Energy Conversion and Management: X 16 (2022) 100294

Contents lists available at ScienceDirect

Energy Conversion and Management: X

journal homepage: www.sciencedirect.com/journal/energy-conversion-and-management-x

Planning a novel regional methane network: Demand forecasting and economic evaluation

Tommaso Barbiero, Carlo Grillenzoni *

IUAV: Institute of Architecture, University of Venice, 30135, Italy

ARTICLE INFO

Keywords:
Autoregressive models
Energy planning
Infrastructure design
Spatial prediction

ABSTRACT

Planning the use of non-renewable energy resources is a crucial issue for social and environmental development, both at global and local scales. Although fossil fuel depletion is still far in time, their rational use is essential to prevent energy shortages and avoid economic and political crises. Natural gas (NG) is one of the preferable sources since it is comparatively clean and cheap, and its energetic power covers a wide range of purposes, especially in domestic heating and electric generation. This paper aims to design and evaluate a novel methane infrastructure in the insular region of Sardegna, which has not yet an NG network. We provide a statistical methodology to forecast gas demand at the municipal level in the absence of historical data on gas consumption, by using spatial autoregressive (SAR) models with exogenous (X) variables. Specifically, the model identification is performed by fitting various SARX systems on the data of the main cities of continental Italy, for which data on gas consumption and its covariates (X) are available. Next, the most suitable model is selected by optimizing the out-of-sample prediction performance on a subset of continental cities which are similar to those of Sardegna from the environmental and social viewpoint and are held outside of the estimation. Then, the coefficients of the model with the smallest prediction errors are applied to the spatial covariates of Sardegna cites, obtaining the predicted per-capita gas demand. Finally, using the global demand and its spatial diffusion, the entire gas network, at central and peripheral levels, is designed and its economic sustainability is evaluated. This analysis considers several categories of fixed and variable costs, together with the revenues of the private companies of gas transportation and distribution, that are involved in the realization of the project. The analysis demonstrates the economic viability of gas infrastructure for Sardegna under 70 % of the market penetration rate, as it allows a break-even point within 15 years.

Introduction

The use of energy is both an economic, an infrastructural and an environmental problem. To prevent energy shortages, rational exploitation of the available sources is necessary; since fossil fuels have no regeneration capacity, their utilization must be carefully planned over time. In particular, taxation policies may modulate the relationships between demand and supply, so as to reduce consumption without introducing rationing measures. A theoretical scheme for non-renewable energy exploitation is provided by Hubbert [34], who modeled the stock of mining resources over time, as a bell-shaped curve. Although the oil peak is almost reached during the CoViD-19 lockdown [35], its decline may be slowed by new deposit discoveries and extraction technologies. This issue is less urgent for natural gas (NG), as its reserves are much greater; however, a *gas peak* may occur in the next decades and global energy planning is necessary to make its decline smoother so as to avoid sudden energy crises. In general, the rational approach to planning should be to use fossil fuels to develop energy technologies that exploit renewable sources.

Among fossil fuels, NG is a comparatively clean and cheap energy resource, owing to its low CO_2 emission, costs of production and transportation, e.g. Economides et al., [19]. Thus, it represents both a solution to the problems of environmental pollution and economic development, e.g. Afgan et al., [1]. Its importance is also stressed by the "European Energy Security Strategy" which states that diversification of NG suppliers is important to prevent shortages of energy; in particular, the European Commission [22] envisages the strengthening of the gas Southern Corridor, connecting Europe to the Middle East and Central Asia. However, in view of the recent crisis Ukraina-Russia, it could also reconsider the project of the GALSI pipeline [26], connecting Europe to

* Corresponding author.
E-mail address: carlog@iuav.it (C. Grillenzoni).

https://doi.org/10.1016/j.ecmx.2022.100294
Received 30 June 2022; Received in revised form 30 August 2022; Accepted 1 September 2022
Available online 7 September 2022

Algeria and Central Africa through Sardegna region.

New technologies, such as heat pumps and photovoltaic systems are important for decarbonization and may delay the depletion of fossil fuels; however, they are currently not affordable to most people, so NG usage must be regulated in a balanced way. This also means that its underuse yields a condition of non-profitability for the companies in the NG supply chain, which may increase the use of more pollutant fuels, such as coal. On the other side, excessive exploitation leads to quick depletion, with possible economic crises, as NG is essential for industrial activities and transportation. Since the costs of extraction and distribution will increase in the future, in the long run the NG should be limited to essential uses. Indeed, various studies, e.g. Kesicki [43], suggest that buildings heating should be provided by renewable technologies only, in agreement with decarbonization targets and the reduction of pollution.

Studies on planning gas networks

According to the previous remarks, planning NG infrastructures is crucial for energy management; this involves the design of transportation pipelines, distribution networks and also liquid natural gas (LNG) facilities (e.g. deposits and regasifiers). The demand for LNG is increasing as it is a flexible form of gas supply, more independent of international political and economic constraints.

Main gas infrastructures include gathering pipelines, trunk pipelines and delivering pipelines, see Wu et al., [74]. Gathering pipelines drive gas from the points of extraction to trunk pipelines. Trunk pipelines are long-distance and high-pressure pipelines, transporting large amounts of fuel; they are also equipped with several compression stations and valves for the delivery of the gas to local distributors [4]. Finally, local distribution networks move gas at relatively low flow and pressures to the final consumers, but pumping plants are still needed to move the gas; for the engineering aspects see Mokhatab et al., [53].

From the design viewpoint, the major issue is to define the optimal number of delivery points and their decompression stations. In general, there is a trade-off between the goal of reaching the largest number of NG consumers and the length of the low-pressure distribution networks. Indeed, the scenario of a network that serves *all* households is impossible, in that the revenues of NG sold in remote areas do not cover the costs. Moreover, trunk pipelines must not cross cities, landslides and seismic areas, in order to avoid risks of accidents. In general, the design and optimization of a gas infrastructure regard technical aspects, spatial coverage and land planning.

The academic literature on NG infrastructure design usually focuses on trunk pipelines [33]; it deals with the optimization of multi-objective functions subject to physical, technical and economic constraints. For example, Vasconcelos et al., [71] calculate the maximum flow in a network with multiple points of reception and delivery by using an algorithm based on linear programming; Percell et al., [57] compute the steady-state flow in a gas network which minimizes the resource consumption and maximize the throughput by means of non-linear programming. Similarly, Ruan et al., [64] minimize the life-cycle costs of pipelines and compressors, and [52] find the optimal flow and compression rate by modeling the gas network as a tree structure. Under the steady operation condition, Martin et al., [48] provide further examples of optimization for costs, revenues and flows, also with progressive expansion of the network [15]. More recently, Wang et al., [73] present a synergy planning method for designing a gas network in a regional system. Flow velocity, pipe diameter, plant capacity, network and station locations are optimized through genetic algorithms (GA); the resulting infrastructure can reduce annual costs by 32 %.

The approach of Dodds et al., [17] is distinctive as treats NG planning from a macroeconomic viewpoint; it considers the behavior of all actors involved in the gas supply chain of the UK. The authors use the Markal allocation model of Loulou et al., [47], which estimates the energy dynamics from a multiperiod perspective. Given certain required energy services, a Markal model computes the minimum global cost by simultaneously taking into account the equipment investments, operating decisions and primary energy supply. These factors are affected by targets and constraints that iteratively modify the demand for energy in the next period, thus influencing future investments. The levels of service of the energy infrastructure are defined under the condition in which the suppliers exactly produce the quantities demanded by the consumers (that is a condition of intertemporal partial equilibrium).

As regards the management of gas infrastructures: Pantaleo et al., [56] develop a mixed integer linear programming to combine biomass and natural gas for heat and power generation in urban areas. The design optimizes the allocation of multi-biomass sources at both spatial and temporal levels, which may result competitive with natural gas in district heating networks. Villicaña et al., [72] present a strategic planning scheme to satisfy the national demand for natural gas through various sources which are sustainable from the economic and environmental viewpoint. A multi-objective optimization provides a solution that allows a 66 % gas autonomy to Mexico based on shale gas. Alves et al., [3] investigate the behavior of the gas market in Brazil in relation to several gas suppliers. Using a dynamic optimization model, they show a reduction of the distribution cost by 18 %. Finally, Soha et al., [66] use geographical information systems (GIS) and multi-criteria decision-making to select suitable locations for power-to-gas (PtG) facilities. The approach shows that centralized solutions are 20 % cheaper than decentralized scenarios in Hungary.

In planning an NG network at the regional level, it is important to analyze the factors which drive the NG demand and the socio-environmental conditions of the region. Forecasting the NG demand and its spatial spread is essential to determine the length of the network, the number of compression stations, pressures and diameters of the pipes, etc.. Classical econometrics use time-series data and dynamic regression models to forecast aggregate demand for gas; the models are linear as Bianco et al., [11], or nonlinear as Szoplik [68]. However, time-series data on NG consumption are not available in regions that do not use methane and where a new gas infrastructure must be built; therefore, other numerical methods are needed to estimate the potential demand in these cases.

The case study and the proposal

Sardegna island is not yet equipped with an operating NG network, but the region is now involved in a new global infrastructure project. After the year 2000, local institutions put effort toward municipal distribution networks (e.g. Regione Sardegna [59]), and a subsequent project named GALSI aimed to connect the gas wells near Hassi R'Mel (Algeria) to continental Italy passing through Sardegna [26]. This project was promoted by the Algerian NG exporter Sonatrach, by Sardegna regional government and Italian gas import companies (Edison, Enel, Wintershall, Hera Group), see Yousfi [75]. Despite these efforts, the GALSI pipeline was set aside by the EU Commission due to the failure of shareholders to reach commercial agreements [23].

Anyway, both the local government [61] and the Italian Minister of Economic Development (MiSE) [50] stressed the importance of a complete regional gas network. This led private companies Snam SpA [67] and Società Gasdotti Italia (SGI) to outline a new project based on liquid natural gas (LNG). This project involves an infrastructure with naval terminals, storage deposits, LNG regasifiers and high-pressure transportation pipelines. Given the complexity and the costs of facilities and raw materials (the LNG gas usually costs 2–3 times the ordinary NG; see [20]), accurate analysis and forecasting of the potential demand are necessary.

In this work, we want to estimate the potential NG demand (both for private consumption and economic activities) of Sardegna without using time series data, because they are not available. Our approach consists of using spatial auto-regressive (SAR) models with a wide range of explanatory variables (X) available at the territorial level. The coefficients of the models are estimated on the data of the provinces of continental Italy whose social and environmental features are similar to Sardegna; next, they are applied to the municipal X-data of the Island. According to Goulard et al., [28], there is *not* a single method for obtaining spatial forecasts; therefore, we consider various SARX systems and select the model which has the best out-of-sample prediction performance on a subset of continental cities, which are homogeneous to Sardegna.

Finally, we use the forecasts of the NG demand to design the spatial pattern of the NG infrastructure (location of LNG terminals and path of the transportation pipeline) subject to conditions of economic sustainability. The economic analysis deals with the costs and revenues of the main pipelines and urban networks. The assessment is based on specific assumptions about the plan of work and considering a realistic penetration rate of NG in the energy market of the region. The article demonstrates the economic sustainability of complete gas infrastructure, as the global equilibrium point may be reached within 15 years.

The plan of the work is as follows: Section 2 illustrates the prediction method based on regression models in presence of spatial autocorrelation. Section 3 describes the variables and the data involved and Section 4 presents the estimated models and the maps of the predicted gas consumption for Sardegna municipalities. Finally, Section 5 deals with the quantitative aspects of gas demand and supply and Section 6 evaluates the economic sustainability of the planned infrastructure.

Methods and models

The nature of spatial data can be summarized by the statement of Tobler: "everything is related to everything else, but near things are more related than distant things" [70]. Socio-economic and environmental variables are common examples, in which proximity effects are present and may influence their causal relationships (Dubin [18]). These effects define the concept of spatial autocorrelation (SACR), which occurs when the level of a variable in a location, depends on its level in neighboring points (LeSage [46]).

A spatial prediction strategy

Owing to the lack of time series on the NG consumption in Sardegna, a necessary strategy is to carry out predictions on spatial data. Thus, let $\{y_i\}_{i=1}^n$ be the (unknown) gas consumption of the n municipalities of Sardegna, and $\{x_{ki}\}_{k=1}^m$ their m socio-economic and environmental covariates. The SARX model which connects these quantities is given by the regression equation:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_m x_{mi} + \rho y_{i-1} + e_i, \quad i = 1, 2, \cdots, n, \tag{1}$$

where β_k,ρ are coefficients, e_i are residuals and y_{i-1} is the nearest-neighbor (NN) of y_i in the *north* direction, i.e. it is the value of y in the city which is closest to the i-th unit in the upper half-plane. This term accounts for the spatial dependence of y_i and avoids autocorrelations of errors e_i, which causes inferential and forecasting problems to the model (1). The north–south (NS) ordering of y_{i-1}, with respect to y_i, is termed *causal* in lattice data (e.g. Grillenzoni [29]), as it is similar to the past-present relationship in time series; hence, it allows standard statistical properties.

