ORIGINAL PAPER

Assessing geographical differences of the COVID-19 impact on fuel consumption: the case of Italy

Riccardo Borgoni[1] · Matteo Denova[1] · Paolo Maranzano[1,2] · Caterina Morelli[1]

Abstract

We propose a spatio-temporal analysis aiming at estimating the variation of gasoline and diesel consumption occurred in the Italian provinces as a consequence of the COVID-19 pandemic during 2020 and 2021. We employ generalized linear models and generalized additive models to predict business-as-usual per capita fuel consumption by means of socio-economic variables, such as demography, tourist stays, temperature and territorial characteristics. The province-specific impact of the restrictions is then estimated as the difference between the observed and model-predicted values. The spatio-temporal dynamics is introduced by means of appropriate spline basis functions allowing both spatial and time trend to be smooth surfaces. Moreover, we include a smooth spatio-temporal interaction that properly capture the seasonal characteristic of different areas. Our findings highlight that the national mobility restrictions imposed to fight the spread of COVID-19 in the first wave (March to May 2020) reduced gasoline and diesel consumption homogeneously nationwide, with some slight differences between northern and southern provinces. However, the second wave lockdown (Winter 2020–2021) determined heterogeneous impacts across Italian provinces. In particular, we highlight the existence of a divergence between North and South Italy in terms of reaction to the pandemic. Indeed, while southern provinces experienced the weakest and less variable reductions, the northern provinces recorded strong and persistent contractions of fuel consumption. Also, southern and tourism-devoted provinces recovered quickly and compact to the post-pandemic reopenings.

Keywords COVID-19 · Fuel consumption · Italy · Spatio-temporal statistics · Generalized additive models

✉ Paolo Maranzano
  paolo.maranzano@unimib.it

[1] Department of Economics, Management and Statistics (DEMS), University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126 Milan, Italy

[2] Fondazione Eni Enrico Mattei (FEEM), Corso Magenta 63, 20123 Milan, Italy

 Springer

1 Introduction

COVID-19 was recognized by the World Health Organization (WHO) as a pandemic on the 11th of March 2020 and most countries introduced a variety of containment strategies in order to reduce the number of infections, such as travel restrictions, social-distancing and shutdowns of schools, non-essential businesses and public offices (Perra 2021; Lai et al. 2020). Consequently to the restrictions, the overall consumption of households and businesses (Zhang and Sogn-Grundvåag 2022), and in particular that associated with the demand for energy, electricity (Li et al. 2022), fuels (Güngör et al. 2021), and heating (Chrulski 2022), suffered large negative shocks. In addition to the reduction in consumption, the environment also benefited from improvements in local and aggregate emissions of pollutants due to the severe reduction in traffic flows (Du et al. 2021).

In Italy, the first case of COVID-19 was officially recognized on the 20th of February 2020, and within weeks it had spread throughout the country, with particular relevance in the North and the Lombardy region (Cereda et al. 2020). To contrast the spread of the pandemic, the Italian government adopted two separate lockdown periods, which had different duration and restrictions and applied differently to the targeted territories. The first lockdown, which coincided with the first virus wave in spring 2020, covered the entire nation (without exception) for 71 days ranging from March 9 to May 18 (Presidenza del Consiglio dei Ministri Italia 2020b; Il Sole Ore 2023). In first period, the countermeasures adopted were uniform and equal for all areas. The second lockdown was adopted between October and November 2020 (Presidenza del Consiglio dei Ministri Italia 2020a) and heterogeneously affected Italian regions. In fact, three different policies (namely, the yellow, orange and red zones) were implemented in regions according to their local risk as measured by pooling 21 indicators (Pelagatti and Maranzano 2021). Each region was assigned a restriction level according to the number of cases and death registered in the last week. Moreover, each province was allowed to impose a more restricted level with respect to the region to which it belongs if necessary. By May 2021, the restrictions are almost completely relaxed and the lockdowns are completed.

We rely on the hypothesis that a direct consequence of the regionalization of pandemic response policies is that restrictions have affected Italian regions and provinces with different magnitudes. In particular, we have reason to believe that fuel consumption has been heterogeneously affected by individual (e.g., how much and where to travel during the lockdown-free periods) and government choices within the Italian territory. However, to the best of our knowledge, only a few researchers considered the heterogeneity in pandemic effects within the same country. For instance, it is worth noting that, Bressan and Zaccomer (2023) stated that there were significant changes in consumption in Italian border regions (e.g., Friuli-Venezia Giulia on the border with Slovenia and Croatia), which saw their boundless mobility advantage compromised. In particular, drivers, who were no longer able to refuel in neighboring regions and abroad because of the mobility restriction, balanced the losses, reporting smaller contractions compared to the national average.

Due to these considerations, in this paper, we propose a spatio-temporal statistical analysis aiming at estimating the variation of gasoline and diesel consumption in Italy at provincial level occurred during the COVID-19 pandemic. In particular, we focus on the years 2020 and 2021, which include the first lockdown wave (i.e., March to May 2020) and second wave of COVID-19 (i.e., November 2020 to May 2021) and the following restarting periods. We are interested in highlighting the spatio-temporal dynamics within and between the Italian provinces and regions during 2020 and 2021. Specifically, we are interested in addressing two research issues:

1. To investigate whether the variations within provinces have followed the evolution of COVID-19 pandemic and the policies countering the virus in the same way. That is, we aim to study the temporal dynamics of fuels consumption within provinces and regions during the pandemic period. By using spatio-temporal regression models in a 'business-as-usual' forecasting framework, we estimate the province-specific variation in gasoline and diesel consumption in 2020 and 2021, pointing out the differences in the impacts over the two lockdowns and the restarts.[1]

2. To assess whether changes in fuel consumption were homogeneous or heterogeneous at the national level. In other words, we want to explore the spatiotemporal dynamics of variations in gasoline and diesel fuel consumption between areas. We then compare the time series of provinces and regions, highlighting which areas reacted more quickly to restarts and whether the magnitudes of the two lockdowns are comparable or not.

The rest of the paper is structured as follows. In Sect. 2, we present a brief literature review reporting some stylized facts regarding the evolution of the pandemic in 2020 and 2021 and the main effects on the economy and energy markets around the world. In Sect. 3 we describe the available data, sources, and conduct a quick exploratory analysis of the statistical characteristics of the data. In Sect. 4, we discuss the methodology used to estimate the impact of restrictions during the pandemic. In particular, we present the predictive out-of-sample methodology and define how it is used in both model selection and optimization of the spatial component of the GAM. In Sect. 5, we report and discuss the empirical results in light of the history of COVID-19 and the policies adopted to counter it. Finally, in Sect. 6, we conclude the paper by summarizing the main results obtained.

---

[1] It is worth noting that, to the best of our knowledge, while in the literature can be found many applications of business-as-usual techniques in the environmental field, there is no trace of studies involving spatio-temporal models within policy/event assessment framework or having as main focus regionalized fuel consumption.

2 Some stylized facts about COVID-19, energy and environment

The literature presents a large number of papers affirming that COVID-19 pandemic caused remarkable and lasting effects on different aspects of individual and collective life around the world. Ferrero et al. (2022) describe recessive effects on the economy due to lockdowns, only partially alleviated by the expansive fiscal policies of the most advanced countries, Zhang and Sogn-Grundvåag (2022) found a decrease of firms revenue and a raising demand for liquidity, resulting in increased financial stress for firms throughout the world and the International Energy Agency reported that road transport activity has dropped between 50–55% in US and Europe (Sung and Monschauer 2020).

Considering the Italian context, the economy fell sharply. Fezzi and Fanghella (2020) estimates a reduction in GDP of 30% during the first 3 weeks of lockdown. According to the Bank of Italy, households faced severe reductions in disposable incomes and thus in their consumption and savings Ferrero et al. (2022). In particular, household spending has shrunk by about 10% between 2019 and 2020, compared with a 3% drop in disposable income (Guglielminetti and Rondinelli 2021; Ercolani et al. 2021) and overall a spiral of unemployment, collapsing business sales and severe liquidity strains had been trigged, affecting small and medium-sized enterprises (Visco 2021).

Focusing on the energy sector, government restrictions and mobility were significantly associated with reductions in electricity consumption in both the initial and recovery time periods of the pandemic (Buechler et al. 2022), in particular in France, Germany, Spain and UK the electricity generation from non-renewables decreased by 21%-25% in the second quarter of 2020 compared to the same period of 2019 (Şahin et al. 2021). Li et al. (2022) show that lockdown and social distancing compelled by governments worldwide resulted in a significant reduction in energy demand and such reduction is statistically correlated with the intensity and severity of the pandemic. In the EU scenario, according to Eurostat (2022) consumption of gas and diesel oil in EU fell by 7% in 2020, while in 2021, the consumption increased but was still below 2019 levels (−6%). Compared with with 2019, gas and diesel oil consumption in 2020 in Italy reduced by about 13% and in 2021 by 2.5%. Specifically for the gas market, in 2020, there was no observed recovery in Italian demand for natural gas, and the level is not comparable to that of 2019 (Chrulski 2022). Even stronger reductions were observed in Luxembourg, Spain and Slovenia. Only some Eastern European countries (e.g., Poland and Romania) recorded positive rebounds in 2021 (Eurostat 2022). Moreover, Dembińska et al. (2022) analyze fuel supply in EU during COVID-19 and they found that without the pandemic approximately 31,495 thousand tons of kerosene-type jet fuel and 11,396 thousand tons of gas oil and diesel oil would have been additionally supplied in 2020 and 2021.

Considering the US, Wang et al. (2022) observed that oil consumption under the epidemic was about 18.14% lower than under the normal epidemic-free situation in the period between January 2020 and March 2021. Furthermore, Ou et al. (2020) estimate that US on-road transportation fuel consumption in April 2020 was 30% lower than in April 2019.

Zhang et al. (2021) examined the variation of gasoline and diesel consumption of fuel vehicles in February 2020 in China, and they found a decrease of 46.99% and 46.12% respectively compared to the same month in the previous year. Aruga et al. (2020) reveal that the energy consumption in India plummeted dramatically by the end of March 2020 because of lockdown however, since the end of April, energy consumption started to recover as the regulation relaxed. Lalas et al. (2021) show that oil consumption for road transport in Greece registered relevant reduction during the lockdown with respect to 2019, in particular −21% in March 2020, −37% in April −20% in May, −10% in June, and returned to pre-COVID levels in July.

