Global Environmental Change 34 (2015) 13–21

Contents lists available at ScienceDirect

Global Environmental Change

journal homepage: www.elsevier.com/locate/gloenvcha

A spatial typology of human settlements and their CO_2 emissions in England

Giovanni Baiocchi [a,*], Felix Creutzig [b,c], Jan Minx [b,d,e], Peter-Paul Pichler [d]

[a] Department of Geographical Sciences, University of Maryland, USA
[b] Mercator Research Institute on Global Commons and Climate Change, Berlin, Germany
[c] Technische Universität Berlin, Berlin, Germany
[d] Potsdam Institute for Climate Impact Research, Potsdam, Germany
[e] Hertie School of Governance, Berlin, Germany

ARTICLE INFO

Article history:
Received 8 July 2014
Received in revised form 28 May 2015
Accepted 2 June 2015
Available online

Keywords:
Human settlement
GHG emissions
Mitigation
Typology

ABSTRACT

Case studies demonstrate that urban greenhouse gas emissions are driven by socio-economic, climatic and urban-form specific characteristics. But neither the interdependence between attributes nor their place-specific context has been well understood. In this paper, we develop a nested typology of human settlements in England, containing both urban and rural environments, that is based on local drivers of emissions from direct energy use in nearly 7000 local areas. We reject the standard hypothesis that settlements obey a global linear model explaining emissions. The emissions of human settlement types are characterized by unique, place-specific combinations of emission drivers. We find that density and income are dominant classifiers of local carbon dioxide emissions. However, their specific impacts are particular to human settlement types as characterized by the place-specific combination of income, household size, and local climate, which are themselves spatially contextualized. Our typology strongly correlates with the geographic distribution of lifestyles. Average household carbon dioxide emissions are highest for very high income households (top 3%) living in low-density settlement areas with large houses, mostly concentrated in outer suburbs. Our results provide a first step towards enabling decision makers to go beyond one-size-fits-all approaches but instead to apply appropriate and specific mitigating measures for each type of human settlement. In turn, successful strategies could be transferred between similar types of human settlements.

1. Introduction

Most CO_2 emissions originate in human settlements, from villages and small towns to megacities. The characterization of emissions of carbon dioxide caused by urban energy consumption has become the focus of a fast growing body of scientific literature (Lenzen et al., 2004; Kennedy et al., 2009; Parshall et al., 2010; Baiocchi et al., 2010; Weisz and Steinberger, 2010; Glaeser and Kahn, 2010; Minx et al., 2011, 2013; Hillman and Ramaswami, 2011; Gurney et al., 2012; Baur et al., 2014). Recent assessments on cities and climate change have highlighted that there is still a lack of understanding concerning what determines emissions in cities and which type of design-features make a difference in the carbon output (Grubler et al., 2012; Seto et al., 2014). A key hurdle is that approaches so far largely rely on simple correlation analysis or single common regression equations to understand the practical and statistical relevance of emission drivers (see, e.g., Newman and Kenworthy, 2006; Min et al., 2010; Karathodorou et al., 2010; Jones and Kammen, 2014; Parshall et al., 2010; Kennedy et al., 2009; Makido et al., 2012; Minx et al., 2013). Such an approach may suppress both spatial context, non-linear effects, and the interdependence of emission drivers.

Local conditions, infrastructures, and lifestyles vary widely. Many of the relevant emission drivers for understanding carbon emissions of human settlements can therefore only be appreciated in their particular context. For example, heating patterns in the UK – as expressed in the heating degree day variable used for this research – vary widely due to differences in the local climate. Equally, it is well recognized that the available housing infrastructure within different parts of a settlement directly determines

* Corresponding author at: Department of Geographical Sciences, University of Maryland, 2181 Samuel J. LeFrak Hall, College Park, MD 20742, USA.
Tel.: +1 301 405 2342.
E-mail address: baiocchi@umd.edu (G. Baiocchi).

http://dx.doi.org/10.1016/j.gloenvcha.2015.06.001
0959-3780/

energy consumption of different areas (Boardman, 2007). Schelling's (1969) seminal work on segregation has helped to explain the clustering of lifestyles in space, which has been observed in cities around the world (Clarke, 1991).

All those local drivers combine differently in different places and there is no reason to assume that they would do so linearly and independently. Today, entire research fields such as geo-demographics or other variants of consumer segmentation are built around the importance of place-specific contextualization for the prediction of consumer behaviour (Harris et al., 2005; Longley, 2012). The local context is equally important for understanding and predicting the carbon output of settlement areas. The limited understanding of emission drivers of cities might, in fact, be related to a lack of appreciation of this heterogeneity.

In this paper we test the hypothesis that the impact of drivers of residential CO_2 emissions can be understood using simple one-size-fit-all models disregarding complex dependencies among each other and local circumstances. We apply a tree regression analysis to devise a human settlement typology of residential CO_2 emissions that reveals the differential impacts of emission drivers and their interactions depending on the particular spatial context. Such typologies help to appreciate the heterogeneity across human settlements. Also, by grouping settlements areas into different clusters they allow for the generalization of insights to similar areas, which is helpful for policy applications. We apply this method to data, which is geographically complete (all human settlements across England) and has a high spatial resolution (sub-city level). The endogenous clustering applied to this data largely avoids the necessity of drawing arbitrary city boundaries as well as any further a priori spatial discrimination.

Our approach allows us to disentangle the interactions between the most important socio-economic, infrastructural and climatic drivers of residential CO_2 emissions of a given location while at the same time revealing the conditions of the wider spatial context of the human settlement. The wider spatial context is relevant as, for example, the emission profiles of two otherwise identical low-density residential areas may substantially differ if the first is part of a small rural town and the second a suburb of London (Büchs and Schnepf, 2013). In fact, we demonstrate that typical emission patterns of human settlements also depend on the spatial context of these settlements. We also show that types of residential CO_2 emissions correlate with observed location-specific lifestyles.