Given the observable regressors x_{ki}, the predictor $\widehat{y}_i$ of the model (1) requires the value of the coefficients. Our strategy is to estimate β_k,ρ on the data of a set of cities of continental Italy that are homogeneous to those of Sardegna, and for which data of gas consumption and covariates Y_j, X_{kj} are available. Thus, let the SARX model of the continental cities be.

$$Y_j = \beta_0 + \beta_1 X_{1j} + \beta_2 X_{2j} + \dots + \beta_m X_{mj} + \rho Y_{j-1} + e_j, \quad j = 1, 2, \cdots, N \tag{2}$$

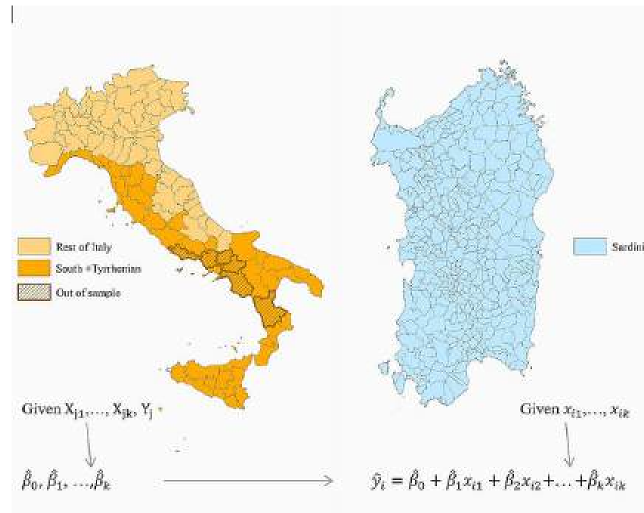

Fig. 1. Model building and spatial prediction strategy based on Eqs. (1)-(3). Sample sizes: N=102 is the total number of Italian provinces with data on natutal gas consumption; N^*=46 is the subset of South and Tyrrhenian provinces; n_o=6 are the provinces for forecast evaluation; n=377 is the total number of Sardegna municipalities to forecast.

this corresponds to model (1); in particular, $e_j \sim \text{IN}(0, \sigma_e^2)$ are independent Normal residuals.

Under the north restriction of the contiguous terms Y_{j-1}, the coefficients of (2) can be consistently estimated with conventional methods, such as ordinary least squares (OLS). Further, the social-based variables Y_j, X_{kj} are divided by the population (to get per capita quantities), so as to avoid dimensionality effects that produce outliers (which make OLS estimates biased). When the estimates $\widehat{\rho}$,$\widehat{\beta}_k$ are obtained, then the recursive predictions of the model (1) are computed as.

$$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{1i} + \widehat{\beta}_2 x_{2i} + \dots + \widehat{\beta}_m x_{mi} + \widehat{\rho} \ \ \widehat{y}_{i-1}, \quad i = 1, 2, \cdots, n \tag{3}$$

starting from suitable initial values, e.g.$\widehat{y}_0 = 0$, on the north border of the region. If p_i is the population of the i-th municipality, then the predicted amount of the regional NG consumption will be given by $\widehat{T}_n = \sum_{i=1}^n \widehat{y}_i p_i$. This is the basis for designing the size of the NG infrastructure.

In the above scheme, the statistical methodology is relatively standard; however, it must be properly *calibrated* in order to be effective. The first issue deals with the selection of the training data for the auxiliary model (2). Our case study considers two groups: The first is the entire continental Italy plus Sicilia, while the second is the set of southern and Tyrrenian cities only. The latter group has social and physical characteristics similar to Sardegna, with mild winters and an agro-industrial-based economy. In this group, a further subset, which is even closer to Sardegna, is kept outside of parameter estimation for out-of-sample forecasting evaluation (see Fig. 1). The second issue deals with the estimation methods of the model (2). Besides the OLS technique, the least absolute deviation (LAD) estimator is more robust (insensitive) to outlying observations (which yield large residuals) and, therefore, is more reliable in forecasting.

Fig. 1 summarizes our building strategy of the model (2), which involves two sets of training data (that of the entire Peninsula and that of Tyrrhenian regions only) and two estimation methods (OLS and LAD). This scheme yields, at least, *four* solutions for $\widehat{\beta}_k, \widehat{\rho}$, whose prediction ability is evaluated on the subset of 6 continental cities which are similar to those of Sardegna and are kept out of the estimation. The coefficients of the best model are then applied to the predictor (3), to infer the total NG consumption and its regional distribution.

The advantage of this strategy is to provide detailed forecasts of all n

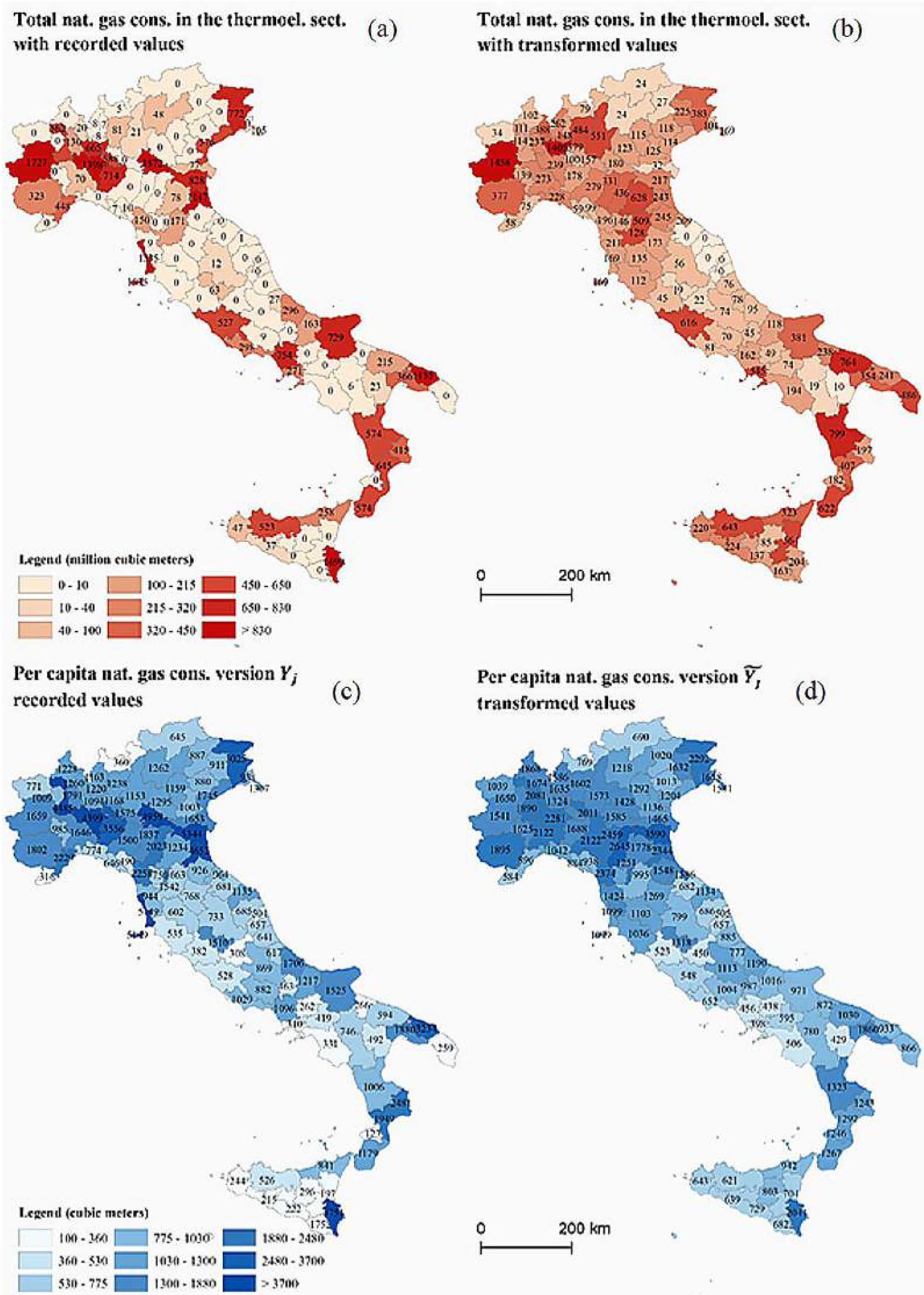

Fig. 2. Spatial distribution, over Italian provinces, of the natural gas consumption in the year 2016: (a,b) Total thermoelectric consumption; (c,d) Per-capita consumption. Computations are made with the formulas (8) (Panels a,c) and (9) (Panels b,d).

= 377 municipalities of Sardegna, without historical data on their gas consumption. This is similar to forecast the sales of a new product in a commercial market, by only knowing the general features of buyers. In the absence of time-series data, this approach has no viable alternative, apart from the naïve solution of using the average per capita consumption $\overline{Y}_o$ of a similar region and multiplying it by the local populations: $\tilde{y}_i = \overline{Y}_o\, p_i$. However, every municipality has its own environmental, urban and social structure, characterized by the X-variables, which determine the gas demand. It is the choice of the training data and estimation algorithms that determine the reliability of forecasts; as in time series, this can be checked on a small portion (6/102) of provinces in Fig. 1.

Technical aspects

Model (2) is a spatial autoregression with exogenous variables (SARX), see LeSage [46]. By defining the vectors $\boldsymbol{y} = [Y_1, Y_2, \cdots, Y_N]'$, $\boldsymbol{\beta} = [\beta_0, \beta_1, \cdots, \beta_m]'$, etc., Eq. (2) can be written as.

$$\boldsymbol{y} = \boldsymbol{X\beta} + \rho \boldsymbol{Wy} + \boldsymbol{e}, \quad \boldsymbol{e} \sim \mathrm{N}\left(\boldsymbol{0}, \boldsymbol{I}\sigma_e^2\right), \tag{4}$$

where $\boldsymbol{X} = \{X_{kj}\}$is an $N \times (m+1)$ matrix of regressors and $\boldsymbol{W}$ is a sparse *contiguity* matrix that connects each Y_j to its nearest neighbor (NN) in the north direction, i.e. $\boldsymbol{Wy} = Y_{j-1}$. With this constraint, $\boldsymbol{W}$ is lower triangular and has a single 1 by row, not by columns as units may share the same north NN. The lagged term Y_{j-1} copes with the spatial dependence, which must be represented in the model (4) to have uncorrelated residuals. Indeed e_j must be unpredictable by definition, so as to allow the efficiency of OLS estimates and consistency of their standard errors.

In *lattice* data as images, the NS ordering of $\boldsymbol{W}$ is termed *causal* as it is similar to the unidirectional (past-present) relationship in time-series [29]. The causal ordering can also be extended to *areal* data and is a necessary condition both for the consistency of OLS estimates of $\boldsymbol{\beta}, \rho$ and recursivity of the predictor (3). These two properties are strictly related, because if $\boldsymbol{W}$ is non-causal, then the elements of $\boldsymbol{Wy}$ and $\boldsymbol{e}$ in Eq. (4) are correlated; this violates the condition of independence errors-regressors and yields bias in OLS estimates (see Copiello and Grillenzoni [16]). To solve this problem in general matrices $\boldsymbol{W}$, various alternative estimators have been proposed, based on instrumental variables (Kelejian et al., [42]), generalized moments (Lee [45]) and indirect inference (Bao et al., [8]). These involve various computational difficulties, such as separate estimation of $\boldsymbol{\beta}, \rho$ (see Appendix F); however, the OLS method is adequate in our case, because we work under the constraint of triangular $\boldsymbol{W}$, which is suitable for spatial forecasting.

To derive the OLS estimator, define $\mathbf{z}_j = \left[1, X_{1j}, X_{2j}, \cdots, X_{mj}, Y_{j-1}\right]'$ and $\boldsymbol{\delta} = [\beta_0, \beta_1, \cdots, \beta_m, \rho]'$, then rewrite Eq. (2) as $Y_j = \boldsymbol{\delta}'\mathbf{z}_j + e_j$ and Eq. (4) as $\boldsymbol{y} = \boldsymbol{Z\delta} + \boldsymbol{e}$; thus, applying OLS we get.

$$\widehat{\boldsymbol{\delta}}_N = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{y}, \qquad \mathrm{V}(\widehat{\boldsymbol{\delta}}_N |\ \boldsymbol{Z}) = \sigma_e^2(\boldsymbol{Z}'\boldsymbol{Z})^{-1}. \tag{5}$$

Under the stated conditions, the estimator (5) is consistent for $\boldsymbol{\delta}$ and its conditional dispersion $\mathrm{V}(\widehat{\boldsymbol{\delta}}_N|\boldsymbol{Z})$ provides unbiased standard errors for testing the significance of $\widehat{\boldsymbol{\delta}}_N$.

