Residential electricity demand projections for Italy: a spatial downscaling approach

Authors: Massimiliano Rizzati[1], Enrica De Cian[2], Gianni Guastella[3], Malcolm N. Mistry[2,4], Stefano Pareglio[3]

Abstract

This work projects future residential electricity demand in Italy at the local (1 km grid) level based on population, land use, socio-economic and climate scenarios for the year 2050. A two-step approach is employed. In the first step, a grid-level model is estimated to explain land use as a function of socio-economic and demographic variables. In the second step, a provincial-level model explaining residential electricity intensity (gigawatt hours [GWh] per kilometre of residential land) as a function of socio-economic and climatic information is estimated. The estimates of the two models are then combined to project downscaled residential electricity consumption. The evidence suggests not only that the residential electricity demand will increase in the future but, most importantly, that its spatial distribution and dispersion will change in the next decades mostly due to changes in population density. Policy implications are discussed in relation to efficiency measures and the design of green energy supply from local production plants to facilitate matching demand with supply.

This is the pre-print manuscript published in: Energy Policy, 160, 2021.

DOI: https://doi.org/10.1016/j.enpol.2021.112639

When citing, please refer to the published version.

JEL: Q47; R11;

Keywords: electricity demand, projections; spatial downscaling; linear mixed models

[1] Fondazione Eni Enrico Mattei, Corresponding Author, massimiliano.rizzati@feem.it, Corso Magenta, 63 – 20123 Milan, Italy, +39 0252036972
[2] Department of Economics, Ca' Foscari University of Venice - Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC) - RFF-CMCC European Institute on Economics and the Environment
[3] Department of Mathematics and Physics, Università Cattolica del Sacro Cuore – Fondazione Eni Enrico Mattei
[4] The London School of Hygiene and Tropical Medicine (LSHTM), UK

1 Introduction

The availability of high-resolution data on residential electricity demand might benefit energy transition. While many scenarios relevant for decision making, such as the Intergovernmental Panel on Climate Change (IPCC) Shared Socioeconomic Pathways (SSPs) (Riahi et al., 2017) for emissions, land use, and energy, are elaborated at the country level, finer depictions of the same are required for data-informed decision making at sub-national administrative levels. Sub-regional data are often either unavailable or available with considerable delay (Murakami and Yamagata, 2019; Höwer et al., 2019), and electricity demand data are seldom published at levels such as the municipal level (Murakami et al., 2015). Such practice is in contrast with the need for downscaled data at even higher resolutions. The availability of research-based evidence at a more refined geographic scale is key for facilitating detailed spatial studies. Indeed, information available on finer grids makes it possible to arrange the evidence on the basis of policy needs, which might differ among political and administrative units (Gaffin, 2004). For instance, Murakami and Yamagata (2019) discuss how important granular resolutions are for the analysis of climate change mitigation and adaptation processes, arguing that relevant features such as urban form typologies can be captured at this level only.

This work aims to contribute to this policy requirement by developing a spatially detailed dataset of residential electricity demand projections in Italy to 2050 at a very fine grid of approximately 1 kilometre by 1 kilometre (1 km x 1 km). The year 2050 is a key target in the 2018 long-term European Union (EU) strategy to be climate-neutral, an objective at the heart of the European Green Deal (European Commission, 2020). The analysis is focused on residential electricity demand, a relevant contributor to greenhouse gas (GHG) emissions that globally account for about 17% of carbon dioxide (CO_2) emissions (Nejat et al., 2015). Residential electricity demand is subject to potential shifts due to the increase in urbanisation processes (United Nations, 2018) that may, however, be decoupled from population growth (Guastella et al., 2019). The focus on a single country, Italy in our

case, allows for the computational feasibility of our approach[4]. In addition, the heterogeneity of the climatic and socio-economic conditions of Italy makes the country a good case study to test this new methodological approach. Replication of the analysis in other countries is undoubtedly possible and warranted for future research, although with some caveats. For instance, extending the framework to a multi-country setting would require controlling for institutional and regulatory differences among countries, differences in energy prices and supply mix, and increased heterogeneity of the geographical conditions.

Our methodology relies on statistical spatial downscaling, a 'top-down' approach based on available aggregate information that has been of growing interest in several fields (Swan and Ugursal, 2009; Van Ruijven et al., 2019). For instance, Murakami et al. (2015) use spatial downscaling to derive disaggregated electricity demand for Japan at 1 km x 1 km grid resolution in relation to sustainability and energy transition. Several other works have employed the downscaling methodology to retrieve local-level information about population, income and emissions. Gaffin et al. (2004) propose a linear downscaling of global population and GDP until 2100 under different Special Report on Emissions Scenarios (SRES) at approximately a 30 km grid resolution. Van Vuuren et al. (2007) obtain global population, GDP and emissions for IPCC- SRES scenarios downscaled to the grid level of approximately 27.75 km. Nam and Reilly (2013) downscale the global population density from The National Aeronautics and Space Administration (NASA)'s Socioeconomic Data and Application Centre (SEDAC) for 1990–2015 to a 27.75 km resolution grid with a rank-size rule-based approach to estimate city size. Jones and O'Neil (2013) downscale the population in the US to a grid of approximately 11 km for selected SRES by employing a population potentials model. The same authors adopt a gravity approach to downscale the global population from all five SSPs (Jones and O'Neil, 2016). Finally, Murakami and Yamagata (2019) downscale the GDP and population scenarios

[4] The reader is referred to Arbia et al. (2019) for a discussion

developed by the SSPs to an approximately 50 km grid map by applying a four-step approach based on city population projections and urban and rural areas potential.

Few works at the global level seem to be dedicated to achieving resolutions lower than 10 km x 10 km, something that instead can be achieved in works at the national level. The main advantage in the latter case would be the availability of a higher number of relevant variables to be considered in the downscaling exercise and the computational feasibility of parametrised models applied to reduced geographical areas (Gao, 2017).

The downscaling approach proposed in this work follows a relatively simple two-step procedure grounded on the intensity method (Yamagata et al., 2015; Seiya et al., 2016). This procedure uses essential information on electricity intensity, residential area and population projections. This feature ensures its replicability in most countries in Europe and other world regions and has the advantage of being computationally feasible. In this approach, the provincial (NUTS-3 in the European Nomenclature Unit for Territorial Statistics) residential electricity intensity (GWh per ha residential area) is regressed on population density, income and cooling degree days (CDDs). The grid-level total residential area is regressed on population, income and geographical factors, such as the distances of each grid from the most populated areas and the road density (Murakami and Yamagata, 2019). The estimated parameters from the grid-level model are used to project urbanisation in 2050 following the changes in a given socio-economic scenario. The projected urbanised area is then combined with population, income and CDD projections for the selected scenario, and the first model estimates are used to project electricity intensity. Finally, the grid-level demand is obtained as the product of NUTS-3 level projected intensity and grid-level projected urbanised area.

The choice to include both weather and income information in the electricity intensity equation is grounded on the existing literature about electricity and energy consumption determinants. Indeed, the effect of the interconnections between local weather, a changing climate, and residential

electricity demand has been well addressed in the literature (Atalla and Hunt, 2016 and references therein).

The literature on electricity demand determinants is vast and covers multiple explanations in addition to the interconnections between socio-economic and climate changes. For instance, Ko (2013) reviews the effects of a city's physical shape on energy use considering housing size and type, density, community layout, planting, surface coverage, building design, heating efficiency, ventilation, air-conditioning systems and dwellers' behaviour. Yoshida et al. (2019) use location activity data to estimate quasi-real-time energy consumption in commercial buildings. Van Ruijven et al. (2019) study the effects of the changing climate on amplifying energy demand. They develop long-term projections of electricity, oil and natural gas demands, considering climate shocks. Testing this model over an array of scenarios, they find that moderate and vigorous warming might increase the world energy demand by 11%–27% and 25%–58%, respectively. Damm et al. (2017) look at the effect of a 2° C temperature increase on electricity demand for 26 European countries. Maintaining current socio-economic variables, they find a general reduction in consumption. Interestingly in terms of the scope of this work, they find that Italy is the only country where this does not apply, with an increase of 0.6%. Hostick et al. (2014) estimate US energy demand in 2050, finding a decrease in electricity intensity for the residential sector caused by the greening of buildings, more stringent building codes, improved appliance and equipment standards and the level of research on ultra-efficient buildings. Schweizer and Morgan (2016) also look at the US and perform a bounding analysis for electricity demand in 2050 based on changes in adaptation and mitigation policies and low or high GDP growth, finding modest or substantial improvements in energy intensities. The projected values for US electricity demand might then span 3100–17000 terawatt hours (TWh). Huang and Hua (2019) study the balance between economic and green growth, focusing on the eco-efficiency of 191 Chinese cities from 2003–2013 using a spatial approach and data envelopment analysis, finding evidence of convergence for efficiency scores. Zaman et al. (2012) study the determinants of electricity

consumption in Pakistan between 1975 and 2010 with an autoregressive distributed lag (ARDL) model, finding a positive and significant effect of GDP per capita, population growth and foreign direct investment.