Apparently in contrast with the above-mentioned results, several research report evidence of increasing energy demand for domestic purpose during the lockdown (Menneer et al. 2021; Rouleau and Gosselin 2021; Edomah and Ndulue 2020). However, we can address this variation to the fact that people spend more time at home due to restrictions. Overall, the total demand of energy seems to decrease (Buechler et al. 2022; Li et al. 2022; Dembińska et al. 2022; Aruga et al. 2020)

The shrinkage of fuel consumption and mobility shutdown led to a significant reduction in the emission of pollutants, thus an overall improvement in air quality. According to Liu et al. (2020), CO_2 global emissions were 8.8% lower between January 1st and June 30th of 2020 as compared to 2019, mainly due to the ground transportation sector and domestic and international aviation. Also, considering 13 stations in 11 European cities Nicolini et al. (2022) found that urban CO_2 emissions were cut by 5% to 87% with respect to the same period in previous years. Moreover Ray et al. (2022) affirm that variation of CO_2 in 2020 was heterogeneous across different countries, depending on the length of the lockdown period.

Rugani and Caro (2020) show that the carbon footprint in the lockdown period in Italy is 20% lower than the mean calculated for the past and this reduction is higher in the Northern regions which are the utmost industrialized areas of Italy and have been the ones mostly affected by the outbreak. Smith et al. (2021) take into consideration 32 different countries, both advanced and emerging economies, and, using a global vector autoregressive (GVAR) model, they predict that fossil fuel consumption and CO_2 emissions will return to, and even exceed, pre-crisis levels within two years, despite the sharp reductions in the first quarter of 2020.

Considering China, Zhang et al. (2021), after estimating the variation of fuel vehicles, found a decrease in CO_2 emission of 46.45% in the same period, also Wang and Su (2020) indicate that the outbreak of COVID-19 improves China's air quality in the short term and significantly contributes to global carbon emission reduction but in the long run, there is no evidence that this improvement will continue. Otmani et al. (2020) show significant reductions in PM10, SO2 and NO during the lockdown period in Salè City (North-Western Morocco). Lalas et al. (2021) estimated the emission reduction in Greece, comparing 2020 to the previous year, and they found a net 11,898 GWh decrease and a resulting emissions reduction of 3.48 MtCO2eq corresponding to 4.1%. Generally, we can confirm that COVID-19 lockdown led to a temporary reduction in energy demand and in the concentration of pollutants, and this effect is heterogeneous across different countries.

3 Data

3.1 Available data and sources

We collected data for fuels from the monthly oil and gas bulletin made available by the Italian Ministry of Ecological Transition (MITE) on its website (Ministero dell'ambiente e della sicurezza energetica 2023). Among the whole set of available information, the bulletin provides monthly data on gasoline, diesel, and LPG consumption (expressed in thousands of tons) for all the 107 Italian provinces. Our analysis takes into account monthly measurements of gasoline and diesel consumption from January 2015 to December 2021.

Data on fuel consumption are supplemented with further socio-economic features connected to the spatial and temporal dynamic of consumption. In Italy, tourism is one of the most important economic activities, accounting for more than 6% of total national value added (Eurostat 2023b). In Europe, the average share is around 4.5% of the aggregate value added. According to Massidda and Etzo (2012), habit formation, relative prices and per capita GDP are the main determinants of tourism. Also, for Italian tourists domestic and international destinations act as substitute goods. According to the literature, tourism acts as a key determinant of fuels consumption. In particular, Bakhat and Rosselló (2013) discuss the major driving factor for energy consumption in tourism, as accomodation, transportation, attractions, pleasure and leisure, and provide evidence of the relationship with gasoline and diesel consumption. Also, Nosheen et al. (2021) show empirical results indicating that tourism is responsible for increasing environmental degradation in the form of carbon emissions in the Asian region. Moreover, Trull et al. (2019) focus on the relation of energy consumption with tourism pressure, finding a higher electricity consumption elasticity in touristic areas and the greater portion of the total electricity consumption due to tourism activities, suggesting the use of tourism indicators to improve the electricity forecast in tourism areas. Indeed, tourism can be particularly relevant to explain the infra-annual dynamic of consumption (i.e., the seasonality), especially in coastal and touristic areas of Italy. In our application, we collected from the Italian Statistical Office (ISTAT) public database (ISTAT 2023) monthly data on the presence (stays) of tourists in accommodation facilities, summing Italian and foreign customers, at the provincial level from January 2015 to December 2021.

Italian provinces are quite heterogeneous both in terms of population and physical dimension. Thus, the total provincial consumption might be influenced by the size of the potential market. To control for the market size of fuels, we consider monthly data on resident population by province, which are provided by ISTAT. Population is then used to obtain three indicators. First, the monthly per capita fuel consumption, measured in kilograms per person; second, the resident population density, calculated as the ratio of the population to the surface (square kilometers) of the province; third, the tourists per capita variable, which measures the number of tourists staying in a given province in a given month relative to the number of residents.

Geographical information for each province, that is the centroid coordinates (i.e. longitude and latitude), administrative boundaries and surface (squared

kilometers) are provided by the NUTS 2021 nomenclature of Eurostat (Eurostat 2021). According to the NUTS classification, Italy counts 107 provinces (i.e., NUTS-3 level representing small regions for specific diagnoses), 21 regions (NUTS-2 level representing basic regions for the application of regional policies), each of them belonging to one of the five macro-regions (NUTS-1 level representing major socio-economic regions), that is North-West, North-East, Center, South, and Islands. In addition, Eurostat provides some classifications of provinces based on territorial characteristics, such as the degree of urbanization (*Urb-Degree*), metropolitan regions (*Metropolitan*), the presence of coastal (*Coastal*) or mountainous (*Mountain*) regions, and the bordering regions (*Border*) between European countries. (See Table 1 for further details).

Finally, to model the fuel consumption due to the heating and cooling of buildings, we consider the monthly average information on heating degree days (HDD) and cooling degree days (CDD) registered in each province. Such indicators are weather-based technical indexes designed to describe the energy requirements of buildings in terms of heating or cooling. Data for the whole period are made available by Eurostat (Eurostat 2023a).

The whole dataset contains information about $T = 84$ months across $N = 107$ provinces. Data collected so far are reshaped according to a long-shape (panel) format with $n = N \times T = 8988$ spatio-temporal observations for each variable. A synthesis of the available information is reported in Table 1.

3.2 Exploratory analysis of fuel consumption before COVID-19 pandemic

We choose the per capita consumption of gasoline and diesel fuel as target variables of our analysis. Before estimating the pandemic-induced changes in fuel consumption, we proceed with an exploratory analysis of the per capita quantities of gasoline and diesel being sold on the Italian market.

Figure 1a shows that the linear correlation between gasoline and diesel is positive and extremely strong (over 0.75 to 0.80 before the pandemic) nation-wide. No remarkable differences can be detected among the five macro-regions. The relationship becomes even stronger during the pandemic period; indeed, the linear correlation goes up to 0.90 and 0.95 in 2020 and 2021, meaning an almost-exact proportional relationship in the consumed volumes. The strong link between the two fuels highlights that in the Italian market the two commodities share common tendencies and that their market dynamics run in parallel. Moreover, although the two phenomena have different scales (per capita consumption of diesel fuel is considerably higher than gasoline consumption), our findings for gasoline and diesel consumption will be very similar and will lead to close suggestions and interpretations.

Furthermore, the aggregate distribution of the target variables (all provinces and all months combined together) during the pre-COVID period (2015–2019) depicted in Fig. 1b, raises several issues. First, a strong positive skewness and some potential outliers of the data, and, second, a strong spatial heterogeneity of consumption across areas. A striking North–South pattern is found where northern regions (yellow and green) structurally experience higher and more variable per capita

Table 1 List of available variables with source, unit of measure and short description

Variable name	Source	Unit of measure	Description
Gasoline total	MITE	Thousands of tons	Total gasoline consumption per month and province
Gasoline per capita	MITE/ISTAT	kg per capita	Per capita gasoline consumption per month and province
Diesel total	MITE	Thousands of tons	Total diesel consumption per month and province
Diesel per capita	MITE/ISTAT	kg per capita	Per capita diesel consumption per month and province
Population	ISTAT	Thousand people	Provincial resident population at the first day of each month
HDD	Eurostat	Celsius degrees below 18°	Monthly heating degree days per province
CDD	Eurostat	Celsius degrees exceeding 21°	Monthly cooling degree days per province
MetropArea	ISTAT	Dichotomous	ID for metropolitan areas (set to 0 for the remaining provinces)
Surface	Eurostat	km^2	Area/surface of the provinces
Density	MITE/Eurostat	People per km^2	Population density
Tourist stays	ISTAT	thousandths of people	Total (provincial) number of tourist stays for month
Tourist stays per capita	ISTAT	People per resident	Average monthly number of tourist stays per resident
Metropolitan	Eurostat	Categorical	$Border = 1$ if the province is classified as metropolitan area; $Border = 0$ if the province is not classified as metropolitan area
UrbanDegree	Eurostat	Categorical	$UrbanDegree = 3$ if perdominantly urban; $UrbanDegree = 2$ if intermediate; $UrbanDegree = 1$ if predominantly rural
Border	Eurostat	Categorical	$Border = 1$ if the province is a national border area; $Border = 0$ if the province is not a national border area
Coastal	Eurostat	Categorical	$Coastal = 3$ if more than 50% of population living within 50 km of the coastline; $Coastal = 2$ if the province own costal lines; coastal = 1 if the province does not own coastal lines
Mountain	Eurostat	Categorical	$Mountain = 4$ if 50% of population and 50% of surface are in mountain areas; $Mountain = 3$ if more than 50% of population live in mountain areas; $Mountain = 2$ if more than 50% of surface are in mountain areas; $Mountain = 1$ if other areas
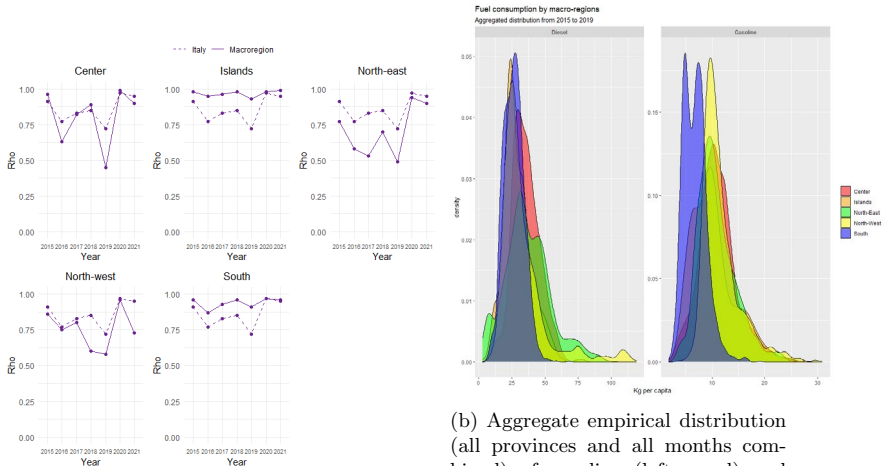
(a) Correlation between annual consumption of gasoline and diesel grouped by Eurostat NUTS-1 macro-regions.

