Challenges in assessing the environmental impact of inbound tourism-related transport in Italian destinations

Extended Abstract

Globally, the transport sector is responsible of 18% of all carbon dioxide (CO2) emissions from human activities (ATAG, 2023). In Europe, transport produces over 25% of all greenhouse gas and is the main source of CO2 emissions. Hence, it is of outstanding importance to develop sustainable mobility systems, for progressing towards climate neutrality and recover environmental quality. The easiest strategy to lower transport-related emissions and depletion of natural resources is to limit mobility. Thanks to information and communication technology (ICT), many journeys can be avoided. For example, business trips can be replaced with online meetings and study visits with virtual tours. Conversely, transport is essential to (non-virtual) tourism, by definition. Thus, for reconciling sustainability and tourism, innovative policies, mobility systems and technologies have been up taken, mainly in overcrowded cities, but also in some rural destinations. For these pilot projects not to remain isolated experiences, but to lead a global change in tourism mobility, it is necessary that their effectiveness in reducing the environmental impact of tourism-related transport is assessed. This way, best practices can be identified and applied globally. To this aim, environmental data at the area-level where innovations have been experimented is required, with a time coverage sufficient to analyze the change occurred. Unfortunately, data are available for the overall transport sector, not disaggregated for subsectors. Thus, the impact of tourism-related transport, and the effectiveness of mitigation strategies, must be estimated, based on simplifying assumptions. However, many countries, like Italy, provide very fragmented information at the sub-national level. Hence, assessing the environmental impact of tourism in Italian destinations is particularly challenging. This study constitutes a first attempt to accept such challenge. To the best of my knowledge, a similar endeavor has been carried out in Latvia (Grizane & Jurgelane-Kaldava, 2019), but in Italy the data availability issue is much more serious, hence the task is even more challenging.

The objective of this study is to propose a strategy to obtain at least a coarse estimate of the environmental impact of inbound tourism-related transport in the main Italian destinations (NUTS 3 level), given the serious data availability issues. The environmental impact consists in the effects of human activities on natural cycles and processes. It is measured using a variety of variables that capture the different dimensions of environmental degradation, including biodiversity loss and land use changes. The most common direct measures of environmental impact from the transport sector are emissions of pollutants and noise levels, while energy consumption can be used as indirect indicator. The Italian National Institute of Statistics (ISTAT) publishes geographically disaggregated environmental data yearly. However, the surveyed environmental variables and disaggregation level change frequently. So, the largest panel dataset that can be built from ISTAT data ranges from 2015 to 2020 and includes only the fourteen provinces. Moreover, energy consumption and emission intensities are not disaggregated by economic sector.

Along with CO2 emissions, the environmental variables from ISTAT considered in this study are the following. Exceedance of threshold values for the concentrations of particulate matter (PM10 and PM2.5), nitrogen dioxide (NO2), and Ozone (O3), per 100 valid measurements and absolute values. These three variables are especially appropriate to measure the environmental impact, because they refer to dangerous concentrations of gases produced (also) by the transport sector. PM10 and PM2.5 emissions result from the combustion of fossil fuels, wear and tear of tires and brakes, and resuspension of road dust. In fact, road transportation contributes to overall particulate pollution up to 35%. Combustion processes, such as those used in internal combustion engines in vehicles and power plants, are the primary source of NO2 in the atmosphere. Road vehicles, particularly diesel-powered ones, contribute up to 60% to urban nitrogen oxides (NOx) emissions. Ozone at ground level is created by a chemical reaction between NOx and volatile organic compounds (VOC) in the presence of sunlight. While O3 is not directly emitted by transportation sources, the NOx and VOCs emitted by vehicles and other combustion processes contribute to ozone formation. We also consider a more indirect indicator of environmental impact: the total energy consumption (Tep per 100 inhabitants), which includes energy for transport. However, the actual impact depends on how energy is produced, from fossil fuels or renewable sources. While this variable refers to consumption of energy from all different sources, so it is inherently a not very accurate indicator. We also consider two land use variables: the availability of urban green spaces (square meters per inhabitant) and areas of urban forestation (square meters per hectare). These indicators help to measure the environmental impact of transport, because roads, trails, airports, and harbors encroach upon natural land, and the higher the tourist turnover, the greater the need of transport infrastructures.