Unlike OLS, the LAD estimator is defined as the minimization of the sum of absolute residuals $\mathrm{A}(\boldsymbol{\delta})$ of the model (4).

$$\check{\boldsymbol{\delta}}_N = \arg\min\left[\mathrm{A}(\boldsymbol{\delta}) = \sum_{j=1}^{N} \left|Y_j - \boldsymbol{\delta}'\mathbf{z}_j\right|\right] \tag{6}$$

The minimization problem (6) is carried out iteratively and its statistical evaluation is performed on the scores of the objective function $\mathrm{A}(\boldsymbol{\delta})$. The resulting LAD estimator $\check{\boldsymbol{\delta}}_N$ is mildly insensitive (robust) to anomalous observations Y_j^*, which yield residuals $|\widehat{e}_j^*| > 2\sigma_e$; these make OLS estimates $\widehat{\boldsymbol{\delta}}_N$ biased, and so unreliable in forecasting, see [16].

Once the estimates of $\boldsymbol{\delta}' = [\boldsymbol{\beta}', \rho]$ are obtained, their forecasting ability is checked on the set $n_o < N$ of data $\{Y_{oj}, X_{okj}\}$ which are kept out (o) of the estimation. The resulting out-of-sample forecasts are $\widehat{Y}_{oj} = \widehat{\boldsymbol{\beta}}'\boldsymbol{X}_{oj} + \widehat{\rho}\,\widehat{Y}_{oj-1}$, with the initial condition $\widehat{Y}_{o0} = 0$. The prediction errors $\widehat{e}_{oj} = (Y_{oj} - \widehat{Y}_{oj})$ are used to compute the mean absolute percentage error (MAPE) statistic:

$$\mathrm{MAPE} = \frac{1}{n_o}\sum_{j=1}^{n_o}\left|\frac{Y_{oj} - \widehat{Y}_{oj}}{Y_{oj}}\right| \tag{7}$$

The best in-sample dataset and estimation method are those which minimize the statistic (7); next, the applicability of estimates $\widehat{\boldsymbol{\beta}}, \widehat{\rho}$ to other regions depends on the homogeneity of the input data $\{X_{okj}\}$ with those of the destination region $\{x_{ki}\}$. This homogeneity must hold at the physical (latitude, altitude, climate, etc.) and social (economic structure, family organization, etc.) levels; for Sardegna, a similar region may be Campania, with the exception of the Napoli area, see Fig. 1. Estimation and prediction in the general case of spatial contiguity are discussed in Appendix F.

Data description

In the outlined framework, the selection of the exogenous variables, which explain the gas consumption, is of crucial importance for the forecasting performance. In the literature, this issue is discussed from various viewpoints: Thomas [69], on behalf of the European Commission, is concerned with the factors that allow energy savings, such as weather fluctuations, households' habits and innovation in the industrial activities; Albert et al., [2] highlight the relevance of income distribution and Mikalauskienė et al., [49] focus on regulatory and government policies. As regards gas consumption, Bianco et al., [11] state that significant variables are the price of energy and the climatic zone classification (heat degrees day).

Dependent variables

Concerning Y_j, the choice of per capita data enables to deal with different indicators net of the scale (mainly the population) of the statistical units. The total gas consumption provided by MiSE [51] is the starting point; the base year is 2016, as the data of the predictor variables $\boldsymbol{x}_i, \boldsymbol{X}_j$ are jointly available only for that period. According to this dataset, the total gas consumption Y is the sum of 3 components: $\mathrm{Y}_1 =$ domestic, $\mathrm{Y}_2 =$ industrial and $\mathrm{Y}_3 =$ thermoelectric generation, expressed in million of cubic meters (M.m^3). Using the population P in 2016 (Istat [37]), the per capita gas consumption of each province is then given by:

$$Y_j = (\mathrm{Y}_{1j} + \mathrm{Y}_{2j} + \mathrm{Y}_{3j})/\mathrm{P}_j, \qquad j = 1, 2, \cdots, N. \tag{8}$$

However, the thermoelectric plants are only located in a few provinces, although they serve all other cities; hence, it seems reasonable to redistribute their gas consumption among all units j according to their population, at least at the regional level. Let k be the index of regions and n_k, N_k the number of provinces with plants and the total number of provinces in the region k; then the recomputed thermoelectric gas consumption is given by.

$$\tilde{\mathrm{Y}}_{3jk} = \mathrm{P}_{jk}\frac{\sum_{j=1}^{n_k}\mathrm{Y}_{3jk}}{\sum_{j=1}^{N_k}\mathrm{P}_{jk}}, \qquad k = 1, 2, \cdots, M, \tag{9}$$

where P_{jk} is the population of province j in the region k, and M is the number of regions.

The inclusion of the transformation (9) into the formula (8) provides $\tilde{Y}_j$; this involves significant spatial differences with respect to the initial Y_j, both at the total and per capita level, see the maps in Fig. 2. This is a consequence of the fact that electric power in Italy is mainly produced by gas turbines and their consumption is a considerable quote of the total.

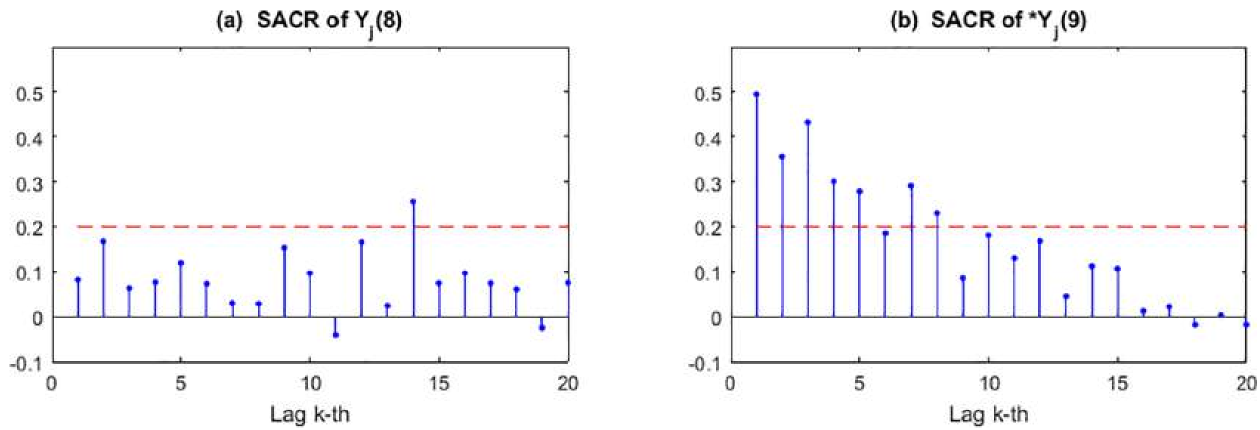

Fig. 3. Spatial Autocorrelations of Y_j Eq.(8) and $\widetilde{Y}_j$ Eq.(9), with respect to their north k-NN.

Independent variables

As regards the explanatory and prediction variables X_j, x_i, the main difficulty is to select indicators that are meaningful and available both at the provincial (j) and municipal (i) scales. For this reason, they are based on the common years 2011 (of the census) and 2016 (of the dependent variable). In general, these co-variates are grouped into four classes:

- *Geo-climate*: It includes physical and environmental features, such as altitude, latitude, longitude and the average temperature of January (the coldest month, when the domestic gas consumption reaches the maximum level).
- *Urban structure*: This group deals with the built environment, such as the average age of buildings, their size (in square meters), the average number of residents per dwelling, the population density and the percentage of dwellings with their own heating system.
- *Socio-economic*: It includes per capita income, the share of houses in ownership, the average size of the family and the share of employees in the industrial sector. Other economic variables are the per capita tourism accommodations and the ratio between employees and residents. These are positively correlated with gas consumption.
- *Energy*: This group regards indicators of energy production and consumption different from natural gas, such as propane, diesel and electricity from renewable sources.

Appendix A provides a detailed description of the covariates, including the year of reference, the level of spatial detail and their sources. The corresponding data can be geographically visualized at the ArcMap online link http://arcg.is/1KTXuf, where they are grouped according to their thematic field. Further, a specific section is devoted to the gas network in Sardegna, including the forecasts, that will be computed in the next Section.

Spatial autocorrelation

Since the models of Section 2 include components of spatial dependence, the analysis of spatial autocorrelation (SACR) of the dependent variable is a necessary descriptive statistic. The SACR is usually computed as a correlation between Y_j and $\overline{Y}_{j-k}$, the average of data Y_i which fall in the ring of width $d > 0$, with external radius kd and center Y_j, see Bivand [12]. Instead, following the causal ordering, we compute $R_k = \mathrm{Corr}(Y_j, Y_{j-k})$, where Y_{j-k} is the k-NN in the north direction; Fig. 3 provides the results for both series (8)-(9) with a 95 % significance band $2/\sqrt{N}$, Grillenzoni [32]. One may see that there is *no* significant autocorrelation in Y_j (8), whereas it increases in $\widetilde{Y}_j$ (9), such that the path of Fig. 3b is consistent with the SAR(1) model $\widetilde{Y}_j = 630 + 0.473\widetilde{Y}_{j-1} + e_j$. However, this dependence may be a consequence of the autocorrelation of the population variable P_j, which is present in Eq. (9) and is widespread as in Fig. 2b,d. It follows that when P_j is used as a regressor X_{kj} in the model (2), then the estimate of parameter ρ could become non-significant.

The use of the population P_j as a regressor for $\widetilde{Y}_j$ is motivated by the fact that per capita gas consumption inversely depends on the size of the cities; in particular, big cities have greater scale economies, hence greater energy efficiency and smaller per capita consumption. Also, the presence of covariates X_{kj} correlated with Y_j may reduce the autocorrelation of residuals and the role of the component ρY_{j-1}. Finally, data transformations, such as logarithm (which is used to hinder the effect of heteroskedasticity, i.e. the spatial variability of the variance σ_e^2), may further contribute to reducing the size of ρ. Indeed, while the estimation of the SAR(1) equation provides $\widehat{\rho} = 0.473$, with z-statistic $z_{\widehat{\rho}} = 6.5$, the OLS estimation of the SARX model $\widetilde{Y}_j = \boldsymbol{\beta}' X_j + \rho\, \widetilde{Y}_{j-1} + e_j$ (with $\widetilde{Y}_{j-1}$ in the north half-plane) provides a negligible $\widehat{\rho} = 0.025$, with $z_{\widehat{\rho}} = 0.48$ (see Appendix B).

Model estimations and spatial predictions

Following the scheme in Fig. 1, there is no model-building strategy that is a-priori optimal in forecasting. The choice of the estimation method (OLS or LAD), the data transformation (logarithm or linear) and the in-sample training zone (Tyrrhenian or whole Italian), must be evaluated with the MAPE statistic (7). More specifically, the model structure depends on 5 settings:

- *Dependent variable:* One may use either the per capita gas consumption Y_j computed in Eq. (8) or its reweighted version $\widetilde{Y}_j$ (9).
- *Data transformations:* The log transformation of Y_j or both Y_j, X_{kj} is suitable when the residuals e_j of the model (2) are non-normal (i.e. their distribution has heavy tails) or are heteroskedastic (i.e. their variance σ_e^2 is spatially-varying). In these cases, the OLS estimates of $\boldsymbol{\beta}$ are not efficient and their standard errors $S_{\widehat{\beta}}$ are biased, thus affecting the statistical inference of the model. The log transformation is also associated to a non-linearity of the relationship between Y_j, X_{kj} of exponential type, as $Y_j = \exp(\boldsymbol{\beta}' X_j)$.
- *In-sample data:* The choice of the data for estimating the parameters is important in forecasting. In time series, only recent data are used; in spatial series, one may use the data closest to the out-of-sample units. From Fig. 1, these are the provinces of Tyrrhenian regions (which

Table 1
Comparison of the prediction performance of various SARX models.

Y_j	Type of model	Estimator ($\widehat{\beta}$)	In-sample (Y_j, X_{kj})	MAPE (Y_j) [%]	max ($\widehat{y}_i$) $[m^3]$	min ($\widehat{y}_i$) $[m^3]$	Total amount $[M.m^3]$
(9)	Linear 95 % signif.	LAD	Italy	**22.4**	3708	364	2756
(9)	Log. 95 % signif.	OLS	Italy	38.8	3903	3	2062
(9)	Log. with all X_k	OLS	Italy	41.5	3920	13	1882
(8)	Linear 95 % signif.	LAD	Italy	43.8	2027	390	1661
(8)	Linear 95 % signif.	LAD	Tyrrh.	56.3	2707	314	1740

have social-physical conditions similar to Sardegna) instead of the whole Italy.

- *Estimation methods:* The choice of the estimation method depends on the structure of the model and on the regularity of data. In the presence of outlying observations (which generate anomalous residuals), OLS estimates of $\boldsymbol{\beta}$ are biased, hence unreliable. The LAD method is less sensitive to big residuals but is computationally demanding even for inference.
- *P-value thresholds*: In classical model selection, the variables X_{kj} which have coefficients $\widehat{\beta}_k$ that are not statistically significant (at the 90–99 % probability level) are discarded from the model. However, in out-of-sample forecasting, such variables may be useful to reduce the MAPE statistic and thus they may be retained in the model.