(b) Aggregate empirical distribution (all provinces and all months combined) of gasoline (left panel) and diesel (right panel) registered during the pre-COVID period (from January 2015 to December 2019) grouped by Eurostat NUTS-1 macro-regions.

Fig. 1 Exploratory analysis of gasoline and diesel per capita consumption

consumption than southern Italy (blue). This clearly supports the hypothesis of spatial heterogeneity across the Italian territory. In the statistical models considered in the following sections, in addition to spatio-temporal dynamics, the presence of positive skewness will be taken into account through the use of GLM class regression models in which the distribution of the response variable is a Gamma variable. With respect to the presence of outliers, the time series up to December 2019 were analyzed on a province-by-province basis. In particular, the Hampel filter was employed on each series to identify values significantly above the bulk of the data. We identified 13 outliers (located in only 3 provinces) for gasoline per capita consumption (0.21% of total observations) and 17 values (located in 5 provinces, 3 of which were in common with gasoline) for motor diesel per capita consumption (0.27% of total observations). Outliers were replaced by values interpolated by Kalman filter for structural time series (trend + seasonality).

4 Empirical strategy

4.1 Out-of-sample predictive methodology

The main goal of the analysis is to quantify the impact of restrictions imposed during the COVID-19 pandemic between 2020 and 2021 on gasoline and diesel consumption at the Italian provincial level. We propose a predictive approach based on a two-folds out-of-sample procedure for spatio-temporal data (Cerqueira et al. 2020, 2017; Bergmeir et al. 2014), i.e., by segmenting the provincial time series into two

non-overlapping blocks (i.e., a training dataset and a test or validation dataset). This hold-out procedure, also referred to as 'last block evaluation' (Bergmeir and Benítez 2012), is commonly used in the time series domain to assess the predictive ability of models. In fact, the procedure requires that an initial part of the available observations is used for fitting a predictive model, while the last part of the time series is held out and used to test the model's forecasting performance (Cerqueira et al. 2020). The main advantage is that the temporal order of observations, thus their temporal dependence structure, is preserved.

In this paper, we employ the 'last block' principle for two purposes:

1. In a first stage, after introducing a variety of specifications of spatio-temporal regression models for provincial gasoline and diesel consumption, we intend to identify the model with the best forecasting capabilities based on various performance metrics;
2. The optimal model identified above is then used to make 'business-as-usual' spatio-temporal predictions of consumption during the pandemic. Here, the model's parameters are estimated using the pre-event observations (training dataset), while the business-as-usual predictions, and the subsequent forecast errors, are computed using the observations within the event (i.e., lockdown) period (i.e., test dataset). The forecast errors in the test sample are then used to quantify the impact of the pandemic within and between Italian provinces.

4.1.1 Prediction accuracy assessment and model selection

In the following sections, we will describe several spatio-temporal regression models used to predict gasoline and diesel consumption. The hold-out predictive strategy described above is used to select the best model among those proposed. Model performance is evaluated by considering only observations from January 2015 to December 2019, thus under the conditions being independent of the event of interest (i.e., the COVID-19 pandemic). Data from 2015 to 2018 are used as the training set, while 2019 is used as test set. The models are evaluated both out-of-sample (metrics are calculated on 2019 observations only) and in-sample (metrics are calculated on data from 2015 to 2018).

The literature offers a large choice of measures of forecasting accuracy based on the difference between predicted and actual values (Chicco et al. 2021). Using the training set, we calculate six in-sample metrics: Akaike Information Criterion (AIC), Adjusted AIC (AICc), Bayesian Information Criterion (BIC), Goodness of Fit Index (R2), Root Mean Square Error (RMSE) and Residual Standard Deviation (Sigma). Eventually, the out-of-sample forecasting accuracy metrics we considered are the following: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R2), Mean Absolute Percentage Error (MAPE), Symmetric Mean Absolute Percentage Error (SMAPE) (Chicco et al. 2021), and Weighted Absolute Percentage Error (WAPE) (Borowski and Zwolińska 2020).

In addition to the above out-of-sample criteria, in order to further evaluate the significance of the forecasting improvement between the estimated models, we perform pairwise Diebol-Mariano (DM) test (Diebold and Mariano 1995) comparing

pairs of models. The null hypothesis of the DM test is that the expected forecast loss is equal for two alternative forecasts, while under the alternative hypothesis we state that the second model provides better predictive accuracy.

4.1.2 Assessing COVID-19 impact on fuels consumption by forecasting

From the empirical literature on the effects of COVID on fuel consumption, it is well known that the reductions of demand generated by restraints was significant during lockdown periods (Zhang et al. 2021; Lalas et al. 2021). We also know taht restarts led to rebounds able to offset losses (Aruga et al. 2020). Therefore, using statistical techniques to study the statistical significance of reductions and following rebounds on the Italian case would lead to a contribution of limited relevance to the literature. Instead, we are interested in studying the spatio-temporal dynamics between and within Italian provinces, quantifying the variations that occurred in fuel consumption during the pandemic, relating them to the timing of pandemic propagation, the policies adopted by the central government, and attempting to explain homogeneity/heterogeneity based on spatial characteristics.

We propose a business-as-usual forecasting approach involving geostatistical regression models with several linear and non-linear predictors, and spatio-temporal components to account for structural features (e.g., seasonality) of fuels consumption. Business-as-usual predictive approaches have been extensively used to assess the impact of COVID-19 on local air quality (González-Pardo et al. 2022; Keller et al. 2021), on unemployment (Roopnarine et al. 2023), on air transport demand (Li et al. 2021). Other contributions tried to investigate the impact of the pandemic on energy issues, such as commercial energy demand prediction (Anik and Rahman 2021) or national oil consumption forecasting (Wang et al. 2022b). However, to the best of our knowledge, no contributions using spatio-temporal models to predict local (provincial) fuel consumption are made available.

To assess the pandemic impact, the training dataset consists of the observations ranging from January 2015 to December 2019 ($t = 1, \ldots, 60$ for each province s, for an overall number of observations equal to 6420), thus unaffected by the pandemic, whereby the parameters of a number of spatio-temporal models are estimated. The test dataset, on the other hand, is composed of monthly observations from 2020 to 2021 ($t = 1, \ldots, 60$ for each province s, for an overall number of observations equal to 2568).[2] Once a given model is estimated using the training set, predictions for the 24 months of the validation set are calculated for each province.

---

[2] It is worth noting that, again due to travel restrictions, tourism also suffered heavy contractions during the pandemic period. In particular, the recovery of pre-COVID attendance numbers of foreign tourists may require external intervention to speed up the recovery process (Provenzano and Volo 2022). Therefore, in order to avoid endogeneity and structural break issues generated by the pandemic period, we adopted a correction when calculating 'business-as-usual' forecasts. We assume that, under normal conditions, the monthly tourism flow of each province in 2020 and 2021 is equivalent to the monthly average calculated between 2015 and 2019.[3]

[3] Example: January 2020 and January 2021 will be equal to the average of January from 2015 to 2019. This is extended to each of the twelve months of the year.

Let $t = 1, \ldots, 84$ indicate a generic month in the entire period and $s = 1, \ldots, 107$ a generic Italian province. The impact of the pandemic is calculated as the difference between the realized (actual) and predicted consumption for each provincial time series. Employing the terminology of statistical learning, the pandemic effect for the generic month t in the test set and the generic province s coincides with the out-of-sample prediction error, i.e.,

$$\hat{\varepsilon}_{st} = y_{st} - \hat{y}_{st} \qquad \forall t \in 61, \ldots, 84 \tag{1}$$

where y_{st} is the actual consumption of gasoline or diesel for province s at time t, and $\hat{y}_{st}$ is the predicted consumption of gasoline or diesel for province s at time t.

4.2 Modelling fuel consumption

Gasoline and diesel per capita consumption are modeled using two classes of regression models. In the first case, we considered the class of Generalized Linear Models (GLMs), whereas in the second we considered the class of Generalized Additive Models (GAMs) (Wood 2006).

The two models are built in such a way to address both the spatial and temporal evolution of the fuel consumption, that is, we implement spatio-temporal GLMs and spatio-temporal GAMs. In addition, both classes take into account the skewed nature of the consumption as they admit that the response variable has a statistical distribution belonging to the exponential family. In particular, given the positive skewness, a Gamma distribution with logarithmic link function was chosen. GLMs hold the assumption of linearity in the relationships between variables (Wood 2006). This means that no matter what value a certain variable takes, a unit increase in it will always have the same effect on the outcome. In many situations this assumption is unrealistic. For example, in the case of spatial or time trends, it is unlikely that the data will show strongly linear tendencies, but will, instead, be subject to smooth but nonlinear patterns. GAMs are GLMs in which the predictor linear is a summation of smoothing (or smooth) functions of the covariates. The presence of smooth functions allows flexibility in modeling the relationship between response and covariates (and, in turn, permits a better approximation of the data since models are not subject to the linearity constraint). However, this implies the need to represent these smooth functions and determine how much they should be smoothed.