Our approach follows Creutzig et al. (2015) in their development of a global typology of cities with respect to their energy use that points to an urbanization wedge for climate change mitigation. While lacking the global scope, this work is geographical explicit and complete, builds on a richer and consistent set of human settlement data, and focuses explicitly on residential CO_2 emissions. As such, our typology is a first important step towards a contextual understanding of local emissions. It suggests that local policies may need to account for these differential impacts of emission drivers and go beyond 'one-size-fits-all' approaches. At the same time it opens up the possibility that settlements of similar types can learn from context-specific best practice and lead towards more systematic impact testing of policies. Those are fundamental prerequisites for the eventual emergence of an evidence based best practice approach for urban sustainability comparable to standard practice in other areas of public policy such as in health care.

2. Developing the typology

In this section we describe the method and data used for devising a typology of human settlements with respect to CO_2 emissions. We define the term human settlement in a general sense as "cities, towns, villages, and other concentrations of human populations which inhabit a given segment or area of the environment" (UNESCO-UNEP, 1983). Larger human settlements can be conceptualized as being composed of a number of smaller, distinct spatial units. We assume that these smaller spatial units can be usefully classified into groups defined by threshold values for a set of common determinants of domestic carbon dioxide emissions. We will refer to these groups as "settlement types". In this sense, the unique arrangement of settlement types mark the structure of a human settlement.

2.1. Method

We describe first the regression model linking emissions with its determinants, then the recursive partitioning method used to identify the different types of settlements (each of which is subject to a separate regression), and finally the procedures used to statistically test and validate our model.

In line with most theoretical and empirical literature on the determinants of emissions typically based on the demand for specific forms for energy, we consider the following regression equation,

$$\ln E_i = \beta_0 + \beta_1 \ln Y_i + \sum_{j=2}^{k-1} \beta_j \ln X_{j_i} + \varepsilon_i, \qquad (1)$$

where $i = 1, \ldots, N$ indexes settlements. Here E_i denotes CO_2 emissions, Y_i is income, and X_{j_i}, for $j = 2, \ldots, k-1$, denote the other socio-economic variables, k the total number of regressors, and ε_i is the classical error term.

For developing our typology of human settlements with respect to drivers of CO_2 emissions, we used the recursive sample partitioning method developed by Loh (2002) and Kim et al. (2007) known as GUIDE (acronym for *Generalized, Unbiased, Interaction Detection and Estimation*). It is a refinement of the classification and regression tree (CART) methods of Breiman et al. (1984), that iteratively partitions the data into ever increasing homogeneous sub-groups, by fitting a separate regression model at each node (Eq. (1) in our case).

At each split the available sample is partitioned into two groups, using binary splitting rules of the form $X_j \leq x$, obeying separate linear regression models that minimize an overall measure of discrepancy between the observed response and the predicted values of the estimated models, over all possible splits for all available independent variables. The resulting model can be conveniently presented as a binary decision tree: the branch on the right (left) of each non-terminal node contains the settlements for which the split variable is greater (lower) than the split value. CART type models can be viewed as computationally efficient strategies for estimating a fully nonparametric regression model. GUIDE has been shown to improve on its predecessors by reducing classification errors and increasing the interpretability of the results (see, e.g., Loh, 2009; Tan, 2010, for details).

To avoid "overfitting" the available sample, a large tree is "grown" first which is then reduced in size by a suitable "pruning" procedure. In practice, because of the flexible nonparametric nature of the approach, it is entirely possible to fit a tree with many parameters that has adapted too well to noisy features of the data (OLS tends to give higher weights to noisy observations) and is therefore unsuitable for generalizations and difficult to interpret. Generally, as the number of splits increases, the variance of the model prediction for a given observation will decrease but the model will tend to overfit the available sample and decrease its accuracy when applied to different data. This is known as the *variance-bias trade-off* as decreasing one comes at the expense of increasing the other. To choose a model that takes both errors into account, variance and bias can be combined to form the mean squared error (MSE). This is at the basis of the well-established cost-complexity pruning methodology, first introduced by Breiman et al. (1984), to determine the size of trees that minimizes miss-classification errors. In practice, the predictive ability of a tree of a particular size is assessed by using a technique known as (10-fold) cross-validation (see, e.g., Clark and Pregibon, 1993). This is performed by randomly splitting the available sample into 10 equally sized parts (folds), leaving one part out for validation, and using the remaining 9 parts as training data to grow a tree. The partitioning procedure is then repeated 10 times, with a different subset of the data reserved for use as the test dataset each time, averaging the performances of the 10 models to yield a *cross-validation* estimate of how well the model performs with unseen data. The size of tree that minimizes the cross-validated error is chosen as the size of the final pruned tree. For more details see the Supporting Information (denoted henceforth as SI) and the references therein (Chaudhuri et al., 1994).

One limitation of the regression splitting approach is the lack of distribution theory useful for inference on splitting variables and thresholds. Hansen (1996, 2000) developed a testing procedure that addresses this issue which we employed to obtain confidence intervals for the main splits (see also, Loh, 2009; Tan, 2010, for details).

2.2. Data

All data used in the analysis is publicly available through national statistics offices. Final gas and electricity consumption data of households were obtained from the Department of Energy and Climate Change (DECC, 2011). We focus on these domestic CO_2 emission sources, because they are shaped by local drivers and can often be directly influenced by policies at the local level – most importantly spatial planning. Emissions from transportation would be of similar importance, but are not available at the required level of detail. The focus of many studies on all emission sources can in fact blur the picture.

The geo-referenced GIS datasets provided by the Office for National Statistics (ONS, 2011) were used for visualization purposes and to obtain heating degree days (HDD) averages from the Met Office UKCIP dataset (Met Office, 2011). All other indicators are part of the ONS Neighbourhood Statistics (ONS, 2011). If possible, the time period of the datasets used was 2005.

