrastructure under the hypothesis of large-scale deployment of BEVs, therefore the number of recharging spots results to be compatible with a fully developed infrastructure, constituting a long-term scenario. However, the insights offered by this approach, as well as the possibility to isolate locations which might work as energy demand hubs to drive the early deployment of the infrastructure itself allow to use the present results also for short-to-mid-term scenarios.

On top of this, a V2G model has been implemented, based on the assumption that each parking event which is not associated to a recharge, not constituting an energy demand, can be used as a potential energy offer event, during which the parked vehicle can release a small amount of the energy stored in its battery back to the grid, i.e. 2% of its nominal capacity, to serve the neighbour vehicles which are charging. The aim of this study is to understand whether the parked BEVs fleet which is not recharging can be used as a distributed energy storage system, filling the demand of those vehicles that are charging and reduce the load on specific locations characterised by high demand. In spite of the simplicity of this algorithm, the results show that the electric energy load offered by the parked vehicles is a substantial share of the demand, which basically depends on the different recharge strategies adopted. The results suggest that the energy offer ranges from 30% to 50% of the energy demand for AC recharges, decreasing from 10% to 30% for the DC recharges. Time dependant analysis show that the V2G can shave peak demands from 25% to 50% in specific hot-spots, suggesting a good potential of such application and opening-up to more refined and better controlled algorithm to further improve the performances of future V2G applications. The details of these results, including an analysis for several scenarios, are also reported in full in [46].