Few works in the literature attempted to project energy and electricity consumption in Italy, and to our knowledge none proposed a spatial downscaling to the 1 km x 1 km level resolution. Bianco et al. (2009) relate national economic and demographic variables, such as population and GDP per capita, to find price and GDP consumption elasticities for the historical energy consumption in Italy, spanning the years 1970–2007. They then attempt to provide an accurate forecasting model up to 2017. In line with the Terna (2007) estimates, they forecast a yearly average of 2% growth. Employing a sample of households' hourly electricity load for 2011, Alberini et al. (2019) study residential electricity demand to isolate the effect of temperature values. They find an irrelevant impact until 24.4° C, after which temperature increases result in a sharp rise in consumption up to 9%–12%. For Italy, the effect of cooling is lower, given the predominance of gas heating. This feature seems to disprove the usual asymmetric V-shaped effect of temperature on electricity use, providing a linear increase instead. Besagni and Borgarello (2018) try to characterise the relationship between household composition and energy demand in Italy, segmenting the Italian population using the Italian Household Budget Survey of 2017. They find that socio-economic characteristics play a more important role compared to dwelling and appliance characteristics. Regarding electrical consumption, they find relevant differences related to geography and floor surface areas, pointing to the importance of evaluating the spatial dimension of electricity consumption models.

The remainder of the article is structured as follows. Section 2 describes the downscaling and projection approach in detail and presents the sources of employed data with summary statistics. Section 3 presents model estimates and projections and discusses the characteristics of projected electricity consumption data highlighting the changes in the spatial distributions of residential

electricity demand in Italy and at the regional level. The implications for policy and decision makers, as well as future research directions, are summarised in Section 4.

2 Methodology and Data

2.1 Empirical strategy

The methodology relies on a two-step approach—one statistical model at the grid level and one at the provincial level. Each step adopts a logarithmic model with a mixed-model specification combining the fixed effects of covariates with hierarchically clustered random effects at the different territorial levels (West et al., 2014). These random effects account for spatially distributed random deviations of the observed values from the predictions. Estimation of these kinds of models is feasible using a restricted maximum likelihood approach, in this work using the R lme4 package (Bates et al., 2018). Operationally, this provides a sizeable advantage in estimating parametric downscaling models with a large number of observations. Consider, for instance, that the grid-level model in this study relies on a dataset with 476,836 spatial units, substantially more than the maximum (approximately 100,000) that the most forgiving spatial econometrics model can handle (Yamagata and Seya, 2019). While the spatial econometrics methods are preferable in the presence of a strong spatial dependence, creating a weight matrix is a computationally demanding task. Considering the resolution in this work, for instance, it would limit the sample of grid cells to a single NUTS-2 region. Moreover, the high number of observations (grids) results in some potential downsides in this step due to the high number of empty units, which causes some concerns about zero inflation and overdispersion, as well as the non-normality of our dependent variable. These concerns are easy to address in a mixed model compared to a spatial econometric model.

The first step projects residential area per grid in 2050. First, a linear model, expressed in natural logarithms, is proposed in equation (1):

$$ln(A_{i,c,n,r,t}) = \beta_0 + \beta_1 ln(P_{i,c,n,r,t}) + \sum_j \beta_j \ln(X^j_{i,c,n,r,t}) + \omega_i + \lambda_c + \gamma_n + \theta_r + \tau_t + u_{i,t}. \quad (1)$$

Here, the total residential area A is indexed[5] for grid i, local administrative unit (LAU) c, NUTS-3 (provincial) level n, NUTS-2 (regional) level r and time t. Among the explanatory variables, there is total grid population $P_{i,c,n,r,t}$ along with grid-specific j geographical factors included in matrix X. These include the road density, RD, and the distances from the closest populated grid cell, located in the grid's i LAU and NUTS 3, DC and DN, respectively. The location-specific determinants of land use are then considered based on the set of geographically fixed effects and the time trend but also including the grid-specific, LAU-specific, province-specific and region-specific random effects, ω_i, λ_c, γ_n and θ_r, respectively, as well as the idiosyncratic error term $u_{i,c,n,r,t}$.

The projected values for the 2050 residential area are computed as the conditional expectation of the model of Equation (1), extended by expected change in the time-varying variable $P_{i,t}$, as shown in Equation (2):

$$ln(\tilde{A}^S_{i,c,n,r,2050}) = \ln(\hat{A}_{i,c,n,r,t}) + \hat{\beta}_1\left(\ln\left(\Delta\tilde{P}^S_{i,c,n,r,2050}\right)\right) + \hat{\tau}\,, \quad (2)$$

where $\tilde{A}^S_{i,c,n,r,2050}$ is the projected residential area per grid in 2050 for each considered scenario S. This is calculated by adding $\hat{A}_{i,c,n,r,t}$, the prediction from the fitted model of Equation (1), and the expected change associated with the population variation from 2018–2050 and the fitted time trend $\hat{\tau}$. Notice that future $\tilde{P}^S_{i,c,n,r,2050}$ values are selected from each different scenario S. These are discussed in the next subsection about data sources.

The second step of the model is at the NUTS-3 level and starts with Equation (3):

$$ln\left(I_{n,r,t}\right) = \alpha_0 + \alpha_1\left(ln(D_{n,r,t})\right) + \alpha_2\left(ln(G_{n,r,t})\right) + \alpha_3\left(C_{n,r,t}\right) + \delta_t + \eta_n + \zeta_r + \varepsilon_{n,r,t}. \quad (3)$$

[5] Notice that the subscripts reflect the nested structure of the dataset; for conciseness, the variables at the grid and provincial level will be presented with just the lowest level indexing outside of equations (1)–(4).

*L*et $I_{n,r,t}$ be the electricity intensity measure indexed at the Italian NUTS-3 (provincial) level n, NUTS-2 (regional) level r and year t, taken in natural logarithm terms. Intensity is measured as the quantity of electricity consumption $E_{n,r,t}$ (in GWh) per residential land use $A_{n,t}$ (in square km), such that $I_{n,r,t} = (E_{n,r,t}/A_{n,r,t})$. The model in Equation (3) links intensity to its selected determinants, namely the population density $D_{n,r,t} = \frac{P_{n,r,t}}{A_{n,r,t}}$, the GDP per capita $G_{n,r,t} = \frac{GDP_{n,r,t}}{P_{n,r,t}}$ taken in natural logarithm terms and the CDDs[6] $(C_{n,r,t})$. In the estimating equation, the time trend δ_t is included to account for common time patterns in electricity intensity over time. The composite error term is made by province-specific and region-specific random effects (η_n, ζ_r) allowing random variations in the observed intensity around the sample mean and an idiosyncratic error term $\varepsilon_{n,r,t}$.

The projection of the electricity intensity in 2050 can be computed as the expected intensity value conditional to the projected scenario value of the independent variables, as declared in Equation (4):

$$ln(\tilde{I}^{S}_{n,r,2050}) = ln(\hat{I}_{n,r,2018}) + \hat{\alpha}_1 ln(\Delta\tilde{D}^{S}_{n,r,2050}) + \hat{\alpha}_2 ln(\Delta\tilde{G}^{S}_{n,r,2050}) + \hat{\alpha}_3 ln(\Delta\tilde{C}_{n,r,2050}) + \hat{\delta}\,, \quad (4)$$

with $\tilde{I}^{S}_{n,r,2050}$ being the projection of the electricity intensity in 2050 for the scenario *S*. $\hat{I}_{n,r,t}$ is the prediction from the fitted model of Equation (3). The latter is projected to 2050 by multiplying the fitted coefficients $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3$ of the selected determinants by the delta of their 2050 projected values compared to the last year available value.

It must be noted that the projection of the population density value $\tilde{D}^{S}_{n,r,2050}$ requires a projected residential area value for each scenario *S*. These values are recovered from the first step of the procedure by aggregating the values of $\tilde{A}^{S}_{i,c,n,r,2050}$.

