TRAFAIR: Understanding Traffic Flow to Improve Air Quality

Conference Paper - October 2019

8 authors, including:

Laura Po
Università degli Studi di Modena e Reggio Emilia
41 PUBLICATIONS 324 CITATIONS

José R.R. Viqueira
University of Santiago de Compostela
72 PUBLICATIONS 259 CITATIONS

Some of the authors of this publication are also working on these related projects:

- H-BOLD View project
- QueryGen: Semantic Interpretation of Keyword Queries View project
TRAFAIR: Understanding Traffic Flow to Improve Air Quality

Laura Po∗, Federica Rollo∗, José Ramón Ríos Viqueira†, Raquel Trillo Lado‡, Alessandro Bigi∗, Javier Cacheiro López§, Michela Paolucci¶, Paolo Nesi∗

∗Enzo Ferrari’ Engineering Department
Via Vivarelli, 10, 42125 Modena, ITALY
Email: name.surname@unimore.it

†Centro Singular de Investigación en Tecnoloxías da Información (CITIUS)
Universidade de Santiago de Compostela (USC), Campus Vida, 15782 Santiago de Compostela, SPAIN
Email: jrr.viqueira@usc.es

‡Aragon Institute of Engineering Research (I3A), University of Zaragoza
C/Mariano Esquillor s/n, 50018 Zaragoza, SPAIN
Email: raqueltl@unizar.es

§Galicia Supercomputing Centre (CESGA)
Avda. de Vigo s/n, 15705 Santiago de Compostela, SPAIN
Email: jlopez@cesga.es

¶DISIT, University of Florence, ITALY
Email: name.surname@unifi.it

Abstract—Environmental impacts of traffic are of major concern throughout many European metropolitan areas. Air pollution causes 400 000 deaths per year, making it first environmental cause of premature death in Europe. Among the main sources of air pollution in Europe, there are road traffic, domestic heating, and industrial combustion.

The TRAFAIR project brings together 9 partners from two European countries (Italy and Spain) to develop innovative and sustainable services combining air quality, weather conditions, and traffic flows data to produce new information for the benefit of citizens and government decision-makers. The project is started in November 2018 and lasts two years. It is motivated by the huge amount of deaths caused by the air pollution. Nowadays, the situation is particularly critical in some member states of Europe that cannot reach the European objectives. In February 2017, the European Commission warned five countries, among which Spain and Italy, of continued air pollution breaches. These countries fail to address repeated breaches of air pollution limits for nitrogen dioxide ($\text{NO}_2$) whose most emissions result from road traffic (see Figure 1). The European Commission urged these Member States to take action to ensure good air quality and safeguard public health.

In this context, public administrations and cit-
The TRAFAIR project aims to estimate the level of pollution on an urban scale by producing the following main results:

- Provision of real-time estimates of air pollution in the city on an urban scale;
- Development of a service for forecasting urban air quality based on weather forecasts and traffic flows;
- Publication of Open Data to describe maps of urban air quality and forecast maps in 6 European cities: Zaragoza (600,000 inhabitants), Florence (382,000), Modena (185,000), Livorno (160,000), Santiago de Compostela (95,000) and Pisa (90,000).

Nine partners are involved in the project: 4 Universities, 3 Public Administrations, 1 HPC center and 1 company. More than 70 persons are involved in the projects with different expertise and backgrounds: they are mainly computer scientists, environmental experts, HPC experts and public administration managers.

This paper describes the first results obtained in the TRAFAIR project in almost one year of work. In particular the implementation of the two models that allow to create predictive urban air pollution dispersion maps.

The rest of the paper is organized as following. An overview of the entire project is provided in Section II, while the software and HPC architecture are described in Section III. The main models implemented in the project are the traffic model and the air pollution dispersion models that are implemented in each city. These models are described in Section IV and Section V. Conclusions and future research direction are summarized in Section VI.

