



# PhD in MODELLI E METODI MATEMATICI PER L'INGEGNERIA / MATHEMATICAL MODELS AND METHODS IN ENGINEERING - 40th cycle

**PNRR 630 Research Field: SOURCE APPORTIONMENT OF FINE PARTICULATE MATTER  
AND RISK OF ADVERSE BIRTH EVENTS**

<b>Monthly net income of PhDscholarship (max 36 months)</b>
<b>€ 1400.0</b>
In case of a change of the welfare rates during the three-year period, the amount could be modified.

<b>Context of the research activity</b>	
<p><b>Motivation and objectives of the research in this field</b></p>	<p>Low birth weight (LBW) and preterm birth (PTB) have been associated in the literature with increased neonatal morbidity and increased risk of health issues and developmental delay later in life [1]. Several studies have found a significant association between exposure to particulate matter (in particular PM<sub>2.5</sub>) and LBW [2] or between PM<sub>2.5</sub> and PTB [3].</p> <p>However, it remains unclear which PM sources mainly contribute to these associations. Identifying the sources contributing to a pollution mixture can be challenging because most observed constituents are generated by multiple sources. Source apportionment (SA) aims at decomposing the observed individual constituents providing information about the sources that contribute to the mixture (e.g. traffic, power generation). Air pollution and associated health issues are a compelling problem in Lombardy [4]. PM levels are particularly high in the Po Valley because of a combination of urban and industrial areas, its geographical shape and climate. The goal of this project is to identify which sources of PM in Lombardy might be associated with LBW, PTB, other adverse birth events or early childhood adverse events, taking into account the data's spatial or spatiotemporal connotation. The objective consists in developing innovative statistical models:</p> <p>(i) to identify the sources of PM in Lombardy, for</p>



	<p>the broader problem of developing general intervention strategies to mitigate PM pollution,</p> <p>(ii) to identify which sources of PM are associated with the health outcomes and estimate their relative contribution.</p> <p>The objective might make available knowledge and tools to the institution of local governance for a better understanding of features and evolution of health outcomes related to air pollution. We would strengthen the administrative capacity in relation to the formulation and design of public policies, by developing diagnostic skills and helping in the identification of solutions, decisions and implementation. In summary, the research activity aims to significantly develop knowledge in the PNRR's areas of interest. The research activity is in collaboration with ATS Milano, Unità di Epidemiologia and ARPA Lombardia, U.O. Qualità dell'Aria. The latter will provide data and expertise about air pollution, while the former has expertise in the epidemiology of air pollution morbidities and mortalities.</p>
<p><b>Methods and techniques that will be developed and used to carry out the research</b></p>	<p>There are many statistical methods and challenges for source apportionment [5]. One of the most flexible is the receptor-based Positive Matrix Factorization (PMF) source apportionment, which assumes a mass balance framework, i.e., there is no constituent mass lost from source to receptor. To model PM over time at different locations in Lombardy, we assume a number of latent functional profiles to represent the sources and combine them to model PM data (or more components of PM data, e.g., sulfur, aluminium, etc.) over time through latent factors [6]. We will also consider the Bayesian approach to the problem, so that we can easily incorporate spatial random effects representing the spatial association of data. The number of latent factors is random and can be inferred from the data. Finally, we aim to jointly model the health outcomes first by marginally modelling PM (or PM components) and then by modelling the conditional distribution of the health outcomes, given the combination of latent factors and source profiles, as a generalized linear model. However, there are key computational</p>