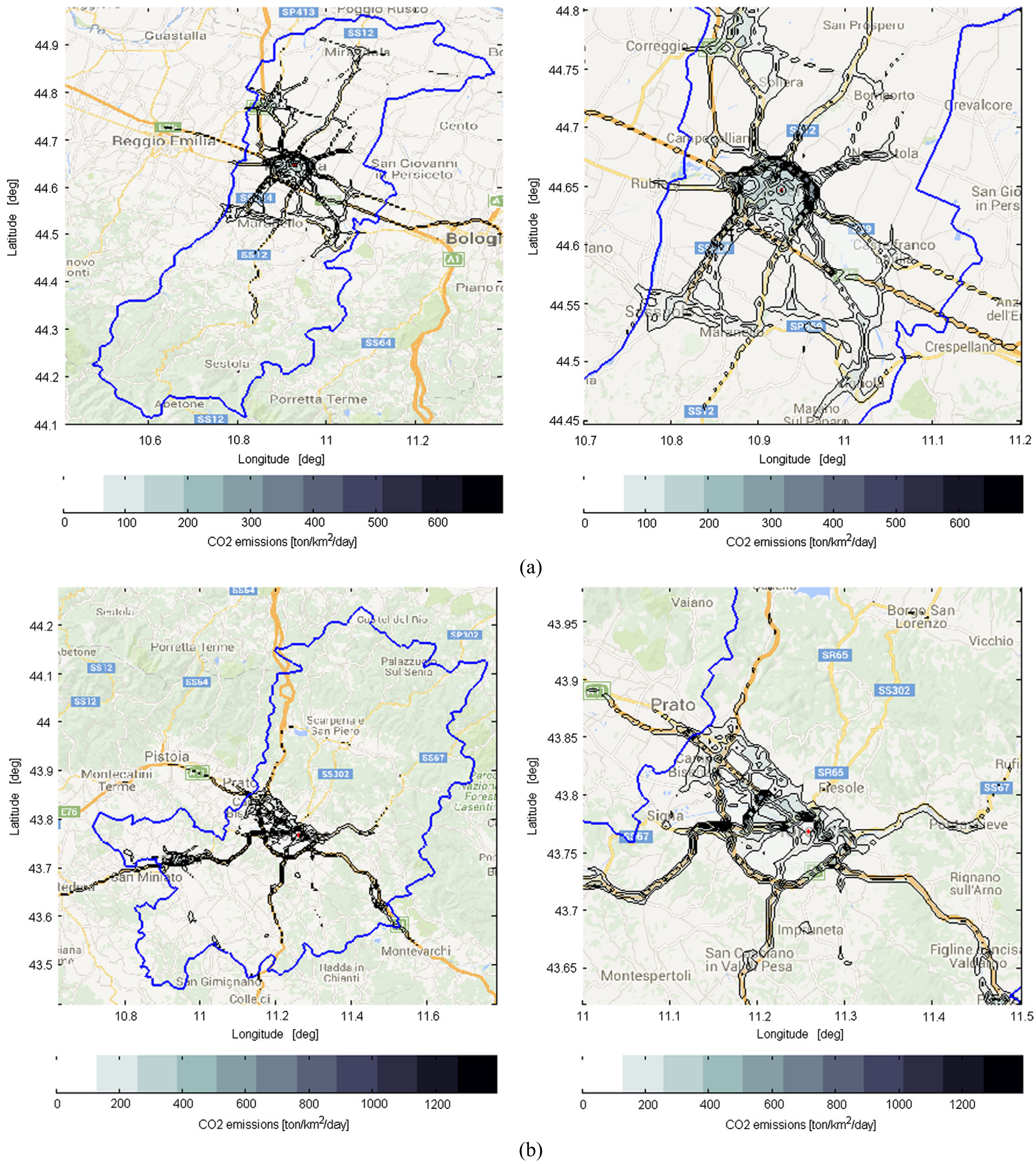


Fig. 4. Geo-referenced CO_2 real-world driving emissions, province of Modena (a) and province of Firenze (b), whole province area (left, blue line indicates the province border) and zoom on the city area (right). The results are calculated from the fuel consumption records of the year 2008 [67] and emission factors from [59,60] estimating 150.6 [g/km] for the passenger cars in the province of Modena and 162.8 [g/km] in the province of Firenze. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. Driving and evaporative emissions results

As described in section 2.2, Module 5 is built to estimate the geo-referenced gaseous emissions based on the real world patterns of vehicles, under different assumptions of the fleet composition. As explained above, three types of emission sources are considered: driving emissions, related to the operation of the internal combustion engine during driving phases, cold-start emissions, related to the release of hydrocarbons happening when the vehicle is operated with cold-catalyst, and evaporative emissions, related to the VOCs emitted when the vehicle is parked. As far as the latter emission source is concerned, this is made of three main components: fuel vapours emitted by the fuel injection system immediately after the combustion engine is switched off (i.e. hot-soak losses), fuel vapours permeating through the plastic material of the fuel tank system (i.e. permeation losses), and the fuel vapours generated in the tank which are not adsorbed by the activated carbon canister (i.e. breathing losses). All these emission sources, handled by a single simulation platform, contribute to an overview of the global atmospheric impact of on-road vehicles as derived from their real world driving and parking patterns, allowing con-