II. THE PROJECT

Air pollution modeling is a mean used to describe the causal relationship between emissions, meteorology, atmospheric concentrations, deposition, and other factors. Air pollution measurements give important, quantitative information about ambient concentrations and deposition, but they can only describe air quality at specific locations and times, without giving clear guidance on the identification of the causes of the air quality problem. Air pollution modeling, instead, can give a complete deterministic description of the air quality problem in an urban context, including an analysis of factors and causes (emission sources, meteorological processes, and physical and chemical changes), and some guidance on the implementation of mitigation measures.

Air pollution models are one of the main methods to quantify the deterministic relationship between emissions and concentrations/depositions, including the consequences of past and future scenarios and the determination of the effectiveness of abatement strategies. This makes air pollution models indispensable in regulatory, research, and forensic applications. The concentrations of substances in the atmosphere are determined by 1) emissions, 2) transport, 3) dispersion and diffusion, 4) physical and chemical transformation, and 5) deposition. The TRAFAIR project will provide both the results of modeling and measuring activities: a daily forecast of the pollution map and a semi-real-time air quality information within the urban area.

Figure 2 shows the way information flows within the project: which input data are collected and by which models they are used, what kind of outputs are made available.

A. Input

The primary input will be the data provided by the sensor networks: traffic flow data, weather forecast, atmospheric pollution condition. Moreover, the atmospheric emission inventory, the vehicular fleet composition, a 3D model of the urban area will be collected for each city (see the left box in Figure 2). These primary inputs will be integrated in one platform [2] and used by the traffic model to produce vehicular traffic flow maps, by the air pollution dispersion model to provide a daily forecast of the air quality within the urban area and by an interpolation function to produce real time urban air quality maps.

Static source of pollutants can be found on open data repositories at a European, national or regional level. At European level, the European Environment Agency (EEA) maintains the European Pollutant Release and Transfer Register (E-PRTR); in Italy,
the reference national emission inventory is distributed by ISPRA, however the municipality of Modena has its own bottom-up emission inventory; in Spain, the national emission inventory (Registro Estatal de Emisiones y Fuentes Contaminantes) is maintained by the Ministry with competences in environmental monitoring and control (Ministerio de Agricultura y Pesca Alimentacion y Medio Ambiente). The inventories are also maintained at a regional level in Spain. Thus, the region of Aragon maintains its PRTR and in the region of Galicia maintains also an inventory of emissions (Rexistro Galego de Emisins -REGADE-). The above national/regional inventories will be complemented in TRAFAIR with local inventories in the scope of each city whenever necessary.

Weather forecast simulations for each city are provided by either national or regional meteorological agencies, in order to have models tailored to the local domain, with a spatial resolution up to 1 km or less: the air pollution dispersion model will provide the vertical profile (20 levels) within the urban domain temperature, humidity, wind speed and wind direction, with an hourly time step. In addition, some main turbulence parameters will be also provided by the model (e.g. friction velocity, Monin-Obukhov length and convective velocity scale) as well as the depth of the mixing layer and the local stability class, which are needed by the air pollution dispersion models used within the project. These latter models will use as input the outcomes by the weather forecast model, the traffic simulation model, eventually, other local atmospheric emission sources and it will provide an hourly dispersion map for the urban area, with a spatial resolution of 2-4 meters for the following day. In Modena, some tests have been made by using the proprietary model Parallel-Micro-Swift-Spray (PMSS): this is implemented on a 6 km x 6 km domain and considers a road network of about 110 sections with a total length of 210 km. An established freely available alternative to PMSS is the GRAL model: this is the model that will be mainly used during the implementation of TRAFAIR. The final goal is to create a short-range predictive modeling chain (24 hours) to simulate the atmospheric dispersion of emissions (for instance NO\textsubscript{x} and CO) for the 6 cities.

B. Output

The TRAFAIR outputs provides a possible, complete, reproducible representation of the urban air quality, easily adaptable at the international level since multi-language, and compliant with the best practices and standards defined by FAIRMODE, W3C, OGC, and INSPIRE. The project, involving several European cities with different size, will provide a flexible and adaptable service to show the current and predictive air quality on an urban spatial scale. The Consortium expects that this service can be adopted by other cities (the number of cities involved will increase) thanks to the dissemination activities.