	<p>challenges in this context that must be solved:</p> <p>(i) the non-identifiability between the unknown source profiles and the latent factors, which might be dealt with by normalizing the factors or the profiles and</p> <p>(ii) the design of the associated Markov Chain Monte Carlo algorithm, in order to compute the posterior distribution of all the parameters to derive the inference.</p> <p>[1] Belbasis, L., Savvidou, M.D., Kanu, C., Evangelou, E., Tzoulaki, I. (2016). Birth weight in relation to health and disease in later life: an umbrella review of systematic reviews and meta-analyses. <i>BMC Med.</i> 14, 147</p> <p>[2] Ng, C., Malig, B., Hasheminassab, S., Sioutas, C., Basu, R., Ebisu, K. (2017). Source apportionment of fine particulate matter and risk of term low birth weight in California: exploring modification by region and maternal characteristics. <i>Science of the Total Environment</i>, 605, 647-654.</p> <p>[3] Alman, B. L., Stingone, J. A., Yazdy, M., Botto, L. D., Desrosiers, T. A., Pruitt, S., ... and the National Birth Defects Prevention Study (2019). Associations between PM2. 5 and risk of preterm birth among liveborn infants. <i>Annals of epidemiology</i>, 39, 46-53.</p> <p>[4] EEA: Air quality in Europe - 2019 Report: Technical Report. Tech.rep., European Environmental Agency (EEA)</p> <p>[5] Krall, J. R., Chang, H. H. (2019). Statistical methods for source apportionment. In <i>Handbook of environmental and ecological statistics</i> (pp. 523-545). Chapman and Hall/CRC.</p> <p>[6] Frigeri, M., Berrocal, V., Guglielmi, A. (2024). Bayesian source apportionment for particulate matter pollution. Manuscript.</p>
<p><b>Educational objectives</b></p>	<p>The PhD student will have the opportunity to increase his/her abilities as a biostatistician or an epidemiologist or a data scientist, with strong expertise in statistical learning for healthcare and air pollution. In particular the student will study, develop and apply innovative statistical techniques for spatial survival analysis, time series and clustering, in order to identify spatio-temporal variability of total and cause-specific morbidities and mortalities, useful</p>



	for understanding the role of the air pollution driving factors worsening health conditions of citizens of Milano metropolitan area and the whole Lombardy.
<b>Job opportunities</b>	The program will allow the student to find a job as a biostatistician or an epidemiologist. Moreover, the expertise in spatio-temporal statistical is nowadays a key methodological factor in many applicative industrial fields such as 5G/6G telecom networks, IoT applications (smartcity, public utility), manufacturing and logistics, climatic studies, and for all new technologies allowing for “data everywhere and everytime”. Hence, the program also allows for a data scientist job.
<b>Composition of the research group</b>	0 Full Professors 1 Associated Professors 1 Assistant Professors 4 PhD Students
<b>Name of the research directors</b>	Prof.Alessandra Guglielmi

<b>Contacts</b>	
alessandra.guglielmi@polimi.it	

<b>Additional support - Financial aid per PhD student per year (gross amount)</b>	
<b>Housing - Foreign Students</b>	--
<b>Housing - Out-of-town residents (more than 80Km out of Milano)</b>	--

<b>Scholarship Increase for a period abroad</b>	
<b>Amount monthly</b>	700.0 €
<b>By number of months</b>	6

<b>National Operational Program for Research and Innovation</b>	
<b>Company where the candidate will attend the stage (name and brief description)</b>	ATS MILANO
<b>By number of months at the company</b>	6
<b>Institution or company where the candidate will spend the period abroad (name and brief description)</b>	
<b>By number of months abroad</b>	6

<b>Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information</b>
Educational activities (purchase of study books and material, funding for participation to courses,



summer schools, workshops and conferences): financial aid per PhD student per year

1st year: max 1.902,40 euros

2nd year: max 1.902,40 euros

3rd year: max 1.902,40 euros

The PhD students are encouraged to take part in activities related to teaching, within the limits allowed by the regulations. 1 individual PC per student + several shared PC.

Access to one cluster with 32 processors and 384 GB RAM, and to several multi-processor servers.

Study and research period visiting "Pubbliche Amministrazioni": 6 months at ATS, Città Metropolitana di Milano, UOC Unità di Epidemiologia

Study and research period: 6 months at a university abroad

The PhD project will promote the use of research results and guarantee the protection of intellectual property, ensuring open access to the public to research results and related data, in the shortest time and with the least possible limitations, according to the principles "Open science" and "FAIR Data".

The principles of PNRR will be guaranteed. In particular, our research program concerns disciplinary and thematic areas consistent with the needs of the country and the regional territories affected by the program in terms of highly qualified figures oriented towards satisfying the innovation needs of the companies referred to in the PNRR.

Missions of the PNRR involved in this research: (a) M1C1, (b) M2C4, (c) M6C2.