[6] CDDs, expressed as degree Celsius days and aggregated to monthly or annual timescales, is a commonly employed metric in climate–energy literature to model the energy use for space cooling (see Appendix A for more details).

Finally, reverting the logarithm terms and multiplying the grid-level value of residential area from Equation (2) by the NUTS-III intensity from Equation (4), the residential electricity demand at the grid level is obtained (5). In

$$\tilde{E}_{i,2050} = \tilde{I}_{n,2050} * \tilde{A}_{i,2050}, \quad (5)$$

$\tilde{E}_{i,2050}$ is the projected electricity demand in GWh for grid i in the target year 2050. This follows from the definition of the projected electricity intensity $\tilde{I}_{n,2050} = (\tilde{E}_{n,t}/\tilde{A}_{n,t})$, which multiplied with the projected residential area $\tilde{A}_{i,2050}$ at the grid level, distributes the residential electricity demand using the residential area share as a weight.

2.2 Data

Table 1[7] provides the descriptive statistics of the variables used to calibrate our models. To begin with, the residential area model and its projections require a spatial grid reference. This work selects the same grid map structure as the Gridded Population of the World (GPW) dataset (CIESIN, 2018), with a resolution of approximately 1 km at the equator (closer to 0.7 km for Italy). To avoid possible distortions in reconciling new grids, the structure is kept as is. The grid map for Italy is recovered by intersecting the national administrative boundary of Italy from the Italian national institute of statistics (Istat) to the aforementioned global dataset, accounting for 476,836 individual grid cells.

The residential area variable A is compiled by aggregating for each grid cell the area in square meters of the Corine Land Cover (CLC) classification of continuous and discontinuous urban fabric (codes 111 and 112) for its available years at the time of writing (2000, 2006, 2012 and 2018). Population counts at the grid level are taken from the GPW (CIESIN, 2018). To avoid distortions from the mismatching of the periods (as the GPW is available in 5-year intervals from 2000), the grid's population values are rescaled to match the land use years using NUTS 3 population census values for the correct one.

[7] Tables are generated using the Stargazer package (Hlavac, 2018).

The remaining auxiliary variables are obtained by pairing each grid's centroid with the closest observation. RD is from the Global Roads Inventory Project (GRIP) total density dataset, with all road types combined (Meijer et al., 2018), which provides meters of roads per square km. The value is then converted to km of roads per square km. DC and DP, the distances to the closest most populated grid cell in the same LAU area and in the same NUTS-3 area, are calculated by aggregating the gridded population counts of the variable P for each LAU and NUTS-3 unit. Each grid is then matched with the one with the closest highest population, and then distances are calculated in degrees.

The projected population counts needed in Equation (2) are recovered from the 2050 population projections by Gao (2020). The latter presents downscaled population counts for each SSP of the approximately 11 km resolution data published in Jones and O'Neill (2016). SSPs provide different quantitative trajectories of socio-economic variables' evolutions in relation to different combinations of climate policies—SSP1: Sustainability, SSP2: Middle of the Road, SSP3: Regional Rivalry, SSP4: Inequality and SSP5: Fossil-fuelled Development (Rihai et al., 2017). This work adopts the SSP2 scenario as the benchmark case, while the others are nonetheless developed to test the sensitivity of the results.

The second step of the procedure, namely the electricity intensity model in Equation (3) and Equation (4) are calibrated as follows. The residential electricity demand at the provincial level in GW/h is recovered from TERNA, the Italian transmission system operator. Changes in NUTS-3 codes occurred in the sample period (e.g. due to the creation of new provinces) required to harmonise data using the share of electricity over provincial level-populations as a weight. Residential areas at the NUTS 3 level are aggregated from the grid-level data. Population at the NUTS-3 level is from the Istat population census data. Population counts have been harmonised to prevent discontinuities as a result of the changing provincial boundaries. The NUTS-3 GDP in EUR per inhabitant at current prices is available from Eurostat. Historical CDDs aggregated at the NUTS-3 level were recovered

from the dataset presented in Mistry (2019). More details about the methodology to obtain the historical CDDs are presented in Appendix A1.

The projections of Equation 4 require the projected values from residential areas, population, GDP and CDDs. Residential areas and population projections are obtained by aggregating the grid-level ones from the previous step. GDP projections for each SSP scenario are provided by the Organisation for Economic Co-operation and Development (OECD) at the country level in constant 2005 USD (Dellink et al., 2017). These values are adjusted for inflation, converted into current EUR, downscaled to the NUTS-3 level by assuming a constant share of growth from the 2018 values and then taken in per capita terms. Finally, projections of the 2050 CDDs are elaborated similar to the historical values but with NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) for a single scenario of moderate warming (Representative Concentration Pathway 4.5). More details about the methodology for constructing the projected CDDs are discussed in Appendix A1.

[Table 1]

In Table 2, we present the descriptive statistics for the variables employed in the first step of our methodology at the 1 km x 1 km grid level, with average values that display an increasing trend in the grid-level population count and residential areas. Projections of grid cell-level population to 2050 for each SSP scenario show an increase in population and spatial dispersion for all scenarios, while SSP3 and SSP5 present the lowest and highest mean value, respectively.

[Table 2]

Table 3 presents the same summary statistics for the variables employed in the NUTS-3 level step of the analysis. Intensity decreases after 2000 due to the residential area increase (as in the case of grid cell level), after which it is stable until 2018, despite increasing NUTS-3 electricity consumption. For all other variables, the observed values increase from 2000–2018. In addition, the three projected

variables for the year 2050 show an increase in 2050, with the different scenarios presenting different mean values due to their narratives.

[Table 3]

In Figure 1, the two main dependent variables are depicted, residential area and electricity intensity for 2018.

[Figure 1]

3 Results and Discussion

This section presents the results of the linear mixed model estimation for both the NUTS-3 level intensity model and the grid-level residential area model and the resulting dataset of projected residential electricity demand for 2050. The results of the grid-level model are presented in Table 4[8]. The fixed effects coefficients of the preferred model, column (4) of Table 4, and the predicted values for 2018, are used to calibrate Equation 2 and obtain the projected 2050 residential area value for each grid cell. As a realistic assumption, the projection is bounded not to exceed the limit of the grid in case of growing areas and assumed to be subject to hysteresis, meaning that in case of a decrease the value of 2018[9] is kept. Residential area increases in percentage due to a rise in population, a feature maintained in all the tested models, and with a positive and significant time trend. The closest most populated cell in the same LAU and NUTS-3 shows negative and positive signs, respectively. These features seem to characterise a cell as more peripheral for the DC variable, while DP might present nonlinear effects. Finally, the RD variable shows a negative and insignificant sign. A

[8] As a robustness check this model is fitted with the R package glmmTMB (Magnusson et al., 2017) allowing for a zero inflation and overdispersion structure conditioned on population levels. Results are robust, at least in terms of slope, magnitude and significance of the population coefficient. It is possible, however, that different assumptions about the dependent variable distribution may impact projections. We leave for future work a refinement in this direction.

[9] Reductions of residential area can be observed in the 2000–2018 series, most probably due to the conversion of urban fabric areas to other classification codes, such as from urban fabric to commercial use but with a very marginal amount of grid cells.

robustness check is presented in Appendix B.1, where reduced samples of the data are fitted with models that accounts explicitly for the spatial structure: a linear spatial lag model with fixed spatial effects and a linear mixed model with spatial autocorrelation in the errors.

[Table 4]

Table 5 reports a summary of the results of this projection compared to the values for 2018 (the last available year). Growth in the projected variable and in dispersion could be observed. This follows the model, allowing for residential areas to be generated in previously empty cells and given the more dispersed nature of the 2050 grid-level population projection scenarios used.

[Table 5]

The best model was selected by iterating the random intercept model, adding the other variables and testing improvements using ANOVA; these models are presented in Table 6. Electricity intensity reports a positive percentage effect of population density and GDP per capita, in line with the results in the literature. A less important but still significant at the 10% level effect is found for CDD, which points to an increase in energy intensity in case of rising temperatures. The time dimension is the only variable that appears not to be significant.

[Table 6]

The final result consists of a spatially explicit dataset of the projected residential electricity demand for the year 2050, at 1 km x 1 km gridded resolution in GWh, with a total of 476,836 grid cells covering the geographical territory of Italy. The summary statistics for the downscaled residential electricity demand (E) at the 1 km x 1 km grid level are presented in Table 7. The variable increases in every scenario, driven by the positive signs of the coefficients.