C. Urban sensor network within the cities

A preliminary survey has been conducted in the cities involved in the project to collect information on the urban sensor network, the monitored measures, and the access to data.

Modena has 2 regulatory air quality monitoring stations (whose data is already accessible through the web portal of the regional meteorological agency, ARPAE) and several traffic sensors spread within the urban area (one for each traffic light
Traffic sensor information is collected by the Traffic Data Center, which provides semi-real-time traffic data, particularly information about the flow (number of transits detected in the selected time span), the employment rate (percentage average) and the average speed in km h\(^{-1}\). During TRAFAIR some new traffic sensors and 14 low-cost air quality monitoring devices will be placed in the urban area. Tuscany counts on 35 air quality stations (7 in Florence, 5 in Pisa, 5 in Livorno) and 800 traffic sensors (300 in Florence, 74 in Pisa and 40 in Livorno). The regulatory air quality monitoring stations can be queried through the web portal of the regional meteorological agency, ARPAT, while the traffic sensors information is available. UNIFI with the support of Tuscany Region will integrate the yet available sensor network by deploying ca. 4 low costs devices for air quality monitoring to be distributed opportune in different areas in the three cities of Florence, Pisa and Livorno. The technical infrastructure for the data acquisition, storage, and big data analysis is Snap4City, the Open Urban Platform for a Sentient Smart City\(^3\) developed by the DISIT Lab of UNIFI (\([3],[4]\)). Snap4City yet ingests and manages environmental data coming from a restricted number of environmental stations in the regional territory and aggregates the data provided by the 800 traffic sensors. It actually ingests about 1,200,000 complex events per day. Santiago de Compostela has 2 regulatory air quality monitoring stations, whose data is already accessible through the web of the regional meteorological agency (MeteoGalicia) and various traffic sensors in the main city entrances. For the project, some traffic sensors and ca. 10 low-cost air quality monitoring devices will be deployed. Zaragoza has 7 regulatory air quality monitoring stations provided by EuroAirnet (European Air Quality monitoring network) and the national network of air quality stations. Besides, 4 weather stations belonging to the Meteorological State Agency (Agencia Estatal de Meteorología -AEMET-) are located in Zaragoza. The data collected by the different stations are either freely accessible as Open Data on the Web or accessible in exchange for a previous payment. For the project, about 10 low-cost air quality monitoring devices will be deployed across the University campuses sited in different areas of the city and representative of the North, South, West and East parts of the city. Besides, how to improve the information about traffic will be also tackled.

III. SOFTWARE AND HPC ARCHITECTURES

The software architecture is organized in four different tiers, as shown in Figure 3.

At the bottom, a Sensor Data Acquisition Tier (the bottom layer in Figure 3) provides functionality to sample the required data from the deployed traffic and air quality sensor networks to populate a database of Latest Sensor Data in the Data Management Tier (the second layer from the bottom in Figure 3). Periodically, data is extracted from the Latest Sensor Database, transformed and loaded in an Observation Data Warehouse, which maintains historical sensor data. The Data Management Tier is completed with a Geospatial Data Warehouse and a Scientific Array Data Warehouse, that record respectively city geospatial data (including streets and buildings) and all environmental models collected and generated in the

\(^3\)https://www.snap4city.org
project. Well-known standards from the OGC and W3C are used to provide open access to the Data Management Tier. The three models implemented in the project are contained in a Data Processing Tier (the third layer from the bottom in Figure 3). Advanced HPC hardware and technologies will be used to implement this Tier, as it is described below. Finally, the User Interaction Tier (the top layer in Figure 3) provides multichannel GUIs for end-user applications, including web and mobile applications. Besides, geographic maps generated from the result model data are published through the well-known OGC Web Map Service interface, following INSPIRE guidance.