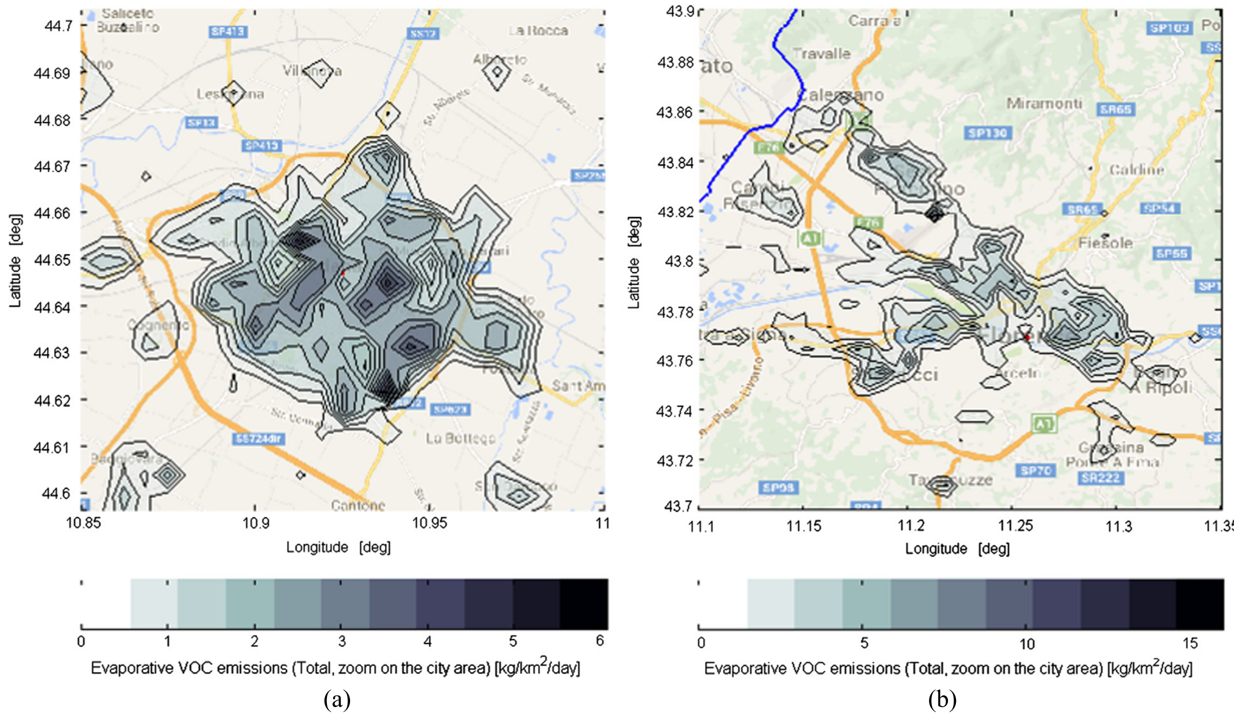


Fig. 5. Extract from the geo-referenced evaporative VOCs emissions results, province of Modena (a) and province of Firenze (b), zoom on the city area. The results are derived by considering the worst emission scenario, i.e. month of July, tank headspace volume of 40 [l], activated carbon mass of 100 [g] and desorption flow rate of [100 l/h], i.e. scenario 28 according to results reported in [47].

sidering the global effect of different fleet scenarios and diverse shares of HEVs and BEVs in the urban fleet.

As far as driving emissions concerns, preliminary results on real world CO_2 emissions are reported in Fig. 4 in the areas of Modena and Firenze. The average fleet share weighted value of CO_2 grams per km have been calculated by considering the yearly fuel consumption statistics in Italy for each fuel type [67], the emission factors as per [59], as explained in section 2.2, the number of vehicles for each vehicle category and the related total km driven per year for each vehicle category as per [68,63]. The classification of vehicles according to their emission control technologies is made on the basis of the legislation they comply with [68]. The calculated average CO_2 emissions in [g/km] for each vehicle type are then scaled up to the province fleet size considering the province fleet share [68]. A weighted fleet share value of CO_2 [g/km] is hence derived to be geo-referenced for the given province area. The results reported in Fig. 4 refer only to the passenger cars (both gasoline and diesel), as per the vehicles in the databases. The simplified assessment foresees also light and heavy duty vehicles, buses, mopeds and motorcycles, not considered in this example. Applying this methodology for the year 2008, and assuming that the driving pattern was the same as that recorded for 2011, a weighted passenger cars fleet emission factor of CO_2 equal to 150.6, [g/km] for the province of Modena and 162.8, [g/km] for the province of Firenze is derived. These values are in line with values reported in literature [69]. These CO_2 emission factors have been then distributed over the geo-referenced driving patterns, scaled up to the surveyed urban fleet of the province, representative of the 30.7% of the total fleet, integrated in time over the analysis period of one month and averaged per day. Fig. 4 shows how CO_2 sources distribute over the major roads in the province, with a peak value beyond 600 [t/km^2/day] in Modena and beyond 1,200 [t/km^2/day] in Firenze. The average values calculated on the roads result to be, in both cases, approximately one-third of the peak values. A similar estimation can be done per other pollutants too, such as carbon monoxide, nitrogen oxide, hydrocarbons and particulate matter.

As far as the evaporative emissions concerns, the work has been developed to support a more effective type approval procedure, according to the article 4 of the regulation (EC) No. 715/2007 [70] and the communication 2008/C 182/08 [71]. The results highlight that the evaporative emissions control system currently fitted into most of the European passenger cars might not effectively work under real world usage condition, especially at low European latitudes (i.e. Mediterranean countries such as Spain, Italy and Greece), and during late spring and summer (i.e. from May to September). In fact the thermal cycle imposed by the average daily temperature at such latitudes, combined with the fragmentation of the trip and parking events, as highlighted by the urban mobility analysis, do not allow an effective desorption of the fuel vapours from the activated carbon canister during the driving phase, resulting in a substantial loss of effectiveness of the emission control system and leading to uncontrolled VOCs emissions significantly above the limit imposed by the current type approval regulation (i.e. 2 [g/day]). The results show that only 2.5% of the of real world trips are driven in a condition which is comparable to the current European type-approval test procedure, and that more than 80% of the evaporative emissions events potentially exceed the emission limit. The emission peak value is estimated to be approximately 4 [g/day] in May and 8 [g/day] in July while the time-dependent results show emission rates up to nearly 15 [g/s] in the province of Modena and 30 [g/s] in the province of Firenze for the gasoline urban fleet, with a cumulative value up to 0.4 and 0.8 tons of VOCs per day in July, respectively in the two provinces. The space-dependent results show a value of the emissions in July of approximately 4-to-8 [kg/km^2/day]. Fig. 5 reports an extract of the geo-referenced evaporative VOCs emissions distribution reported in [47], in the city area of Modena and Firenze. These results refer to the condition characterised by the highest emissions among the 36 scenarios considered, i.e. scenario no. 28, month of July, tank headspace volume of 40 [l], activated carbon mass of 100 [g] and desorption flow rate of [100 l/h]. Such results both confirm the need of revising the current normative designing a type approval test which is more representative of real world conditions as well as the possibility of reverse-engineering the emissions control system to keep the VOCs emissions below the type approval limit, given the real world use conditions. The results also suggest that combining an activated carbon canister mass with an increased mass of activated carbon, i.e. 400 [g], with a canister desorption volume flow rate of 400 [l/h] might significantly reduce real world evaporative emissions. Note that the vehicles certified under the current type approval test for evaporative emissions are equipped with a carbon canister with a mass of activated carbon ranging from 150 to 250 [g] and a desorption flow rate of approximately 200 [l/h]. More details on the topic and a comprehensive scenario analysis, with the details of the models implemented, are provided in [45] and [47].