In both models, spatial and temporal trends are modeled by the same covariates, but with different functional shapes. Regarding the temporal component, the GLM includes the following covariates: a fixed effect for the *Month* of the year (reference value is January) and a linear trend for *Year* (taking value from 1 to 7). The spatial effect is introduced in the model using a fixed-effect term built using a set of 107 dummy variables, one for each province.

The GAM specifies the temporal component by including two smooth terms: seasonality is captured by the variable *Month*, which is included as a cyclic P-splines (Eilers and Marx 1996), whereas the medium-run temporal trend, i.e. *Year*, is modeled by a cubic B-spline basis with integrated squared second-derivative penalty (Wood 2017). The inclusion of a spline for the annual trend is justified by the fact

that over the training period (2015–2019) the fuel consumption in some regions (see Figure S5) shows smooth but not necessarily linear tendencies and therefore may not be well approximated by a strictly linear function. Considering the spatial component, it is useful to notice that the conventional smoothers have constraint based on the Euclidean distance between observations, which is not a proper measure of spatial distance (Miller and Wood 2014). Duchon Splines (Duchon 1977), also called D-splines, have been developed as a method of smoothing with respect to generalized distances and they can be defined as a generalization of the thin-plate smoothers (See Sect. 5.5 in Wood 2006). Regarding the computation, D-splines aim to obtain the generalized distances between points and then use a multidimensional scaling to find a configuration in Euclidean space that approximate the true distance, then they perform the smoothing Miller and Wood (2014). Therefore, we considered bivariate Duchon splines for handling the smooth spatial trend surface.

Recall that y_{st} is the actual consumption of gasoline or diesel for province s at month t. The GLM, hereinafter also referred as $M0$, can be expressed as follows:

$$\log[E(y_{st}|\mathbf{x}_{st})] = \mathbf{x}'_{st}\boldsymbol{\beta} \qquad (2)$$

where y_{st} is *Gamma* distributed and $\mathbf{x}_{st}$ is a vector containing the following exogenous covariates: *Year*, *Month*, *TouristStaysPerCapita*, *HDD*, *CDD*, *Density*, *Surface*,*Metropolitan*, *UrbDegree*, *Coastal* and *Border*, one dummy for each *Province*.

The GAM, hereafter also referred as $M1$, can be expressed as follows:

$$\begin{aligned}
\log[E(y_{st}|\mathbf{z}_{st}, \mathbf{v}_{st})] = \ &\mathbf{z}'_{st}\boldsymbol{\beta} + f_1(TouristStaysPerCapita_t) + f_2(Month_t) + f_3(Year_t) \\
&+ f_4(CDD_t) + f_5(HDD_t) + f_6(Density_t) + f_7(Surface_t) \\
&+ f_8(Latitude_s, Longitude_s)
\end{aligned}$$
$$(3)$$

where $y_{st} \sim Gamma$ and $\mathbf{z}_{st}$ is a vector of exogenous variables composed by *Metropolitan*, *UrbDegree*, *Coastal* and *Border*. Note that the spline terms representing the spatial and temporal trends, i.e., f_2, f_3 and f_8, are described above, whereas the other spline components are cubic B-spline basis with integrated squared second-derivative penalty.

4.3 Model tuning for GAM

In order to improve the spatio-temporal fitting of the GAM to the greatest extent possible, while not running into overfitting, we perform a parameter tuning exercise on the smooth function associated with the geographic coordinates. The goal is to determine the optimal number of nodes of the Duchon spline (i.e., f_8 in Eq. 4) that would guarantee the best fitting and predictive ability under equal other conditions (i.e., leaving other smooth functions unchanged). Thus, we estimate several GAM specifications with an increasing number of nodes for the bivariate D-spline using the following grid: $k = [30, 40, 50, 60, 70, 80, 88 : 1 : 106]$, where $a : 1 : b$ states an arithmetic progression of step 1 from a to b. The choice of the optimal model is

conducted by considering both the in-sample and out-of-sample forecasting performance criteria described in Sect. 4. Model tuning is performed for both diesel and gasoline per capita consumption separately. In the following, we will refer to the optimized model as *M*2.

4.4 GAM with spatio-temporal interaction

Finally, we propose a fourth model including a space-time (seasonal) non-linear interaction, similarly to Augustin et al. (2009); Nigussie et al. (2023). This component allows the seasonal smooth to varying smoothly within the space dimensions and can be interpreted as each province has different seasonality component in fuel consumption across the Italian territory. Recalling the specification of model *M*1, we specify a two-dimensional marginal smooth for *Latitude* and *Longitude* using a Duchon spline, and a marginal cyclic P-splines basis expansion for *Month*. Here, in addition, we add a three-dimensional tensor product interaction of coordinates and month. In this context, it is important to emphasize the role of tensor product smoothers. Considering coordinates and month jointly we can not assume that the relationship between fuel consumption and coordinates should be roughly as "wiggly" as the relationship between fuel consumption and month, using the same smoothing parameter, because they have different natural scales. Tensor product smooths assume that each variable have its own smooth term, creating multi-dimensional smooth terms by multiplying the values of each one-dimensional basis by each value of the second. In the next sections, this model with space-time interaction will be referred to as the *M*3 model, which can be expressed as follows:

$$\log[E(y_{st}|\mathbf{z}_{st}, \mathbf{v}_{st})] = \mathbf{z}'_{st}\boldsymbol{\beta} + f_1(TouristStaysPerCapita_t) + f_2(Month_t) + f_3(Year_t)$$
$$+ f_4(CDD_t) + f_5(HDD_t) + f_6(Density_t) + f_7(Surface_t)$$
$$+ f_8(Latitude_s, Longitude_s) + f_9(Latitude_s, Longitude_s, Month_t)$$
$$(4)$$

where the space-time seasonal smooth term, i.e., f_9, is a three-dimensional tensor product interaction.

5 Results and discussion

5.1 GAM tuning for GAM

The results of GAM fine tuning for gasoline are shown in Figure S6 (in-sample metrics) and Figure S7 (out-of-sample metrics) of the Supplementary Materials, while results for diesel are reported in Figures S18 and S19. Similar conclusions about the optimal number of nodes are supported by either gasoline or diesel.

As expected, in-sample metrics suggest that model fitting improves significantly as the number of nodes, hence the complexity, of the spline component representing the spatial trend increases. All the metrics show monotonically decreasing patterns. Also, R^2 is maximized around $k = 92$, while all the other indices are minimized

around the same value. Considering out-of-sample indices although their pattern is not exactly monotonic, all the metrics agree about the optimal value, which can be detected at $k = 91$ or $k = 92$. To preserve a balance between predictive power, fitting to the data and model simplicity (Hastie et al. 2009) (i.e., the parsimony criterion), we choose an optimal value $k^* = 92$. Indeed, when considering a number of nodes greater than $k = 92$, none of the considered statistics provide remarkable improvements. Note, in addition, two points. First, in line with expectations (see Sect. 7 of Hastie et al. 2009), the in-sample and out-of-sample RMSE are approximately the same order of magnitude for low number of nodes (low complexity), while their distance grows as complexity increases, and in particular the out-of-sample error tends to be higher than the in-sample error. Second, generally speaking, gasoline models perform much better than diesel models. In fact, both percentage and absolute errors (MAE and RMSE) are always lower for gasoline. This latter aspect can be traced to the greater complexity of the diesel data, which leads to less accurate predictions and thus higher errors overall. In particular, we refer to the fact that the diesel time series (see, for instance, Figures S5 and S17) are much less regular (less seasonal) and suffering from several structural breaks, while the gasoline series are very regular in pattern and not subject to shocks.

5.2 Predictive model selection

In Fig. 2, we show the out-of-sample forecasting performance of the four estimated models for gasoline (the corresponding figure for diesel is Figure S8). Recall that $M0$ is the GLM, $M1$ is the non-optimized GAM, $M2$ is the optimized GAM with $k^* = 92$ nodes for the spatial Duchon spline, and $M3$ is the optimized GAM with space-time seasonal interaction.

On the right side, we show the estimated out-of-sample metrics computed using the four models. On the left side, we represent the ratio of the metrics (i.e., the relative metrics) using as benchmark model M0 (i.e., the GLM). The purpose of the latter is to show the improvement obtained passing from a purely linear to a semi-parametric approach.

Indeed, for what concern the gasoline, from Fig. 2, it is clear that the simple GAM is not enough to improve the prediction accuracy, while it requires at least a preprocessing step to find out the optimal order of spline basis expansion. In fact, while $M1$ seems to perform worse than the GLM, in turns $M0$ performs worse than the optimized GAM ($M2$ - blue) and the GAM with space-time interaction ($M3$ - red) models. In particular, the gain in RMSE and MAE moving from $M0$ to $M3$ is -5.4% and -8.2%, respectively. Whereas, when moving from $M0$ to $M2$, the gains for RMSE and MAE are, respectively, -3.5% and -6%. This improvement can be explained by the spatial spline optimization (model $M2$) and by inclusion of non-linear smooth space-time interaction. In fact, the choice of a large number of bases for the spatial spline, but still less than the maximum number (corresponding to the number of provinces, i.e., 107) allows for the estimation of a spatial surface similar
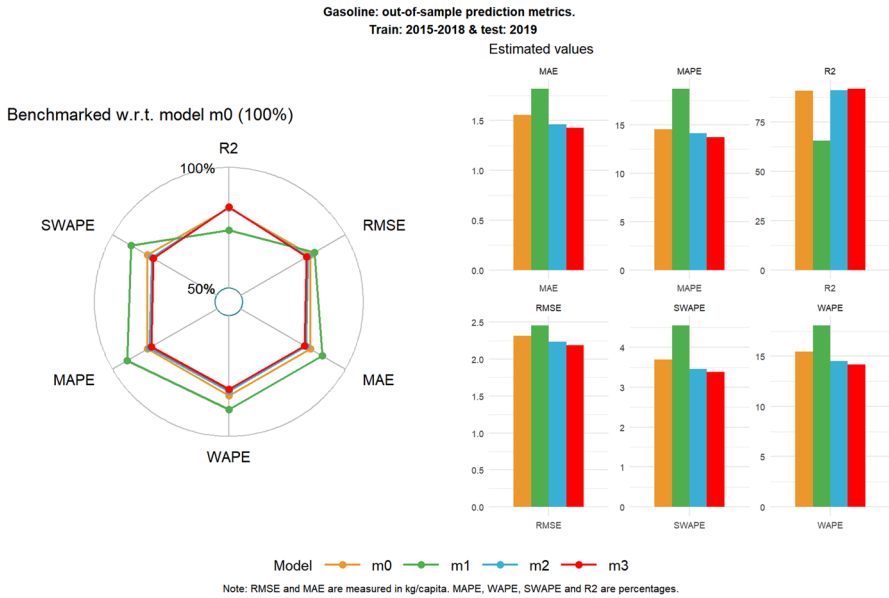
Fig. 2 Gasoline: out-of-sample prediction metrics computed using 2019 as test set and 2015–2018 as training set

to a fixed effect, but with a greater degree of flexibility and less complexity.[4] On the other hand, the inclusion of the seasonal-spatial interaction further improves the fit of the $M3$ model, which can better capture the spatiotemporal characteristics of the phenomenon. In support of this, Fig. 3, which shows the estimated smooth spatial surface for each month of the year, is iconic. The interaction term demonstrates that in the winter months, southern Italy experiences decreases in gasoline consumption per capita, while the north (the Northeast in particular) experiences some increases. In contrast, in the summer months (July and August, especially), driven by strong tourist presences, consumption in the south and islands experiences large increments.