The spatial units underlying the analysis in this study are the *Middle Layer Super Output Areas* (MSOA). England is partitioned into 6780 MSOAs, which are part of the layered statistical geography hierarchy developed by the Office for National Statistics (ONS, 2011). They are constructed to contain an average population of 7000 and are constrained to the administrative *Local Authority* (LA) boundaries of 2003. The particular partitioning of a given geographical area used for spatial analysis has major impacts on the results of the analysis. The layered statistical geography hierarchy has been specifically designed to improve the usefulness of small area statistics and to promote the validity and comparability of different studies (ONS, 2011).

The following list specifies the origin of the various input data.

- *Energy consumption (EC)* data is based on domestic electricity and gas (weather corrected) consumption data provided by the Department of Energy and Climate Change (DECC, 2011). The data is based on detailed meter readings provided by electricity suppliers and gas transporters which have been matched to the MSOA levels to generate georeferenced datasets. All datasets are subject to a number of validation and quality assurance procedures (DECC, 2011). We convert energy consumption data into carbon dioxide emissions following the reporting guidelines by the Department of the Environment, Food and Rural Affairs (DEFRA, 2007). The resulting CO_2 emission account therefore comprises scope 1 (all direct emissions in the domestic sector) and scope 2 emissions (indirect emissions that arise in the production of electricity consumed by the domestic sector).
- *Income* represents average weekly total household income in pounds estimates (unequivalised) as reported in the ONS Neighbourhood Statistics (ONS, 2011). The modelling methodology applied enables a combination of survey data with census and administrative data in order to improve the quality of estimates at small area level.
- *Heating degree days* (HDD) have been calculated by overlaying the GIS data for the MSOA layer with the 5 km gridded HDD data for 2005 available from the Met Office UKCIP dataset (Met Office, 2011).
- *Housing in poor condition* and *Central heating* are two sub-indicators of the underlying indicators of the Living Environment domain, one of the seven domains contributing to the Index of Multiple Deprivation 2007 published by the Office for National Statistics. The data are published on the LSOA level and have been aggregated to the MSOA level for this analysis.
- *Density* in persons per hectare is calculated by dividing the population of an MSOA by its area.

3. Results

The regression tree for predicting CO_2 emissions from household energy consumption in England is displayed in Fig. 1. The tree presents a nested and contextualized description of emission drivers of human settlements. The 15 terminal nodes contain the estimated subsamples and represent the human settlement types of the typology. The preliminary threshold estimation produced a tree with 75 terminal nodes that was reduced to 15 using the cross-validation procedure described above to prevent overfitting. Each type is characterized by its specific sequence of attributes in the regression tree. Individual regressions reveal the impact of attributes on CO_2 emissions for each human settlement type (SI Tables S2, S3, and S4. The respective descriptive statistics are provided in Tables S5 and S6). We compare our typology with the UK national statistics' geo-demographic area classification at the MSOA level (ONS, 2001) in the following called ONS lifestyle groups (see, for details, Harris et al., 2005; Baiocchi et al., 2010). The overlap between human settlement type, as driven by CO_2 emissions, and these lifestyle groups is coded in Fig. 2. The spatial distribution of settlements is displayed in Figs. 3 and 4.

The results of our analysis demonstrate that the non-linear combination of CO_2-emission drivers uniquely characterize human settlements (Fig. 1) together with their attribute values. The 15 distinct types of human settlement can only be properly understood by taking a closer look at the non-trivial relationships between the characteristics of the settlement types and their respective CO_2 emissions (see Fig. 1) but, altogether, not by autocorrelation statistics (S4 and Table S7). We observe six key insights.

First, density and income are the key drivers of the typology. Fig. 1 shows that density is the best discriminator for a typology of human settlements, as characterized by CO_2-emission drivers. The split based on density occurs at the threshold of 50 persons per ha (82nd percentile). The *high* density types (12–15, 56–58, 59), comprise an area which amounts to less than 1% of the area of England but includes about 18% of its population. For Greater London, high density nodes cover about 36% of its area and 66% of its population. Income is the second most important driver of CO_2 emissions, particularly for *low* density types where it correlates more strongly with CO_2 emissions within each regression node and appears higher up in the regression hierarchy: In the *low* density