MAPE statistic (computed on 6 out-of-sample provinces), is not the sole indicator of the forecasting performance of the models. The forecasts $\widehat{y}_i$ (3) can also be evaluated directly as regards their consistency with the Sardegna environment. In particular, the models which provide negative or excessive point forecasts $\widehat{y}_i$, or a regional average $n^{-1}\sum_{i=1}^{n}\widehat{y}_i$ that is too different from the Italian (or Tyrrhenian) mean $N^{-1}\sum_{j=1}^{N}Y_j$, must be discarded. The application of these criteria does not affect the validity of the outlined methodology, as its main goal is to provide a regional map of potential demand. Table 1 presents the models with *admissible* forecasts $\widehat{y}_i$, obtained from various model settings; computational details are in Appendix B and Barbiero [9, Appendix 3].

We focus on the first model of Table 1 as it has the lowest MAPE statistic and the best compromise between min and max $\widehat{y}_i$. The model is

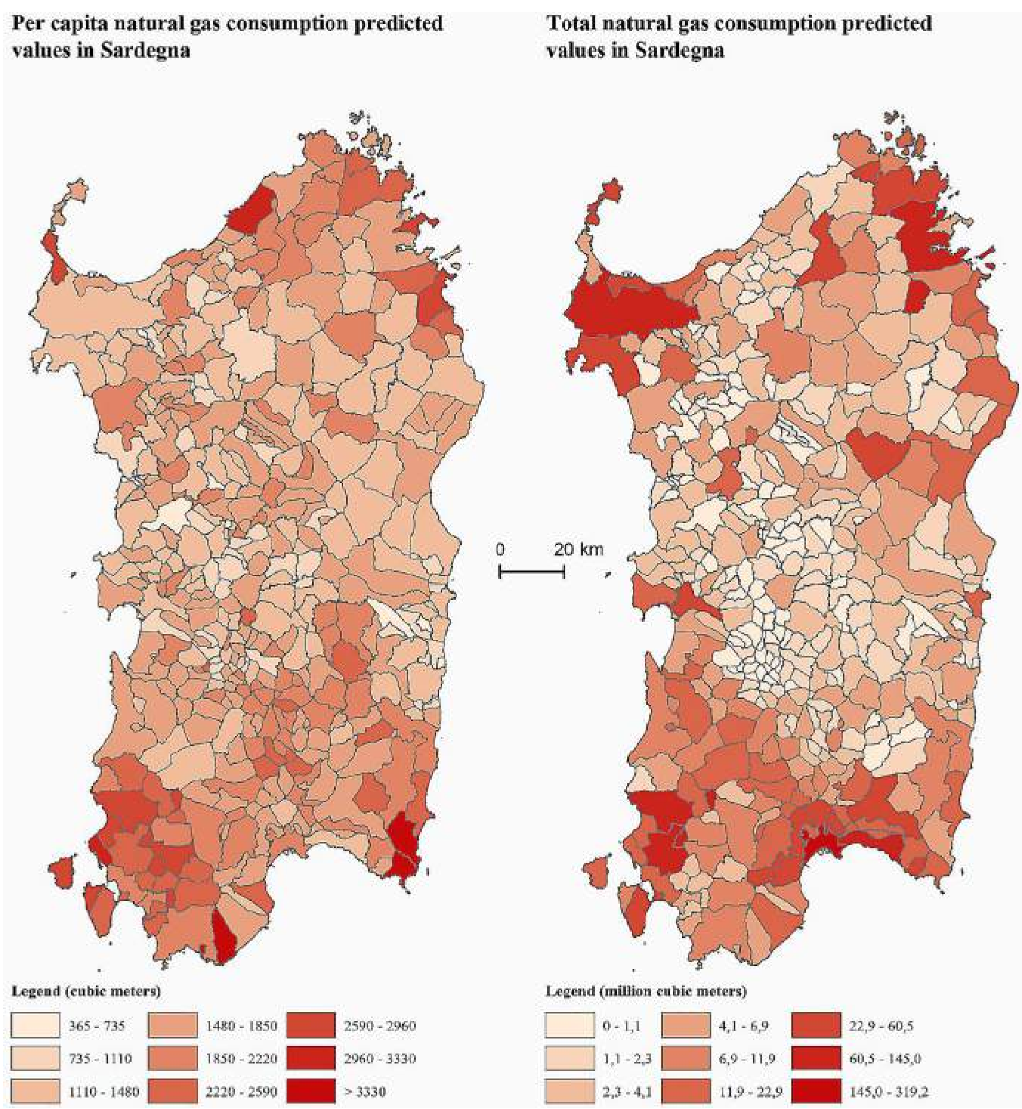

Fig. 4. Predicted municipal gas consumption of Sardegna based on model 1 of Table 1: a) Per capita values $\widehat{y}_i$; b) Total amount $\widehat{y}_i\ p_i$.

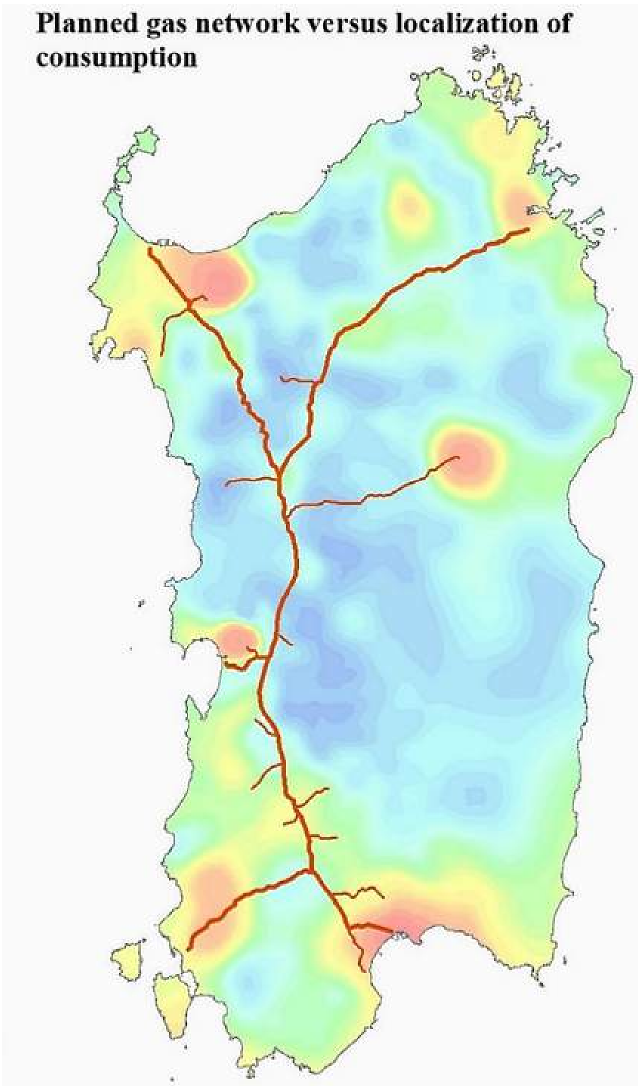

Fig. 5. Kriging smoothing of Fig. 4b and outline of the gas transportation network.

linear, estimated on $\widetilde{Y}_j$ of the whole Italy, with the LAD method and discarding 95 % non-significant variables X_{kj}. The predicted total gas consumption of Sardegna $\sum_{i=1}^{n}\widehat{y}_i\mathrm{p}_i$ (where p_i is the population of the i-th municipality) amounts to 2.756 Billion m^3, under the assumption of full spatial coverage of the gas network in Sardegna.

Fig. 4 shows the spatial distribution in Sardegna of the predicted per capita $\widehat{y}_i$ and total gas consumption $\widehat{y}_i\mathrm{p}_i$ based on model 1 of Table 1. The highest per capita values are in tourist areas, in the north-east and south-west, because tourists consume gas but are not part of the population. The lowest values are located in inland areas at moderate altitudes, with a typical rural economy and modest social development. As regards Fig. 4b, the biggest amounts are spread across the chief municipalities: Cagliari 319 M.m^3, Sassari 145, Olbia 89, Oristano 39 and Nuoro 42. These maps provide the spread of potential demand but are also useful to define suitable paths for the gas network; in order to identify such paths, the spatial smoothing of point forecasts $\widehat{y}_i\mathrm{p}_i$ is useful.

Smoothed maps

In order to understand the spatial trends of the gas consumption of Sardegna, the map in Fig. 4b must be properly smoothed and/or the point values $z_i = \{\widehat{y}_i\mathrm{p}_i\}$, with spatial coordinates r_i, s_i (longitude and latitude of the municipality center), must be interpolated in the continuous plane as $\widehat{z} = g(r, s)$. There are various techniques of spatial interpolation (e.g. Grillenzoni [30]), which, in general, assume the form of weighted averages $\widehat{z}(r,\ s) = \sum_{i=1}^{n} w_i(r,\ s)z_i$, where $w_i(.)$ are local weights that also depend on the coordinates r_i, s_i. In the inverse distance weighting (IDW) method $w_i(r,s) \propto 1/\left[(r_i - r)^2 + (s_i - s)^2\right]^{1/2}$ are proportional to the inverse of Euclidean distance. Instead, in the Kriging method, the weights depend on a model of spatial covariance function, which is assumed stationary; however, as in IDW, the key smoothing factor is represented by the number of nearest points $n' < n$ which is considered in the sum. The right choice of n' is a compromise between smoothness and variability of the surface.

Fig. 5 provides the Kriging smoothing of Fig. 4b obtained with the Gaussian variogram and$n' = 100$ neighbors, see ArcGIS [5]. It enhances the spatial trends of the areas with major gas consumption, thus allowing to identify the suitable path for the gas transportation pipeline. This path could be detected in an automatic way with ridge clustering techniques (e.g. [31]); however, a supervised design is necessary to balance earnings and costs of the infrastructure, i.e. to maximize the number of users and minimize the pipeline length. A key corridor for gas transportation is the Campidano valley, connecting the south-west areas of Cagliari and Oristano with the planned LNG terminal of Porto Vesme. The pipeline outlined by Regione Sardegna [62] is consistent with the scheme in Fig. 5, although in the north there is still uncertainty. In any event, the city of Sassari, and other populous areas on the east side, must be connected to the Oristano-Cagliari section. A final remark concerns the high-pressure pipes and the decompression stations, that cannot be located in inhabited areas.

Physical sizing of the new gas network

There are various projects concerning the gas network in Sardegna. In order, to evaluate their physical and economic sustainability, one must compare the various scenarios of supply and demand of natural gas, and check which ones approach an equilibrium point.

The gas supply scenario

The most challenging project was named GALSI [26], and had to connect the gas wells of Hassi R'Mel (Algeria) to the north Italian network, passing through Sardegna. This project was set aside by European Commission [23], but with the present Ukraina-Russia crisis, it could be reconsidered. Meanwhile, the incoming project is based on LNG terminals, regasifiers and related deposits, see Castellarin [14] and documents of MiSE and [62]. The infrastructure has a main pipeline of

Table 2
Planned LNG terminals and deposits in Sardegna ($1m^3$LNG $\approx 600\,m^3$gas).

Type	Location	Storage $[liquid\ \mathrm{M}.m^3]$	Storage $[gas\ \mathrm{M}.m^3]$	Manager	Status (Nov 2019)
LNG terminal	PortoTorres	0.160	96.0	Eni	Planned
LNG terminal	Cagliari	0.022	13.2	Isgas	Authorized
Deposit	Oristano	0.009	5.4	Ivi	Plannned
Deposit	Oristano	0.010	6.0	Edison	Authorized
Deposit	Oristano	0.009	5.4	Higas srl	Authorized

Source: geoportal of Sardegna Energia [62].

