Downscaling Aggregate Urban Metabolism Accounts to Local Districts

Isabel M. Horta and James Keirstead

Summary

Urban metabolism accounts of total annual energy, water, and other resource flows are increasingly available for a variety of world cities. For local decision makers, however, it may be important to understand the variations of resource consumption within the city. Given the difficulty of gathering suburban resource consumption data for many cities, this article investigates the potential of statistical downscaling methods to estimate local resource consumption using socioeconomic or other data sources. We evaluate six classes of downscaling methods: ratio-based normalization; linear regression (both internally and externally calibrated); linear regression with spatial autocorrelation; multilevel linear regression; and a basic Bayesian analysis. The methods were applied to domestic energy consumption in London, UK, and our results show that it is possible to downscale aggregate resource consumption to smaller geographies with an average absolute prediction error of around 20%; however, performance varies widely by method, geography size, and fuel type. We also show how mapping these results can quickly identify districts with noteworthy resource consumption profiles. Further work should explore the design of local data collection strategies to enhance these methods and apply the techniques to other urban resources such as water or waste.

Introduction

Over 50% of the world’s population currently lives in urban areas, a figure that is expected to rise to 66% by 2050 driven largely by growth in less-developed regions (UN 2014). The popularity of urban living has many explanations, including the provision of physical security, opportunities for commerce and employment, and cultural attractions (Kotkin 2005; Glaeser 2008). But regardless of an individual’s motivation for living in a city, it is clear that they would not stay long if the city was unable to satisfy basic resource requirements such as those for water, energy, food, and materials. A city must also expel its wastes effectively in order to avoid inhibiting urban activities. These resource flows—“the materials and commodities needed to sustain the city’s inhabitants at home, at work and at play”—are known as a city’s urban metabolism (UM) (Wolman 1965, 179) and they are of significant global and local importance. For example, it is estimated that 75% of global final energy use can be attributed to cities (Grubler et al. 2012) and cities account for 76% of global energy-related carbon dioxide (CO₂) emissions (Dhakal et al. 2008).

There have been many studies of UM representing cities from around the world (see Kennedy et al. [2007] for a review), and the metabolic data presented therein are typically reported at the scale of a whole city on an annual basis. For many purposes, this coarse resolution is sufficient. Cities can learn which resource flows are the most significant by mass or energy content, how resource consumption links to urban growth processes, how the city is connected with its hinterland, and compare the resource efficiency of a city with its rivals. However, in order to plan specific design interventions or assess the impact of new policies, local authorities need more detailed resource
consumption information, both for its own sake (i.e., to plan the operation of related infrastructures) and to establish links between socioeconomic activities and resource consumption. For example, identifying areas of fuel poverty—the inability of households to afford adequate warmth—requires a combination of data about the built environment, sociodemographics, and energy consumption (Fahmy et al. 2011; DECC 2013). Similarly, industrial eco-parks require an awareness of the detailed spatial geography of resource consumption and waste production in order to achieve overall efficiency gains (Jacobsen 2008; Chertow 2000).

In this article, we investigate approaches for downscaling aggregate UM data to smaller spatial scales within a city. Using the case of domestic energy consumption in London, we seek to answer three related questions: (1) How can socioeconomic data be used to estimate resource consumption within the city?; (2) Which statistical methods provide the most accurate estimates?; and (3) What is the significance of these findings for urban metabolism research and policy? First, we review previous attempts to estimate urban resource consumption at fine spatial scales with a particular focus on energy consumption. Then, we discuss the statistical methods and data sources for our analysis before presenting the results. The last section concludes by noting that although the methods presented here achieve a level of accuracy suitable for basic insights, additional data collection and more detailed assessments are likely to be necessary before specific investments or decisions are made.

Background

Downscaling data from one geography to another is a common problem in the scientific literature: for example (Van Vuuren et al. 2007), downscale environmental change scenarios of population, economic growth, and greenhouse gas (GHG) emissions from global regions to countries and a smaller 0.5° × 0.5° grid. They draw a methodological distinction between conditional modeling—where data for one geography constrains a model of the other geography, common in climate-change modeling (Wilby and Wigley 1997)—and “clearly defined algorithms,” which we have termed statistical downscaling methods in this article. For UM, unlike climate-change modeling, there is no equivalent of a global circulation model for use within a conditional modeling framework; we therefore focus on statistical downscaling methods.

Energy Flows in Urban Metabolism Studies

The study of UM arguably began with Wolman’s analysis of a hypothetical American city (Wolman 1965). Although this first study did not focus on the use of energy within cities, the literature has grown since then to include dozens of studies in which urban energy consumption plays an important role. Here, we consider the ways in which energy flows have been characterized within UM studies in order to understand the strengths and weaknesses of current practice.

Kennedy and colleagues (2011) review the UM literature and note the presence of two broad traditions. The first is based on H. T. Odum’s notion of emergy accounting. Odum defines emergy as “the energy of one type required in transformations to generate a flow or storage” and, by convention, solar energy is used as the reference energy type (Odum 1988, 1135). The concept has been used to understand material-use cycles including urban metabolic flows (Tilley 2004; Odum 1983). However, emergy accounting is controversial, particularly within the energy engineering community who prefer the use of exergy, which is based on the second law of thermodynamics rather than difficult-to-calculate solar energy “transformities” (Sciubba and Ulgiati 2005). The other major tradition in the UM literature is mass-based accounting, for example, tonnes of water imported or waste disposed. Because these studies are often concerned with the flows of specific materials (e.g., steel or cement), the use of mass units is appropriate; however it would be problematic if energy were to be accounted in a similar framework. Of course, analysts could use tonnes of oil (or coal) equivalent, as in many national energy accounting frameworks, but, unlike a national economy where such imported flows might be physically measured at a port or other border point, cities are open systems connected into national grids and mass-less electricity is a significant flow. Consequently, most UM studies now report both mass and energy flows; Kennedy and colleagues (2011, 1967) suggest that this is largely a practical choice, in order to use units that “local government officers would use, recognize and understand.”