[Table 7]

Figure 2 maps the results of the downscaling analysis at different scales for illustrative purposes.

[Figure 2]

The results are compared with an external forecast. In the Stated Policy scenario, the International energy Agency (IEA) estimates a 2.1% yearly growth rate for global electricity demand and a global residential electricity demand in the order of 6000 TWh in 2019 growing to 11,000 TWh in 2040 (an increase of 83.3%, compared with this work with an SSP2 value of 87.47%) (IEA, 2019). Referring to these forecasts, this work's average annual growth rates (aggregated at the national level in Table 7) are more in line with developing countries. This is to be expected, given that the proposed methodology is neutral in terms of efficiency measures. Future research might start from the generated data to test different efficiency evolution scenarios as a policy evaluation exercise. Besides considering changes in the magnitude for the projected variables across each grid cell, other interesting information can be extracted from the results. For instance, whether the downscaling of residential electricity demand shows evidence of some significant change in the underlying spatial distribution is one example.

A simple measure of inequality, the Gini Index for the distribution of GWh, is computed across the grid cells for 2018 and all SSP scenario projections (Table 8). Overall, the value is high, driven by the presence of a great number of grid cells devoid of residential areas. However, it seems that the rise and dispersion of electricity demand reduce (albeit very slightly) the inequality index. Because the Gini index is an a-spatial measure of inequality, the analysis is complemented with Moran's indicator of global autocorrelation (Moran, 1950) to test the spatial unevenness of the electricity demand distribution in 2018 and 2050. The results are also reported in Table 8. As expected, the p-values (< 0.05) reject the hypothesis of no spatial autocorrelation in all cases. The value of the statistics is, however, lower only for SSP3, which follows this scenario characterisation of lower population growth, lower GDP growth and eventually lower projected electricity demand.

[Table 8]

In Figure 3, in the benchmark scenario SSP2 the absolute change from 2018 to 2050 in the regional concentration (calculated using a Herfindahl–Hirschman index) of residential electricity demand is compared and plotted against the change in regional electricity intensity calculated by aggregating grid-level data for residential areas and residential electricity demand. The size of the markers is weighted to reflect the respective levels of residential electricity in 2018 and 2050. This plot seems to suggest a slightly positive effect of the change of electricity intensity on the change in the concentration index; however, the fitted linear model is not significant. Regions with a greater residential electricity demand saw a decrease in concentration but of a smaller magnitude.

[Figure 3]

4 Conclusion and Policy Implications

In this work, a dataset with a 1 km x 1 km grid resolution for residential electricity demand in Italy in 2050 is constructed. The data are assembled by projecting electricity intensity at the provincial level, a step itself based on projections of residential areas at the grid level. The projections are obtained by identifying the effects of different socio-economic and geographical variables on historic residential areas and electricity intensity and combining these effects with the 2050 projections of the explanatory variables.

The methodology projects an increase in residential areas, an increase in electricity intensity and an increase in the average downscaled residential electricity demand per grid. The work quantifies the effect of different socio-economic and geographical variables on residential area and electricity intensity. Residential area depends positively on population, while the effect of geographical location is mixed. Being projections focused on 2050, population scenarios with positive growth result in a residential area increase. For electricity intensity, a positive effect of population density, GDP per

capita and CDDs is found. Combining these results with the 2050 projections of the aforementioned variables provides the final positive increase in projected 2050 electricity intensity.

Finally, the 2050 spatial distribution of the results is investigated. The spatial distribution of the projected electricity demand seems to be a more dispersed version compared to its 2018 values. The increase in 2050 residential areas, itself dispersed by the underlying projections of population in 2050, is a likely explanation. Various tests for concentration seem to confirm this take.

It is possible to have different takeaways about the policy implications of this work. First, policies that target the growth of residential areas might prove particularly effective to contain the increase in the electricity demand's growth rates. Second, the expansion of residential land use invites reflection about the role of efficiency-enhancing policy measures. The changing spatial distribution of residential electricity demand in this work implies that the increase in land use outside of the current built-up area could be one of the leading factors in sizeable demand growth rates. Policies aiming at decarbonising the energy mix should then foster the efficiency measures of new buildings (Aroonruengsawat et al., 2012). In this work, however, the efficiency is fixed. Acting only on new buildings might not then suffice to reduce electricity demand below current levels. The final policy implication, therefore, confirms the need for more aggressive efficiency policies (Reyna and Chester, 2017). Datasets such as the one developed in this work could help in the decision about the allocation of resources towards the latter or to legacy buildings, providing high-resolution data on which to test different efficiency scenarios.

The quantified link between residential areas and population counts' distribution and their residential electricity demand projections could indeed help in the assessment of future scenarios besides the SSPs. A relevant example is the current COVID-19 pandemic; due to lockdown restrictions and increased preferences towards smart working and personal transportation methods (Moslem and Campisi, 2021), the residential choice of households might change, with potential impacts on their

residential electricity demand (Bielecki et al., 2021). For instance, Beria and Lunkar (2021) study the mobility and spatial distribution dynamics of Italian residents during the first lockdown wave in March 2020, finding an increased preference towards urban belts. Future research could then further generalise the projections of this work to different degrees of dispersion of the future population according to quantified changes in location preferences. Another valuable analysis would be to update the results to account for the direct and current effects of the pandemic. Santiago et al. (2021) report that economic signal to electricity demand due to restrictions has been observed for many countries, with a decrease of 25% electricity use in Italy. However, the same work reports that this is a consequence of the decrease in the industrial sector, while in Italy an increase in the demand by households was observed. This also happened in other countries (Krarti and Aldubyan, 2021) and follows the increase in cooking and use of digital devices (Snow, 2020). This seems to lead to a hypothesised increase in the electricity intensity for the year of the pandemic, at least in the case of Italy and assuming all the other drivers did not vary. This could revise the results of the current work upwards. However, a thorough assessment would require the development of pandemic-tailored 2050 projections for the time-varying variables in each step of the procedure.

The approach we propose has a clear limitation. The electricity price, considered a standard determinant of electricity consumption, is missing from the analysis, both in terms of current information and projected change. Filipović, Verbič, and Radovanović (2015) study the energy intensity in the EU at the national level from 1990–2012 using a panel model, finding energy prices, energy taxes and GDP per capita to have an effect, with the price effect being the most significant. The price information is not available at the NUTS-3 level in Italy or, to the authors' knowledge, in other European countries at the same administrative level. This lack of information prevents embedding the price effect in the proposed model for residential electricity intensity. Despite the evident limitation, we deem the proposed approach consistent with the overall goal of the statistical approach, that is, to provide a framework for projections rather than make inferences about the

determinants of electricity demand. Notwithstanding the lack of price information, the electricity intensity estimates proposed in the paper are, in fact, rather precise. Moreover, past works suggest that price might not be a relevant explanatory variable in the specific Italian context (Bianco et al., 2009).

Future refinements of this analysis would ideally develop a third stage for parametric 2050 population projections at the grid level. Moreover, as computational resources become more available it should be possible to refine the analysis with spatial econometric approaches and with mixed models beyond the linear specifications to account for the peculiar distribution of residential areas. However, the current approach seems quite flexible, as the same procedure could be iterated for different target years' projected variables to perform scenario analysis, as well as to other sectors, for instance, substituting residential areas and residential electricity demand with industrial areas and industrial electricity demand. Moreover, further tests including more than one country, especially with different energy mixes, should also be performed to generalise the results.

Finally, the changing spatial distribution of residential electricity demand, conditioned on future projections and current relevant parameters, could be linked with geographically dependent information about electricity generation from renewable sources. The latter is often based on spatially disaggregated and heterogeneous processes, which need to be reconciled with the local demand (Ramachandra and Shruthi, 2007; Aydin et al., 2010; Barrington-Leigh and Ouliaris, 2017). For instance, given information on solar radiation and wind, and especially in the case of decentralised and autonomous generation, the methodology could provide a valuable guide for policymakers in assessing future trends and areas to prioritise with tailored interventions.

Acknowledgements: We gratefully acknowledge the comments and suggestions from the participants in the ninth Italian Association of Environmental and Resource Economists (IAERE) Annual Conference (online, April 21-23, 2021) and from the two anonymous referees. Errors remain our own.

M. Rizzati, G. Guastella and S. Pareglio did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. E. De Cian and M. Mistry received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 756194.