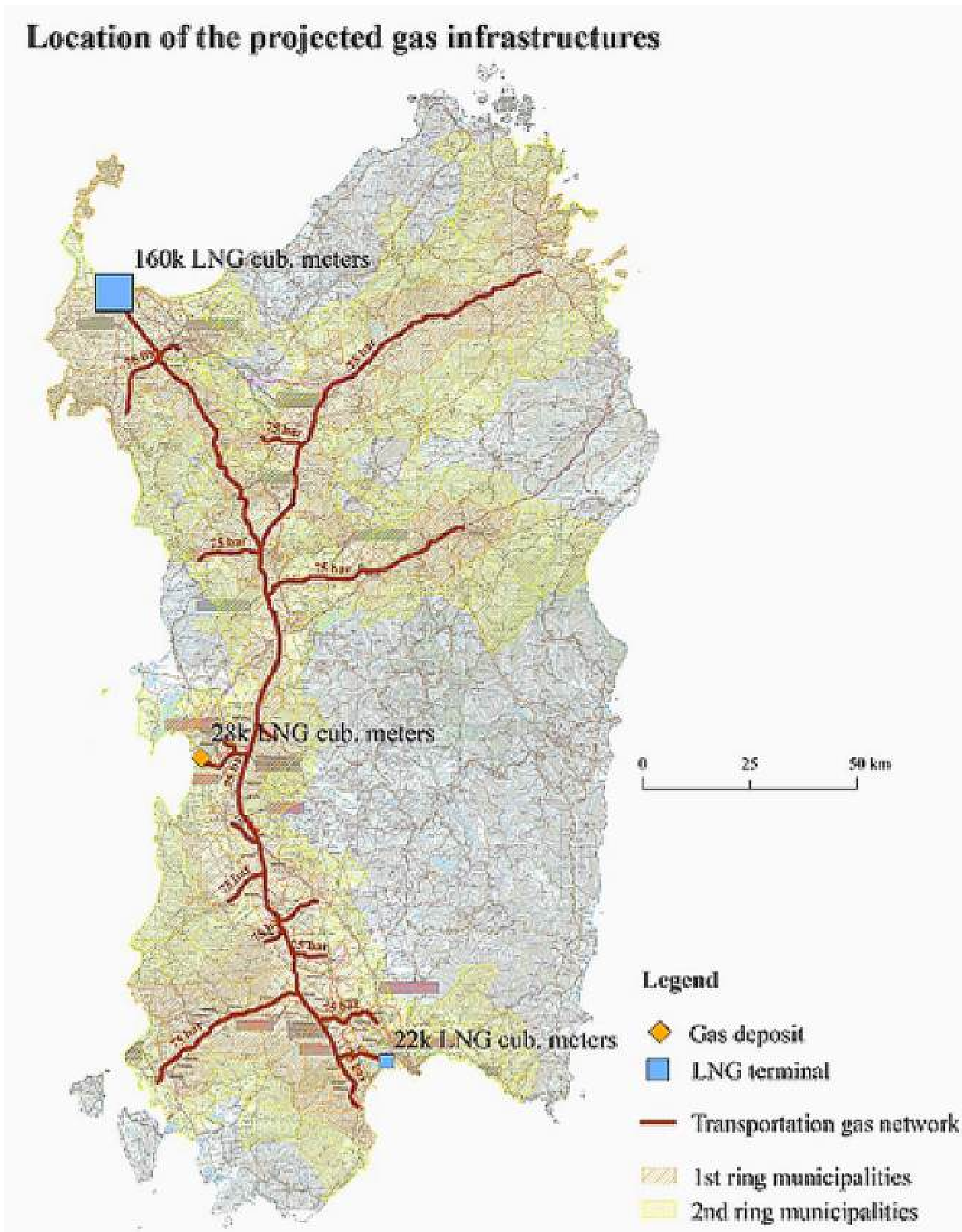

Fig. 6. Planned gas infrastructure in Sardegna.

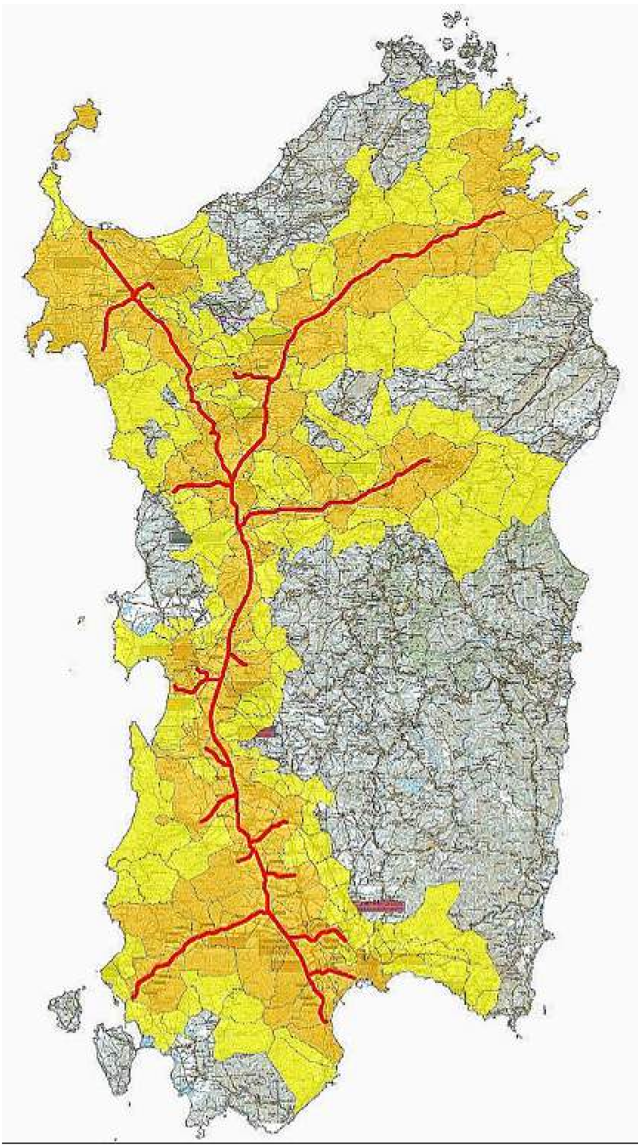

Fig. 7. Municipalities involved in the last two scenarios of Table 3.

277 km, that connects Porto Torres (Sassari), Oristano and Sarroch (Cagliari), two secondary branches with lengths of 57 and 75 km respectively and many local pipes, see Fig. 5. The system is powered by two LNG terminals with a storage capacity of 182,000 m^3 of liquid gas, which correspond to 109.2 M.m^3 of natural gas. In addition, there are 3 storage deposits near Oristano, with an equivalent capacity of 16.8 M.m^3 gas. For details on LNG terminals and deposits see Table 2, while their location is shown in Fig. 6.

At the planning level, Regione Sardegna [59] has organized the land into 38 catchment areas and has financed grants up to 50 % of the cost of the networks. The incentives are addressed to local associative companies, that realize the pipelines and must entrust their management to a single gas operator. Subsequently, Regione Sardegna [60] has extended the grants to the gas distribution companies (urban networks), by setting constraints on the expected cost per meter and the length of the pipes. Regione Sardegna [62] provides data on the realization of this infrastructure and its current state is shown in Appendix C. At present, 103 municipalities (27 %) have started local networks; 64 % have adhered to the project, but have not started the works yet; finally, 9 % have not adhered at all. The main Italian gas distributor has allocated 500 million euros in Sardegna (Italgas [38]), and has completed 160 km of the major pipeline (Italgas [39]). The majority of 103 municipalities are placed around Cagliari and Sassari and are crossed by the main pipeline; instead, those who have not adhered are in central-eastern areas and will hardly be connected in the future.

Table 4
Analysis of the fixed costs of the new gas infrastructure.

Cost items	References	Physical size	Unit cost per m	Total cost [M.€]
Transportation network C_1	[65]	576 km	722 €/m	415.7
Distribution network C_2	[58]	4317 km	148 €/m	613.4
Distribution paid by operators $(1-g_S)C_2$	[61]	4317 km	76 €/m	315.9
LNG terminal C_3	[44]	210,000 m^3	676 €/m^3	142.0
Infrastructure costs C_T				1171.1
Infrastructure paid by operators C_S (INF)				873.6

The gas demand scenario

On the demand side, we have the following scenarios:

- *Maximum consumption*. This is the amount estimated by the SARX model in Table 1 and consists of 2756 M.m^3 per year. However, this forecast assumes the complete spatial coverage of the gas network, which is not realistic; e.g. isolated areas will not be reached.
- *Italian average*. It assumes that the per capita gas consumption in Sardegna is equal to that of Italy: 1244.4 m^3 (per capita/year), see MiSE [51]. This provides a total consumption of 2050.8 M.m^3, on which the SARX municipal forecasts in Fig. 4 could be re-proportionated.
- *Minimum consumption*. It is based on the present consumption of LPG and diesel in Sardegna (Gherardini [27]) and their replacement with natural gas, yielding 521.3 M.m^3.
- *Realistic coverage*. This scenario adjusts SARX forecasts by dropping the municipalities which are *not* reachable by the gas network, as they are too distant and have little population. In this context, we define the first and second bands around the main pipeline (see Fig. 7), as the actual basin of gas consumption. This approach involves about 50 % of municipalities, 58 % of the surface and 79 % of the population, for a gas amount of 2144.1 M.m^3.
- *Realistic penetration*. This scenario is similar to the previous one but assumes a penetration rate of 70 %, which is the average rate in Italy (Istat [36]). So, the total amount of gas demanded is 70 % of the previous one, namely 1511.6 M.m^3 per year.

Table 3
Scenarios of natural gas consumption in Sardegna.

Scenarios	Method	Surface cov. [%]	Municipality cov. [%]	Penetrat rate [%]	Population cov. [%]	Population amount	NG demand [M.m^3]
Minimum	NG equiv.	100	100	18.9	100	208 k	521.3
Italian avg.	Heuristic	100	100	70.5	100	1161 k	2034.0
Maximum	SARX	100	100	100	100	1648 k	2756.0
Realistic coverage	Contiguity	58.2	49.9	100	78.8	1300 k	2144.1
Realistic penetration	Contiguity	58.2	49.9	70.5	55.5	916 k	1511.6

Table 5
Composition of the gas price paid by finals consumers in Sardegna.

Components	References	Unit cost [€/m^3]
Taxes and levies (τ)	[51 7 24]	0.168
Raw material costs (q)	[6]	0.281
Revenues for providers and concessionaires (r)	($p-\tau-q$)	**0.154**
Average price (p)	[6]	0.603

Table 3 summarizes the various scenarios and Fig. 7 provides the geographical representation of the most realistic one.

Economic sustainability of the new gas network

To complete the analysis, we evaluate the economic sustainability of the new gas infrastructure; that is, the conditions under which the cost of investments will be repaid by the revenues, in a suitable number of years. Under the assumption of a gas penetration rate for Sardegna equal to the Italian average (70.5 %, Istat, 2011) [36], the budget performance depends on public grants and interest rates, costs and prices of the gas, returns and taxes.

Cost analysis

The total infrastructure cost C_T is the sum of the costs of the transportation pipeline C_1, the distribution network C_2 and LNG terminals C_3. By considering the average rate of public grants $g_S = 0.485$ for C_2, the actual cost of the infrastructure becomes:

$$C_T = C_1 + (1 - g_S)C_2 + C_3 \tag{10}$$

Following [65], the estimated cost C_1 should be 415.7 M.€ and refers to the pipeline in Fig. 6 (with a length $L_1 = 576$ km and a unit cost of 722 €/m), and its compression stations.

The evaluation of C_2 is more difficult, as it regards the peripheral network, whose length L_2 is unknown. However, using the average ratio of Italy $L_2/L_1 = 7.49$ (see ARERA [6, p.195]), the estimated length in Sardegna becomes $\widehat{L}_2 = 4317$ km. Next, assuming a unitary cost of 147.9 €/m, see the standards of Regione Sardegna [58], which include decompression stations, connections to final users and upgrading of existing networks [25], the estimated cost is $\widehat{C}_2 = 613.4$ M.€. Further, owing to the grants of Regione Sardegna [60], the actual cost becomes $(1-g_S)\widehat{C}_2 = 315.9$ M.€.

Finally, the cost C_3 of LNG terminals and deposits is obtained from similar projects in Northern Europe which features a unit cost of 676 €/m^3 (King et al., [44]). Owing to the overall capacity of 210,000 m^3 of the 3 plants in Sardegna, it follows $C_3 = 142$ M.€. Table 4 summarizes the main aspects of the analysis developed so far.

Given the total fixed cost (about 1 billion in Table 4), to complete the budget analysis of investors it is necessary to estimate the variable costs and earnings per m^3 of gas. This means reviewing taxation (τ), cost of raw materials (q) and returns (r), which compose the final price paid by customers. The taxes range from 0.299 €/m^3 for households (ARERA [7]), to 0.053 €/m^3 for firms (Eurostat [24]); since the domestic sector absorbs 46.6 % of total consumption (see MiSE, 2017b) [51], the average tax is $\tau = 0.168$ €/m^3. Similarly, averaging the final prices paid by households and firms reported by ARERA [6], one obtains $p = 0.603$ €/m^3; while the cost of the gas from LNG is $q = 0.281$ €/m^3. It follows that the unit revenue ($r = p - \tau - q$) of the infrastructure companies is 0.154 €/m^3, see Table 5.