Since 1990, the Higher Institute for Protection and Environmental Research (ISPRA) has published, just every four or five years and currently until 2019, very fine-grained environmental information. However, the database made publicly available by ISPRA contains so many missing data, that a panel dataset cannot be built with reference to the transport sector only. Data about aggregate emissions of benzene, CO2, methane (CH4) and non-methane organic compounds (NMOCs) are available for twenty-five provinces only, with different time gaps for each. Benzene is a VOC and a known carcinogen. It is emitted from transportation sources primarily through the combustion of gasoline and diesel fuels. Evaporative emissions from fuel systems and refueling operations also contribute to atmospheric benzene levels. Methane emissions occur through incomplete combustion processes and from the evaporation of fuel. Thus, although less commonly associated with transportation compared to CO2, CH4 is still emitted by vehicles, particularly those powered by natural gas. To a lesser extent, gasoline engines produce also CH4, which is a component of exhaust emissions from diesel engines too. NMOCs encompass a wide range of organic chemical compounds, generated from transportation means through the incomplete combustion of fuel and from evaporative losses. These compounds contribute to the formation of ground-level ozone and secondary organic aerosols, both of which are significant air quality concerns. Hence, these variables are good indicators of transport-related environmental impact.

Unluckily, not all the fourteen destinations for which environmental data from ISTAT are available, are also covered by the ISPRA dataset. However, the two databases are linked by the presence of CO2 emissions data. Moreover, all the environmental variables considered are closely tied, since some compounds react to form others (e.g. VOCs react with NOx to form O3) and most pollutants result from the same process (e.g. fuels combustion produce PM10 and PM2.5, NO2, and VOCs). Similarly, energy consumption is the main responsible of the emissions measured by ISPRA. While green spaces have a mitigating effects of air pollution. In fact, the surface area of vegetation, including leaves, stems, and trunks, acts as a physical filter, trapping particulate matter. Vegetation absorbs NO2, ozone, and many VOCs (including benzene) through stomata (small openings on the leaves). Some plants can metabolize and break down VOCs in their biological processes. Certain soil bacteria associated with plant roots can even consume methane. Hence, we propose to exploit the close relationships between the ISTAT and ISPRA environmental variables, in the imputation of the excessive missing values, to fill the information gaps.

In view of the very limited environmental data availability, imputed values are characterized by high uncertainty, whichever imputation method is used. In fact, uncertainty is caused by the very small number of data points compared to the many missing values across both time and destinations, and to the small number of destinations compared to that of the variables of interest. Thus, for increasing the precision of imputation, we also employ two-hundred auxiliary variables, in a two-step imputation procedure. Auxiliary variables are measured with no error nor missing values by ISTAT. Such auxiliary information regards demographic, socio-economic and territorial characteristics of each destination. In both steps of the imputation procedure, we rely on Random Forests (RF). RF is a very flexible non-parametric machine learning algorithm, that can be used for both continuous and categorical variables, linked by linear or non-linear relations, also in the presence of complex interactions. We have chosen a non-parametric method for avoiding possible misspecification and the consequent bias. However, non-parametric methods can yield very volatile estimates, if data are not missing at random and the sample size is too small. Compared to parametric methods, the uncertainty of data imputed with non-parametric algorithms is generally more difficult to quantify, precisely due to the lack of a probability model. A notable advantage of RF is that the spread of imputed values for a missing data point across different trees can be considered a reliable uncertainty measure. Moreover, through ensemble machine learning, where each imputed value depends on multiple decision trees, RF inherently reduces the prediction error's variance. This implies a decrease in the uncertainty associated with the imputed data, if the cause of missingness is not related to these values.