Whereas UM may have begun as a theoretical thought experiment, it now has practical application in several areas: the preparation of urban sustainability indicators and associated reports, the calculation of urban GHG emissions inventories, policy analysis, and to inform sustainable urban design (Kennedy et al. 2011). The accurate accounting of urban energy flows is central to all of these applications. For example, sustainability strategies for major cities like New York and London emphasize the role of energy (The City of New York 2011; Mayor of London 2011). In some cases, aggregate statistics are sufficient for such analyses: Weisz and Steinberger (2010), for example, illustrate how energy consumption might vary by income groups within cities by drawing on national data sets. Alternatively, these energy flows are often reported by end-use sector, such as buildings or transport, and such data can be particularly helpful for emissions inventory studies (e.g., Sovacool and Brown 2010). However, the variability of energy consumption in space and time is rarely examined. 1 Kennedy and colleagues (2007) suggest that such analyses could be valuable, for example, to understand the detail of resource storage processes within cities.

We can therefore conclude that the aggregate accounting of urban energy flows can be valuable for some applications of UM, such as the calculation of GHG inventories. However, when trying to plan detailed policy responses, or how such flows are shaped by processes within the city, other data sources may be beneficial.
Modeling Intraurban Energy Consumption

If one widens the definition of UM studies to include all studies that examine urban resource demands, regardless of whether they use the "metabolism" label, then a range of additional sources can be found. Here, we review recent studies that have attempted to model energy consumption within cities at a disaggregate scale (e.g., per building or for whole districts).

A significant part of urban energy consumption can be attributed to energy use in buildings; in London, it is estimated that this fraction is 75% (GLA 2011). Heiple and Sailor (2008) present an approach that simulates the demand for electricity and natural gas for prototypical buildings in an urban area, providing estimated demand data at up to hourly temporal resolutions and individual parcel spatial scales. The intention is to use these results to support analyses of urban heat island effects and air quality. A similar analysis has also been applied to New York (Howard et al. 2012). In this case, energy demand was assumed to be related to a building's function (as opposed to age or construction type), and the results are given as annual average energy intensities (e.g., kilowatt-hours per square meter) at tax lot scale; the intention is to inform energy efficiency and retrofit programs. In the UK, Choudhary (2012) focuses on the nondomestic building stock in London and uses Bayesian statistical techniques to understand the variability of energy demand in different end-use sectors (e.g., retail, hotel, and so on) and individual local authorities within Greater London.

These approaches build on knowledge of existing buildings and their likely energy consumption. In contrast, Minx and colleagues (2009) use consumption data and input-output analyses to understand the variability of carbon emissions across London (a similar approach could be used to extract energy consumption, both direct and indirect). Another method is to simulate the activities of individuals as they move about the city and to couple this to building energy models to capture energy demands at high spatial and temporal resolutions (Keirstead and Sivakumar 2012). Finally, Parshall and colleagues (2010) use a high-spatial-resolution database of GHGs and work back to the energy consumption of urban areas. This approach enables them to capture both building energy demand, but also on-road transport energy consumption.

All of these techniques require extensive data inputs, on the built environment, measured energy consumption, activity patterns, or other sources. A notable exception is Taylor and colleagues (2014), who use statistical downscaling to take official government statistics on energy consumption and combine it with census data to increase the spatial resolution to a 1-square-kilometer grid. The results were used to examine future bioenergy scenarios for the UK, which rely on high-spatial-resolution demand data.

Our goal in this article is to expand upon this latter approach. Taylor and colleagues (2014) do not specifically consider alternative methods for statistical downscaling in their article. But an evaluation of these methods is important in order to understand their data requirements, accuracy, and hence suitability to different applications. In particular, we are interested in the ability of statistical downscaling methods to be applied in data-poor cities, which may lack the detailed bottom-up data sources used in many of the studies cited in this review; this might be because the data have not been collected or it may have been collected as part of commercial operations and not be made publicly available. An effective means of converting aggregate UM accounts of energy consumption to spatially resolved maps of likely demand would therefore enable these locations to conduct a range of analyses, including assessments of urban heat island effects, to plan local energy efficiency programs, evaluate the sustainability of local energy supplies, assess alternative energy supply strategies, and more.

Methods and Data

Overview

This section presents a number of methods for the statistical downscaling of UM data on energy consumption. Statistical downscaling is the process whereby a statistical relationship is used to relate observed data at a larger geography to other variables at a smaller geography. For example, in our case, we know energy consumption for the city as a whole, but we want to know how much energy has been consumed in a specific district within that city, perhaps as a function of the population of the district.

The following analyses test these methods in London. London has been chosen because data on energy consumption and other socioeconomic variables are available at multiple geographies, thereby allowing us to validate our models against observed data as an indication of how the methods might perform in general. Our data set contains four nested geographies: (1) London, a single geographic region describing the Greater London Authority (GLA) as a whole, which contains; (2) 33 local authority districts (LADs). Each LAD can be divided into a number of (3) middle super output areas (MSOAs). London contains 983 MSOAs, each with a population between 5,000 and 15,000. Finally, MSOAs are further divided into (4) lower super output areas (LSOAs). There are 4,835 in London, with a population between 1,000 and 5,000. A map showing these geographies can be seen in figure 1.

Energy consumption data for all geographies examined here are compiled by the Department of Energy and Climate Change (DECC) and have been downloaded using the decctools R software package (Keirstead 2014). For the regression analysis, additional socioeconomic statistics have been assembled from the Office for National Statistic's Neighbourhood Statistics website and the Annual Survey of Hours and Earnings. Additional information on the methods and data used here can be found in Keirstead and Horta (2015).

In order to validate our methods, we require common energy consumption statistics at all levels; this restricts our analysis to domestic electricity and gas consumption. Each downscaling method is evaluated by comparing the true value of energy consumption for fuel f in geography g, \( \hat{E}_{f,g} \), with the estimated value calculated by the downscaling method in question, \( \hat{E}_{f,g} \). The two
measures we use are the MAPE (mean absolute percentage error) and RMSPE (root mean square percentage error). These are defined as follows:

\[
PE_{i,g} = \frac{E_{i,g} - \hat{E}_{i,g}}{E_{i,g}} \times 100
\]

\[
\text{MAPE}_f = \frac{1}{n} \sum_{g=1}^{n} |PE_{i,g}|
\]

\[
\text{RMSPE}_f = \sqrt{\frac{1}{n} \sum_{g=1}^{n} \left(PE_{i,g}\right)^2}
\]

Where appropriate, repeated random cross-validation is used to evaluate the impact of sample selection. In these cases, the error rates reported are the average values of all simulations.