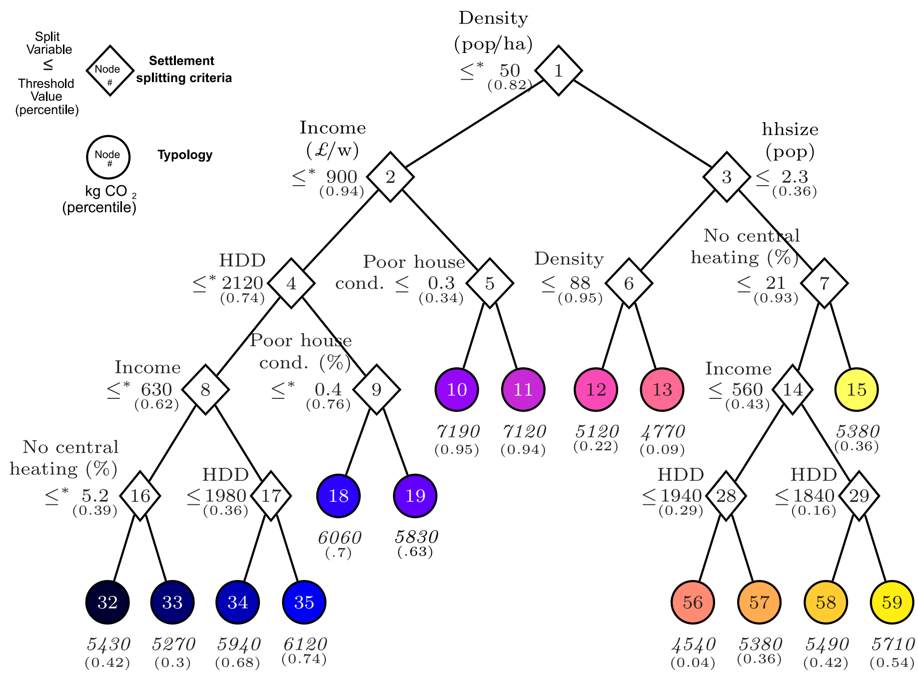

Fig. 1. Human settlement types in England as determined by their CO_2 emissions drivers. CO_2 emissions drivers split spatial units recursively to produce maximally distinct settlement types. The results are highly robust (S1-3). Diamonds indicate the splitting criteria in terms of splitting variable and threshold value of splitting variable; circles are terminal nodes which represent the different settlement types and contain the estimated subsamples. A human settlement whose attribute satisfies a splitting criterion, it is assigned to a node down the left branch; otherwise it goes down the right branch. Beneath each node, the average predicted emission, in italics, with associated sample pth quantiles, in parenthesis, are reported. The symbol $\leq^*$ denotes $\leq$ or missing. There are only 64 missing values, all for the income variable, less than 1% of the 6780 observations (these belong to rural and isolated areas). The 95% confidence interval for the most important first and second splits are (44, 51) for density, (880, 910) for income, and (2.3, 2.35) for household size. For details see the SI.

part of the tree (nodes 10, 11, 18, 19, 32–35), income is the dominant discriminatory attribute splitting clusters at about 900 £/week. Crucially, the relationship between emissions and determinants changes significantly for high income (94th percentile) and low density highlighting the non-linear relation between emission attributes and resulting domestic CO_2 emissions. In the high-density (right) part of the tree, income is less relevant in discriminating between human settlement types (only occurring at node 14).

Second, the impact of emission drivers is highly context dependent. We reject the hypothesis that emission drivers of human settlements can be adequately explained by a unique global model. Instead, we find highly differential impacts and combinations of emission drivers depending on the particular settlement

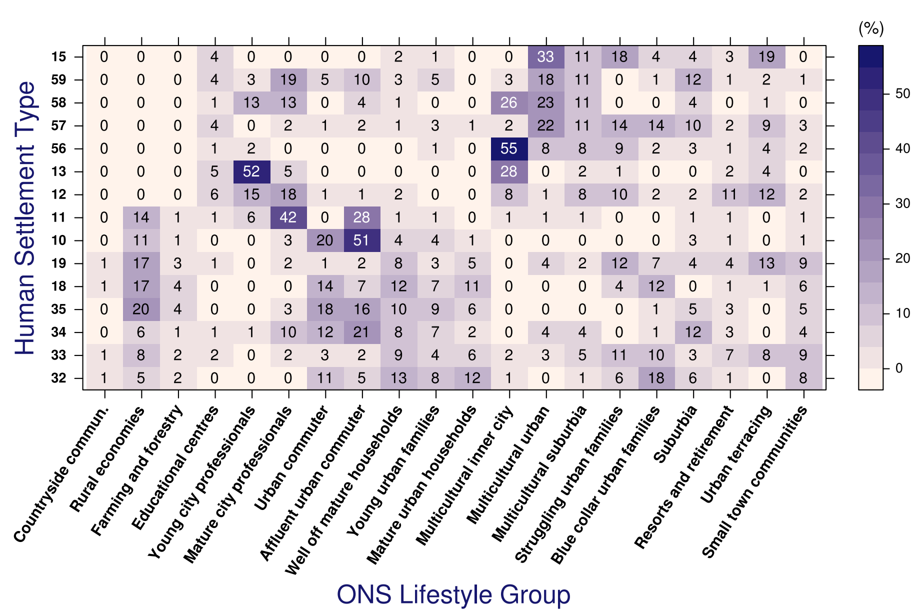

Fig. 2. Frequency of lifestyle group membership occurring in each human settlement type, as characterized by its CO_2 emission drivers. A few settlement types map into ONS lifestyle categories quite well. For instance, the prevalently London type 13 (83%) of high density higher income and smaller household size, has a majority of 'young city professionals'. The other almost London specific (80%) type 56, with low income and high density, has a majority belonging to the 'multicultural inner city' group. The lifestyles of richest and largest emitters are also clearly identified: 10, most distant type from centers, with 'affluent urban commuter', and 11 with 'Mature city professionals'. See the text for more details.