The implementation of the different data management and processing components of the above architecture requires HPC technologies. In particular, a combination of SQL and NoSQL technologies has been used for the implementation of large-scale datasets. Besides, the Air Pollution Spatial Interpolation and the Traffic Model demands large-scale data processing technologies based on the map-reduce paradigm. The different cities make use of HPC technologies by using different solutions:

- in the local area of Tuscany, in the cities of Florence, Pisa and Livorno, the HPC platform of University of Florence (UNIFI) is used,
- Modena and Santiago de Compostela are using the HPC platform of CESGA,
- Zaragoza is using HPC platform available in University of Zaragoza (UNIZAR), in particular, in I3A.

[1] In Florence, Pisa, Livorno and Santiago, a real time traffic flow reconstruction has been realized with the Disit Lab traffic Model by exploiting the DISIT HPC resources [5].

IV. TRAFFIC MODEL

The large number of vehicles complicates urban planning and often leads to traffic congestion and areas of increased air pollution. Traffic management solutions typically require the use of simulators able to capture in detail all the particular characteristics and dependencies associated with real-life traffic.

In Italy, traffic models are provided every 3-5 years by Public Administrations for the main municipalities, this type of models usually provides traffic information during the peak hour of an average working day of the year. These type of traffic models do not aim to describe the variation of traffic but an average situation. These models will give an approximation over the year of the traffic flows with an error around the 8-9%. For making this model dynamic, it is needed the collection and the integration of real-time traffic data. These information can be earned from several types of traffic sensors, usually traffic cameras or induction loops sensors positioned around the city or alternatively, source of traffic flow can be produced by users such as insurance black boxes places inside the cars or GPS tracks stored and used in applications such as Google traffic, INRIX traffic, Open Traffic.

For the scope of the project, we intend to catch real-time traffic information. In Florence, UNIFI has experience in real time traffic reconstruction map for the main streets of Florence. The map is constructed based on the data collected from 60 sensors. The model is able to manage the two-way roads (with a higher resolution than Google traffic). Within the TRAFAIR project this real time traffic model will be expanded to the other involved cities. To reach a good real time traffic estimation, the cities of Santiago and Modena will increase their traffic sensor networks.

Currently, Zaragoza is monitoring traffic in real time to create more realistic traffic models in the future. The current traffic monitoring system is mainly based on the use of Bluetooth receivers installed in some traffic lights along the city. Thanks to Bluetooth technology, it is possible to identify not only the amount of vehicles crossing a specific point but also traffic trends (the routes followed by the vehicles). However, this system only allows an estimation of the real traffic, but not precise measurements (a single car could be counted more than once if several passengers have Bluetooth activated in their personal devices or mobile phones, or not counted at all if there is no device with Bluetooth activated in the car). So, this system needs to be improved by considering data curation techniques.

In Modena, a traffic flow model has been implemented based on traffic sensors real data [6]-[8]. We used the Simulation of Urban MOBility (SUMO) tool5 that is a microscopic, collision-free, space-continuous and time-discrete simulator. Every step of the simulation corresponds to one second in reality. We used as input the data collected by more than 300 traffic sensors placed near traffic junctions. The output of the model was calculated every day from November 1st 2019 and stored in roads (with a higher resolution than Google traffic). The model is able to manage the two-way roads (with a higher resolution than Google traffic). Within the TRAFAIR project this real time traffic model will be expanded to the other involved cities. To reach a good real time traffic estimation, the cities of Santiago and Modena will increase their traffic sensor networks.

Currently, Zaragoza is monitoring traffic in real time to create more realistic traffic models in the future. The current traffic monitoring system is mainly based on the use of Bluetooth receivers installed in some traffic lights along the city. Thanks to Bluetooth technology, it is possible to identify not only the amount of vehicles crossing a specific point but also traffic trends (the routes followed by the vehicles). However, this system only allows an estimation of the real traffic, but not precise measurements (a single car could be counted more than once if several passengers have Bluetooth activated in their personal devices or mobile phones, or not counted at all if there is no device with Bluetooth activated in the car). So, this system needs to be improved by considering data curation techniques.