3.6. Generalisation of the results

The results presented in the previous sections show the different capabilities of the TEMA platform, both in managing and analysing large sets of data as well as in supporting policy assessment, design, and verification, identifying non-obvious relations among large datasets. The generalisation of these results is a key aspect in big data studies. The present work illustrates the results obtained considering data from two Italian provinces, and, given the representativeness of the sample discussed in section 2.1, results are certainly valid at a regional level in the analysed areas. Additional datasets will be analysed referring to other European countries both to extend the analysis as well as to highlight the differences in the mobility behaviour across Europe.

To assess the potential of the GPS data in contributing to mobility studies, the results for the two analysed provinces have been compared with relevant statistics from other European countries and from USA, as per [42]. The comparison shows a homogeneity of the results in EU, and a substantial difference with USA, where people move in a different way, given the larger distances that characterise US urban areas. Another aspect that must be taken into account is the fact that the statistical data is in general acquired with different means (i.e. geo-coded Computer-Assisted Telephone Interviews (CATI), face to face or written survey), in respect to the data of the present paper, acquired via GPS, resulting in more accurate information, not affected by personal interpretation as it can occur with survey interviews. The survey data can be easily biased by personal evaluations errors and short journeys, viewed as insignificant by interviewed persons, can be omitted. This is likely to decrease the number of trips per day as well as increase average trip distances being long trips overrepresented and short trips underrepresented. The results presented in this work are representative of average medium-sized urban areas, and they can presumably be extended to most of the European cities. They cannot be extended to areas characterised by significantly larger population (e.g. large EU cities like Paris or London), as well as to US and Chinese megalopolis, where mobility demand and infrastructural networks are significantly different than those represented in this work.

4. Conclusions and future applications

This paper presents the results of the data processing platform TEMA designed for supporting EU transport policies assessment via big data. The platform is natively conceived for multi-purpose applications, and it is made of 6 modules: pre-processor, statistical mobility module, modal shift and vehicle usability module, energy demand module, infrastructure design module, V2G applications and gaseous emission module.

TEMA has the objective to investigate the activity patterns of large scale samples of conventional fuel vehicles or GPS data to understand the mobility in urbanised areas, evaluating the potential of deploying low-carbon vehicles, i.e. HEVs and BEVs, and deriving the modal-shift introduced by these technologies. Their impact in term of energy demand on the electric energy distribution grid is also addressed, together with the smart design of the infrastructure and V2G solutions. Additionally a gaseous emissions model is also implemented, to evaluate the driving, cold-start and evaporative emissions of the current conventional fuel vehicles fleet, as well as the decrease of the emissions introduced by shifting shares of this fleet to HEVs and BEVs.

TEMA has been already used to support the revision of the type approval procedure for evaporative emissions test, subsequent to the article 4 of the regulation (EC) No. 715/2007 [70] and the communication 2008/C 182/08 [71] and for supporting the directive on the deployment for alternative fuel infrastructure in EU [72] and technical guidelines for eco-innovation [73].