Equilibrium point

Given the amount of gas predicted (TY, Table 3) and the composition of unit prices (p, Table 4 and 5), one can compute total revenues (TR) and total costs (TC); however, they depend on the percentage π_t of realization of the infrastructure at year t. Assuming constant prices p and market penetration rate $P_R \leq 1$, the total gross revenue of the gas companies at year t, is given by:

$$TR_t = p(TY \cdot P_R)\ \pi_t\ , \tag{11}$$

Instead, the total cost TC_t at time t, depends on the costs of raw materials Raw_t, taxes Tax_t, operating activities Ope_t, passive interests Int_t and investments in the infrastructure Inv_t (see the scheme in Appendix E), minus the public grants (Regione Sardegna [60]). Hence:

$$TC_t = Raw_t + Tax_t + Ope_t + Int_t + Inv_t, \tag{12}$$

where Raw_t and Tax_t depend on the amount of gas sold and market factors.

$$Raw_t = q(TY \cdot P_R)\ \pi_t\ , \tag{13}$$

$$Tax_t = \tau(TY \cdot P_R)\ \pi_t\ . \tag{14}$$

Instead, Ope_t depends on the operating costs at full capacity $O_{\max}$ (see Appendix D) and the rate of completion of the network, regardless of the population served and market factors; hence.

$$Ope_t = O_{\max}\,\pi_t. \tag{15}$$

Further, the amount of passive interests Int_t is obtained by multiplying the interest rate i_r and the invested capital K_{t-1}(or loans repaid). Given the annual loan INF/n (where INF is the total infrastructure cost,

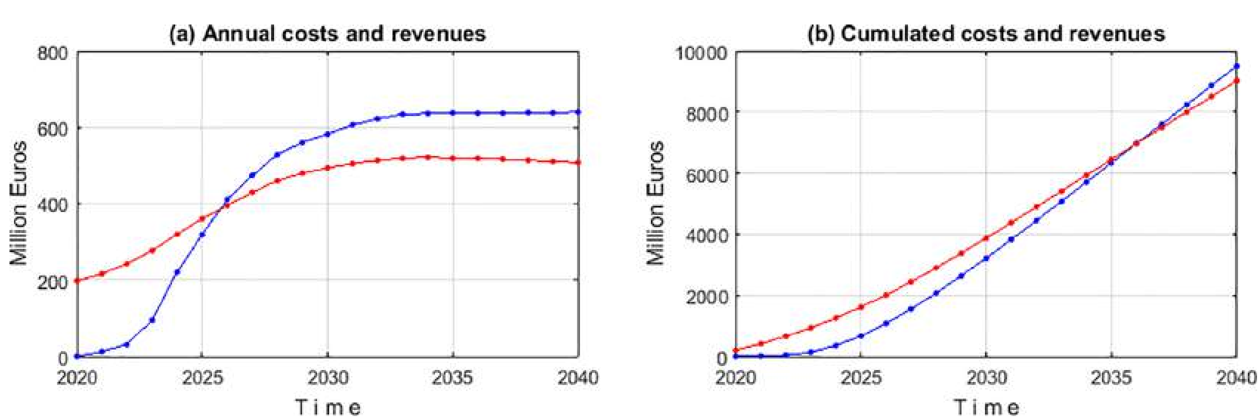

Fig. 8. Trends of annual and cumulative costs (red) and revenues (blue) of the gas network project in Sardegna, assuming the average Italian gas penetration rate 70.5%.

Table 6
Variable costs and revenues.

	Overall penetration	Penetration first band	Penetration second band	Popul.	Break even	Gas sold in 2031 $[\mathrm{M}.m^3]$	Net revenues in 2031 [M.€]
Full penet.	100 %	100 %	100 %	1300 k	2029	1511.6	232.8
Part. penet.	70.5 %	73.2 %	65.9 %	916 k	2035	1065.7	164.1

defined in the last row of Table 4), with a duration of $n = 25$ years at the interest rate $i_r = 0.04$, the sought cost becomes:

$$Int_t = i_r \ K_{t-1} = i_r(t-1)\frac{INF}{n}. \quad (16)$$

Finally, the actual infrastructure cost Inv_t may vary according to the working plan in Appendix E, which assumes that main facilities (LNG terminals, high-pressure pipeline and distribution network around main cities) will be completed within 10 years and the full operation capacity will be reached in 2040. This also provides the sequence of annual percentage π_t of infrastructure realization.

Having reviewed total revenues and costs, the year n' of the "break-even" is given by.

$$n' : \sum_{t=1}^{n'}(TR_t - TC_t) > 0. \quad (17)$$

This point is a function of the gas penetration rate P_R, whose value in Italy is 70.5 % (Istat [36]); a further assumption is that P_R in the second band of Fig. 7 is inferior by 10 % with respect to the first band municipalities. Under these conditions, Fig. 8 shows the annual and cumulative trends of revenues and costs over time and shows that the "break-even" may be reached in $n' = 2035$, with 1066 $\mathrm{M}.m^3$ of gas sold, see Barbiero [9], Chapter 4. Moreover, Table 6 provides the net revenues $R_t = r(TY \cdot P_R)\ \pi_t$ at year $t = 2031$, when the gas infrastructure will become fully operational, with two P_R scenarios.

The analysis developed so far, on the basis of the data of official documents and realistic hypotheses, has treated the new gas infrastructure as a single investment. However, the companies involved in the project have different budgets, in which the value of the assets is spread over various financial years as depreciation costs. In the case of public-sector infrastructures, there are methodological difficulties to account for depreciation costs, due to the long period of realization of the works, the presence of significant non-repayable public grants, the long life span of facilities and the incoming maintenance costs (Burns [13]). For this reason, the "equilibrium" point computed in Fig. 8 is not a budget indicator, but is a measure of the economic sustainability of the project.

Conclusions

Planning the use of non-renewable resources is a topic of growing importance, especially for oil and natural gas. Their depletion is still far, as many scenarios indicate, since the oil peak point should occur in 2020/22 and its breakdown after 2100, [63]. As regards natural gas, the present world consumption is 3.96 trillions m^3 [54], but the amount of proven reserves is greater than 206.68 trillions m^3 [55], and increases every year; therefore, the date of the Hubbert's peak for natural gas is still unknown. Despite this, the *rational* use of fossil fuels is essential to postpone their depletion and reduce the social and environmental risks involved in their usage. From a planning perspective, the main goals are energy-saving and a fast transition toward renewable energy systems; here, oil and gas should primarily be used for building facilities for renewable energies.

In the long run, fossil fuels should essentially be employed where there are no viable alternatives, such as in metal foundries and in airplane transportation, where strong unitary powers are required. Instead, in private home heating and in urban mobility they can be easily replaced by energy systems based on biomasses, heat pumps and photovoltaic electricity. However, in Italy more than 95 % of households still employ fossil fuels [36], which means that the transition to renewable sources will be slow. On the side of energy-saving in the housing sector, thermal isolation of buildings is hindered by the high refurbishment costs [10]. However, the conversion to the natural gas of coal and oil-based systems and the use of efficient boilers and turbines is still subsidized, in view of the low pollution emissions and the efficiency in electric power generation [21].

This work has proposed an original way to design and evaluate an incoming gas infrastructure in a region (Sardegna) where a natural gas network is currently absent. The analysis has grounded on the predictions, at the municipal level, of spatial regression models estimated on the province data. Other studies considered time-series models and rely on aggregate data and physical variables, such as climate features. Instead, our work has considered a wide range of covariates that may determine the behavior of gas consumption, including Urban structure, Socio-Economic, Energy sources and Geo-climate (see Appendix A). The estimates presented in Sect. 4 reveal that models which have a good forecasting performance (in terms of MAPE statistic) on the training sample, may not be reliable for the target region. Thus, the selection of the predictions has avoided models which provide negative or excessive municipal values. The final results are the estimated per capita and total gas consumption for 377 Sardegna municipalities.

The economic analysis has been performed on the municipalities nearby the main gas infrastructure depicted in Fig. 4. These units are the most likely to be involved in the natural gas supply, instead of remote and depopulated areas of Sardegna. The analysis has considered infrastructural costs, variable costs (raw materials, taxes, operating costs), passive interests and the revenues of the companies in charge of the transportation pipeline (Snam SpA) and the distribution network (Italgas SpA). The estimated costs and revenues depend on the plan of scheduled works and on the share of households that will adopt natural gas. The analysis in Sect. 6 has shown that the gas infrastructure may be sustainable from an economic viewpoint, despite the high fixed costs. Indeed, the cumulative revenues may overtake the cumulative costs within 11 years, provided a 70 % market penetration rate for natural gas. As regards the life-cycle assessment, notice that the equilibrium point will occur much earlier than the periods of refurbishment or rebuilding of the infrastructure (approximately 60–70 years); hence, its amortization plan can be established with confidence. In this period, Sardegna may take advantage of a cleaner energy system, comparable with the rest of Italy, which may improve its social and economic conditions.

Credit author statement

The paper "Planning a Novel Regional Methane Network: demand forecasting and economic evaluation", submitted for publication in "Energy Conversion and Management: X" is the result of the joint collaboration of the two authors: C. Grillenzoni and T. Barbiero. The original idea arose in 2017 and was developed in the subsequent years.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

Appendix A. Description of explanatory variables

The prediction variables for natural gas consumption are divided into four groups: Physical (geo-climate), Urban and housing, Socio-economic and Energy supply. Given the model-building strategy in Fig. 1, the main difficulty of selection is to find variables that are available both at the provincial and municipal scales; the recent periods where this is possible are 2011 and 2016. The collected data can be visualized at the spatial level at the ArcMap link http://arcg.is/1KTXuf.

Table A1.

Table A1
List of explanatory variables of the Italian gas consumption considered in the study.

Type	Exogenous variables	Scale	Year	Ssource
	Heating degree days	Municipal	n.a.	Presidential Decree 412 of 1993
	Altitude	Municipal	n.a.	Presidential Decree 412 of 1993
	Latitude	Municipal	n.a.	Shp files from Istat
Physical	Longitude	Municipal	n.a.	Shp files from Istat
	Average high temperature in January	Municipal	2011, 2016	3B meteo
	Average temperature in January	Municipal	2011, 2016	3B meteo
	Rain days in January	Municipal	2011, 2016	3B meteo
	Annual hours of activation of heating system	Municipal	2011, 2016	Presidential Decree 412 of 1993
	Residents per inhabitation ratio	Municipal	2011	Istat (2011 pop. Census)
	Independent of hot water system rate	Municipal	2011	Istat (2011 pop. Census)
	Independent of heat system rate	Municipal	2011	Istat (2011 pop. Census)
	Share of buildings built before 1960	Municipal	2011	Istat (2011 pop. Census)
	Share of buildings built after 1990	Municipal	2011	Istat (2011 pop. Census)
Urban	Average building age	Municipal	2011	Istat (2011 pop. Census)
	Share of non-occupied buildings	Municipal	2011	Istat (2011 pop. Census)
	Rate of population in urban area	Municipal	2011	Istat (2011 pop. Census)
	Population density	Municipal	2011, 2016	Istat (2011 pop. Census)
	Average habitable area per resident	Municipal	2011	Istat (2011 pop. Census)
	Average size of dwellings	Municipal	2011	Istat (2011 pop. Census)
	Share of houses owned	Municipal	2011	Istat (2011 pop. Census)
	Share of houses rented	Municipal	2011	Istat (2011 pop. Census)
	Average family size	Municipal	2011, 2016	Istat (2011 pop. Census)
	Per capita industrial employees	Municipal	2011	Istat (2011 indust. Census)
Social	Per capita income	Municipal	2011, 2016	Ministry of Economy & Finance
	Share of employees in the industrial sector	Municipal	2011	Istat (2011 indust. Census)
	Tourist accommodations per capita	Municipal	2011, 2016	Istat
	Share of the graduated population	Regional	2011, 2016	Istat
	Rate of enterprises making innovations	Regional	2016	Istat
	Rate of employees in innovative sectors	Regional	2011, 2016	Istat
	Per capita water consumption	Municipal	2012	Istat
	Share of recycled waste	Municipal	2011, 2016	Ispra
	Per capita recycled waste production	Municipal	2011, 2016	Ispra
	Per capita urban waste production	Municipal	2011, 2016	Ispra
	Per capita electric consumption in thermoelectric	Regional	2011, 2016	Terna
	Per capita diesel sold	Provincial	2011, 2016	Ministry Economic Development
Energy	Per capita LPG sold	Provincial	2011, 2016	Ministry Economic Development
	Per capita electric consumption in industrial sector	Provincial	2011, 2016	Terna
	Per capita electric consumption in domestic sector	Provincial	2011, 2016	Terna
	Per capita electric energy consumption	Provincial	2011, 2016	Terna
	Per capita energy produced by PV systems	Provincial	2011, 2016	Gse
	Per capita energy produced by wind systems	Provincial	2011, 2016	Gse
	Per capita energy by all renewable energy sources	Provincial	2011, 2016	Gse

Appendix B. Results of the SARX model estimation

We present the OLS estimates of a SARX model (2) on the data of Italian provinces in the year 2016. The dependent variable $\widetilde{Y}_j$ is the per-capita gas consumption obtained by redistributing at the provincial level the gas employed by thermoelectric plants of the regions. Only 95 % significant coefficients $\widehat{\beta}_k$ are presented, together with their z-statistics. These results weaken (as the number of significant regressors and fitting performance) by estimations with the LAD method, the data in logarithm and the reduction of the sample size (to the Tyrrhenian provinces; see Fig. 1). Notice that the lagged spatial term $\widetilde{Y}_{j-1}$ (in the north quadrant) is not significant, while a dummy (0–1) variable is introduced to cope with the outlying unit $j = 40$ (Ferrara, which has many thermoelectric plants).

Table B1.

Table B1

OLS estimates of the SARX model of $\widetilde{Y}_j$ on the Italian province data 2016.