References

Alberini, A., Prettico, G., Shen, C., & Torriti, J. (2019). Hot weather and residential hourly electricity demand in Italy. *Energy*, *177*, 44-56.

Arbia, G., Ghiringhelli, C., & Mira, A. (2019). Estimation of spatial econometric linear models with large datasets: How big can spatial Big Data be? *Regional Science and Urban Economics*, 76, 67-73.

Aroonruengsawat, A., Auffhammer, M., & Sanstad, A. H. (2012). The impact of state level building codes on residential electricity consumption. *The Energy Journal*, 33(1), 31-52.

ASHRAE (2009). 2009 American Society of Heating, Refrigerating and Air-Conditioning Engineers Handbook: Fundamentals. ASHRAE, Atlanta.

Atalla, T. N., & Hunt, L. C. (2016). Modelling residential electricity demand in the GCC countries. *Energy Economics*, *59*, 149-158.

Aydin, N. Y., Kentel, E., & Duzgun, S. (2010). GIS-based environmental assessment of wind energy systems for spatial planning: A case study from Western Turkey. *Renewable and Sustainable Energy Reviews*, *14*(1), 364-373.

Barrington-Leigh, C., & Ouliaris, M. (2017). The renewable energy landscape in Canada: A spatial analysis. *Renewable and Sustainable Energy Reviews*, *75*, 809-819.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Green, P. (2018). Package 'lme4'. Version, 1, 17.

Beria, P., & Lunkar, V. (2021). Presence and mobility of the population during the first wave of Covid-19 outbreak and lockdown in Italy. *Sustainable Cities and Society*, *65*, 102616.

Besagni, G., & Borgarello, M. (2018). The determinants of residential energy expenditure in Italy. *Energy*, *165*, 369-386.

Bianco, V., Manca, O., & Nardini, S. (2009). Electricity consumption forecasting in Italy using linear regression models. *Energy*, *34*(9), 1413-1421.

Bielecki, S., Skoczkowski, T., Sobczak, L., Buchoski, J., Maciąg, Ł., & Dukat, P. (2021). Impact of the lockdown during the COVID-19 pandemic on electricity use by residential users. *Energies*, *14*(4), 980.

[dataset] Center for International Earth Science Information Network - CIESIN - Columbia University, (2018). Gridded population of the world, version 4 (GPWv4): Population count, revision 11. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY. https://doi.org/10.7927/H4JW8BX5. Accessed 01/10/2020.

Damm, A., Köberl, J., Prettenthaler, F., Rogler, N., & Töglhofer, C. (2017). Impacts of+ 2 C global warming on electricity demand in Europe. *Climate Services*, *7*, 12-30.

Dellink, R., Chateau, J., Lanzi, E., & Magné, B. (2017). Long-term economic growth projections in the Shared Socioeconomic Pathways. *Global Environmental Change*, *42*, 200-214.

European Commission, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing the framework for achieving climate neutrality and amending Regulation (EU) 2018/1999 (European Climate Law) COM/2020/80 final

[dataset] © European Union, Copernicus Land Monitoring Service 2000, 2006, 2012, 2018 European Environment Agency (EEA).

Filipović, S., Verbič, M., & Radovanović, M. (2015). Determinants of energy intensity in the European Union: A panel data analysis. *Energy*, *92*, 547-555.

Gaffin, S. R., Rosenzweig, C., Xing, X., & Yetman, G. (2004). Downscaling and geo-spatial gridding of socio-economic projections from the IPCC Special Report on Emissions Scenarios (SRES). *Global Environmental Change*, *14*(2), 105-123.

Gao, J. (2017). Downscaling global spatial population projections from 1/8-degree to 1-km grid cells. National Center for Atmospheric Research, Boulder, CO.

[dataset] Gao, J. (2020). Global 1-km downscaled population base year and projection grids based on the Shared Socioeconomic Pathways, revision 01. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY. https://doi.org/10.7927/q7z9-9r69.

Guastella, G., Oueslati, W., & Pareglio, S. (2019). Patterns of urban spatial expansion in European cities. *Sustainability*, *11*(8), 2247.

Hlavac, Marek (2018). Stargazer: Well-formatted regression and summary statistics tables. R Package Version 5.2.2. https://CRAN.R-project.org/package=stargazer

Hostick, D. J., Belzer, D. B., Hadley, S. W., Markel, T., Marnay, C., & Kintner-Meyer, M. C. (2014). *Projecting electricity demand in 2050* (No. PNNL-23491). Pacific Northwest National Lab (PNNL), Richland, WA.

Höwer, D., Oberst, C. A., & Madlener, R. (2019). General regionalisation heuristic to map spatial heterogeneity of macroeconomic impacts: The case of the green energy transition in NRW. *Utilities Policy*, *58*, 166-174.

Huang, J., & Hua, Y. (2019). Eco-efficiency convergence and green urban growth in China. *International Regional Science Review*, *42*(3-4), 307-334.

IEA (2019). World Energy Outlook 2019. IEA, Paris. https://www.iea.org/reports/world-energy-outlook-2019

Jones, B., & O'Neill, B. C. (2013). Historically grounded spatial population projections for the continental United States. *Environmental Research Letters*, *8*(4), 044021.

Jones, B., & O'Neill, B. C. (2016). Spatially explicit global population scenarios consistent with the Shared Socioeconomic Pathways. *Environmental Research Letters*, *11*(8), 084003.

Ko, Y. (2013). Urban form and residential energy use: A review of design principles and research findings. *Journal of Planning Literature*, *28*(4), 327-351.

Krarti, M., & Aldubyan, M. (2021). Review analysis of COVID-19 impact on electricity demand for residential buildings. *Renewable and Sustainable Energy Reviews*, 143, 110888.

Magnusson, A., Skaug, H., Nielsen, A., Berg, C., Kristensen, K., Maechler, M., ... & Brooks, M. M. (2017). Package 'glmmTMB'. R Package Version 0.2.0.

[dataset] Meijer, J. R., Huijbegts, M. A. J., Schotten, C. G. J., & Schipper, A. M. (2018). Global patterns of current and future road infrastructure. *Environmental Research Letters*, 13-064006, 1-10. www.globio.info

[dataset] Mistry, M. N. (2019). Historical global gridded degree-days: A high-spatial resolution database of CDD and HDD. *Geoscience Data Journal*, *6*(2), 214-221.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17-23.

Moslem, S., Campisi, T., Szmelter-Jarosz, A., Duleba, S., Nahiduzzaman, K. M., & Tesoriere, G. (2020). Best–worst method for modelling mobility choice after COVID-19: Evidence from Italy. *Sustainability*, *12*(17), 6824.

Murakami, D., & Yamagata, Y. (2019). Estimation of gridded population and GDP scenarios with spatially explicit statistical downscaling. *Sustainability*, *11*(7), 2106.

Murakami, D., Yamagata, Y., & Seya, H. (2015). Estimation of spatially detailed electricity demands using spatial statistical downscaling techniques. *Energy Procedia*, *75*, 2751-2756.

Nam, K. M., & Reilly, J. M. (2013). City size distribution as a function of socio-economic conditions: An eclectic approach to downscaling global population. *Urban Studies*, *50*(1), 208-225.

Nejat, P., Jomehzadeh, F., Taheri, M. M., Gohari, M., & Majid, M. Z. A. (2015). A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries). *Renewable and Sustainable Energy Reviews*, *43*, 843-862.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). Package 'nlme'. Linear and Nonlinear Mixed Effects Models, Version, 3(1).

Ramachandra, T. V., & Shruthi, B. V. (2007). Spatial mapping of renewable energy potential. *Renewable and Sustainable Energy Reviews*, *11*(7), 1460-1480.

Reyna, J. L., & Chester, M. V. (2017). Energy efficiency to reduce residential electricity and natural gas use under climate change. *Nature Communications*, *8*(1), 1-12.

[dataset] Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'Neill, B. C., Fujimori, S., ... & Tavoni, M. (2017). The shared socio-economic pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Global Environmental Change*, *42*, 153-168. ISSN 0959-3780, DOI:110.1016/j.gloenvcha.2016.05.009

Santiago, I., Moreno-Munoz, A., Quintero-Jiménez, P., Garcia-Torres, F., & Gonzalez-Redondo, M. J. (2021). Electricity demand during pandemic times: The case of the COVID-19 in Spain. *Energy Policy*, *148*, 111964.

Schweizer, V. J., & Morgan, M. G. (2016). Bounding US electricity demand in 2050. *Technological Forecasting and Social Change*, *105*, 215-223.