Finally, to support the analysis on the choice of the optimal forecasting model, in Figure S9 we report the results of the Mariano-Diebold tests calculated on the out-of-sample time series pairs. The graph suggests that, for each level of significance and for each horizon, model $M3$ is better than the others (it produces statistically more accurate predictions), while there is substantial predictive equivalence between model $M0$ and model $M1$.

The statements asserted above for gasoline only partly apply to diesel. In fact, results for gasoline and diesel appear to be slightly different, but comparable. It is noticeable from the materials in the Supplementary Materials that there is a

---

[4] Remember that the number of parameters to be estimated in the GLM for the spatial effect is about 107, while it is 92 in the $M2$ and $M3$ models.
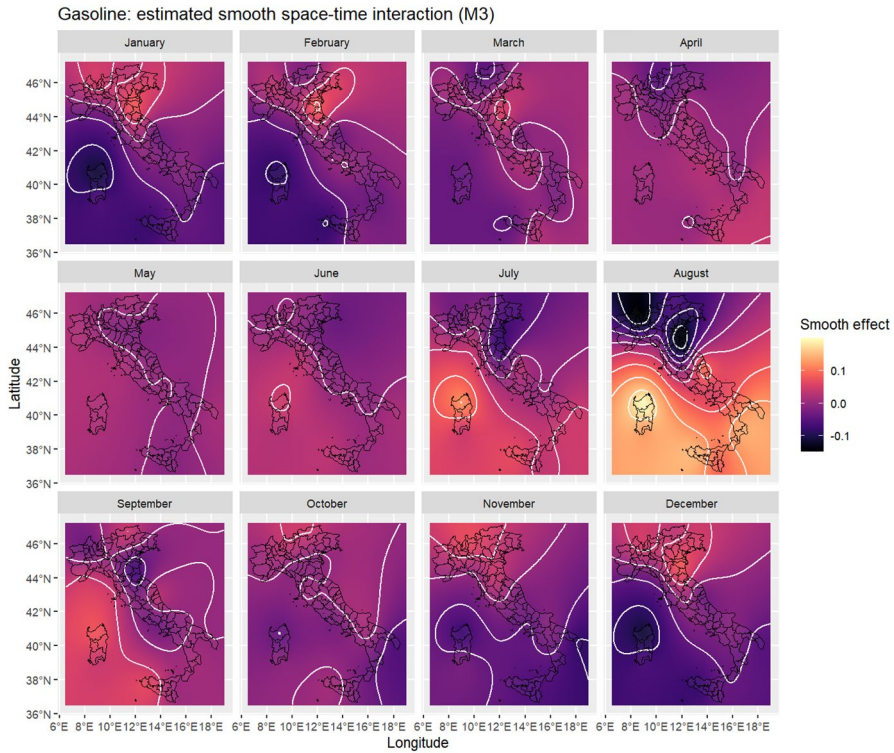
Fig. 3 Gasoline: estimated monthly smooth spatial surface (M3)

substantial predictive equivalence between the $M0$, $M2$ and $M3$ models. Figure S8 shows that the out-of-sample metrics are virtually identical, and the Mariano-Diebold test (Figure S21) also shows that for each horizon and for each significance level, all four models are statistically equivalent (with the exception of $M3$, which outperforms $M2$). However, the estimated spatio-temporal interaction with $M3$ (Figure S22) is similar in both space-time distribution and magnitude to the corresponding for gasoline.

In summary, while the results on gasoline suggest a distinctly better predictive ability of the GAM model with spatio-temporal interaction ($M3$) over the others, for diesel, each of the four models is identical. In order to maintain interpretive homogeneity, we choose to employ $M3$ as the predictive model for business-as-usual fuel consumption (and thus for variations generated by COVID-19 restrictions). Also, without leading to loss in the contents, we only report the relevant results for per capita gasoline consumption during the pandemic period. However, the two fuels will be extensively discussed through a comparison of relevant economic and statistical findings. Detailed results for gasoline are provided in Section S1 of the Supplementary Information, while detailed results of the diesel consumption analysis are provided Section S2.
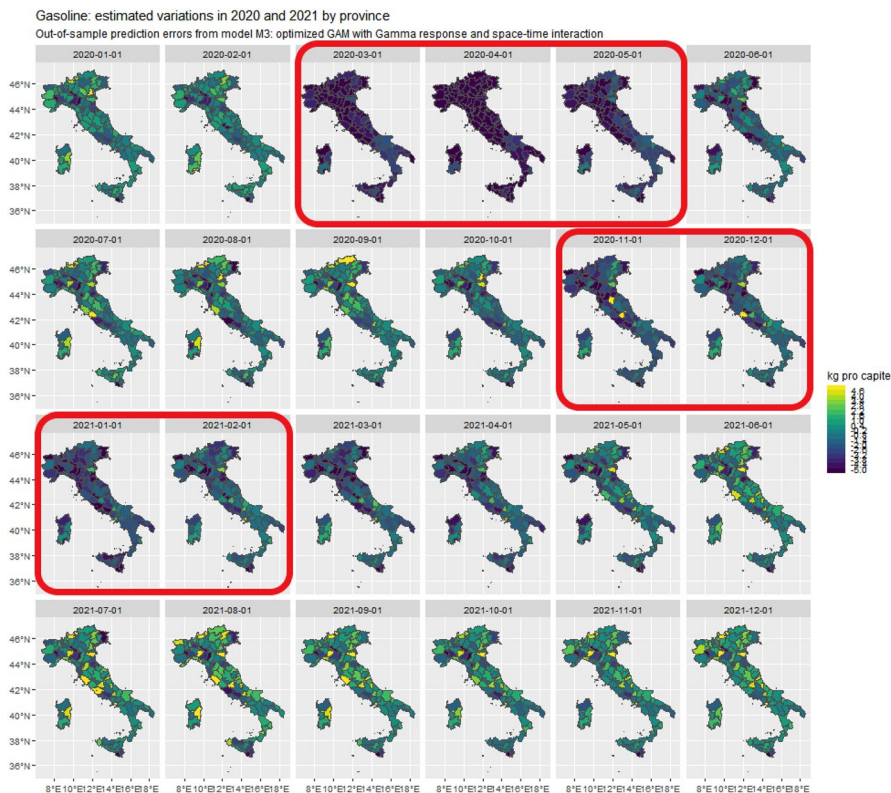
Fig. 4 Maps of estimated variations (out-of-sample prediction errors) of per capita gasoline consumption in 2020 and 2021 by province. Estimates are computed using Model M2

5.3 Evolution of fuels during the pandemic

In Fig. 4, we depict the spatial distribution of the difference between the observed amounts of gasoline being consumed and the business-as-usual predictions obtained by the optimized GAM with space-time seasonal interaction (*M*3) during 2020 and 2021. The maps are useful to interpret the variations in gasoline consumption following the evolution of COVID-19 restrictions in Italy. In particular, we focus on the first lockdown period (March to May 2020) and on the second lockdown period (November 2020 to February 2021), both highlighted by red boxes surrounding the maps.

In January and February 2020 the deviation seems to be very small, indeed the prevailing color scale is green, that is, the color associated with a null variation (actual and predicted values coincide). This indicates that before the pandemic, actual consumption was aligned with that predicted by the model. From March 2020 the whole Italian territory has been affected by the COVID. Starting from March 4th, the Italian central government establishes the full lockdown only for Lombardy

region (located at the center of the North) and a few days later, the lockdown is then extended throughout Italy. Consumption of fuels immediately fell quite uniformly in all the Italian provinces during March and April. This collapse seems to be slightly more pronounced in Northern and Central Italy, where the reductions reached values between −5 kg/person and −10 kg/person (regarding diesel, the drops reach in some provinces around −17 kg/person to −19 kg/person). The North–South gap, albeit moderate, can be justified by the uneven evolution of the epidemic in the first few weeks. In fact, despite the level of restrictions was the same nation-wide, the regions of Northern Italy were the most affected by the COVID-19 virus, causing the collapse of the health and welfare system, and leading to high casualties in the population.

In the following two months (May and June 2020), there is a gradual recovery in fuel consumption, which can easily be explained by the progressive lessening of restrictions. As for the lockdown period, the recovery seems faster in southern provinces and islands. During summer 2020 (from June to September), the general level of per capita consumption returns to be very close to that predicted by the model. However, in several provinces (mainly in the north-central) there were still negative or at most nil variations. In fact, during that period, although the restrictions imposed were weaker (such as the obligation to wear a mask indoors), human activities, industries and businesses were still limited. Moreover, it can be noticed the presence of several specific provinces that show remarkably high rebounds in consumption compared to the others. In particular, considering touristic areas near the Mediterranean see or the Alps, fuels consumption appears to be higher than that estimated by the models, with positive differences above +4.5kg/person for gasoline and +10 kg/person for diesel. Specifically, gasoline consumption is moderately elevated in eastern Sardinia (Western island), Trentino-Alto Adige (North-East) and Tuscany coast (Central Italy), while diesel consumption seems very high in Sardinia, Tuscan coast and several provinces in Apulia, Calabria (Southern Italy) and Sicily. These provinces traditionally have been destinations for domestic and international summer tourism. Italian households, being unable to choose foreign destinations, have turned to the main domestic tourist hubs, fostering a growth in fuel consumption.