**Ratio-based Normalization**

In the first set of analyses, we estimate the energy consumption at a lower level as a constant ratio, that is, normalized by a common predictor at the upper level:

\[
\hat{E}_{i,1} = \frac{E_{i,u}}{x}
\]

where \(\hat{E}_{i,1}\) is the estimated energy consumption of fuel \(f\) at the lower level \(1\) (\(u = \text{upper level}\)) and \(x\) is the metric of interest, available at both upper and lower levels. The upper level here is taken to be London as a whole and all the three lower geographies are tested (LAD, MSOA, and LSOA). Four methods of normalization were compared: per unit area; per capita; per household; and per unit wealth (the product of a zone’s median income and number of households).

**Linear Regression (External Calibration)**

The next model is a linear regression that predicts energy consumption as a function of multiple variables. Mathematically, it can be stated as:

\[
E_{i,u} = \beta X_u + \varepsilon
\]

where \(X_u\) is a vector of independent variables at the upper level, \(E_{i,u}\) is the known energy consumption at the upper level, \(\varepsilon\) is a error term assumed to be normally distributed, \(\varepsilon \sim N(0, \sigma^2)\), and \(\beta\) is a vector of model coefficients determined by a least-squares fit. \(X_u\) might also include a constant intercept term.

The lower-level energy consumption is then predicted by:

\[
\hat{E}_{i,1} = \beta X_i
\]

where \(X_i\) are the same independent variables, but measured at the lower level.

The normalization mode presented above can be thought of as a linear regression model, but with only one independent variable and no intercept. As a consequence, the model can be “fitted” using only the single geography of London as a whole and ignoring the error term. However, if we want to introduce multiple explanatory variables, then we need multiple data points to calculate the model coefficients. Clearly, there is only one London region, so we must look for alternative data sources. For the “external calibration” case, we consider the case where data are available for a set of similar cities outside the target city. (A second “internal calibration” case where the model is fit to a selection of the current city’s lower geographies is presented in the next section.) Here, we take all of the non-London local authority districts in England and Wales, with the implicit assumption that the energy consumption of subgeographies within London follows the same pattern. After removing missing values from the data set, there are 309 complete data points.

Before performing the analysis, one must first select the dependent variables \(E_{i,u}\). Because our aim is to use the model to predict energy consumption from another data set, it is important that the dependent variable is on a similar scale in both the training and validation data sets. By inspecting box plots of
both total and per capita energy consumption, it was decided to model per capita energy consumption because these values are of a similar magnitude at all geographies.

The next step is to identify the potential independent variables. Based on the review of residential energy literature (Kelly et al. 2013; Kavousian et al. 2013; Min et al. 2010; Boulaire et al. 2014), we identified a number of socioeconomic variables to predict energy consumption. The candidate variables were sought from Office of National Statistics sources and had to be available at all of the geographies studied. The final set of variables includes:

- Age: the percentage of children (< = 16) and elderly people (> = 74) with the hypothesis that these age groups will be home more often and therefore will have greater energy service demands.
- Tenure: the proportion of owner-occupiers
- Educational attainment: the proportion of the population with no qualifications
- Employment status: the percentage of the population in full-time employment
- Heating type: the proportion of households with gas heating
- Dwelling type: the proportion in detached homes
- Other: including median household income, population, household size, and area

All of the proportional variables (i.e., those measured on a 0 to 1 scale) are transformed with the logit transformation so that the values span from $-\infty$ to $\infty$. Owing to declining marginal utility, income is log-transformed. Population is also transformed to be measured in thousands of people to make it a similar order of magnitude.

With ten predictors, there is a risk of multicollinearity affecting the regression results. We conducted a variance inflation factor (VIF) analysis to test for collinearity. For a given independent variable, a VIF value greater than 10 suggests that multicollinearity may complicate the model estimation and interpretation; however, the VIF values obtained for all variables were below 4 here. Thus, multicollinearity is not an issue with the present data set and the regressions were performed using all of the input variables. Separate regressions were conducted for each fuel and the results are shown in table 1.

### Linear Regression (Internal Calibration)

In some locations, it may be that a suitable set of alternative cities does not exist for fitting the regression model. In this case, the alternative approach is to randomly sample some of the districts within the city and fit the model to these observations. This scenario might arise if a local authority decides to fund a small data collection exercise in order to improve its understanding of local energy use.

The model formulation used here is the same as in the previous selection. However, the difference is that rather than having "upper" and "lower" data sets, we essentially divide the lower geography into two data sets: one for training the regression model, and the rest for validation (the "testing" data set). A single analysis therefore has limited meaning because the precise numerical results will depend on which districts were assigned to which group. This problem can be avoided using cross-validation. Here, we use repeated random cross-validation, which means that we select a random group of observations to use for training with the remaining observations used for testing. We then repeat the sampling, model fitting, and validation process ten times and report the average of these trials.

There is also a question of how many districts to include in the training sample. With traditional k-fold cross-validation, the analyst might use as much as 90% of the data set for training with only 10% for testing and validation. This would be impractical for a local authority seeking to implement this method because such a large training set implies a substantial data gathering exercise. We therefore evaluated a range of partition sizes to see how the model performance changes; the results reported here use a 40% training fraction.

### Table 1 Results of linear regression fit to data from non-London LADs

| Dependent variable | Electricity | Gas |
|--------------------|-------------|-----|
| Population         | -0.054      | 0.497* | 584.093*** 753.948* |
| Age                | -525.420*** | -1,327.810*** |
| Household size     | -113.173*** | 134.455 |
| Area               | -0.0001     | 0.0004 |
| Employment status  | -64.998     | -796.267*** |
| Household income   | 372.738***  | 2,177.010*** |
| Gas heating        | -225.127*** | 1,498.188*** |
| Tenure             | 141.310***  | 327.687** |
| Dwelling type      | 40.490**    | 75.113 |
| Constant           | 8.492       | -14,638.180*** |
| Observations       | 309         | 309 |
| Adjusted $r^2$     | 0.808       | 0.800 |

Note: See text for description of variables.
*p < .1; **p < .05; ***p < .01.
LADs = local authority districts.
Murakami (2014) downscale the aging ratio of the population to the municipalities of the North Kanto region in Japan based on a spatial lag model.