Seya, H., Yamagata, Y., & Nakamichi, K. (2016). Creation of municipality level intensity data of electricity in Japan. *Applied Energy*, *162*, 1336-1344.

Snow, S., Bean, R., Glencross, M., & Horrocks, N. (2020). Drivers behind residential electricity demand fluctuations due to COVID-19 restrictions. *Energies*, *13*(21), 5738.

Swan, L. G., & Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, *13*(8), 1819-1835.

Terna (2007). Forecasting of electricity demand and power needs in Italy for the years 2007–2017. http://www.terna.it/default/
Home/SISTEMA_ELETTRICO/statistiche/dati_statistici/tabid/418/Default.aspx

United Nations, Department of Economic and Social Affairs, Population Division (2018). World Urbanization Prospects: The 2018 Revision, Online Edition.

van Ruijven, B. J., De Cian, E., & Wing, I. S. (2019). Amplification of future energy demand growth due to climate change. *Nature Communications*, *10*(1), 1-12.

van Vuuren, D. P., Lucas, P. L., & Hilderink, H. (2007). Downscaling drivers of global environmental change: Enabling use of global SRES scenarios at the national and grid levels. *Global Environmental Change*, *17*(1), 114-130

West, B. T., Welch, K. B., & Galecki, A. T. (2014). Linear Mixed Models: A Practical Guide Using Statistical Software. Crc Press, New York

Yamagata, Y., Murakami, D., & Seya, H. (2015). A comparison of grid-level residential electricity demand scenarios in Japan for 2050. *Applied Energy*, *158*, 255-262.

Yamagata, Y., & Seya, H. (Eds.). (2019). Spatial analysis using big data: Methods and urban applications. Academic Press, Cambridge.

Yoshida, T., Yamagata, Y., & Murakami, D. (2019). Energy demand estimation using quasi-real-time people activity data. *Energy Procedia*, *158*, 4172-4177.

Zaman, K., Khan, M. M., Ahmad, M., & Rustam, R. (2012). Determinants of electricity consumption function in Pakistan: Old wine in a new bottle. *Energy Policy*, *50*, 623-634.

APPENDIX

A

A.1 Cooling Degree Days

Cooling Degree Days (CDDs; measured as °C days) is a commonly used indicator to model the potential impacts of increased air temperature on the energy demand for space cooling (ASHRAE, 2009; Mistry 2019). CDDs are defined as the cumulative sum of the positive differences between daily mean surface air temperature (Td in °C) and a base threshold temperature (Tb in °C) accumulated over a certain time period (e.g. a month or a year) (Mistry 2019) (Equation A1).

$$\mathrm{CDD} = \sum_{i=1}^{n} (T_d - T_b)^+ \qquad \text{(A1)}$$

The annual CDDs in the historical period in our study are drawn from Mistry (2019)'s dataset, assembled using the daily minimum (Tmin) and maximum (Tmax) surface temperatures[10] at a high-spatial resolution (0.25° gridded, about 27 km x 27 km at the equator) from the NASA Global Land Data Assimilation System (GLDAS). At each grid cell, the CDDs are calculated using the American Society of Heating, Refrigerating and Air Conditioning (ASHRAE) method (Equation A1) and by fixing the Tb at 24 °C. For our study, we extract the data covering the grid cells within mainland Italy, using R packages—sp, raster and rgdal.

A.2 Projections

Projections of future CDDs are assembled using similar methodology as in the historical period but using bias-corrected Tmin and Tmax from NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP)[11]. NEX-GDDP is a large ensemble of downscaled and biased-corrected 0.25 gridded daily meteorological fields from 21 global climate models (GCMs) that simulate vigorous (RCP 8.5) and moderate (RCP 4.5) warming under the Coupled Model Intercomparison, Phase V (CMIP5) climate modelling exercise. For our projections, we use the median CDDs computed using the data from 21 GCMs.

B

A linear spatial lag with fixed spatial effects and a fixed time effects model for residential area was run on a NUTS-3 (ITC4C, Milan) subsample of the original dataset with 10,428 observations. The results are displayed in Table B.1. The log of the population per grid shows a positive and significant coefficient. The only other significant variables are the distance from the most populated cell in the

[10] The T_d used in the computation of CDDs in Mistry (2019) is computed as $T_d = (T_{max} + T_{min})/2$.
[11] Data accessed from https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp in August 2020.

same LAU, DC and the year. The spatial autoregressive coefficient Lambda is positive and significant. In Table B.2, the results for an estimation using a linear mixed model extended with a spherical correlation structure based on the grid's coordinates are presented. The model, fitted with the *nmle* R package (Pinheiro et al., 2017), was able to obtain results only for a sample restricted for one year, 2018; this removes the year fixed effect and the individual grid random effect from the estimation. The coefficients signs are consistent with the results presented in Table 6, except for the road density variable RD. In Table 6 however RD ceases to be significant with the inclusion of the most complete structure of nested random effects, an effect that could be lost here.

[Table B1]

[Table B2]

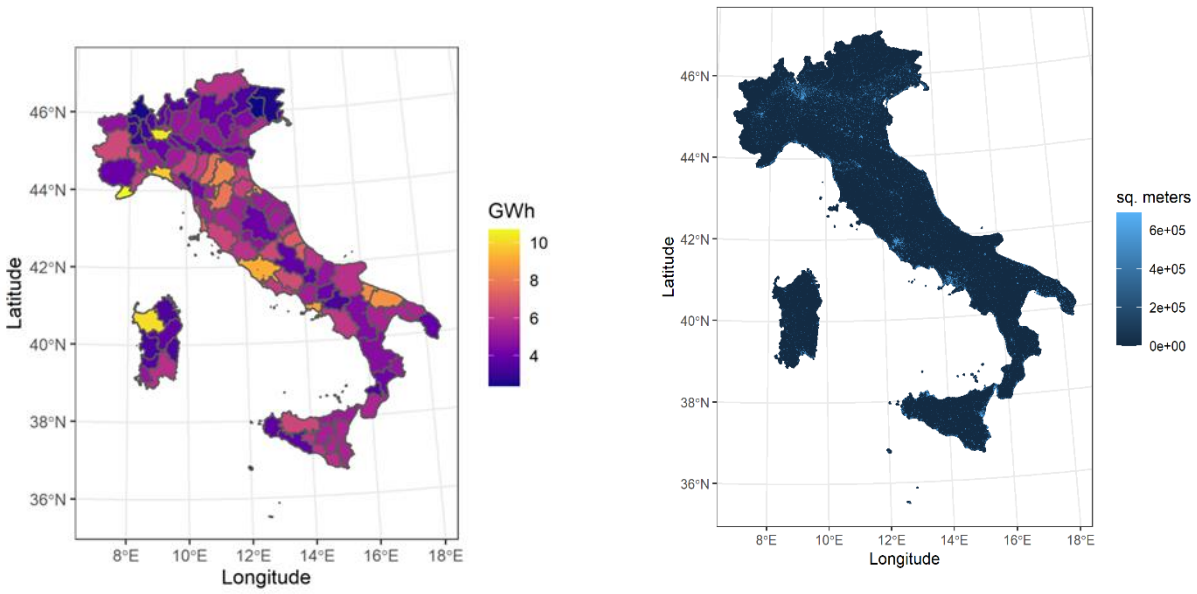

Figure 1 Mapping of the model's two-step dependent variables. The left panel maps residential electricity intensity at the NUTS-3 level for 2018 in GWh. The right panel, maps residential areas in 2018 in square meters.