In November 2020, fuel consumption begins to slow down again due to a new increase in COVID infections and the introduction of new restrictive measures. The figure shows that from November 2020 to April 2021 the prediction errors for both gasoline and diesel are negative and most of the provinces tend to assume colors close to dark blue (i.e., strongly negative variations). This period coincides with the second wave of COVID-19 pandemic. In this phase the heterogeneity of the unexpected variation in consumption across Italy is remarkable. Such heterogeneity is consistent with the implementation of the different levels of restrictions (white, yellow, orange or red areas) at the regional and sometimes at the provincial level. This restriction system has been in effect form November 6th until the end of the health emergency, but starting from the month of June the level of restrictions throughout Italy becomes minimal.

During the last six months of 2021 we assist to a second recovery in fuel consumption, represented by higher discrepancies among actual values and predictions.
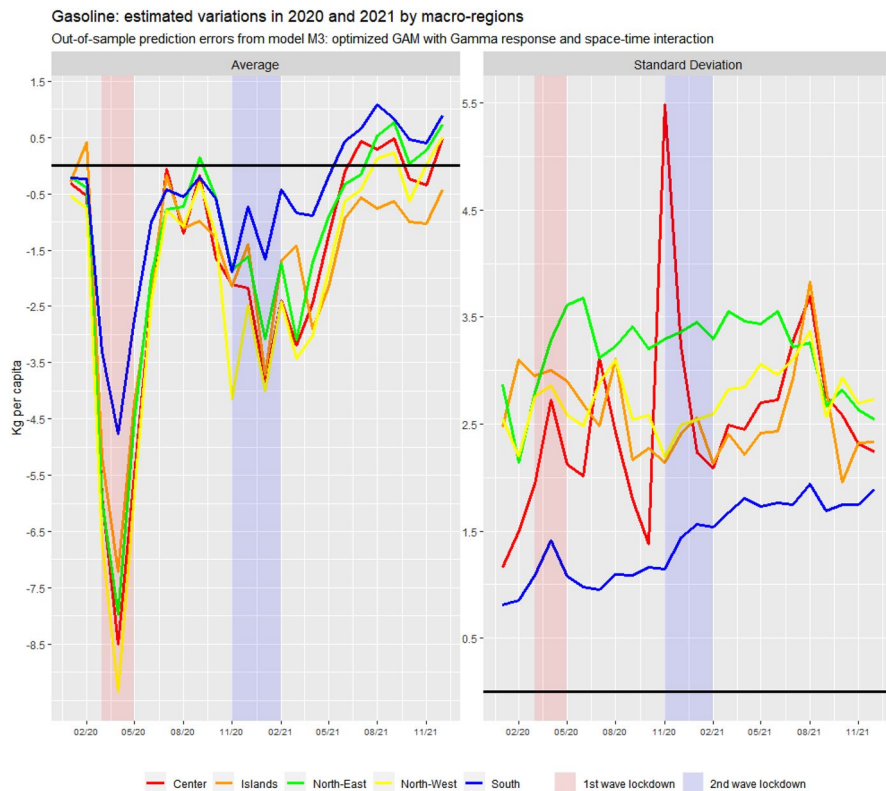
Fig. 5 Estimated variations in 2020 and 2021 of per capita gasoline consumption by macro-regions. Left panel: estimated monthly average variation per macro-region. Right panel: estimated monthly standard deviation per macro-region

As in 2020, also during 2021 the rebound effect is particularly evident during the summer period. However, if compared to the first recovery period, one can notice that the unexpected increase in consumption seems to be stronger (mostly for gasoline) and more diffuse across the provinces. Actually, considering the months of July, August and September 2021, provinces with colors tending toward yellow are numerous in Central Italy and coastal areas (Sardinia and Tuscany, mainly) or Alpine areas. In contrast, some northern provinces record consumption below forecasts for the entire year. Moreover, the heterogeneity among provinces is more accentuated, especially for diesel consumption.

The contrast between North and South during both the first and second waves of restriction is further emphasized by Fig. 5 (see Figure S13 in the Supplementary Information for diesel). The first plot summarizes the mean (left panel) and the standard deviation (right panel) of the out-of-sample prediction error, aggregating provinces into macro-regions. For both gasoline and diesel we can observe that the collapse in consumption that coincides with the first wave of COVID is actually less accentuated in Southern Italy and Islands, while it is

Corr. in 2020 and 2021

	Center	Islands	North-East	North-West	South
South	0.96	0.91	0.96	0.97	1
North-West	0.98	0.94	0.98	1	0.97
North-East	0.98	0.93	1	0.98	0.96
Islands	0.94	1	0.93	0.94	0.91
Center	1	0.94	0.98	0.98	0.96

Corr. in 2020

	Center	Islands	North-East	North-West	South
South	0.97	0.98	0.98	0.99	1
North-West	0.98	0.97	0.98	1	0.99
North-East	0.99	0.97	1	0.98	0.98
Islands	0.98	1	0.97	0.97	0.98
Center	1	0.98	0.99	0.98	0.97

Corr. in 2021

	Center	Islands	North-East	North-West	South
South	0.98	0.92	0.95	0.98	1
North-West	0.98	0.87	0.98	1	0.98
North-East	0.97	0.78	1	0.98	0.95
Islands	0.85	1	0.78	0.87	0.92
Center	1	0.85	0.97	0.98	0.98

Var. corr. from 2020 to 2021

	Center	Islands	North-East	North-West	South
South	0	-0.06	-0.03	-0.02	0
North-West	0	-0.11	0	0	-0.02
North-East	-0.01	-0.19	0	0	-0.03
Islands	-0.13	0	-0.19	-0.11	-0.06
Center	0	-0.13	-0.01	0	0

Cos. dist. in 2020 and 2021

	Center	Islands	North-East	North-West	South
South	0.95	0.88	0.96	0.94	1
North-West	0.99	0.97	0.98	1	0.94
North-East	0.99	0.94	1	0.98	0.96
Islands	0.96	1	0.94	0.97	0.88
Center	1	0.96	0.99	0.99	0.95

Cos. dist. in 2020

	Center	Islands	North-East	North-West	South
South	0.99	0.99	0.99	1	1
North-West	0.99	0.99	0.99	1	1
North-East	0.99	0.98	1	0.99	0.99
Islands	0.99	1	0.98	0.99	0.99
Center	1	0.99	0.99	0.99	0.99

Cos. dist. in 2021

	Center	Islands	North-East	North-West	South
South	0.76	0.45	0.8	0.69	1
North-West	0.98	0.91	0.97	1	0.69
North-East	0.97	0.78	1	0.97	0.8
Islands	0.86	1	0.78	0.91	0.45
Center	1	0.86	0.97	0.98	0.76

Var. cos. dist. from 2020 to 2021

	Center	Islands	North-East	North-West	South
South	-0.22	-0.54	-0.19	-0.31	0
North-West	-0.01	-0.08	-0.02	0	-0.31
North-East	-0.02	-0.2	0	-0.02	-0.19
Islands	-0.13	0	-0.2	-0.08	-0.54
Center	0	-0.13	-0.02	-0.01	-0.22

Eucl. dist. in 2020 and 2021

	Center	Islands	North-East	North-West	South
South	7.34	6.67	5.85	9.06	0
North-West	2.95	5.04	4.06	0	9.06
North-East	2.54	4.33	0	4.06	5.85
Islands	4.09	0	4.33	5.04	6.67
Center	0	4.09	2.54	2.95	7.34

Eucl. dist. in 2020

	Center	Islands	North-East	North-West	South
South	5.83	3.94	4.74	7.32	0
North-West	2.56	4.12	3.37	0	7.32
North-East	1.94	2.1	0	3.37	4.74
Islands	2.48	0	2.1	4.12	3.94
Center	0	2.48	1.94	2.56	5.83

Eucl. dist. in 2021

	Center	Islands	North-East	North-West	South
South	4.46	5.38	3.43	5.34	0
North-West	1.48	2.91	2.27	0	5.34
North-East	1.64	3.78	0	2.27	3.43
Islands	3.25	0	3.78	2.91	5.38
Center	0	3.25	1.64	1.48	4.46

Var. Eucl. dist. from 2020 to 2021

	Center	Islands	North-East	North-West	South
South	-1.37	1.43	-1.32	-1.98	0
North-West	-1.08	-1.22	-1.1	0	-1.98
North-East	-0.29	1.68	0	-1.1	-1.32
Islands	0.76	0	1.68	-1.22	1.43
Center	0	0.76	-0.29	-1.08	-1.37

Fig. 6 Gasoline: (dis)similarity measures for out-of-sample residuals by macro-regions

more pronounced in the Northern provinces. Note that North-West macroregion includes the Lombardy region, which was the epicenter of the first COVID-19 wave. During the first lockdown, the Northern provinces clearly suffered the greatest impact, with an average reduction of −48 kg/person for gasoline and −23 kg/person for diesel, respectively. On the opposite, the Southern provinces suffered losses around −4.5kg/person for gasoline and −15 kg/person for diesel.

The second lockdown was milder for all areas, with reductions ranging from −3 kg/person to −13 kg/person for diesel and up to −4 kg/person for gasoline. The fastest macro-region to achieve again neutrality is Southern Italy (blue line), both considering gasoline and diesel. In fact, on average, it will always show non-negative changes from May 2021 to December. Gasoline forecasts for the Northern and Central provinces align with the values observed throughout the second half of 2021, alternating from positive to negative signs. The islands, on the other hand, succeed in returning (on average) to breakeven only at the end of 2021 and only for diesel. In addition, it is interesting to note that all the diesel predictions for non-Southern provinces report stable negative average values for the whole of 2021, showing the difficulty in recovering pre-pandemic consumption levels.