Type	X_{kj}	$\widehat{\beta}_k$	$\widehat{z}_k$
	Constat	8120.6	5.51
Various	Spatial lag $\widetilde{Y}_{j-1}$	*0.025*	*0.48*
	Dummy (j = 40)	1508.9	20.1
	Altitude	−0.643	−2.89
	Latitude	162.5	3.91
Physical	Longitude	−42.8	−2.42
	Average temperature in January	−68.3	−3.53
	Rain days in January	−18.7	−2.21
	Independent of hot water system rate	927.8	2.68
Urban	Independent of heat system rate	−1010	−2.81
	Share of buildings built before 1960	3422	3.52
	Average building age	−31.8	−3.03
Social	Per capita income	0.034	2.11
	Share of employees in the industrial sector	1462.4	2.82
	Per capita water consumption	3.43	2.39
	Share of recycled waste	−1242.5	−3.12
Energy	Per capita recycled waste production	3.65	4.89
	Per capita electric consumption industrial sector	0.123	8.91
	Per capita electric consumption domestic sector	0.378	4.20
	Per capita electric energy consumption	−0.230	−2.72
Statistics:	$\max\lvert\widehat{e}_j/\widehat{\sigma}_e\rvert = 2.71$, p-value (homosked. test) = 0.492, R^2= 0.876.	.	.

Legend: $\widehat{z}_k$ are Student-type statistics of the estimated regression parameters.

Appendix C. State of implementation of the gas distribution network

Fig. C1.

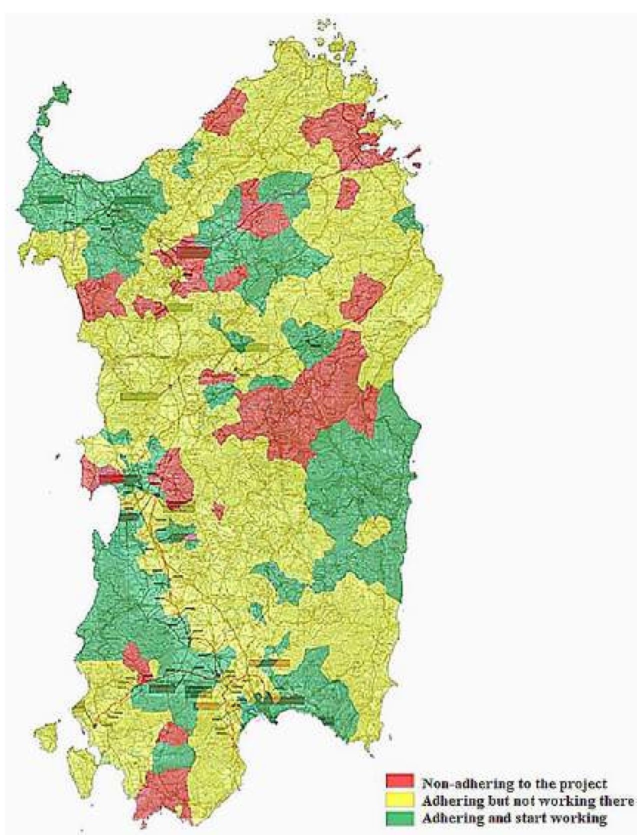

Fig. C1. State of implementation of the gas distribution network: 1) Green: Municipalities that have started the works; 2) Yellow: Municipalities that have adhered to the project but have not started the works yet; 3) Red: Municipalities that have not adhered to the project at all.

Appendix D. The operating costs of the gas infrastructure in Sardegna

The total operating cost O_{max} is calculated as the sum of the costs related to the transportation pipeline O_1, the distribution network O_2, and the LNG import terminals O_3, such that $O_{max} = O_1 + O_2 + O_3$. As regards O_1, an amount of 9 million Euros per year should be suitable (see MiSE [50]); the unit cost $o_2 = 5487$ €/km per year (Italgas [40]) and the length $L_2 = 4317$ km (see Table 4), provide $O_2 = 23.7$ M.€ per year. Table D1 summarizes these operating costs in the case of full functioning of the network.

Table D1
Annual operating costs in the case of network completed.

Type of cost	References	Cost (M.€)
Transport operating cost (O_1)	[50]	9.0
Distribution operating cost (O_2)	[40]	23.7
LNG terminal operatingcost (O_3)		4.4
Total operating cost (O_{max})		37.1

Appendix E. Plan of investments of the gas infrastructure in Sardegna

We outline a feasible plan of work for building a complete Gas infrastructure in Sardegna. The priority is the realization of storage facilities and regasification plants within 2023; this should take about 2–3 years, see [41]. Meanwhile, the construction of the transportation pipeline may be carried out at a constant rate for 5 years and the distribution network proceeds at an average annual rate of 10 %. As a result, in the early years revenues will be low, against high investment costs.

In this context, it is expected that percentage of the population connected to the network in South Sardegna will rise rapidly (hopefully 50 % by 2025), and Snam company completes the entire pipeline by 2025. With a delay of 1–2 years, also the quote of customers in North Sardegna will grow and around 2027 about 75 % of the population nearby the major pipeline may have access to natural gas. The final phase will be characterized by a reduction in investments and an increase of revenues, the entire infrastructure should be fully operating in 2031; this scenario is summarized in Table E1.

Table E1
Annual scheduled costs to complete the gas network.

Year	Invest. on LNG terminals and deposits		Invest. on transportation pipeline		Invest. on distribution Network		Total investments on infrastructure [M.€]
	Percentage	Cost [M.€]	Percentage	[M.€]	Percentage	Cost [M.€]	
2020	25	35.5	20	83.1	10	31.6	150.2
2021	35	49.7	20	83.1	10	31.6	164.4
2022	20	28.4	20	83.1	15	47.4	158.9
2023	20	28.4	20	83.1	15	47.4	158.9
2024	0	0	20	83.1	10	31.6	114.7
2025	0	0	0	0	10	31.6	31.6
2026	0	0	0	0	10	31.6	31.6
2027	0	0	0	0	5	15.8	15.8
2028	0	0	0	0	5	15.8	15.8
2029	0	0	0	0	5	15.8	15.8
2030	0	0	0	0	5	15.8	15.8
2031	0	0	0	0	0	0.0	0.0
Total	100	142	100	415.7	100	315.9	873.6

Appendix F. Spatial predictions in the general case

When the contiguity matrix involves neighbors of Y_j in *any* direction of space (i.e. the rows of $\boldsymbol{W}$ have multiple 1 s), then the spatial predictor cannot be cast in the recursive form (3). A typical example is the *queen* contiguity, where Y_j depends on the average of its direct neighbors as $\overline{Y}_{j-1} = n_j^{-1}\sum_{i=1}^{n_j} Y_{ij-1}$ and the rows of $\boldsymbol{W}$ has fractional elements $1/n_j$. In this case, algebraic manipulations are necessary to derive the prediction formulae (e.g. LeSage and Pace [42]). From Eq. (4).

$$(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e} \tag{18}$$

and applying OLS to this model, the estimator (5) becomes.

$$\widetilde{\boldsymbol{\beta}}_N = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{I} - \rho\mathbf{W})\ \boldsymbol{y} \tag{19}$$

this requires pre-estimation of the coefficient $\widetilde{\rho}$ (say), which affects the properties of OLS.

If the $\boldsymbol{N} \times \boldsymbol{N}$ matrix $(\boldsymbol{I} - \rho\boldsymbol{W})$ is invertible (a necessary condition is $|\rho| < 1$), one has.

$$\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{e} \tag{20}$$

hence, taking the conditional expectation of $\mathbf{y}$ one has the prediction formula.

$$\widehat{\mathbf{y}} = \mathrm{E}(\mathbf{y}|\mathbf{X}) = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta} \tag{21}$$

This formula may be implemented with the estimates $\widetilde{\boldsymbol{\beta}}|\widetilde{\rho}$ and its dispersion is given by.

$$\mathrm{V}(\widehat{\mathbf{y}}|\mathbf{X}) = (\mathbf{I} - \rho\mathbf{W})^{-1}\ \left((\mathbf{I} - \rho\mathbf{W})^{-1}\right)^{'}\sigma_e^2 \tag{22}$$

which provides confidence intervals of predictions. This formula can also be used for the recursive forecasts in Eq. (3); however, under the causal constraint (Y_{j-1} in the north direction), the variance becomes similar to that of an AR(1) process, namely $\mathrm{V}(\widehat{Y}_j|X_j) = \sigma_e^2/(1-\rho^2)$.

Finally, there are more general models than Eq. (4); these include spatially lagged regressors $\mathbf{WX}$ and an AR dynamic for the residuals, namely.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W}\mathbf{y} + \mathbf{W}\mathbf{X}\boldsymbol{\alpha} + \mathbf{u} \tag{23}$$

$$\mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \mathbf{e}, \qquad \mathbf{e} \sim \mathrm{N}\left(\mathbf{0}, \mathbf{I}\sigma_e^2\right) \tag{24}$$

We do not apply this scheme because our data $\mathbf{X}$ mostly regard physical and social characteristics which are not *transferable* (such as altitude, latitude, buildings, housing, etc.); hence, $\mathbf{X}_{j-1}$ does not influence Y_j. Further, a second-order SARX model $Y_j = \boldsymbol{\beta}'X_j + \rho_1 Y_{j-1} + \rho_2 Y_{j-2} + e_j$, (where Y_{j-2} is the second NN of Y_j in the north direction), is preferable to $\mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \mathbf{e}$, for improving the forecasts of Y_j.

References

[1] Afgan NH, Carvalho MG, Pilavachi PA, Martins N. Evaluation of natural gas supply options for south east and central Europe. Part 1: indicator definitions and single indicator analysis. Energy Convers Manag 2007;48:2517–24. https://doi.org/10.1016/j.enconman.2007.03.022.

[2] Albert A, Gebru T, Ku J, Kwac J, Leskovec J, Rajagopal R. Drivers of variability in energy consumption. Stanford (CA): Stanford University; 2016. Available at: http://ai.stanford.edu/~tgebru/papers/drivers_variability.pdf.

[3] Alves Jr O, Fontes C. Modeling and optimization of natural gas distribution networks for new supplier projects. Energy Conversion and Management: X 2022; 15:100240. https://doi.org/10.1016/j.ecmx.2022.100240.

[4] ANL - Argonne National Laboratories (2008). Natural Gas Pipeline Technology Overview. Environmental Science Division. Chicago (IL). Available at: https://corridoreis.anl.gov/documents/docs/technical/APT_61034_EVS_TM_08_5.pdf.

[5] ArcGIS (2019). ArcMap 10.3: How kriging works. ESRI, Redlands (CA). Available at: http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm.

[6] ARERA - Autorità Regolazione Energia Reti Ambiente (2019a). Relazione Annuale Stato dei Servizi. Volume 1, Milano. https://www.arera.it/allegati/relaz_ann/19/RA19_volume1.pdf.

[7] ARERA - Autorità Regolazione Energia Reti Ambiente (2019b). Composizione percentuale del prezzo del gas naturale per un consumatore domestico tipo. https://www.arera.it/it/prezzi.htm.

[8] Bao Y, Liu X, Yang L. Indirect inference estimation for spatial autoregressions. Econometrics 2020;8(3):34. https://doi.org/10.3390/econometrics8030034.

[9] Barbiero T. Planning the Use of a non-Renewable Energy Resource: the case study of the natural gas in Sardegna. University IUAV. Venice, Italy: Dept. of Planning; 2019. Master Thesis.

[10] Barbiero T, Grillenzoni C. A statistical analysis of the energy effectiveness of building refurbishment. Renew Sustain Energy Rev 2019;114:109297. https://doi.org/10.1016/j.rser.2019.109297.

[11] Bianco V, Scarpa F, Tagliafico LA. Analysis and future outlook of natural gas consumption in the Italian residential sector. Energy Convers Manag 2014;87:754–64. https://doi.org/10.1016/j.enconman.2014.07.081.

[12] Bivand R. Spatial econometrics functions in R: classes and methods. J Geogr Syst 2002;4:405–21. https://doi.org/10.1007/s101090300096.

[13] Burns, P. (2020). Depreciation of infrastructure assets. Talking Infrastructures. https://talkinginfrastructure.com/wp-content/uploads/2020/11/Depreciation-of-Infrastructure-Assets-Sep-1994-copy-1.pdf.

[14] Castellarin, R. (2018). Mediobanca, Italgas e Snam pronte per il metano in Sardegna. *Milano Finanza*. https://www.milanofinanza.it/news/mediobanca-italgas-e-snam-pronte-per-il-metano-in-sardegna-201811261033047402.

[15] Chebouba A, Yalaoui F, Smati A, Amodeo L, Younsi K, Tairi A. Optimization of natural gas pipeline transportation using ant colony optimization. Comput Oper Res 2009;36(6):1916–23. https://doi.org/10.1016/j.cor.2008.06.005.

[16] Copiello S, Grillenzoni C. Robust space-time modeling of solar photovoltaic deployment. Energy Rep 2021;7:657–76. https://doi.org/10.1016/j.egyr.2021.07.087.

[17] Dodds PE, McDowall W. The future of the UK gas network. Energy Policy 2013;60:305–16. https://doi.org/10.1016/j.enpol.2013.05.030.