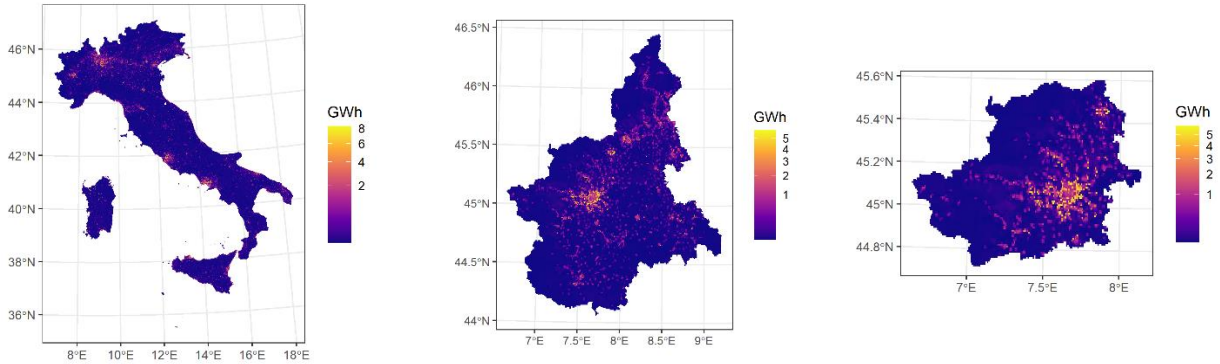

Figure 2 Differences in the downscaled residential electricity demand between in the 2050 projections (SSP2) and the 2018 values. These are mapped with an extent at different resolutions: Left panel NUTS-0, middle panel NUTS-2, only ITC1 (Piedmont) depicted, right panel NUTS-3, only ITC11 (Turin) depicted

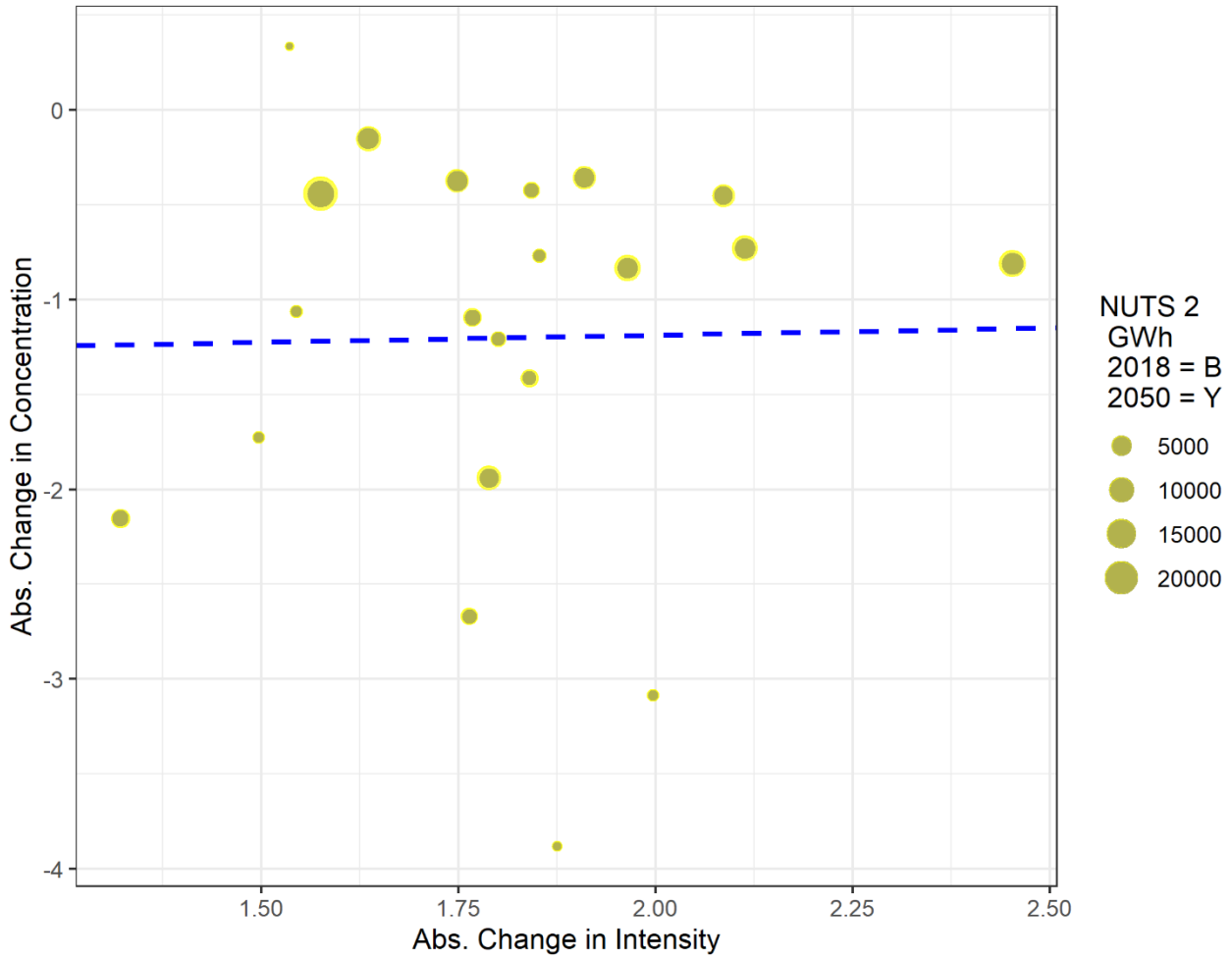

Figure 3 Regional changes in Concentration of residential electricity vs. the regional change in Electricity Intensity, 2050-2018, markers size weighted by residential electricity in GWh in 2050 SSP2 scenario (yellow) and in 2018 (blue)

Variable	Description	Source
Residential Area model and 2050 projections, $\widetilde{A}_{i,2050}$		
RA	Residential Area in square meters,	CORINE Land Cover, Copernicus Land Monitoring Service (© European Union, 2000-2018)
P	Population counts per grid	Gridded Population of the world (CIESIN, 2018)
RD	Road density (kilometres per square km)	Global Roads Inventory Project, GLOBIO (Meijer et al., 2018)
DC	Distances to the closest most populated cell in same LAU	Istat Administrative boundaries shapefile
DP	Distances to closest most populated cell in same NUTS-3	Istat Administrative boundaries shapefile
${P^S}_{i,2050}$	2050 projected population counts for different SSP scenarios	Gao (2020)
Electricity intensity model and 2050 projections, $\tilde{I}_{n,2050}$		
NUTS-3	NUTS-3 administrative boundaries	Istat Administrative boundaries
E	Provincial residential electricity consumption in GW/h	Terna
RA	Residential Area in square meters,	CORINE Land Cover, Copernicus Land Monitoring Service
P	NUTS-3 level population	Istat census data
G	GDP at the NUTS 3 level	Eurostat (nama_10r_3gdp)
C	Historical CDDs aggregated at the NUTS-3 level	Mistry (2019)
P_N^S 2050	Projected population at NUTS-3 level	Gao (2020)
G 2050	GDP SSP projections for 2050	Dellink et al., (2017)
C 2050	CDDs projections for 2050	Discussed in Appendix

Table 1 Description and sources of the variables employed in the two-step model and in the 2050 projections.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
A_{2000}	6,840	46,575	0	0	0	687,400
A_{2006}	23,776	84,782	0	0	0	687,400
A_{2012}	24,613	86,235	0	0	0	687,400
A_{2018}	24,651	86,329	0	0	0	687,400
P_{2000}	118.140	566.949	0	0.214	28.017	24,025
P_{2006}	120.501	570.790	0	0.211	28.665	23,511
P_{2012}	123.130	575.852	0	0.209	29.393	23,008
P_{2018}	126.051	582.317	0	0.206	30.175	22,516
RD	0.906	0.959	0	0.400	1.100	47
DP	0.291	0.153	0	0.177	0.383	2.072
DC	0.039	0.026	0	0.019	0.053	0.484
P_{2050}^{SSP1}	132.296	377.935	0.000	10.967	107.197	11,189.920
P_{2050}^{SSP2}	127.902	348.680	0.000	11.977	108.606	10,353.340
P_{2050}^{SSP3}	108.711	278.251	0.000	11.981	98.637	8,291.429
P_{2050}^{SSP4}	122.189	335.692	0.000	11.300	102.911	9,963.015
P_{2050}^{SSP5}	147.908	414.821	0.000	12.819	122.817	12,284.110