As further insights for examining the spatio-temporal dynamics of variation, in Fig. 6 (see Figure S12 for diesel), we report some measures of global similarity and dissimilarity suggested by Cassisi et al. (2012) and calculated on the time series of out-of-sample residuals by macro-regions. We took Pearson's linear correlation index and cosine similarity as similarity metrics, and Euclidean distance as dissimilarity. The measures were calculated both on the total sample (2020 and 2021) and on the two separate years, allowing us to assess how they changed over the two periods. All metrics confirm the previous argument: (1) during 2020, the regions are very similar to each other (correlation and cosine close to 1), showing that the reaction to the restrictions of spring 2020 were very similar across the country; (2) from 2020 to 2021, the gap between the Northern provinces and the South and Islands increases significantly (both Euclidean distance increases, while cosine decreases); (3) the gap between South and Islands increases significantly (recall that the South recovered much faster to pre-COVID consumption than the others).

Eventually, the estimation of the standard deviation by macro-regions provides us with insights into within-area variability. It clearly emerges that northern regions show strong and stable inter-province variability compared to the South. Interestingly, provinces of Central Italy alternate period with high and low variability. For the entire period, southern Italian provinces show both lower average impact and lower inter-province variability, confirming the hypothesis of a territorial homogeneity in the fall and recovery. When comparing the two fuels, one can also notice that after the first wave of COVID-19, the unexpected changes in each macro-region is more homogeneous for gasoline than for diesel consumption (i.e., the variability for gasoline is always lower than variability for diesel). However, while the variability of gasoline in the North increases around the restrictions period, when it comes to diesel, the fuel the inter-provincial variability is systematically higher.

6 Conclusions

In this paper, we proposed a geostatistical analysis aiming at estimating the variation of gasoline and diesel consumption occurred in the Italian provinces as a consequence of the COVID-19 pandemic during 2020 and 2021. We employed generalized linear models (GLMs) and generalized additive models (GAMs) to model the per capita fuel consumption as functions of several exogenous variables, such as the touristic presence, and demography (used as proxy of market size). By means of regression splines, we allowed the spatial and temporal trends to be smooth surfaces whose functional form were optimized in order to guarantee the best predictive and fitting abilities. Moreover, we include a smooth spatio-temporal interaction that properly capture the seasonal characteristic of different areas. The main findings can be summarized as follows:

1. The national mobility restrictions imposed to fight the spread of COVID-19 in the first wave (March to May 2020) affected homogeneously the whole country. Strong reduction in fuel consumption were detected for both gasoline and diesel. However, diesel fuel registered the most impressive reductions;

2. The second wave lockdown (Winter 2020–2021), characterized by weaker and local restrictions, determined heterogeneous impacts across the territory: southern provinces experienced the smallest average reductions, while northern provinces (especially the North-West) recorded the strongest. In addition, southern Italy and some mountainous (alpine and appenini), pushed primarily by domestic tourism, reacted more quickly to re-openings leading to major rebounds;

3. While southern provinces showed more concentrated variations (less inter-regional variability), northern Italy shows strong infra-regional heterogeneity. In particular, some of the more tourist-oriented areas achieved in reaching pre-pandemic levels during 2021, while for other provinces permanent fuel consumption reductions have been estimated throughout 2021.

These results might be relevant for national and local policy makers interested in mobility and energy markets. The presence of strong socioeconomic heterogeneity among Italian areas is something well known and historically validated (De Philippis et al. 2022). Our results confirm the existence of a divergence between North and South of Italy related to the pandemic evolution. Indeed, northern regions were characterized by extended and heterogeneous losses in the levels of fuel consumption, while southern provinces showed more reactivity in recovering from the restrictions and provided an infra-regional variability less pronounced.

7 Supplementary information

Extended results for gasoline and diesel consumption can be found in the Supplementary Information (SI) material.

Declarations

Conflict of interest None.

References

Anik, A.R., Rahman, S.: Commercial energy demand forecasting in Bangladesh. Energies **14**(19), 6394 (2021)

Aruga, K., Islam, M.M., Jannat, A.: Effects of covid-19 on indian energy consumption. Sustainability (Switzerland), **12**(14). https://doi.org/10.3390/su12145616 (2020)

Augustin, N.H., Musio, M., Von Wilpert, K., Kublin, E., Wood, S.N., Schumacher, M.: Modeling spatiotemporal forest health monitoring data. J. Am. Stat. Assoc. **104**(487), 899–911 (2009). https://doi.org/10.1198/jasa.2009.ap07058

Bakhat, M., Rosselló, J.: Evaluating a seasonal fuel tax in a mass tourism destination: a case study for the balearic islands. Energy Econ. **38**, 12–18 (2013). https://doi.org/10.1016/j.eneco.2013.02.009

Bergmeir, C., Benítez, J.M.: On the use of cross-validation for time series predictor evaluation. Inf. Sci. **191**, 192–213

Bergmeir, C., Costantini, M., Benítez, J.M.: On the usefulness of cross-validation for directional forecast evaluation. Comput. Stat. Data Anal. **76**, 132–143. https://doi.org/10.1016/j.csda.2014.02.001 (2014)

Borowski, M., Zwolińska, K.: Prediction of cooling energy consumption in hotel building using machine learning techniques. Energies, **13**(23). https://doi.org/10.3390/en13236226 (2020)

Bressan, G., Zaccomer, G.P.: New geographies of fuel purchase during the covid-19 pandemic: evidence from the italian northeastern border zone. (pp. 179–196). Springer Nature, Switzerland (2023)

Buechler, E., Powell, S., Sun, T., Astier, N., Zanocco, C., Bolorinos, J., Rajagopal, R.: Global changes in electricity consumption during covid-19. iScience, **25**(1). https://doi.org/10.1016/j.isci.2021.103568 (2022)

Cassisi, C., Montalto, P., Aliotta, M., Cannata, A., Pulvirenti, A.: Similarity measures and dimensionality reduction techniques for time series data mining. In: InTech (ed.) Advances in Data Mining Knowledge Discovery and Applications (chap. 3) (2012)

Cereda, D., Tirani, M., Rovida, F., Demicheli, V., Ajelli, M., Poletti, P., Barone, A.: The early phase of the covid-19 outbreak in Lombardy, Italy. arXiv preprint arXiv:2003.09320 (2020)

Cerqueira, V., Torgo, L., Smailović, J., Mozetič, I.: A comparative study of performance estimation methods for time series forecasting. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA) (pp. 529–538) (2017)

Cerqueira, V., Torgo, L., Mozetič, I.: Evaluating time series forecasting models: an empirical study on performance estimation methods. Mach. Learn. **109**(11), 1997–2028 (2020). https://doi.org/10.1007/s10994-020-05910-7

Chicco, D., Warrens, M.J., Jurman, G.: The coefficient of determination rsquared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. PeerJ Comput. Sci. **7**, 1–24 (2021). https://doi.org/10.7717/PEERJ-CS.623

Chrulski, T.: Statistical review of the italian gas transmission system operator under conditions of the covid-19 pandemic and the supply restriction from the russian federation. Gases **2**(4), 134–145 (2022)

De Philippis, M., Locatelli, A., Papini, G., Torrini, R.: La crescita dell'economia italiana e il divario nord-sud: trend storici e prospettive alla luce dei recenti scenari demografici [italian economic growth and the north-south gap: Historical trends and future projections in light of the recent demographic scenarios] . Bank of Italy Occasional Paper (683) (2022)

Dembińska, I., Barczak, A., Szopik-Depczyńska, K., Dul, I., Koliński, A., Ioppolo, G.: The impact of the covid-19 pandemic on the volume of fuel supplies to eu countries. Energies, **15**(22). https://doi.org/10.3390/en152284390 (2022)

Diebold, F., Mariano, R.: Comparing predictive accuracy. J. Bus. Econ. Stat. **13**(3), 253–63 (1995)

Du, J., Rakha, H.A., Filali, F., Eldardiry, H.: Covid-19 pandemic impactson traffic system delay, fuel consumption and emissions. Int. J. Transport. Sci. Technol. **10**(2), 184–196. https://doi.org/10.1016/j.ijtst.2020.11.003 (2021)

Duchon, J.: Splines minimizing rotation-invariant semi-norms in sobolev spaces. In: W. Schempp & K. Zeller (Eds.), Constructive Theory of Functions of Several Variables (pp. 85–100). Springer, Berlin (1977)

Edomah, N., Ndulue, G.: Energy transition in a lockdown: an analysis of the impact of covid-19 on changes in electricity demand in Lagos Nigeria. Global Transitions **2**, 127–137 (2020). https://doi.org/10.1016/j.glt.2020.07.002

Eilers, P.H.C., Marx, B.D.: Flexible smoothing with B-splines and penalties. Stat. Sci. **11**(2), 89–121. https://doi.org/10.1214/ss/1038425655 (1996)

Ercolani, V., Guglielminetti, E., Rondinelli, C.: Fears for the future: saving dynamics after the covid-9 outbreak (2021)

Eurostat Energy statistics - cooling and heating degree days (nrg chdd) [Online Database]. Retrieved from https://ec.europa.eu/eurostat/cache/metadata/en/nrgchddesms.htm (2023)

Eurostat: Fuel consumption still affected by covid-19 in 2021. Retrieved from https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220407-1 (2022)

Eurostat: Nuts 2021 - nomenclature of territorial units for statistics [Online Database]. Retrieved from https://ec.europa.eu/eurostat/web/nuts/background (2021)

Eurostat: Tourism: €572 billion gross value added in the eu (Report). Author. Retrieved from https://ec.europa.eu/eurostat/web/products eurostat-news/w/WDN-20230414-1 (2023)

Ferrero, G., Pisani, M., Tasso, M.: Policy mix during a pandemic crisis: a review of the debate on monetary and fiscal responses and the legacy for the future. Report Banca d'Italia, 267–320. https://doi.org/10.32057/0.QEF.2021.623 (2022)

Fezzi, C., Fanghella, V.: Real-time estimation of the short-run impact of covid-19 on economic activity using electricity market data. Environ. Resource Econ. **76**(4), 885–900 (2020). https://doi.org/10.1007/s10640-020-00467-4