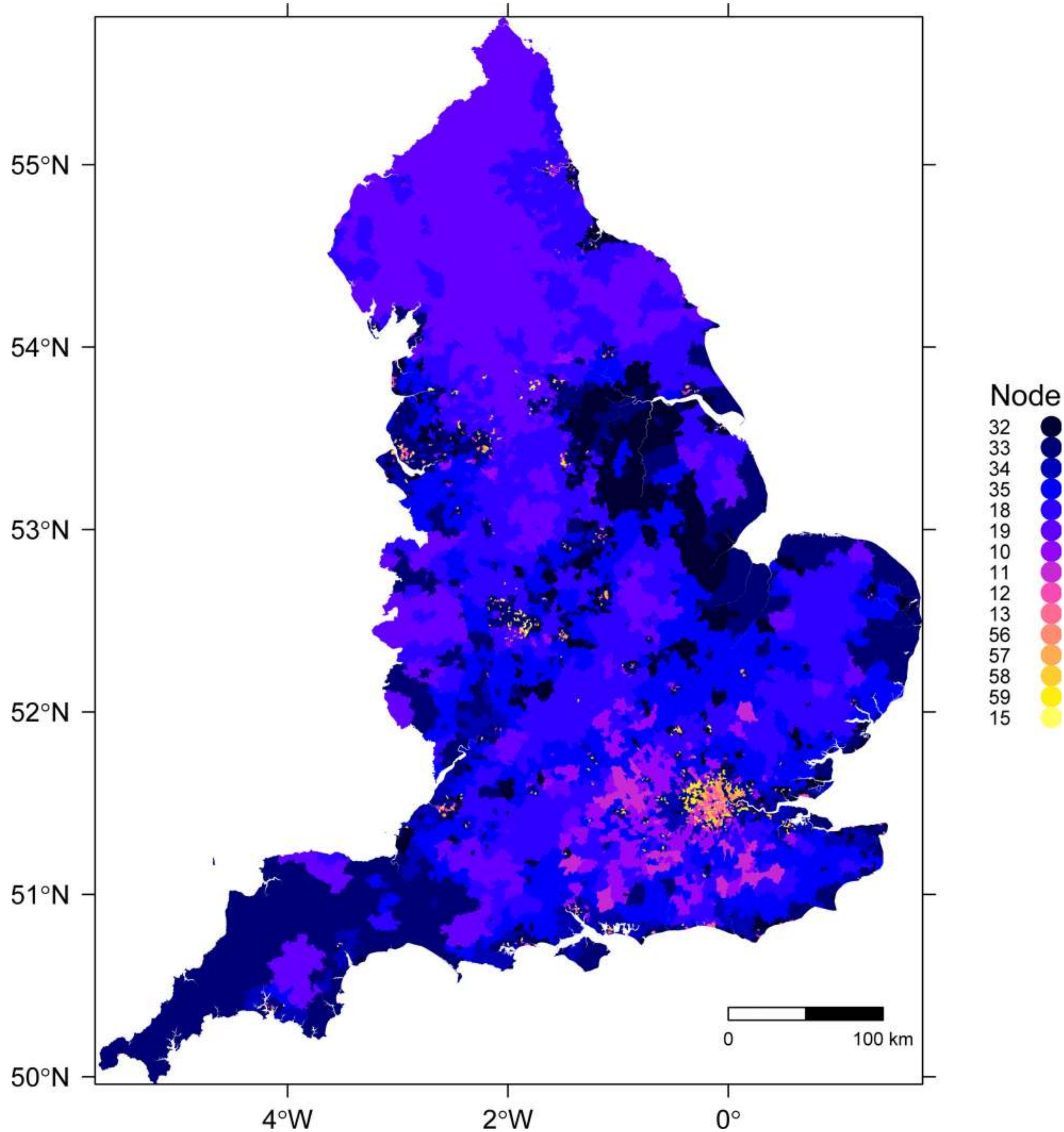

Fig. 3. Human settlement type, as characterized by its CO_2 emission drivers in 6780 English human settlements (Middle Layer Super Output Areas, MSOA). Each human settlement is colored according the corresponding node from the tree regression results it belongs to (Fig. 1). Settlement types can be compared with the ONS areas classification in terms of lifestyle (Fig. 2). Nodes 10, 11, 34 and 35, have the highest proportion of the ONS group 'Affluent urban commuter', 51%, 28%, 21%, and 16%, respectively. Node 10, the node with largest average distance from centres (see Fig. 4), has also a high proportion of 'Urban commuter' (20%) and Node 11, which are, on average, slightly closer to the city centres, of 'Mature city professionals' (42%). These very high income, high emissions, and lower density settlement types tend to have large houses, with four or more bedrooms, often in Georgian and Victorian terraces and are found in many urban areas of the UK, but particularly in London (Wandsworth, Ealing, Barnet, Richmond and Merton), Manchester (Trafford), and Oxbridge.

type and its particular place-specific context. Table S4 presents the regression results for the aggregated low and high density areas (nodes 2 and 3 in the tree regression picture) and the pooled OLS. Estimates differ greatly across typologies, both statistically and substantively. The spatial Chow test (Anselin, 1990) very strongly rejects the stability of coefficients across the identified typologies both jointly and individually (see Supporting information). For instance, population density is statistically significant for most typologies but has a large negative impact only for the higher density areas, above 50 persons per ha (82nd percentile). For higher density areas a doubling of density decreases emissions on average by about 18% as opposed to only 2.2% for all settlements (pooled OLS). Density appears very important for high density settlement types (12, 13, 58, and 59), but not for the two poorest settlement types (57 and 15). For the lower density settlements income has a much higher impact for the richest top 6% of the settlements (nodes 10 and 11). A 10% increase in income, increases average emissions by almost 7% more than twice the overall effect for all settlements. Analogous heterogeneity is present in all other variables. Estimating one equation to explain emissions for all

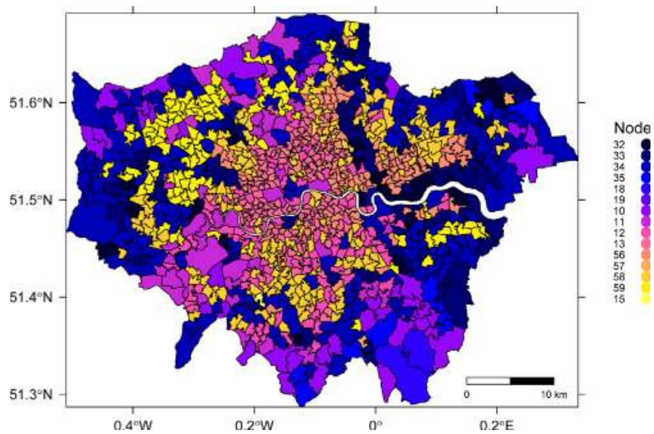

Fig. 4. Human settlement type, as characterized by CO_2-emission drivers, of 982 MSOAs of Greater London. Some human settlements are endemic to London. London specific Node 13 includes 52% of ONS group 'Young city professionals' and 28% of group 'Multicultural inner city'. The former are young, highly qualified people in their late 20s living mostly on their own in urban areas flats primarily in Inner London, the latter are populated by young multi-ethnic communities living in flats, mostly in their 20s and 30s with only a few very young children that are primarily found in London boroughs.

settlements greatly misrepresents the impact of various determinants by ignoring nonlinearities interactions between them and their spatial context.