We applied spatial regression to downscale energy consumption, again assuming that some local data are available with which to calibrate the model. However, before applying a spatial regression model, we checked which model form is the most appropriate. We used the Lagrange Multiplier test to select between the use of a spatial error model or a spatial lag model. The spatial lag model captures the spatial dependence through the introduction of a spatial term in the model formulation. The spatial error model captures the spatial dependence through the error term of the model. The p values obtained for the spatial lag models were statistically significant and lower than the p values obtained for the spatial error models in all geographies analyzed. Thus, it is possible to conclude that the spatial lag model is the most suitable spatial regression to be used. The formulation of this model is:

$$E_{h,i} = \rho WE_{h,i} + \beta X_i + \varepsilon$$

where $\rho$ is the spatial autoregressive parameter and $W$ is the spatial weights matrix.

We use two different procedures for the construction of the matrix $W$. One is the sphere of influence (SOI), a graph-based procedure to identify neighbors by joining relatively close centroids (i.e., geometric center of the regions). Centroids are SOI neighbors if circles centered on the centroids, of radius equal to the centroid’s nearest neighbor distances, intersect in two places. The other approach is the k-nearest neighbors method that uses the k-closest centroids as neighbors (we used $k = 6$). The $\varepsilon$ is the error term with a $N(0, \sigma^2 I)$ distribution, where $I$ represents the identity matrix.

**Multilevel Regression Model**

In the previous analyses, it was assumed that all of the variables were available at both the upper and lower geographies (or within the “training” and “testing” data sets). However, for some cities, it may be that only certain variables are available at the lower geography; in this case, the regression approach used above is not applicable. Similarly, if we rely only on common variables, for example, as in the basic ratio-scaling technique, then we may be discarding significant additional information.

An in-between solution is multilevel or hierarchical regression modeling, whereby a wider set of regression variables is used at the upper level to shape the coefficients of a lower-level regression. Here, we will assume that only the variables household size and income are known at the lower level, but that all of the other variables from the previous linear regression will be kept as group-level predictors. The model formulation is given by Gelman and Hill (2007, 280, equation 13.2), and it allows for group-level predictors of both intercept and slope.

In the current analysis, we are trying to estimate the energy consumption of three nested geographies within London: LADs, which contain MSOAs, which contain LSOAs. Because there is no higher geography (e.g., “cities”), it is not possible to use the LAD geography as a lower level within the model formulation. We will therefore assume that the LAD data are available for use as upper-level geographies and will calculate the MSOAs and LSOAs as separate lower levels within those boundaries (i.e., we do not look at LSOAs nested within MSOAs nested within LADs in a single model). The data are otherwise prepared as in the previous regression analysis.

**Bayesian Analysis**

The final approach is to use Bayesian statistics as an efficient way of incorporating partial or uncertain data into an estimated quantity. In our case, a practical application of this method might be to use a series of small local metabolism surveys to refine our estimate of local energy consumption. The previous models presented above could be written with Bayesian formulations, but this section focuses on a simple two-parameter model. This analysis is performed using the STAN modeling language (Stan Development Team 2014a, 2014b).

The basic model is that the energy consumption (per capita) for some geography, $E_h^c$, is a random variable with a log-normal probability density function with unknown $\mu$ and $\sigma$ parameters. Further, we assume that, through data collection, we can measure some of the $E_h$ values, but not all $g \in G$. But we can refer to similar data from related geographies to inform our belief about the likely distribution of these values.

We can therefore express this relationship using Bayes’ rule. In the formulation below, $\theta$ stands for the unknown $\mu$ and $\sigma$ parameters in the log-normal distribution of $E_h$ and $y$ represents the (partial) measured data that we have on $E_h^c$.

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}$$

The prior distribution $p(\theta = \mu)$ is taken to be log-normal. It can either be completely uninformative (which given the log-normal constraints, STAN will model as a uniform distribution) or we can estimate its likely distribution from other national data sets. In the analysis that follows, this is the per capita energy consumption of all urban local authorities in the UK excluding London. We also need a prior distribution for the $\sigma$ parameter. The uninformative prior is taken to be STAN’s default uniform distribution, but restricted to the positive domain only, and the informative one will be an Inverse-$\chi^2$ distribution with $v = 1/\sigma_{E_h}^2$ degrees of freedom, where $\sigma_{E_h}$ is the standard deviation of the energy consumption in the other geographies ($h \in H$).

To begin the analysis, we examined the distribution of per capita energy consumption for all urban local authority districts in the UK. Data are retrieved from the decctools package, and the MASS package’s fitdistr function is used to fit an appropriate distribution (Venables and Ripley 2002). A log-normal shape was chosen by trial and error, and the calculated parameters are: $\mu_{E_h} = (0.54, 1.72)$, $\sigma_{\mu_{E_h}} = (0.087, 0.118)$, $\sigma_{E_h} = (0.0069, 0.0093)$ for electricity and gas, respectively. This distribution can then be used as a prior for the actual per capita energy consumption in London as described above. To explore the effect of improved knowledge, we use STAN to simulate how our estimate of the underlying log-normal
### Table 2  Summary of all model errors