This work shows the effective development and implemantation of a methodology capable to handling large amount of data, identifying non-obvious relations among it and performing customised analyses. The presented application consists in processing data from two large GPS databases from urban vehicles, to assess the impact of low-carbon road transport technologies and policies. The first level is the analysis of the urban mobility that results to be fragmented in a high number of short trips with an average trip distance of approximately 5 [km] and trip duration of 10 minutes associated to a mean parking duration of approximately 2 hours during the day, versus a trip distance of 20 [km], a trip duration 20 minutes and parking duration of 12 hours during the night. Half of the vehicles in the sample make less than 6 trips and 20 [km/day] and 30 trips and 200 [km/week], being parked for more than 90% of the time. Approximately 78% of the vehicles in the sample travel up to 50 [km/day] and approximately 9% of the vehicles in the sample exceeds 100 [km/day], reducing to 3% exceeding 150 [km/day], a value that is compatible with the driving range of most of the BEVs available on the market. These results suggest a large potential for deploying BEVs in cities without significantly affecting the real world activity patterns, together with the possibility to establish synergies between fast recharge during the day and slow recharge during the night. In numbers, approximately from 10% to 25% of the urban fleet can be targeted for an early deployment of BEVs, being its activity fully compatible with the BEVs constraints, increasing to approximately 50%, i.e. the 2030 EU target value as per [1], by accepting a rather limited modal shift up to one trip per week. The electric energy demand increase derived from this shift ranges from 0.4% to 5.1% of the total electric energy demand in the analysed provinces, i.e. from 2.2% to 19.6% of the domestic electric energy demand, with a power demand ranging from 0.4 [MW] to more than 4 [MW] depending on the infrastructure, i.e. AC and/or DC, and on the assumptions of the model. This results show how a large set of GPS data can support mobility studies and the assessment of innovative transport technology deployment in urban areas. Additionally to this assessment, a smart recharge infrastructure design algorithm has been implemented. It derives a number of recharging spots from two-to-four times higher than the circulating BEVs with the possibility to implement V2G applications using BEVs fleet as a system of distributed energy storage devices for decreasing the energy demand in recharging hubs up to a value ranging from 30% to 50% for AC recharges, and from 10% to 30% for the DC recharges.

The results from the emissions model show how driving and evaporative emissions from conventional fuel vehicles can be quantified and geo-referenced based on real world driving data, and therefore how such models can be used to understand how, and to which extent, the deployment of low-carbon vehicles can effectively decrease pollutant emission and GHGs in urban areas, under different scenarios and constraints.

Together with these results this paper aims to underline how challenging is the development of big data applications, but also the opportunities offered by this approach, providing unprecedented insights on urban mobility, simulating, with detail never reached before, the consequences of deploying low-carbon vehicle technologies on large scale based with real world activity patterns, geo-referencing the impact of the oil-to-electricity energy demand shift, suggesting the layout of a customised alternative fuel infrastructures and the development of V2G applications for future smart grids and smart cities.

TEMA is natively conceived to operate on a regional/national scale, and the authors foresee extending the analysis to several European regions, with the objective of addressing low-carbon mobility policies on a continental scale. The results achieved and the potential offered by big data suggest how this discipline can substantially increase the effectiveness of future policies in the field of transport and energy, ultimately changing the policy making development processes and promoting the active participation of the citizens to data collection campaigns for shaping the Europe of tomorrow.

References

[1] European Commission, White Paper, Roadmap to a single European transportation area – towards a competitive and resource efficient transport system, 2011.

[2] European Commission, Communication from the Commission of 25 February 2015 on Energy Union Package – a framework strategy for a resilient energy union with a forward-looking climate change policy COM (2015) 80, 2015 [online].

[3] J. Zhou, E. Murphy, Y. Long, Commuting efficiency in the Beijing metropolitan area: an exploration combining smartcard and travel survey data, J. Transp. Geogr. 41 (2014) 175–183.

[4] S. Tao, J. Corcoran, I. Mateo-Babiano, D. Rohde, Exploring bus rapid transit passenger travel behaviour using big data, Appl. Geogr. 53 (2014) 90–104.

[5] S. Galland, L. Knapen, A. Yasar, N. Gaud, D. Janssens, O. Lamotte, A. Koukam, G. Wets, Multi-agent simulation of individual mobility behavior in carpooling, Transp. Res., Part C, Emerg. Technol. 45 (2014) 83–98.

[6] X. Liu, W.Y. Yan, J.Y. Chow, Time-geographic relationships between vector fields of activity patterns and transport systems, J. Transp. Geogr. 42 (2015) 22–33.