Third, average carbon dioxide emissions are highest for low-density, high-income and lowest for high-density, low-income human settlement types. This pattern is generally in line with the notion that increasing density is associated with lower emissions (Newman and Kenworthy, 2006). However, our results go substantially beyond insights from these simple relationships and show that emissions are particular to human settlement types as characterized by the place-specific combination of emission drivers, including income, household size and local climate. For example, income becomes relatively more important as driver of CO_2 emissions in low-density settlements than in high-density settlements. Specifically, the highest emissions are found for low density, high income settlement types (10, 11). The settlement type with the highest mean density in the high density part of the tree (13) has the second lowest mean carbon dioxide emissions. Types 10, 11, 34 and 35, have the highest proportion of 'Affluent urban commuter', 51%, 28%, 21%, and 16%, respectively (Fig. 2). Settlements of type 10 have the largest average distance from the closest major urban center (see Fig. 4) and also a high proportion of 'Urban commuter' (20%). In comparison, settlements of type 11 are, on average, slightly closer to major urban centers and have 'Mature city professionals' (42%) as the dominant demographic (Fig. 2). These settlement types have populations with very high income, far above average higher education qualification, and relatively low density. They tend to have large, often detached houses, with four or more bedrooms and are found in many urban areas of the UK, but particularly in London (Wandsworth, Ealing, Barnet, Richmond and Merton), Manchester (Trafford), and Oxbridge. Overall, the six types with highest mean emissions have low density and the types with the lowest mean emissions have high density.

The smaller impact of income on emissions in high density areas comes with a comparatively larger impact of other drivers. At the highest sub-node (3) the sample is split by household size at a threshold of 2.3 persons per household. Subsequent nodes above that threshold contain types with the lowest and highest mean CO_2 emissions for high density areas. In this context, household size, heating degree days, poor housing condition, and central heating become important in explaining the fine structure of urban typology. For example, settlements located in warmer areas are generally associated with lower emissions. If the final split is based on HDD, warmer areas are always associated with lower emissions than colder areas. However, if a split based on HDD occurs higher up in the tree, the effect of the HDD variable on emissions can be outweighed by the influence of other socio-economic determinants. For example, types 34 and 35 are typically located in warmer parts of England than types 18 and 19, which have the highest average HDD. Due to these differences in climate it takes about 20% more energy to heat similar homes in most areas of types 18 and 19, yet the emissions of type 34 and 35 are higher owing to higher income in these areas.

Fourth, variance in emissions is higher for low-density than for high-density types. The mean emissions of the fifteen settlement types vary by a factor of 1.6. CO_2 emissions vary slightly more across low density (factor of 1.36) than high density (factor of 1.25) settlement types. This can partially be explained by more diverse incomes in low density settlements. Also, size and structure of dwellings tend to be more similar in high-density settlement types, for example, due to tighter space constraints. (This heterogeneity problem highlights the dangers of a "one size fits all" regression approach. This problem is expected also because low density settlement types have higher income and therefore more of what is know in economics as "discretionary income", i.e., scope for more choice about consumption and therefore more variability (see, e.g., Gujarati, 2004).

Fifth, the Greater London area is notable in that it contains a number of settlement types which are mostly endemic (Types 13, 56 and 58). Most MSOAs in London's extremely dense core belong to type 13 and account for 82% of all MSOAs of this type. Node 13 includes 52% of ONS group 'Young city professionals' and 28% of group 'Multicultural inner city' (Fig. 2). The former are young, highly qualified persons in their late 20s living mostly on their own in urban flats primarily in Inner London (Westminster, Camden, Islington, Wandsworth, Hammersmith, and Fulham), the latter are populated by young multi-ethnic communities living in urban flats, mostly in their 20s and 30s with only a few very young children that are primarily found in Inner and Outer London boroughs. Type 58, which occurs almost exclusively in the London area ($\approx$94%), is characterized by high income and the lowest average HDD. Low HDD settlement types occur in London in more than 50% of all cases. The different types and frequency of human settlements occurring in London re-emphasize that London differs radically from the rest of England, not only in finance, culture and real estate (Sassen, 2001), but also in terms of its CO_2 emission drivers.

Sixth, the same settlement types tend to cluster in space. The maps of the spatial distribution of human settlement types, presented in Fig. 3 (England) and Fig. 4 (London) provide further insights on the relationship among settlement types and the characterization of cities. Crucially, areas of the same type cluster together which demonstrates that features covary spatially. An *adjacency* analysis substantiates this insight. Fig. 5 reports the relative frequencies of settlements of different types being neighbors. For settlements of most types, their most likely neighbor is of the same type. Settlement type 33 is common and hence neighbors to many other types. Types 13, 56, and 58 occur often together, and represent the neighborhoods in London that are inhabited by mostly 'Young city professionals', 'Multicultural inner city', and 'Multicultural urban' inhabitants, all with relatively low emissions. This effect is due to the different scales on which the underlying drivers vary spatially. Climatic conditions usually vary on a comparatively large scale unless influenced by factors like the urban heat island effect. In large cities, residential neighborhoods sharing similar energy

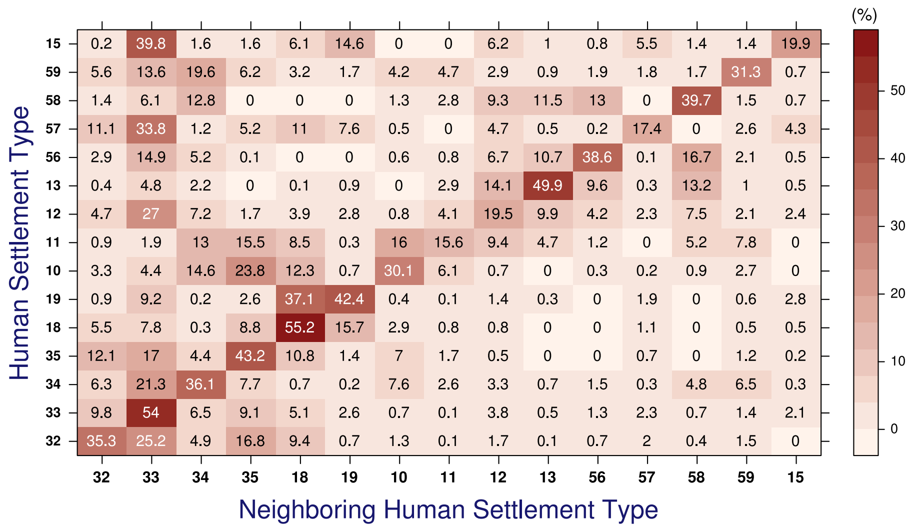

Fig. 5. Neighboring relationship between human settlement types. Settlements of the same type are neighboring each other and effectively cluster together in space (the dark diagonal). But also some human settlements of different types tend to occur more often together than expected by chance.

consumption patterns due to similar construction periods might span multiple MSOAs. Also other properties like income or infrastructural features are not randomly distributed at the MSOA level. Income related clusterings on the city level (e.g., within London) could be signs of segregation dynamics (Schelling, 1969).