| Model                      | Electricity |          | Gas          |          |
|----------------------------|-------------|----------|--------------|----------|
|                            | MAPE        | RMSPE    | Note         | MAPE     | RMSPE    |
| Ratio-based normalization  | LAD         | 38.1     | 50.4         | LAD      | 37.8     | 49.7     | Area      |
|                            | LAD         | 8.6      | 10.6         | LAD      | 17.3     | 23.9     | Household |
|                            | LAD         | 10.4     | 15.1         | LAD      | 13.1     | 19.3     | Population|
|                            | LAD         | 34.9     | 50.5         | LAD      | 54.2     | 96.4     | Wealth    |
|                            | MSOA        | 56.7     | 111.2        | MSOA     | 56.9     | 124.2    | Area      |
|                            | MSOA        | 14.0     | 24.0         | MSOA     | 24.7     | 35.4     | Household |
|                            | MSOA        | 16.7     | 25.0         | MSOA     | 23.1     | 32.2     | Population|
|                            | MSOA        | 27.3     | 44.4         | MSOA     | 32.4     | 53.2     | Wealth    |
|                            | LSOA        | 71.4     | 487.7        | LSOA     | 164.5    | 172.6    | Area      |
|                            | LSOA        | 20.1     | 193.3        | LSOA     | 33.2     | 90.6     | Household |
|                            | LSOA        | 22.3     | 180.4        | LSOA     | 29.8     | 75.0     | Population|
|                            | LSOA        | 31.2     | 192.3        | LSOA     | 38.4     | 135.7    | Wealth    |
| Linear regression (external)| LAD         | 4.3      | 6.6          | LAD      | 12.5     | 16.1     |          |
|                            | MSOA        | 15.1     | 24.4         | MSOA     | 23.3     | 27.4     |          |
|                            | LSOA        | 17.8     | 196.4        | LSOA     | 26.6     | 37.2     |          |
| Linear regression (internal)| LAD         | 8.1      | 11.4         | LAD      | 11.7     | 17.4     |          |
|                            | MSOA        | 9.2      | 20.1         | MSOA     | 7.8      | 11.0     |          |
|                            | LSOA        | 14.7     | 179.1        | LSOA     | 11.4     | 22.4     |          |
| Spatial regression         | LAD         | 11.1     | 14.1         | LAD      | 13.5     | 17.3     | Sphere of influence |
|                            | LAD         | 19.9     | 24.0         | LAD      | 12.2     | 16.1     | k-nearest neighbors |
|                            | MSOA        | 9.9      | 22.7         | MSOA     | 7.9      | 11.0     | Sphere of influence |
|                            | MSOA        | 10.1     | 22.9         | MSOA     | 8.0      | 11.0     | k-nearest neighbors |
|                            | LSOA        | 14.6     | 155.3        | LSOA     | 11.4     | 20.7     | Sphere of influence |
|                            | LSOA        | 15.0     | 151.7        | LSOA     | 11.4     | 20.8     | k-nearest neighbors |
| Multilevel regression      | MSOA        | 17.2     | 25.4         | MSOA     | 18.8     | 27.1     |          |
|                            | LSOA        | 18.6     | 113.0        | LSOA     | 20.7     | 39.9     |          |
| Bayesian analysis          | LAD         | 12.5     | 16.3         | LAD      | 17.1     | 23.6     | Informative prior |
|                            | LAD         | 12.5     | 16.9         | LAD      | 14.4     | 19.4     | Uninformative prior |
|                            | MSOA        | 18.4     | 27.7         | MSOA     | 24.1     | 34.0     | Informative prior |
|                            | LSOA        | 22.5     | 153.2        | LSOA     | 30.9     | 74.6     | Informative prior |
|                            | LSOA        | 21.5     | 146.4        | LSOA     | 29.8     | 75.8     | Uninformative prior |

Note: k-nearest neighbors uses the six closest centroids as neighbors.
MAPE = mean absolute percentage error; RMSPE = root mean square percentage error; LAD = local authority district; MSOA = middle super output area; LSOA = lower super output area.

distribution of per capita domestic energy consumption changes as we gather more information about the districts within London. For contrast, we compare the results of using the informative national per capita energy distribution with STAN’s default uninformative uniform prior.

## Discussion

The results obtained by applying these statistical downscaling methods are, of course, only applicable to the case of domestic energy consumption in London. Additionally, there are other techniques that have not been considered here, as well as alternative formulations of the examples given. Nevertheless, we believe that sufficient results have been gathered to allow for an initial evaluation of these methods and their use in urban metabolism research. This section therefore considers three specific topics: an overall comparison of the methods demonstrated here; an evaluation of the spatial distribution of the prediction errors; and an assessment of the suitability of these methods for different kinds of cities.

### Summary of Results

We predicted the energy consumption for three different geographies within London using six statistical downscaling methods as described above. In each case, the accuracy of the technique was assessed by calculating the MAPE and RMSPE of the predicted value versus the known actual value. Table 2 gives the detailed results of each model by fuel type.

Ideally, we would like to be able to accurately downscale urban metabolic flow data using the minimum number of covariates. Both the ratio measures and Bayesian analysis methods meet this latter criterion, but their performance is mixed. Regarding the MAPE, the ratio estimates based on population
and household numbers were good predictors (MAPE = 19.4 [8.6–33.2]% vs. MAPE = 45.3 [27.3–71.4]% for wealth- and area-based normalizations); this is not unexpected given Bettencourt’s empirical and theoretical work on household urban resource consumption scaling linearly with population (Bettencourt et al. 2007; Bettencourt 2013). The Bayesian analysis also gave a very similar result (MAPE = 20.2 [12.5–30.9]%). However, when measuring the RMSPE, there are substantial large errors with these methods (RMSPE = 60.4 [10.6–193.3]% for household and population ratio measures vs. RMSPE = 130.7 [44.4–487.7]% for wealth- and area-based normalizations); the Bayesian analysis has slightly lower errors (RMSPE = 53.6 [16.3–153.2]%).

Multivariate models were observed to offer better predictions. The most accurate method is linear regression calibrated on data from within the city with a MAPE of 10.5 (7.8–14.7)%. Adding the spatial lag term does not appear to make a substantial difference to the MAPE (12.1 [7.9–19.9]%) or the RMSPE (40.6 [11–155.3] vs. 43.6 [11–179.1] for the nonspatial version). The linear regression calibrated on analogous cities outside of London also performed well (MAPE = 16.6 [4.3–26.6]%, RMSPE = 51.4 [6.6–196.4%]). The multilevel formulation had similar, and in some cases better, performance, particularly at smaller geographies (MAPE = 18.8 [17.2–20.7]%, RMSPE = 51.3 [25.4–113]%).

Geography has a clear impact on the results (table 3). As expected, the models perform better when predicting the consumption values of larger geographies (i.e., where the exceptional demands of single areas are more likely to be smoothed out). This is particularly true for RMSPE, which shows some wildly inaccurate predictions at very small geographies.