[7] S. Hasan, S.V. Ukkusuri, Urban activity pattern classification using topic models from online geo-location data, Transp. Res., Part C, Emerg. Technol. 44 (2014) 363–381.

[8] F.A. Marcianó, G. Musolino, A. Vitetta, Signal setting optimization on urban road transport networks: the case of emergency evacuation, Saf. Sci. 72 (2015) 209–220.

[9] Q. Xu, Z. Zu, Z. Xu, W. Zhang, T. Zheng, Space P-based empirical research on public transport complex networks in 330 cities of China, J. Transp. Syst. Eng. Inf. Technol. 13 (1) (2013) 193–198.

[10] E. Ortega, I. Otero, S. Mancebo, TITIM GIS-tool: a GIS-based decision support system for measuring the territorial impact of transport infrastructures, Expert Syst. Appl. 41 (16) (2014) 7641–7652.

[11] M. Zhang, B. Wiegmans, L. Tavasszy, Optimization of multimodal networks including environmental costs: a model and findings for transport policy, Comput. Ind. 64 (2) (2013) 136–145.

[12] EU-FP7 REDUCTION Project, Cordis webpage: reducing environmental footprint based on multi-modal fleet management system for eco-routing and driver behaviour adaptation [online], available: http://cordis.europa.eu/project/rcn/100009_en.html [accessed March 2016].

[13] EU-FP7 PEACOX Project, Cordis webpage: persuasive advisor for CO_2-reducing cross-modal trip planning [online], available: http://cordis.europa.eu/project/rcn/100231_en.html [accessed March 2016].

[14] EU-FP7 eCOMPASS Project, Cordis webpage: eCO-friendly urban multi-modal route planning services for mobile users [online], available: http://cordis.europa.eu/project/rcn/100714_en.html [accessed March 2016].

[15] EU-FP7 ICT-EMISSIONS Project, Development of a methodology and tool to evaluate the impact of ICT measures on road transport emissions [online], available: http://cordis.europa.eu/project/rcn/100769_en.html [accessed March 2016].

[16] Forbes.com, Gartner: top 10 strategic technology trends for 2013 [online], available: http://www.forbes.com/sites/ericsavitz/2012/10/23/gartner-top-10-strategic-technology-trends-for-2013/ [accessed December 2014].

[17] Forbes.com, Gartner: 10 critical tech trends for the next five years [online], available: http://www.forbes.com/sites/ericsavitz/2012/10/22/gartner-10-critical-tech-trends-for-the-next-five-years/ [accessed December 2014].

[18] J.W. Tukey, The Future of Data Analysis, 1962.

[19] European Commission Website [online], available: http://ec.europa.eu/clima/policies/transport/vehicles/index_en.htm [accessed November 2015].

[20] European Environment Agency (EEA), Kyoto burden-sharing targets for EU-15 countries, 2005.

[21] United Nations Framework Convention on Climate Change, Kyoto Protocol to the United Nations Framework Convention on Climate Change, 2011.

[22] European Commission Joint Research Centre, EDGAR home page: emission database for global atmospheric research [online], available: http://edgar.jrc.ec.europa.eu/ [accessed November 2015].

[23] European Commission, Climate action: reducing emissions from transport [online], available: http://ec.europa.eu/clima/policies/transport/index_en.htm [accessed November 2015].

[24] European Commission, Climate action: road transport: reducing CO_2 emissions from vehicles [online], available: http://ec.europa.eu/clima/policies/transport/vehicles/index_en.htm [accessed November 2015].

[25] Eurostat, Passenger transport statistics [online], available: http://ec.europa.eu/eurostat/statistics-explained/index.php/Passenger_transport_statistics [accessed November 2015].

[26] Eurostat, Freight transport in the EU-28 modal split of inland transport modes [online], available: http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Figure_1_Freight_transport_in_the_EU-28_modal_split_of_inland_transport_modes.png [accessed November 2015].

[27] European Environmental Agency, Energy efficiency and specific CO_2 emissions [online], available: http://www.eea.europa.eu/data-and-maps/indicators/energy-efficiency-and-specific-co2-emissions/energy-efficiency-and-specific-co2-5 [accessed November 2015].

[28] McKinsey Company, The road to 2020 and beyond: what's driving the global automotive industry?, 2013.

[29] McKinsey Company, Bigger, better, broader: a perspective on China's auto market in 2020, 2012.

[30] International Transportation Forum, Transport outlook: scenarios to 2050 [online], available: http://2014.internationaltransportforum.org/outlook-scenarios-2050 [accessed November 2015].