Seventh, settlement types are spatially correlated to the closest density hot spot (urban center). We have examined the frequencies of different area types with respect to density hotspots to obtain density-contextualized spatial correlations. As density hotspots we used all 31 cities with a population larger than 150,000 inhabitants and calculated the distance of each MSOA to the closest one. Fig. 6 shows the smoothed frequencies for each area type. The frequencies have been normalized by dividing the occurrences in each distance bin of the individual area type histograms by the total number of MSOAs within the same distance bin because in a two-dimensional centrality metric the expected frequencies increase as a quadratic function of distance. Settlement types exhibit very distinct profiles regarding their frequency of occurrence at different distances from urban centres. For the types mostly endemic to London type 13 makes up the inner city core, followed by type 56, type 11 and type 10 moving outwards. Regarding their emission profiles the two London inner city types (13, 56) have the lowest emissions and the two outer types (11, 10) by far the highest emissions overall. The observation that more central types have lower emissions hold in general. Emissions are significantly correlated with the average distance to the closest city center (1km distance translates to 85.5 kg of CO_2, $p = 0.003$, adj. $R^2 = 0.465$). On the aggregate level, this analysis reveals a number

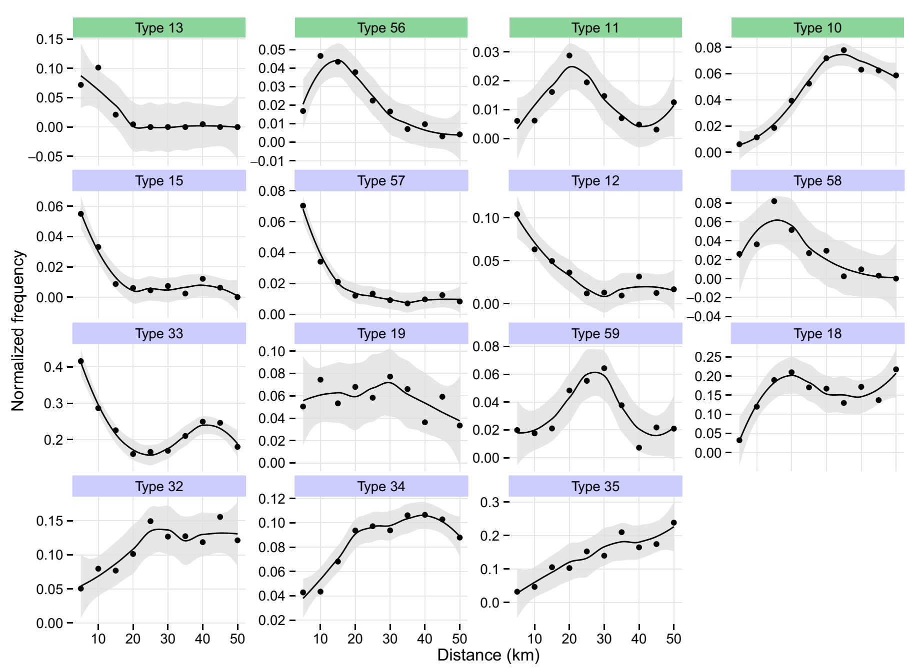

Fig. 6. Normalized frequency of settlement types with respect to distance to the next city center with population in excess of 150,000 inhabitants. First row (green) shows the types mostly endemic to the greater London area. Types are sorted by average distance to nearest city center. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of spatial patterns shown in Fig. 6: (a) smaller households tend to live closer to city center than larger ones; (b) within higher density areas, higher incomes tend to live slightly further away from the city center; (c) in lower density areas, lower incomes tend to live closer to townships than higher incomes. On the level of specific types, the composition of ONS lifestyle groups within each type is well reflected in this metric for the most rural (18, 19, 35), suburban (10, 34, 59) and inner city (12, 13, 15, 56, 57) types, respectively (Fig. 2).

4. Discussion

We here provide a spatially explicit CO_2 emission-driven typology of human settlements that is objective, explicit, and reproducible, where the impact of specific emission drivers is unmasked by its contextualization with other emission drivers. The typology is (a) geographically complete (including all English settlements); (b) non-discriminatory as settlement types are endogenously determined by emission drivers, but not via pre-classification of urban and rural areas; and (c) spatially fine-grained drivers of GHG emissions in human settlements at unprecedented spatial resolution.

We reject the hypothesis that global models are useful to understand emission drivers at the local scale, because they do not sufficiently allow for heterogeneity to appreciate the specific impacts. We demonstrate that local emission drivers are often relevant in a particular context, but less so in others. By pooling the entire sample, averaging effects can overshadow this important place-specific relevance. This could be the reason why studies not always find a very strong impact of context-specific variables like density or HDD (see, e.g., Minx et al., 2011; Kennedy et al., 2009; Makido et al., 2012).

Our results imply that in mostly urbanized countries, like the UK, the rural/urban distinction is becoming less useful compared to countries with lower urbanization levels. Notably, low density, "rural" areas have driver/emission profiles similar to low density "urban" areas. But it holds that low density areas tend to have higher emissions, while higher density areas (and not necessarily meaning densities referring to high-rise buildings) have lower emissions.