Fuel type is also important. Table 4 shows that, whereas MAPE values are similar between both electricity and gas, the RMSPE error is much larger in the case of electricity. One possible explanation for this is that gas consumption is determined largely by the built fabric, whereas electricity consumption is shaped more by the behavior of (more heterogeneous) occupants.

Overall, these results indicate that it is possible to use statistical downscaling methods to estimate UM variables such as energy consumption at high spatial resolution with reasonable accuracy (~20%). Accuracy is determined primarily by the variable being predicted and the granularity of the geography for which the prediction is being made, and prediction errors are remarkably consistent across methods. This suggests that, for many applications, an analyst may be able to use simpler methods, such as traditional linear regression, without resorting to the more complex Bayesian or spatial regression methods. However, if costly decisions are to be based on these results, for example, investing in a major redevelopment of an urban district, then the analyst should choose a method with lower prediction errors or complement the downscaling analysis with ground-truth surveys or other methods.

### Spatial Distribution of Model Errors

The statistical methods used here all rely on the assumption that a given district within a city has some degree of similarity to other districts. For example, in the regression model, the analysis assumes that the model coefficients calibrated on a training data set are still valid when used to make out-of-sample predictions of energy consumption. One way to validate these assumptions is to examine the range of errors, and from table 2, a number of outliers can be seen. On their own, such figures do not enable us to say whether these exceptional results occur randomly throughout the city or are concentrated in specific areas.

To explore this question, figure 2 shows the spatial distribution of absolute percentage errors for predictions of electricity and gas consumption within MSOA geographies using three example methods. The colors have been chosen so that absolute errors are limited at 100%; this enables variations within “average” districts to be seen. For example, in the per capita normalization, we can see clear income effects with overestimates of energy consumption in lower income areas of East London and underestimates in the more-affluent suburbs. The other two methods do not show these general patterns as clearly, although they are still visible in areas like Kensington and Chelsea, but outliers are easily visible, such as Canary Wharf and Millwall (predominantly commercial areas with very few households).

These results suggest that a good understanding of local geography is required to interpret the results of any model. Without it, one may be unable to distinguish between poor model performance and unique local districts that cannot, and arguably should not, be shoehorned into a modeling framework for

---

**Table 3** Summary of errors by error type and downscaling geography

| Error  | Geography | Mean (range) |
|--------|-----------|--------------|
| MAPE   | LAD       | 18.2 (4.3–54.2) |
|        | MSOA      | 20.9 (7.8–56.9) |
|        | LSOA      | 26.3 (11.4–71.4) |
| RMSPE  | LAD       | 23.7 (6.6–96.4) |
|        | MSOA      | 35 (11.0–124.2) |
|        | LSOA      | 132.5 (20.7–487.7) |

*Note: MAPE = mean absolute percentage error; RMSPE = root mean square percentage error; LAD = local authority district; MSOA = middle super output area; LSOA = lower super output area.*

**Table 4** Summary of errors by error type and fuel

| Error  | Fuel     | Mean (range) |
|--------|----------|--------------|
| MAPE   | Electricity | 20 (4.3–71.4) |
| MAPE   | Gas      | 23.8 (7.8–64.5) |
| RMSPE  | Electricity | 85.5 (6.6–487.7) |
| RMSPE  | Gas      | 45.7 (11.0–172.6) |

*Note: MAPE = mean absolute percentage error; RMSPE = root mean square percentage error.*
RESEARCH AND ANALYSIS

Figure 2  Spatial distribution of absolute percent errors for three downscaling techniques at MSOA resolution, based on (a) per capita normalization, (b) internally calibrated linear regression, and (c) linear regression with spatial autocorrelation (sphere of influence). For the internal linear regression and spatial regression methods, the maps represent one of the ten random cross-validation samples and a training fraction of 40% was used; these zones are shown in gray. MSOA = middle super output area.

general predictive purposes. This, in turn, affects that application of a model’s predictions to guide local decision making.

Applying the Methods in Different Cities

Much of the discussion has, so far, focused on the technical performance of each analysis method. But for urban authorities seeking to understand their urban metabolism at finer scales, such arguments boil down to a simple question: What is the right method for the problems facing my city?

The choice of an appropriate downscaling method for a specific problem will depend on a number of issues. Data availability is arguably the most important. Some methods require no local data at all. For example, if national-level data or other proxies are available, then Bayesian analysis can at least provide an estimate of the likely distribution of energy consumption values and, as new local data become available, these estimates can be updated to become increasingly accurate. Ratio-based normalization methods require some data, but nothing too onerous; for example, we demonstrated that by simply knowing the population of a local area, local domestic energy consumption can be estimated with 10% to 20% accuracy. For many purposes, this may be sufficient, but if additional data are available
locally, for example, socioeconomic data from a census, then these crude estimates can be refined using multivariate regression models. However, regardless of the present data availability, we recommend that cities compile urban metabolic flows at suburban spatial scales. Such data are invaluable to validating these methods, and even partial surveys can be used to improve the accuracy of predicted consumption in other parts of the city. Consequently, we also recommend that, where resource consumption data are proprietary (e.g., held by commercial operators), efforts should be made to release this information publicly.

Another major consideration is the goal of the analysis, namely, whether the aim is parameter inference or predictive inference. In other words, do we want to estimate the effect of a variable on resource energy consumption in the city generally (parameter inference), or do we wish to predict the energy consumption in an unmeasured district (or estimate how it might change in light of socioeconomic trends)? All of the methods discussed here provide insight for parameter inference, but if the goal is to predict energy consumption under different conditions, then the univariate Bayesian analysis used here is not applicable. With the other techniques, one could use new data for population size, wealth, and so on, and a prediction of resource demand could be made. Of course, this assumes that the underlying models are responsive to these variables (e.g., population size or household income), and decision makers may prefer models that capture factors within their control.

In all of these cases, though, an estimated resource demand assumes that the modeled relationship holds for the new circumstance. Whereas this might hold for short-term analyses (i.e., what might happen next year, or what is currently happening in another zone of the city?), it is unlikely to provide a good assessment of long-term transformative changes in UM. One promising alternative not investigated here would be the use of a multilevel model were the “levels” are time periods; however, this would require time-series data on the dependent and independent variables of interest.