[31] European Aviation Safety Agency, Annual Safety Review, 2013.

[32] M. Krzyzanowski, B. Kuna-Dibbert, J. Schneider, Health Effects of Transport Related Air Pollution, World Health Organization Europe, 2005.

[33] Health Effects Institute, Traffic-related air pollution: a critical review of the literature on emissions, exposure and health effects, Boston, Massachusetts, 2010.

[34] K. Zhang, S. Batterman, Air pollution and health risks due to vehicle traffic, Sci. Total Environ. 450–451 (2013) 307–316.

[35] P. Romero-Lankao, H. Qin, M. Borbor-Cordova, Exploration of health risks related to air pollution and temperature in three Latin American cities, Soc. Sci. Med. 83 (2013) 110–118.

[36] J.K. Vanos, C. Hebbern, S. Cakmak, Risk assessment for cardiovascular and respiratory mortality due to air pollution and synoptic meteorology in 10 Canadian cities, Environ. Pollut. 185 (2014) 322–332.

[37] L. Li, H. Li, X. Zhang, L. Wang, L. Xu, X. Wang, Y. Yu, Y. Zhang, C. Guan, Pollution characteristics and health risk assessment of benzene homologues in ambient air in the northeastern urban area of Beijing, China, J. Environ. Sci. 26 (1) (2014) 214–223.

[38] E. Brunelle-Yeung, T. Masek, J.J. Roko, J.I. Levy, S. Arunachalam, S.M. Miller, S.R.H. Barrett, S.R. Kuhn, I.A. Waitz, Assessing the impact of aviation environmental policies on public health, Transp. Policy 34 (2014) 21–28.

[39] Official Journal of the European Union, Directive 2008/101/EC of the European Parliamente and of the Council of 19 November 2008 amending Directive 2003/87/EC so as to include aviation activities in the scheme for greenhouse gas emission allowance trading within the Community [online], available: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32008L0101&from=EN [accessed July 2015].

[40] EPA, EPA takes first steps to address GHG emissions from aircraft engines [online], available: http://www.epa.gov/otaq/aviation.htm [accessed July 2015].

[41] Cleansky 2 JU, Cleansky 2 JU: mission and objectives [online], available: http://www.cleansky.eu/content/article/mission-objectives, 2015 [accessed July 2015].

[42] M. De Gennaro, E. Paffumi, G. Martini, H. Scholz, A pilot study to address the travel behaviour and the usability of electric vehicles in two Italian provinces, Case Stud. Transp. Policy 2 (3) (2014) 116–141.

[43] E. Paffumi, M. De Gennaro, G. Martini, H. Scholz, Assessment of the potential of electric vehicles and charging strategies to meet urban mobility requirements, Transportmetrica A: Transp. Sci. 11 (1) (2015) 22–60.

[44] M. De Gennaro, E. Paffumi, H. Scholz, G. Martini, GIS-driven analysis of e-mobility in urban areas: an evaluation of the impact on the electric energy grid, Appl. Energy 124 (2014) 94–116.

[45] G. Martini, E. Paffumi, M. De Gennaro, G. Mellios, European type-approval test procedure for evaporative emissions from passenger cars against real-world mobility data from two Italian provinces, Sci. Total Environ. 487 (2014) 506–520.

[46] M. De Gennaro, E. Paffumi, G. Martini, Customer-driven design of the recharge infrastructure and Vehicle-to-Grid in urban areas: a large-scale application for electric vehicles deployment, Energy Volume 82 (2015) 294–311.