González-Pardo, J., Ceballos-Santos, S., Manzanas, R., Santibáñez, M., Fernández-Olmo, I.: Estimating changes in air pollutant levels due to covid-19 lockdown measures based on a business-as-usual prediction scenario using data mining models: a case-study for urban traffic sites in Spain. Sci. Total Environ. **823**, 153786, https://doi.org/10.1016/j.scitotenv.2022.153786 (2022)

Guglielminetti, E., Rondinelli, C.: Consumption and saving patterns in italy during covid-19. Report Banca d'Italia. https://doi.org/10.32057/0.QEF.2021.620 (2021)

Güngör, B.O., Ertuğrul, H.M., Soytaş, U.: Impact of covid-19 outbreak on turkish gasoline consumption. Technol. Forecast. Soc. Change, **166**, 120637. https://doi.org/10.1016/j.techfore.2021.120637 (2021)

Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer, New York (2009)

Il Sole Ore: Cose che noi umani - storia del covid-19 (Report). Retrieved from https://lab24.ilsole24ore.com/storia-coronavirus/ (2023)

ISTAT: Retrieved from http://dati.istat.it/index.aspx?queryid=7059 (2023)

Keller, C.A., Evans, M.J., Knowland, K.E., Hasenkopf, C.A., Modekurty, S., Lucchesi, R.A., Pawson, S.: Global impact of covid-19 restrictions on the surface concentrations of nitrogen dioxide and ozone. Atmos. Chem. Phys. **21**(5), 3555–3592 (2021). https://doi.org/10.5194/acp-21-3555-2021

Lai, S., Ruktanonchai, N.W., Zhou, L., Prosper, O., Luo, W., Floyd, J.R., Du, X.: Effect of non-pharmaceutical interventions to contain covid-19 in China. Nature **585**(7825), 410–413 (2020)

Lalas, D., Gakis, N., Mirasgedis, S., Georgopoulou, E., Sarafidis, Y., Doukas, H.: Energy and ghg emissions aspects of the covid impact in greece. Energies, **14**(7). https://doi.org/10.3390/en14071955 (2021)

Li, X., de Groot, M., Bäck, T.: Using forecasting to evaluate the impact of covid-19 on passenger air transport demand. Decision Sci. n/a(n/a). https://doi.org/10.1111/deci.12549 (2021)

Li, Z., Ye, H., Liao, N., Wang, R., Qiu, Y., Wang, Y.: Impact of covid-19 on electricity energy consumption: a quantitative analysis on electricity. Int. J. Electr. Power Energy Syst. **140** (2022). https://doi.org/10.1016/j.ijepes.2022.108084

Liu, Z., Ciais, P., Deng, Z., Lei, R., Davis, S.J., Feng, S., Schellnhuber, H.J.: Near-real-time monitoring of global co2 emissions reveals the effects of the covid-19 pandemic. Nat. Commun. **11**(1), https://doi.org/10.1038/s41467-020-18922-7 (2020)

Massidda, C., Etzo, I.: The determinants of Italian domestic tourism: Apanel data analysis. Tourism Manag. **33**(3), 603–610, https://doi.org/10.1016/j.tourman.2011.06.017 (2012)

Menneer, T., Qi, Z., Taylor, T., Paterson, C., Tu, G., Elliott, L.R., Mueller, M.: Changes in domestic energy and water usage during the UK covid-19 lockdown using high-resolution temporal data. Int. J. Environ. Res. Public Health **18**(13). https://doi.org/10.3390/ijerph18136818 (2021)

Miller, D.L., Wood, S.N.: Finite area smoothing with generalized distance splines. Environ. Ecol. Stat. **21**(4), 715–731 (2014). https://doi.org/10.1007/s10651-014-0277-4

Ministero dell'ambiente e della sicurezza energetica, S.e.e.m.: Bollettino petrolifero italiano [Online Database]. Retrieved from https://dgsaie.mise.gov.it/bollettino-petrolifero (2023)

Nicolini, G., Antoniella, G., Carotenuto, F., Christen, A., Ciais, P., Feigenwinter, C., Papale, D.: Direct observations of co2 emission reductions due to covid-19 lockdown across european urban districts. Sci. Total Environ. **830** (2022). https://doi.org/10.1016/j.scitotenv.2022.154662

Nigussie, T.Z., Zewotir, T.T., Muluneh, E.K.: Seasonal and spatial variations of malaria transmissions in northwest ethiopia: evaluating climate and environmental effects using generalized additive model. Heliyon **9**(4). https://doi.org/10.1016/j.heliyon.2023.e15252 (2023)

Nosheen, M., Iqbal, J., Khan, H.U.: Analyzing the linkage among co2emissions, economic growth, tourism, and energy consumption in the Asian economies. Environ. Sci. Pollut. Res. **28**(13), 16707–16719 (2021). https://doi.org/10.1007/s11356-020-11759-z

Otmani, A., Benchrif, A., Tahri, M., Bounakhla, M., Chakir, E.M., El Bouch, M., Krombi, M.: Impact of covid-19 lockdown on pm10, so2 and no2 concentrations in salé city (Morocco). Sci. Total Environ. **735** (2020). https://doi.org/10.1016/j.scitotenv.2020.139541

Ou, S., He, X., Ji, W., Chen, W., Sui, L., Gan, Y., Bouchard, J.: Machine learning model to project the impact of covid-19 on us motor gasoline demand. Nat. Energy **5**(9), 666–673 (2020). https://doi.org/10.1038/s41560-020-0662-1

Pelagatti, M., Maranzano, P.: Assessing the effectiveness of the italian risk-zones policy during the second wave of covid-19. Health Policy, **125**(9), 1188–1199. https://doi.org/10.1016/j.healthpol.2021.07.011 (2021)

Perra, N.: Non-pharmaceutical interventions during the covid-19 pandemic: a review. Phys. Rep. https://doi.org/10.1016/j.physrep.2021.02.001 (2021)

Presidenza del Consiglio dei Ministri Italia, Decreto del presidente del consiglio dei ministri 3 novembre 2020 (Report). Gazzetta Ufficiale della Repubblica Italiana. Retrieved from https://www.gazzettaufficiale.it/eli/gu/2020/11/04/275/so/41/sg/pdf (2020)

Presidenza del Consiglio dei Ministri Italia, Decreto del presidente del consiglio dei ministri 8 marzo 2020 (Report). Gazzetta Ufficiale della Repubblica Italiana. Retrieved from https://www.gazzettaufficiale.it/eli/id/2020/03/08/20A01522/sg (2020)

Provenzano, D., Volo, S.: Tourism recovery amid covid-19: the case of Lombardy, Italy. Tourism Econ. **28**(1), 110–130 (2022). https://doi.org/10.1177/13548166211039702

Ray, R., Singh, V., Singh, S., Acharya, B., He, Y.: What is the impact of covid-19 pandemic on global carbon emissions? Sci. Total Environ. **816** (2022). https://doi.org/10.1016/j.scitotenv.2021.151503

Roopnarine, P.D., Abarca, M., Goodwin, D., Russack, J.: Economic cascades, tipping points, and the costs of a business-as-usual approach to covid-19. Front. Phys. **11**, 183 (2023)

Rouleau, J., Gosselin, L.: Impacts of the covid-19 lockdown on energy consumption in a Canadian social housing building. Appl. Energy **287** (2021). https://doi.org/10.1016/j.apenergy.2021.116565

Rugani, B., Caro, D.: Impact of covid-19 outbreak measures of lockdown on the Italian carbon footprint. Sci. Total Environ. **737** (2020). https://doi.org/10.1016/j.scitotenv.2020.139806

Şahin, U., Ballı, S., Chen, Y.: Forecasting seasonal electricity generation in european countries under covid-19-induced lockdown using fractional grey prediction models and machine learning methods. Appl. Energy **302** (2021). https://doi.org/10.1016/j.apenergy.2021.117540

Smith, L.V., Tarui, N., Yamagata, T.: Assessing the impact of covid-19 on global fossil fuel consumption and co2 emissions. Energy Econ. **97** (2021). https://doi.org/10.1016/j.eneco.2021.105170

Sung, J., Monschauer, Y.: Changes in transport behaviour during the Covid-19 crisis, IEA: International Energy Agency. France. https://policycommons.net/artifacts/1343388/changes-in-transport-behaviour-during-the-covid-19-crisis/1955530/. Accessed 18 Oct 2023. CID: 20.500.12592/zkxzf1 (2020)

Trull, O., Peiró-Signes, A., Carlos García-Díaz, J.: Electricity forecasting improvement in a destination using tourism indicators. Sustainability (Switzerland), **11**(13). https://doi.org/10.3390/su11133656 (2019)

Visco, I.: Considerazioni finali del governatore - relazione annuale anno 2020, https://www.bancaditalia.it/pubblicazioni/interventi-governatore/integov2021/cf_2020.pdf (2021)

Wang, Q., Li, S., Zhang, M., Li, R.: Impact of covid-19 pandemic on oil consumption in the united states: a new estimation approach. Energy **239** (2022). https://doi.org/10.1016/j.energy.2021.122280

Wang, Q., Li, S., Zhang, M., Li, R.: Impact of covid-19 pandemic on oil consumption in the united states: a new estimation approach. Energy, **239**, 122280. https://doi.org/10.1016/j.energy.2021.122280 (2022)

Wang, Q., Su, M.: A preliminary assessment of the impact of covid-19 on environment . a case study of china. Sci. Total Environ. **728**. https://doi.org/10.1016/j.scitotenv.2020.138915 (2020)

Wood, S.N.: Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC, 2nd Edition (2017)

Wood, S.N.: P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data. Stat. Comput. **27**(4), 985–989 (2017). https://doi.org/10.1007/s11222-016-9666-x

Zhang, X., Li, Z., Wang, J.: Impact of covid-19 pandemic on energy consumption and carbon dioxide emissions in China's transportation sector. Case Stud. Thermal Eng. **26** (2021). https://doi.org/10.1016/j.csite.2021.101091

Zhang, D., Sogn-Grundvåag, G.: Credit constraints and the severity of covid-19 impact: empirical evidence from enterprise surveys. Econ. Anal. Policy **74**(337), 349 (2022). https://doi.org/10.1016/j.eap.2022.03.005

Springer