In Modena, a traffic flow model has been implemented based on traffic sensors real data [6]-[8]. We used the Simulation of Urban MOBility (SUMO) tool5 that is a microscopic, collision-free, space-continuous and time-discrete simulator. Every step of the simulation corresponds to one second in reality. We used as input the data collected by more than 300 traffic sensors placed near traffic junctions. The output of the model was calculated every day from November 1st 2019 and stored in a local database for further analyses. This model was used to analyze the real-time traffic situation in the city of Modena and to highlights the more congested zones [8].

V. AIR QUALITY MODEL

In Italy and in Spain local Regional Environmental Protection Agencies provide weather and air

---

4http://www.disit.org/traffic_24h/
5https://sumo.dlr.de/index.html
quality analysis for their respective region, along with their forecasts for the following days.

In the Emilia-Romagna region ARPAE publishes daily average of PM$_{10}$ and PM$_{2.5}$, hourly concentration of NO$_2$, NO, O$_3$, CO, BTX and other regulatory pollutants, and summarize them in an Air Quality Index (AQI). ARPAE also makes use of the NINFA air quality predictive system based on the combination of the chemical transport model Chimere [9] and the meteorological model COSMO [10]. Air quality analysis at a regional scale by the PESCO (Postprocessing and Evaluation with Statistical techniques of the Chimere Output) tool, combining NINFA products with ground-based observations [11]. Other national and European forecast modeling system are also available, an example is the QuaLeAria\(^6\) project developed by ARIANET (consulting firm operating in the environmental field) and ENEA (Italian Agency for New Technologies, Energy and Sustainable Economic Development), which main purpose is to forecast regional scale air pollution over the Italian peninsula. It is based on FARM (Flexible Air quality Regional Model) [12], a 3D Eulerian model simulating dispersion and chemical reactions of atmospheric pollutants, managed by F-Air (ARIANET Integrated Forecast System Manager), a software environment where the data sources and the modeling components are connected. QuaLeAria provides for the Italian and European countries hourly average, daily and daily average concentrations for regulatory interest pollutants: SO$_2$, NO$_2$, CO, PM$_{10}$, O$_3$ and benzene.

Anthropogenic emissions considered by the above mentioned systems are based upon national (ISPRA) and European (e.g. TNO-MACC) emission inventories, and they estimate seasonal and diurnal source variability through ancillary data, leading to an approximated description of time-dependent sources, e.g. traffic. Moreover these models are performed on large scale domains (about 300-1000 km) with an horizontal resolution of few kilometers, while the focus of TRAFAIR is to provide analysis and prediction on an urban scale. Therefore in TRAFAIR we will analyzed a comprehensive set of traffic flow simulations based on real time data to generate emission estimates tailored on the city and the time of the year.

Two are the dispersion pollutant models that we would use for our purposes. The former is Parallel-Micro-Swift-Spray (PMSS) [13] a commercial software developed by Aria Technologies in collaboration with MOKILI and ARIANET. It is composed by PSWIFT, a parallelized mass-consistent 3D wind field model for complex terrain, and PSPRAY, a parallelized three-dimensional Lagrangian particle dispersion model, able to reproduce transport, dispersion, dry and wet deposition of airborne chemically inert species released in meteorological complex conditions. PMSS has been designed to decrease the computational time on very large (tens of kilometers), high resolution (few meters) domain with high number of emitted particles. Parallelization is implemented both for splitting geographically big domains and to gain speed up of calculation, and MPI is used to allow parallel computations both on multicore laptop and on High Perfromance Computing clusters. The latter model is the Graz Lagrangian Model GRAL [14] which can simulate the dispersion of chemically non-reactive pollutants, dry deposition and sedimentation over the full range of wind speed without any lower threshold and for all atmospheric stability conditions. The model allows different type of case studies, with domains down to a region of few square kilometers with an horizontal resolution of few meters. Both, PMSS and GRAL are able to account for the presence of buildings inside the computational domain and can simulate emissions from point, area and line sources, in continuous or discontinuous way. By comparing simulated concentration obtained with the applications of PMSS model and a full CDF model well adapted for

\(^6\)http://www.aria-net.it/qualearia/it/
the planetary boundary layer, final concentration differ in general by 3-5% [15], [16]. Instead, for the GRAL model the temporal correlations between simulation and observation pollutant concentrations are generally in the range 0.5-0.8 at hourly scale and 0.7-0.9 at daily scale [17].