In the first step, we run cross-section RF year by year, only for these years for which both ISPRA and ISTAT datasets contain at least some observation at the NUTS 3 level per variable. This way, we impute values by destination. Then, we create a panel dataset seamlessly over time and apply RF again. This way, we fill time gaps. We use the resulting 'full' measures of local overall environmental impact as dependent variables in a panel model, to estimate the contribution of inbound tourism-related transport. However, information about tourism-related transport in Italian destination is also very limited. Census data about tourist mobility or NUTS 3-level official estimates are not publicly available. Microdata about inbound visitors' transport, spanning 25 years (from 1997 to 2022), can be retrieved from the Survey on International Tourism by the Bank of Italy. The indication of the Italian destination is available only for inbound tourists, so a limitation of this analysis is not to consider domestic tourism, which impacts the local environment as well. For international tourists, disposable information regards: the transport mode employed to reach Italy, the expenditure for transport tickets for routes between Italy and abroad (round trip), the amount spent for moving within Italy, including purchases of fuel therein, as well as the expenditure for tourist activities like guided tours, vehicles rentals and excursions. Although the survey sampling design is very complex, (almost) unbiased estimates can be obtained by weighting observations by the very accurate design weights provided by the Bank of Italy. However, the number of surveyed tourists by destination is small and many interviewees do not answer questions about mobility. Thus, also design estimates are highly uncertain.

For improving the precision of the survey-based estimates of the proportions of visitors who arrived in Italy by different transport modes and of their average expenditure for transport services, an autoregressive Fay-Herriot (FH) model is employed. The FH model allows us to exploit information from tourism-related auxiliary variables from ISTAT, in the same spirit as we used them in RF, but assuming a normal distribution for the error terms. The model also includes a random effect for each destination, to account for the variability of tourist mobility behaviors between areas, that is not explained by the auxiliary variables. This approach allows model permits areas with the smallest number of observations to borrow information from those with a larger sample size, by shrinking the design estimates towards the regression prediction. The autoregressive FH model also accounts for the temporal correlation of tourist mobility variables within destinations across time. This way, model-based estimates borrow strength across years through the autoregressive structure. Hence, through an autoregressive model the limited available information is used more efficiently than by estimating a separate model for each year. In fact, tourist mobility choices depend largely on each destination's supply of transport services, which does not change quickly over time.

Once imputed missing data about the overall environmental impact (dependent variables) and reduced the uncertainty about international visitors' mobility (explanatory variables), a fixed effect model for each environmental measure is estimated, to gauge the contribution of inbound tourists' transport. The fixed effect model is chosen because it accounts for the unobserved time-invariant heterogeneity in environmental quality and energy consumption between destinations. To 'isolate' the contribution of inbound tourists, the following control variables are added to the specification: motorization rate (number of cars per 1000 inhabitants), index of the polluting potential of circulating cars (number of cars with mid-high polluting potential per 100 cars with mid-low polluting potential); percentage of active resident population working in the agriculture sector, in the industry extraction and energy production sector, in the manufacturing industry, in the construction industry, as well as in the transport and communications sector. Since the interest is in the models' coefficients, all the variables not already ranging between 0 and 1 are standardized, so that differences in the magnitude of estimated coefficients do not reflect differences in measurement scales.

Tentative results confirm that inbound tourism-related transport in Italian destinations exacerbates the local environmental impact of human activities. In particular, road vehicles are responsible of a large share of air emissions and are correlated with smaller green spaces. Notwithstanding the statistical effort, the obtained results are very uncertain, due to the irreplaceability of 'first-hand' information. Nonetheless, this study offers a novel contribution to the transport and tourism literature, by proposing a strategy to obtain a coarse estimate of the international tourism-related transports’ environmental impact, where data availability issues make the attempt almost desperate. In addition, final results will bear important implications for the considered destinations’ management organizations, policy-makers and mobility managers. Overall, this study represents a call to Italian institutions to improve the collection and publication environmental data, by ensuring consistency over time in the definition of variables, territorial comprehensiveness, and transparency. We expect that, beyond fulfilling the needs of scientific research, such an improvement could gain to the cause of environmental protection the consensus and collaboration of the most reluctant part of the population. In fact, by sharing with the people the empirical evidence that forces policy-makers to ask people to scarify their mobility habits, the population should understand the concern and the urgency to adopt sustainable behaviors, starting from transport choices.

References

Air Transport Action Group (ATAG, 2023). Facts and Figures. Available at: https://atag.org/facts-figures.

Grizane, T., & Jurgelane-Kaldava, I. (2019). Tourist Transportation Generated Carbon Dioxide (CO2) Emissions in Latvia. *Environmental and Climate Technologies*, 23, 274-292. https://doi.org/10.2478/rtuect-2019-0095.