Figure 3 is an attempt to summarize these choices. In all cases, local authorities should carefully interrogate the results and ask whether the predicted resource demands are consistent
with their understanding of their cities. Unexpected results do not necessarily mean that the prediction is “wrong,” or that the modeling technique was incorrectly applied, but it does encourage the analyst to think why the result is occurring and to ask whether the underlying model successfully captured local realities.

As a final note, many of these methods require a degree of statistical sophistication to implement and interpret correctly. Expert users, for example, engineering or economic consultants and researchers, may be able to select from the full menu of techniques for applications such as detailed technical assessments of urban redevelopment proposals or evaluations of the environmental impact of socioeconomic trends; an example of such a scenario analysis can be found in Keirstead and Horta (2015). Nonexpert users, however (e.g., some local authorities), may be restricted to a more limited set of metrics and simpler descriptive applications. We note that even the most basic methods, such as per capita normalization, provide low error rates that may be suitable for such purposes. Conversely, the results also show that sophisticated methods do not guarantee accurate predictions; users must therefore exercise caution when selecting an appropriate downscaling method for their application.

**Conclusion**

This article was motivated by the growing need to have detailed accounts of urban metabolic flows for policy, infrastructure planning, and public awareness purposes. For many cities, aggregate metabolic accounts are the only data available; the question is therefore whether additional data, such as socioeconomic information, can be used to estimate the consumption of key resources at smaller spatial scales.

There are a variety of statistical methods that can be used for this downscaling of an aggregate metric to smaller units, that is, estimating energy consumption within a single district from the consumption of the city as a whole. Here, we examined six such methods: ratio-based normalization; linear regression (both calibrated on external data sets as well as a sample of the city’s own districts); linear regression with spatial autocorrelation; multilevel linear regression; and a basic Bayesian analysis. Our results showed that, using these methods, it is possible to downscale aggregate resource consumption to smaller geographies with a typical prediction error of around 20%. However, performance varies widely by method, geography size, and fuel type. In the best case (externally calibrated linear regression predicting electricity demand for the larger local authority districts), the MAPE was 4.3% and the RMSPE was 6.6%. However, in the worst case (per unit area normalization of electricity demand for the smallest LSOA geographies), the MAPE was 71.4% and the RMSPE was 488%.

Although these specific errors rates are only applicable to the formulations and data tested here, we proposed a generic flow chart for selecting an appropriate family of methods given the data constraints for a given city. We also showed how maps of error rates can be used to better understand the limitations of any statistical method for capturing the unique features of local geography.

Of course, the ideal situation would be to have official accounts of local resource consumption similar to the London energy data used in this report. With the introduction of smart metering technology, this may become more common in the future. However, the results suggest that, even in cities without substantial local data, downscaling may offer a way of identifying districts with patterns of resource of interest. Some methods, such as the Bayesian analysis, for example, can explicitly incorporate partial knowledge into the estimation process. Similarly, the internally calibrated linear regression showed how a partial survey of local resource consumption could be used to fit an appropriate statistical model. Further research should therefore investigate the detailed design of a long-term data collection strategy for local urban metabolic account data. In the interim, it would be useful to apply these techniques to other resource flows, such as those for water or waste.

**Acknowledgments**

J.K. was supported by the Enel Foundation as part of the “Metabolism of Megacities” project coordinated by the University of Toronto. I.M.H. thanks the Portuguese Foundation of Science and Technology for the financial support through the scholarship SFRH/BPD/86294/2012.

**Notes**

1. The urban metabolism study of Paris by Barles (2009) is a minor exception, but it investigates three different definitions of the city’s overall boundary, rather than looking at detailed flows within urban wards or districts.

2. $\log(x) = \log(x/(1 - x))$.

3. Evidence for the log-normal form is presented below. We adopt the STAN convention that a log-normal distribution with parameters $\mu$ and $\sigma$ represents data $y$ where $\log(y)$ is distributed with a normal distribution with mean $\mu$ and standard deviation $\sigma$.

4. In this section, errors are reported as mean (min-max).

**References**

Barles, S. 2009. Urban metabolism of Paris and its region. *Journal of Industrial Ecology* 13(6): 898–913.

Bettencourt, L. M. A. 2013. The origins of scaling in cities. *Science* 340(6139): 1438–1441. www.sciencemag.org/cgi/doi/10.1126/science.1235823. Accessed 20 June 2013.

Bettencourt, L. M. A., J. Lobo, D. Helbing, C. Kühnert, and G. B. West. 2007. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences of the United States of America* 104(17): 7301–7306.

Boulaire, F., A. Higgins, G. Foliente, and C. McNamara. 2014. Statistical modelling of district-level residential electricity use in NSW, Australia. *Sustainability Science* 9(1): 77–88.
Chertow, M. R. 2000. Industrial symbiosis: Literature and taxonomy. *Annual Review of Energy and the Environment* 25(1): 313–337.

Choudhary, R. 2012. Energy analysis of the non-domestic building stock of Greater London. *Building and Environment* 51: 243–254.

DECC (Department of Energy and Climate Change). 2013. Fuel poverty report: Updated August 2013. London: DECC. www.gov.uk/government/statistics/fuel-poverty-report-2013. Accessed 19 March 2016.

Dhakal, S., L. Guimaraes, S. A. Hammer, S. Kenihan, L. Parshall, M. Ruth, N. Schulz, and W. Wescott. 2008. Energy in cities. In *World Energy Outlook*. Paris: IEA. www.worldenergyoutlook.org/media/weowebsite/2008-1994/weo2008.pdf. Accessed 19 March 2016.

Fahmy, E., D. Gordon, and D. Patsios. 2011. Predicting fuel poverty at a small-area level in England. *Energy Policy* 39(7): 4370–4377.

Gelman, A. and J. Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge, UK: Cambridge University Press.

GLA (Greater London Authority). 2011. *London energy and greenhouse gas emissions inventory 2011*. London: Greater London Authority. www.london.gov.uk/priorities/environment/publications/delivering-londons-energy-future-the-mayors-climate-change-mitigation-and-energy-strategy. Accessed 29 March 2013.

Glaeser, E. L. 2008. Cities, agglomeration and spatial equilibrium. Oxford, UK: Oxford University Press.