In Modena, we elaborated an average traffic model for weekdays, Saturdays and Sundays based on the traffic flows registered from the 1st till the 30th of November. The output of the traffic model is used by a customized version of the open source model Vehicular Emissions Inventory (VEIN) [18] to calculate the emission generated in each road of the city for each hour of the day. Finally, using the weather 48 hour forecast produced daily by ARPAE7, the GRAL air pollution model is executed to create predictions for the next 48 hours. A similar approach was followed in Santiago de Compostela, using meteorological forecast provided by MeteoGalicia8. In Zaragoza, a traffic model was also developed by considering data about traffic flows registered in 2018 and the first three months of 2019 and by using the SUMO tool and VEIN. After that, by considering the meteorological forecast (provided by MeteoGalicia), GRAL is also executed to create predictions for the 24 hours of the day after tomorrow, as currently performing those computations takes around 30 hours. Figure 4 shows a snapshot of 1 hour prediction for air pollution in Santiago (on the left) and Zaragoza (on the right). In Firenze, Pisa, Livorno, a real-time visual self-adaptive solution for the vehicular traffic flow reconstruction developed by the DISIT Lab, has been used [5]. This traffic model allows to visualize in a map the real-time state of traffic and produces the output files containing the reconstructed traffic flow data for the road-segments considered. The output of the traffic model is used in a customized version of the open source model Vehicular Emissions Inventory (VEIN) [18] to calculate the emission generated in each road of the city for each hour of the day. Finally, using the weather 48 hour forecast produced every 3 hour a day, by OpenWeatherMap9, the GRAL air pollution model is executed to create predictions for the next 48 hours.

In these first months of the project, some predictive urban air quality maps were generated. A set of videos showing these results are available on the TRAFAIR website10. We are still working on the creation of spatial interpolation maps based on real time data provided by the low cost air quality monitoring units. An R script was created to generate an interpolation map using air pollution data. The script was tested in Zaragoza and Modena using data from the local regulatory air quality stations (Figure 5); in the future, it will make use of several of the low cost air quality monitoring units, once their calibration is completed and they will be deployed across town.

VI. CONCLUSION

The TRAFAIR project aims to estimate the level of pollution at an urban scale by producing real-time air pollution estimates in the project cities, by developing a service for forecasting urban air quality based on weather forecasts and tailored traffic flow estimates. This two-year project started the 1st November 2018 and is implemented in 6 European cities of different sizes: Zaragoza

---

7https://www.arpae.it/
8https://www.meteogalicia.gal
9https://openweathermap.org/
10http://trafair.eu/models/
The results of the TRAFAIR project will enrich the European Data Portal adding a new kind of air quality datasets: along with the concentration time series of atmospheric pollutant at specific sites by the low cost monitors, we will produce time series of urban air quality maps for the 6 TRAFAIR cities. The maps will describe the impact of traffic on air quality within each project city, raising the awareness of the citizens that work, live or transit in that specific area.

The project, affecting different European cities of different sizes, will provide a flexible and adaptable service to show the current and predictive air quality on an urban scale. We expect that this service can be adopted by other cities thanks to the dissemination activities.

ACKNOWLEDGMENT

This research has been supported by the TRAFAIR project 2017-EU-IA-0167, co-financed by the Connecting Europe Facility of the European Union. The contents of this publication are the sole responsibility of its authors and do not necessarily reflect the opinion of the European Union. The authors would like to thank all the partners involved in the project: the City of Modena, the Tuscany Region, the City of Santiago de Compostela, Lepida S.c.p.A. Also a special thanks goes to the Italian Institute for Environmental Protection and Research (ISPRA) and the Italian Alliance for Sustainable Development (ASviS) and the regional meteorological agencies of the Galicia (MeteoGalicia) and of the Emilia-Romagna (ARPAE) that have expressed their interest in the project.

REFERENCES