[47] M. De Gennaro, E. Paffumi, G. Martini, Data-driven analysis of the effectiveness of evaporative emissions control systems of passenger, Atmos. Environ. 129 (2016) 277–293.
[48] A. Chaouachi, E. Bompard, G. Fulli, M. Masera, M. De Gennaro, E. Paffumi, Assessment framework for EV and PV synergies in emerging distribution systems, Renew. Sustain. Energy Rev. 55 (2016) 719–728.
[49] E. Paffumi, M. De Gennaro, G. Martini, Innovative technologies for smart cities: towards customer driven infrastructure design for large scale deployment of electric vehicles and Vehicle-to-Grid applications, Transp. Res. Proc. 14C (2016) 4505–4514, http://dx.doi.org/10.1016/j.trpro.2016.05.373.
[50] Octo Telematics Italia S.R.L. [online], available: www.octotelematics.it, 2013 [accessed June 2013].
[51] UNECE, Big data in official statistics [online], available: http://www1.unece.org/stat/platform/display/bigdata/Big+Data+in+Official+Statistics [accessed 2016].
[52] United Nations, UN big data for official statistics [online], available: http://unstats.un.org/unsd/bigdata/ [accessed 2016].
[53] M. De Gennaro, E. Paffumi, G. Martini, H. Scholz, H. Lacher, H. Kuehnelt, D. Simic, Experimental investigation of the energy efficiency of an electric vehicle in different driving conditions, SAE Technical Paper 2014-01-1817, 2014, Detroit, Michigan, US.
[54] M. De Gennaro, E. Paffumi, G. Martini, U. Manfredi, S. Vianelli, F. Ortenzi, A. Genovese, Experimental test campaign on a battery electric vehicle: laboratory test results (Part 1), SAE Int. J. Alt. Power. 4 (1) (2015).
[55] E. Paffumi, M. De Gennaro, G. Martini, U. Manfredi, S. Vianelli, F. Ortenzi, A. Genovese, Experimental test campaign on a battery electric vehicle: on-road test results (Part 2), SAE Int. J. Alt. Power. 4 (2) (2015).
[56] G. Martini, U. Manfredi, M. De Gennaro, Gaseous emissions from Euro 3 motorcycles and Euro 5 passenger cars measured over different driving cycles, SAE Technical Paper 2013-01-2619, 2013.
[57] European Parlament, Mapping Smart Cities in the EU, Directorate general for internal policies, Policy department A: economic and scientific policy, IP/A/ITRE/ST/2013-02 – PE 07.480, 2014.
[58] TU Wien, Smart cities, ranking of European medium-sized cities [online], available: http://www.smart-cities.eu/, 2007 [accessed 2015].
[59] European Environmental Agency, EEA, EMEP/CORINAIR Emission Inventory Guidebook – 2007, Group 7: road transport, http://www.eea.europa.eu/publications/EMEPCORINAIR5, 2007.
[60] European Commission, CORINAIR: European Inventory of emissions of pollutants into the atmosphere, 1998.
[61] S. Eggleston, D. Gaudioso, N. Gorißen, R. Joumard, R.C. Rijkeboer, Z. Samaras, K.-H. Zierock, CORINAIR Working Group on Emissions Factors for Calculating 1990 Emissions from Road Traffic. Volume 1: Methodology and Emission Factors, European Commission, ISBN 92-826-5771-X, 1993.
[62] European Environmental Agency, EEA, COPERT III computer programme to calculate emissions from road transport – user manual [online], available: http://www.eea.europa.eu/publications/Technical_report_No_50.
[63] EC4MACS – European Consortium for Modelling of Air Pollution and Climate Strategies, TREMOVE economic transport and emissions model [online], available: http://www.tmleuven.be/methode/tremove/home.htm [accessed 2016].
[64] Google Inc. [online], available: https://maps.google.com/, 2013 [accessed June 2013].
[65] POIplaza website [online], available: http://poiplaza.com/ [accessed December 2014].
[66] Terna, Consumi di energia elettrica in Italia, 2011.
[67] Unione Petrolifera, Data Book Energia e Petrolio [online], available: http://www.unionepetrolifera.it [accessed 2016].
[68] Automobile Club d'Italia, ACI, Open data [online], available: http://www.aci.it/laci/studi-e-ricerche/dati-e-statistiche/open-data.html.
[69] ANPA, Agenzia nazionale per la protezione dell'ambiente, Le emissioni in atmosfera da trasporto stradale, Ser. Stato Ambiente 12 (2000).
[70] European Parliament, REGULATION (EC) No 715/2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial, 2007.
[71] European Parliament, 2008/C 182/08 Communication on the application and future development of Community legislation concerning vehicle emissions from light-duty vehicles and access to repair and maintenance information (Euro 5 and 6), 2008.
[72] European Parliament, Directive 2014/94/EU of the European Parliament and Council of 22 October 2014 on the deployment of alternative fuels infrastructure [online], available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0094&from=EN, 2014 [accessed December 2014].
[73] Technical Guidelines for the preparation of applications for the approval of innovative technologies pursuant to Regulation (EC) No 443/2009 and Regulation (EU) No 510/2011 [online], available https://circabc.europa.eu/sd/a/bbf05038-a907-4298-83ee-3d6cce3b4231/Technical%20Guidelines%20October%202015.pdf, 2015.