Grubler, A., X. Bai, T. Buettner, S. Dhakal, D. Fisk, T. Ichinose, J. Keirstead, et al. 2012. Urban energy systems. In *The global energy assessment*, edited by GEA Writing Team, pp. 1307–1400. Cambridge, UK: Cambridge University Press.

Heiple, S. and D. J. Sailor. 2008. Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles. *Energy and Buildings* 40(8): 1426–1436.

Howard, B., L. Parshall, J. Thompson, S. Hammer, J. Dickinson, and V. Modi. 2012. Spatial distribution of urban building energy consumption by end use. *Energy and Buildings* 45: 141–151.

Jacobsen, N. B. 2008. Industrial symbiosis in Kalundborg, Denmark: A quantitative assessment of economic and environmental aspects. *Journal of Industrial Ecology* 10(1–2): 239–255.

Kavousian, A., R. Rajagopal, and M. Fischer. 2013. Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants’ behavior. *Energy* 55: 184–194.

Keirstead, J. 2014. dectools: Tools for accessing UK energy statistics. http://github.com/jkeirstead/dectools. Accessed 19 March 2016.

Keirstead, J. and I. Horta. 2015. *Downscaling London’s domestic energy consumption*. Project report. London: Imperial College London.

Keirstead, J. and A. Sivakumar. 2012. Using activity-based modeling to simulate urban resource demands at high spatial and temporal resolutions. *Journal of Industrial Ecology* 16(6): 889–900.

Kelly, S., M. Shipworth, D. Shipworth, M. Gentry, A. Wright, M. Pollitt, D. Crawford-Brown, and K. Lomas. 2013. Predicting the diversity of internal temperatures from the English residential sector using panel methods. *Applied Energy* 102: 601–621.

Kennedy, C., J. Cuddihy, and J. Engel-Yan. 2007. The changing metabolism of cities. *Journal of Industrial Ecology* 11(2): 43–59.

Kennedy, C., S. Pincetl, and P. Bunje. 2011. The study of urban metabolism and its applications to urban planning and design. *Environmental Pollution* 159(8–9): 1965–1973.

Korkin, J. 2005. *The city: A global history*. London: Weidenfeld & Nicolson.

Li, T., J. Corcoran, D. Pullar, A. Robson, and R. Stimson. 2009. A geographically weighted regression method to spatially disaggregate regional employment forecasts for South East Queensland. *Applied Spatial Analysis and Policy* 2(2): 147–175.

Mayor of London. 2011. Delivering London’s energy future: The Mayor’s climate change mitigation and energy strategy. London, UK: Greater London Authority. www.london.gov.uk/who-runs-london/mayor/publication/climate-change-mitigation-energy-strategy. Accessed 29 March 2013.

Min, J., Z. Hausfather, and Q. F. Lin. 2010. A high-resolution statistical model of residential energy end use characteristics for the United States. *Journal of Industrial Ecology* 14(5): 791–807.

Minx, J. C., T. Wiedmann, R. Wood, G. P. Peters, M. Lenzen, A. Owen, K. Scott, et al. 2009. Input-output analysis and carbon footprinting: An overview of applications. *Economic Systems Research* 21(3): 187–216.

Odum, H. T. 1983. *Systems ecology, an introduction*. New York: Wiley Interscience.

Odum, H. T. 1988. Self-organization, transformity, and information. *Science* 242(4882): 1132–1139.

Parshall, L., K. Gurney, S. A. Hammer, D. Mendoza, Y. Zhou, and S. Geethakumar. 2010. Modeling energy consumption and CO2 emissions at the urban scale: Methodological challenges and insights from the United States. *Energy Policy* 38(9): 4765–4782.

Sciubba, E. and S. Ulgiati. 2005. Energy and exergy analyses: Complementary methods or irreducible ideological options? *Energy* 30(10): 1953–1988.

Sovacool, B. K. and M. A. Brown. 2010. *Twelve metropolitan carbon footprints: A preliminary comparative global assessment*. *Energy Policy* 38(9): 4856–4869.

Stan Development Team. 2014a. RStan: The R interface to Stan, version 2.5.0. http://mc-stan.org/interfaces/rstan.html. Accessed 19 March 2016.

Stan Development Team. 2014b. Stan modeling language users guide and reference manual, version 2.5.0. http://mc-stan.org/interfaces/srstan.html. Accessed 19 March 2016.

Taylor, S., D. Fan, and M. Rylatt. 2014. Enabling urban-scale energy modelling: A new spatial approach. *Building Research & Information* 42(1): 4–16.

The City of New York. 2011. PlaNYC. www.nyc.gov/html/planycc2030/html/home/home.shtml. Accessed 29 March 2013.

Tilley, D. R. 2004. Howard T. Odum’s contribution to the laws of energy. *Ecological Modelling* 178(1–2): 121–125.

Tsutsumi, M. and D. Murakami. 2014. New spatial econometrics-based areal interpolation method. *International Regional Science Review* 37(3): 273–297.

UN (United Nations). 2014. *World urbanization prospects: The 2014 revision*. New York: United Nations Population Division. http://esa.un.org/unpd/wup/. Accessed 19 March 2016.

Venables, W. N. and B. D. Ripley. 2002. *Modern applied statistics with S*. Fourth edition. New York: Springer. www.stats.ox.ac.uk/pub/MASS4. Accessed 19 March 2016.

Van Vuuren, D. P., P. L. Lucas, and H. Hilderink. 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.

Weisz, H. and J. K. Steinberger. 2010. Reducing energy and material flows in cities. Current Opinion in Environmental Sustainability 2(3): 185–192.

Wilby, R. L. and T. M. L. Wigley. 1997. Downscaling general circulation model output: A review of methods and limitations. Progress in Physical Geography 21(4): 530–548.

Wolman, A. 1965. The metabolism of cities. Scientific American 213(3): 178–190.

About the Authors

Isabel M. Horta is an invited assistant professor and postdoctoral researcher in the Faculty of Engineering at the University of Porto in Porto, Portugal. James Keirstead was a lecturer in the Department of Civil and Environmental Engineering at Imperial College London in London, United Kingdom at the time this article was written. He is now Principal Data Scientist at Cambridge Assessment.