[18] Dubin RA. Spatial autocorrelation: a primer. Jour Housing Economics 1998;7(4):304–27. https://doi.org/10.1006/jhec.1998.0236.

[19] Economides MJ, Wood DA. The state of natural gas. J Nat Gas Sci Eng 2009;1:1–13. https://doi.org/10.1016/j.jngse.2009.03.005.

[20] EIA – US Energy Information Administration (2022). Natural Gas Data. https://www.eia.gov/dnav/ng/ng_pri_sum_dcu_nus_a.htm.

[21] EPA - Environmental Protection Agency (2014). Emission Factors for Greenhouse Gas Inventories. Washington (DC). Available at: https://www.epa.gov/sites/production/files/2015-07/documents/emission-factors_2014.pdf.

[22] European Commission (2014). Communication from the Commission to the European Parliament and the Council. European Energy Security Strategy. Brussel. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014DC0330&from=EN.

[23] European Commission (2017). Amending Regulation N. 347/2013 of the European Parliament and of the Council as regards the Union list of projects of common interest. Brussels. Available at: https://ec.europa.eu/energy/sites/ener/files/documents/annex_to_pci_list_final_2017_en.pdf.

[24] EuroStat (2019). Gas prices for non-household consumers - bi-annual data (from 2007 onwards). https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nrg_pc_203&lang=en.

[25] Franci T. Nuove fonti energetiche e nuovi criteri ambientali: Opportunità per lo sviluppo della Sardegna. Sassari: ENER.LOC; 2017. Available at: https://www.promopa.it/files/ENERLOC%202017/Franci_Tommaso.pdf.

[26] GEM - Global Energy Monitor (2021). GALSI Pipeline. https://www.gem.wiki/GALSI_Pipeline.

[27] Gherardini G. Una nuova strategia per il problema energia. Available: Sardegna Industriale; 2011. at: http://www.sardegnaindustriale.it/article.asp?id=129&IDmagazine=2001002.

[28] Goulard M, Laurent T, Thomas-Agnan C. About predictions in spatial autoregressive models: optimal and almost optimal strategies. Spatial Economic Analysis 2017;12(2–3):304–25. https://doi.org/10.1080/17421772.2017.1300679.

[29] Grillenzoni C. Adaptive spatio-temporal models for satellite ecological data. Jour Agric Biol Environ Stat 2004;9(2):158–80. https://doi.org/10.1198/1085711043541.

[30] Grillenzoni C. Design of kernel M-smoothers for spatial data. Statistical Methodology 2008;5(3):220–37. https://doi.org/10.1016/j.stamet.2007.08.003.

[31] Grillenzoni C. Detection of tectonic faults by spatial clustering of earthquake hypocenters. Spatial Statistics 2014;7:62–78. https://doi.org/10.1016/j.spasta.2013.11.003.

[32] Grillenzoni C. Spatial autocorrelation function of a process in the plane. Available: Matlab code; 2022. at: https://it.mathworks.com/matlabcentral/profile/authors/24113311.

[33] Haddad J, Behbahani R. Optimization of a natural gas transmission system. Int J Comput Appl 2013;66:35–42. https://www.ijcaonline.org/archives/volume66/number11/11131-6205.

[34] Hubbert MK. Nuclear Energy and the Fossil Fuels. Houston (TX): Shell Development Company; Exploration and Production Research Division; 1956. Available at: http://www.hubbertpeak.com/hubbert/1956/1956.pdf.

[35] IEA - International Energy Agency. World Energy Outlook 2020. Available at: https://www.iea.org/reports/world-energy-outlook-2020; 2020.

[36] Istat – Istituto Nazionale di Statistica (2011). Private households by conventional dwelling characteristics with goods and services. Population Housing Census, Private households. Available at: http://dati-censimentopopolazione.istat.it/Index.aspx?lang=en&SubSessionId=23cdacf4-64fa-4d62-88fe-58809f006f58&themetreeid=-200.

[37] Istat (2017). Popolazione per Provincia, anno 2016. http://demo.istat.it/bil2016/dati/province.zip.

[38] Italgas (2019). Acquisizioni mirate, mezzo miliardo di investimenti e una strategia chiara per costruire una solida leadership in Sardegna. https://www.italgas.it/period/ottobre-2019/.

[39] Italgas (2019). Italgas Strategic Plan 2019-2025. https://www.italgas.it/export/sites/italgas/italgas-gallery/Documenti_it/STRATEGY-ITALGAS-2025-2019_DEF.pdf.

[40] Italgas (2020). Relazione Finanziaria Annuale 2019. Avaliable at: https://www.italgas.it/export/sites/italgas/italgas-gallery/Documenti_it/06-investitori/07-debito-e-rating/01-programma-emtn/Documenti-incorporati-mediante-riferimento-Relazione-Finanziaria-Annuale-2018.pdf.

[41] Petrolifera IVI. Impianto di stoccaggio, rigassificazione e distribuzione GNL nel Porto di Oristano-Santa Giusta. Available: Cronoprogramma del lavori; 2019. at: https://va.mite.gov.it/it-IT/Oggetti/Documentazione/6882/9625.

[42] Kelejian, H.H., Prucha, I.R., Yuzefovich, Y. (2004). Instrumental variable estimation of a spatial autorgressive model with autoregressive disturbances: large and small sample results. In LeSage, J.P., Pace, K.R. (Ed.), *Spatial and Spatiotemporal Econometrics* (Advances in Econometrics, 18), Emerald, Bingley, 163-198. https://doi.org/10.1016/S0731-9053(04)18005-5.

[43] Kesicki F. Costs and potentials of reducing CO2 emissions in the UK domestic stock from a systems perspective. Energy Build 2012;51:203–11. https://doi.org/10.1016/j.enbuild.2012.05.013.

[44] King & Spalding LLP. LNG in Europe 2018: An Overview of Import Terminals in Europe. Available at: https://globallnghub.com/wp-content/uploads/2018/09/King.pdf; 2018.

[45] Lee L-F. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. Journal of Econometrics 2007;137:489–514. https://doi.org/10.1016/j.jeconom.2005.10.004.

[46] LeSage, J. P. (1998). *Spatial Econometrics*. Department of Economics, University of Toledo. Toledo (OH). Available at: http://www.spatial-econometrics.com/html/wbook.pdf.

[47] Loulou R, Goldstein G, Noble K. Documentation for the MARKAL Family of Models. IEA: Energy Technology Systems Analysis Programme; 2004. Available at: http://unfccc.int/resource/cd_roms/na1/mitigation/Module_5/Module_5_1/b_tools/MARKAL/MARKAL_Manual.pdf.

[48] Martin A, Möller M, Moritz S. Mixed integer models for the stationary case of gas network optimization. Math Program 2006;105:563–82. https://doi.org/10.1007/s10107-005-0665-5.

[49] Mikalauskienė A, Štreimikienė D, Alebaite I. The main drivers of energy consumption in households. Available at: Inter Proc Econ Devel Research 2012;42(21):105–13. http://ipedr.com/vol42/021-ICKCS2012-K10011.pdf.

[50] MiSE - Ministero dello Sviluppo Economico (2017a). Strategia Energetica Nazionale 2017. Available at: https://www.mise.gov.it/images/stories/documenti/Testo-integrale-SEN-2017.pdf.

[51] MiSE - Ministero dello Sviluppo Economico (2017b). Gas naturale totale distribuito per Provincia, anno 2016. https://dgsaie.mise.gov.it/pub/gas_naturale/consumi/provinciali/ Gas_Distribuito_Province_2016.xlsx.

[52] Misra S, Fisher MW, Backhaus S, Bent R, Chertkov M, Pan F. Optimal compression in natural gas networks: a geometric programming approach. IEEE Trans Control Network Syst 2015;2(1):47–56. https://doi.org/10.1109/TCNS.2014.2367360.

[53] Mokhatab S, Poe WA, Mak JY. *Handbook of Natural Gas Transmission and Processing*, 4-th Ed. Elsevier 2019. https://doi.org/10.1016/C2017-0-03889-2.

[54] OPEC - Organization of Petroleum Exporting Countries (2021b). World proven natural gas reserves by country, Table 9.1. Available at: https://asb.opec.org/data/ASB_Data.php.

[55] OPEC - Organization of Petroleum Exporting Countries (2021a). World marketed production of natural gas by country, Table 9.3. At: https://asb.opec.org/data/ASB_Data.php.

[56] Pantaleo AM, Giarola S, Bauen A, Shah N. Integration of biomass into urban energy systems for heat and power. Part I: An MILP based spatial optimization methodology. Energy Convers Manage 2014;83:347–61. https://doi.org/10.1016/j.enconman.2014.03.050.

[57] Percell PB, Ryan MJ. Steady-state optimization of gas pipeline network operation. In: Proceedings of the 19th PSIG Annual Meeting. Tulsa (OK). 1987.

[58] Regione Sardegna (2004). Deliberazione n. 21/20 del 3 Maggio 2004. Available at: http://www.regione.sardegna.it/documenti/1_39_20050214095057.pdf.

[59] Regione Sardegna (2005). Deliberazione n. 54/28 del 22.11.2005, Sviluppo della rete di distribuzione del metano. Direttive, criteri e modalità per il primo bando di intervento. Attachment 2. Available at: https://www.regione.sardegna.it/documenti/1_72_20051124104609.zip.

[60] Sardegna R. Graduatoria dei bacini risultati idonei al Bando Regionale per lo sviluppo delle reti di distribuzione del Metano. Available at: https://www.regione.sardegna.it/documenti/1_19_20060615133056.pdf; 2006.

[61] Regione Sardegna (2016). Piano Energetico Ambientale della Regione Sardegna 2015-2030. Rapporto Ambientale. https://www.regione.sardegna.it/documenti/1_274_20160804102906.pdf.

[62] Regione Sardegna (2019). Sardegna Energia geoportal. Available at: http://www.regione.sardegna.it/sardegnaenergia/metanizzazionedellasardegna.

[63] Roper LD. World Crude-Oil Future. Blacksburg (VA): Virginia Polythecnic Institute; 2015. Available at: http://www.roperld.com/science/minerals/WorldCrudeOilFuture.pdf.

[64] Ruan Y, Liu Q, Zhou W, Batty B, Gao W, Ren J, et al. A procedure to design the mainline system in natural gas networks. Appl Math Model 2009;33:3040–51. https://doi.org/10.1016/j.apm.2008.10.008.

[65] SGI - Società Gasdotti Italia (2017). Sistema Trasporto Gas Naturale Sardegna: Computo Metrico Estimativo e Quadro Economico delle Opere. https://va.mite.gov.it/it-IT/Oggetti/MetadatoDocumento/218715.

[66] Soha T, Hartmann B. Complex power-to-gas plant site selection by multi-criteria decision-making and GIS. Energy Conversion and Management: X 2022;13:100168. https://doi.org/10.1016/j.ecmx.2021.100168.

[67] SNaM - Società Nazionale Metanodotti (2017). Snam e Sgi: Progetto Comune per la Metanizzazione della Sardegna. Available at: http://www.snam.it/it/media/comunicati-stampa/2017/snam_sgi_progetto_comune_metanizzazione_sardegna.html.

[68] Szoplik J. Forecasting of natural gas consumption with artificial neural networks. Energy 2015;85:208–20. https://doi.org/10.1016/j.energy.2015.03.084.

[69] Thomas S. Drivers of recent energy consumption trends across sectors in EU28. Report prepared for the European Commission: Energy Consumption Trends Workshop; 2018. Available at: https://ec.europa.eu/energy/sites/ener/files/energy_consumption_trends_workshop_report-september_2018.pdf.

[70] Tobler W. A computer movie simulating urban growth in the Detroit region. Economic Geography 1970;46:234–40. https://doi.org/10.2307/143141.

[71] Vasconcelos CD, Lourenço SR, Gracias AC, Cassiano DA. Network flows modeling applied to the natural gas pipeline in Brazil. J Nat Gas Sci Eng 2013;14:211–24. https://doi.org/10.1016/j.jngse.2013.07.001.

[72] Villicaña-Garcí E, Ponce-Ortega JM. Sustainable strategic planning for a national natural gas energy system accounting for unconventional sources. Energy Convers Manage 2019;181:382–97. https://doi.org/10.1016/j.enconman.2018.12.023.

[73] Wang Y, Wang J, Gao M, Zhang D, Liu Y, Tan Z, et al. Cost-based siting and sizing of energy stations and pipeline networks in integrated energy system. Energy Convers Manage 2021;235:113958. https://doi.org/10.1016/j.enconman.2021.113958.

[74] Wu X, Li C, Jia W, He Y. Optimal operation of trunk natural gas pipelines via an inertia-adaptive particle swarm optimization algorithm. J Nat Gas Sci Eng 2014;21:10–8. https://doi.org/10.1016/j.jngse.2014.07.028.

[75] Yousfi M. Gasdotto GALSI: un'opportunità per la Sardegna. Available: Galsi SpA; 2008. at: https://www.regione.sardegna.it/documenti/1_5_20080214111839.pdf.