Table 2 Summary statistics for the residential area grid model's variables (residential area A, population counts P, road density RD, distance from the most populated grid cell in the same province DP, and distance from the most populated grid cell in the same LAU DC) and the different SSP projections for the 2050 population values. Areas are in square meters, population in counts, distances in degrees. Road density is in kilometres per square kilometre.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
I_{2000}	18.1	6.5	8.2	13.2	22.0	39.9
I_{2018}	5.4	1.7	2.4	4.0	6.2	10.7
D_{2000}	18,107	6,631	5,363	13,202	22,195	40,470
D_{2018}	4,990	1,642	1,973	3,687	5,885	9,805
G_{2000}	20,120	5,487	10,800	14,925	24,000	36,300
G_{2018}	26,456	7,680	15,000	19,200	31,975	55,900
C_{2000}	24.5	29.4	0.0	4.1	35.4	132.8
C_{2018}	82.7	70.4	0.0	31.2	116.0	366.8
P_{2050}^{SSP1}	1,434,449	1,810,166	154,468	794,458	2,074,439	2,714,430
P_{2050}^{SSP2}	1,353,695	1,685,853	161,617	757,656	1,949,735	2,545,774
P_{2050}^{SSP3}	1,112,712	1,361,194	150,202	631,457	1,593,967	2,075,222
P_{2050}^{SSP4}	1,295,069	1,617,357	151,425	723,247	1,866,891	2,438,713
P_{2050}^{SSP5}	1,595,607	2,001,934	180,025	887,816	2,303,397	3,011,187
G_{2050}^{SSP1}	49,342	16,870	22,699	35,929	57,527	112,447
G_{2050}^{SSP2}	44,226	14,548	20,529	31,136	51,676	95,889
G_{2050}^{SSP3}	38,399	12,518	17,972	25,871	46,272	78,109
G_{2050}^{SSP4}	48,277	15,851	22,464	34,504	56,933	103,741
G_{2050}^{SSP5}	56,269	19,306	25,817	40,357	65,437	131,335
$C_{2050}^{RCP\ 4.5}$	58.0	1.8	56.7	57.3	58.6	59.3

Table 3 Summary statistics for the NUTS-3 electricity intensity model's variables, electricity intensity (I - residential electricity demand in GWh per residential area), population density (D - population per residential area), GDP per capita (G - values in in current prices EUR) and CDD (C) for the years 2000 and 2018. The lower panel reports the 2050 projections values of the NUTS-3 population, GDP per capita and CDD according to the selected SSP and RCP scenarios.

	(1)	(2)	(3)	(4)
log P	1.129***	1.306***	1.306***	1.306***
	(0.002)	(0.002)	(0.001)	(0.002)
log DC	-2.221***	-1.396***	-1.327***	-0.1.322***
	(0.151)	(0.183)	(0.182)	(0.182)
log DP	1.236***	1.083***	-1.290***	1.284***
	(0.032)	(0.093)	(0.090)	(0.090)
log RD	-0.157***	-0.004	0.005	0.005
	(0.010)	(0.010)	(0.010)	(0.010)
year	0.059***	0.058***	0.058***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-1.722***	-2.165***	-2.269***	-2.282***
	(0.014)	(0.028)	(0.084)	(0.151)
Groups variance:				
ID	4.816	3.717	3.718	3.717
LAU	-	1.662	0.946	0.946
NUTS 3	-	-	0.6917	0.284
NUTS 2	-	-	-	0.378
Marginal R^2	0.405	0.459	0.458	0.460
Conditional R^2	0.757	0.794	0.793	0.792
Observations	1,907,344	1,907,344	1,907,344	1,907,344
Akaike Inf. Crit.	8,614,571	8,531,263	8,527,885	8,527,843
Notes:	***Significant at the 1 percent level.			
Std. Error in parenthesis	**Significant at the 5 percent level.			
	*Significant at the 10 percent level.			

Table 4 Estimation results for the model in equation (1). Column (1) reports a model with an individual grid random effect. Column (2) adds nested LAU random effects. Column (3) adds nested NUTS 3 random effects. Finally column (4) adds nested NUTS 2 random effects.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
A_{2018}	24,651.6	86,329.6	0	0	0	687,432
$\tilde{A}_{2050}^{SSP1}$	35,055.0	115,121.8	0.0	56.2	942.2	687,431.6
$\tilde{A}_{2050}^{SSP2}$	34,789.5	114,231.1	0.0	62.1	1,007.0	687,431.6
$\tilde{A}_{2050}^{SSP3}$	33,388.3	110,397.8	0.0	60.4	940.2	687,431.6
$\tilde{A}_{2050}^{SSP4}$	34,366.9	113,201.8	0.0	58.3	938.3	687,431.6
$\tilde{A}_{2050}^{SSP5}$	36,158.6	117,722.8	0.0	67.7	1,131.1	687,431.6

Table 5 Projected residential area summary statistics at the grid level for each SSP scenario, comparison with 2018 values. Areas are in square meters.

	(1)	(2)	(3)
log D	0.948***	0.959***	0.959***
	(0.006)	(0.007)	(0.007)
log G	-	0.111***	0.114***
		(0.033)	(0.033)
C	-	-	0.0001*
			(0.000)
year	0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.001)
Constant	-6.414***	-7.620***	-7.646***
	(0.060)	(0.361)	(0.360)
Groups variance:			
NUTS 3	0.004	0.004	0.004
NUTS 2	0.007	0.007	0.007
Marginal R^2	0.964	0.966	0.966
Conditional R^2	0.994	0.994	0.995
Observations	440	440	440
Akaike Inf. Crit.	-1,158.324	-1,167.586	-1,169.406
Notes:	***Significant at the 1 percent level.		
Std. Error	**Significant at the 5 percent level.		
In parenthesis	*Significant at the 10 percent level.		

Table 6 Estimation results for Equation (3). All models include nested random effects at the NUTS 2 and NUTS 3 level. Column (1) includes only the intercept and population density D. Column (2) adds GDP G. Finally column (3) adds CDD C.

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max	National Average AAGR
E_{2018}	0.137	0.530	0	0	0	7	-
$\tilde{E}_{2050}^{SSP1}$	0.260	0.917	0.000	0.000	0.007	8.721	2.803 %
$\tilde{E}_{2050}^{SSP2}$	0.257	0.906	0.000	0.000	0.007	8.699	2.733 %
$\tilde{E}_{2050}^{SSP3}$	0.244	0.866	0.000	0.000	0.006	7.945	2.426 %
$\tilde{E}_{2050}^{SSP4}$	0.254	0.897	0.000	0.000	0.006	8.689	2.654 %
$\tilde{E}_{2050}^{SSP5}$	0.271	0.944	0.000	0.000	0.008	8.776	3.048 %

Table 7 Summary statistics for the projected 2050 residential Electricity demand in GWh at the grid level, comparison with 2018 downscaled values using actual electricity intensity data. The last column shows the aggregate (national level) average annual growth rate (AAGR) from 2018 to 2050.

Variable	Gini Index	St. deviate	p-value	Moran I statistic	Expectation	Variance
E_{2018}	0.9362	789.34	< 2.2e-16	5.752e-01	-2.097e-06	5.310e-07
$\tilde{E}_{2050}^{SSP1}$	0.9235	797.44	< 2.2e-16	5.811e-01	-2.097e-06	5.311e-07
$\tilde{E}_{2050}^{SSP2}$	0.9231	792.52	< 2.2e-16	5.775e-01	-2.097e-06	5.311e-07
$\tilde{E}_{2050}^{SSP3}$	0.9239	786.89	< 2.2e-16	5.734e-01	-2.097e-06	5.311e-07
$\tilde{E}_{2050}^{SSP4}$	0.9237	793.32	< 2.2e-16	5.781e-01	-2.097e-06	5.311e-07
$\tilde{E}_{2050}^{SSP5}$	0.9219	796.11	< 2.2e-16	5.801e-01	-2.097e-06	5.311e-07

Table 8 Gini Index and Moran spatial autocorrelation test on downscaled residential electricity demand for 2018 values and for the 2050 projected values for each SSP

	Estimate	Std. Error	t-value	Pr(>|t|)
Lambda	0.234	0.015	15.971	< 2.2e-16 ***
log P	1.400	0.020	67.821	< 2e-16 ***
log DC	-42.028	4.181	-10.051	< 2e-16 ***
log DP	1.617	1.233	1.311	0.189
log RD	-0.051	0.097	-0.527	0.598
Year	0.218	0.005	43.080	< 2e-16 ***
Observations	10428			
R^2	0.6643			

Table B.1 Estimation results for a model with linear spatial lag with fixed spatial effects and fixed time effects run on a subset of the data (sample limited to the NUTS-3 ITC4C).

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	-1.191	0.162	-7.345	0***
log P	1.320	0.020	450.055	0***
log DC	-9.863	0.401	-24.586	0***
log DP	0.956	0.125	7.628	0***
log RD	-0.042	0.009	-4.722	0***
Group St. Dev.				
LAU	0.638			
NUTS-3	0.643			
NUTS-2	0.803			
Marginal R^2	0.470			
Conditional R^2	0.557			
Observations	476836			
Akaike Inf. Crit.	2123284			

Table B.2 Results for a linear mixed model with a spatial spherical correlation structure in the errors. The sample is reduced to 2018 values.