The fine-scaled nature of our results has implication for the spatial resolution of climate mitigation policies. City-wide studies are hampered in policy conclusions as they cannot resolve the large heterogeneity within any city; resulting policy recommendations remain at the level of an one-size fits-all approach. In contrast, as our analysis elucidates drivers of emissions at small-scale human settlement scale, we provide appropriate data for designing mitigation policies that comprehensively cover human settlements of a whole country but nonetheless always integrate locality-specific context information. Our analysis enables both specialization and generalization of policies. Policies can specialize: policy interventions can be targeted to the sub-city level, acknowledging the different combinations of GHG emission drivers in different human settlements. But polices can also generalize: our analysis of settlement types allows for a formulation of potential policy interventions targeting similar settlement types across different cities. Hence, settlements of similar types can learn from context-specific best practice experience. It also improves the potential for systematic impact testing of polices. Based on these comparable units, a systematic, evidence based best-practice approach could be implemented, as it is standard in other areas of public policy such as in health care.

The typology is enabling, but alone insufficient for systematic policy assessment. A policy assessment would also investigate costs of mitigation options and inter-temporal dynamics, e.g., path dependency issues associated with residential density and urban form. But a few hints for targeted policies already emerge. For example, settlements with poor housing conditions (types 11 and 19) could be prioritized for retrofitting measures; and types with high HDD (types 18, 19, 35, 57 and 59) could be prioritized for advanced insulation measures.

The literature on how CO_2 emissions depend on city population size and density is rapidly developing. While there is modest agreement that CO_2 emissions per capita are lower in denser cities compared to less dense suburbs (Glaeser and Kahn, 2010; Jones and Kammen, 2014), urban density differences play out most visibly in the transport sector and between the urban form differences of world regions (Newman and Kenworthy, 1989; Baur et al., 2014), which can be explained, in an urban economics framework, by the long-term differences in transport prices (Creutzig, 2014). This perspective is complicated by the effect of population size: some literature points to linear scaling between population size and CO_2 emissions (Fragkias et al., 2013), whereas a recent study finds that CO_2 emissions from transport scale supra-linearly with city population size: Larger city have larger per capita transport emissions per capita than smaller cities (Louf and Barthelemy, 2014).

Our study adds an additional perspective to this debate. We look only at residential CO_2 emissions, excluding transport CO_2 emissions. But our analysis looks at finer spatial scale, finding that emissions are systematically lower above a certain population density (50 persons per ha) on district level. Hence, rather than debating average city densities, it might be more important to increase densities over this threshold density, but possibly not much more, in as many city quarters and settlements as possible.

Importantly, this density value is roughly of the same magnitude as the critical threshold in population densities needed to support minimal modal choice (Frank and Pivo, 1994; Rickwood et al., 2008). In other words, residential CO_2 emissions change their patterns at similar levels than transport CO_2 emissions. This might hint towards potential synergies in these two key areas of spatial planning, substantiating the claim that climate change adaptation and mitigation should be mainstreamed with urban planning (Creutzig et al., 2012; Viguie and Hallegatte, 2012).

The ex-post analysis demonstrates that emission patterns are location-dependent and a function of centrality. This results hints to the potential of marrying our empirical results with causal urban economic analysis. The concept of centrality is as old as the spatial economic literature (Christaller and Baskin, 1966). Thuenen emphasized the interaction between market value, transport costs and agricultural productivity in making best use of land surrounding a city (Von Thünen, 1875). A century later the model structure was extended to explain residential patterns as a function of the distance to the city center (Alonso, 1964). In this study, we find that residential CO_2-emission types of human settlements (excluding transport emissions) demonstrate a clear spatial signal as a function of distance to the next city center. This is most likely an indirect effect of the building vintage and socio-economic characteristics, which themselves are more directly a consequence of centrality (Fujita, 1989).

Identifying commonalities across space while allowing for place-based specificities, rather than assuming a rigid distinction between urban and nonurban environments, is a precondition for a better scientific understanding of the role human settlements can play in climate change mitigation. Our typology is one contribution in this direction: Common settlement types might face common mitigation challenges, with greater potential for intervention; and mitigation policies might target specific settlement types, rather than relying on one-size-fits-all approaches. A global data-driven typology of cities and their GHG emissions should follow suit.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.gloenvcha.2015.06.001.

References

Alonso, W., 1964. Location and land use. Toward a general theory of land rent.
Anselin, L., 1990. Spatial dependence and spatial structural instability in applied regression analysis. J. Regional Sci. 30 (2), 185–207.
Baiocchi, G., Minx, J., Hubacek, K., 2010. The impact of social factors and consumer behavior on carbon dioxide emissions in the United Kingdom. J. Ind. Ecol. 14 (1), 50–72.
Baur, A., Thess, M., Kleinschmit, B., Creutzig, F., 2014. Urban climate change mitigation in Europe: looking at and beyond the role of population density. J. Urban Plan. Dev. 140 (1), 04013003.
Boardman, B., 2007. Examining the carbon agenda via the 40% house scenario. Build. Res. Inform. 35 (4), 363–378.
Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth, Belmont.
Büchs, M., Schnepf, S.V., 2013. Who emits most? Associations between socio-economic factors and UK households' home energy, transport, indirect and total CO_2 emissions. Ecol. Econ. 90, 114–123.
Chaudhuri, P., Huang, M., Loh, W., Yao, R., 1994. Piecewise-polynomial regression trees. Stat. Sin. 4, 143–167.
Clarke, W.A.V., 1991. Residential preferences and neighborhood racial segregation: a test of the Schelling segregation model. Demography 18 (1), 1–19.
Christaller, W., Baskin, C.W., 1966. Die Zentralen Orte in Süddeutschland (Central Places in Southern Germany); Translated by Carlisle W. Baskin. Prentice-Hall.
Clark, L.A., Pregibon, D., 1993. Tree-based models. In: Chambers, J.M., Hastie, T. (Eds.), Statistical Models in S. Chapman & Hall, London, (Chapter 9), pp. 377–419.
Creutzig, F., 2014. How fuel prices determine public transport infrastructure, modal shares and urban form. Urban Clim. 10, 63–76.
Creutzig, F., Mühlhoff, R., Römer, J., 2012. Decarbonizing urban transport in European cities: four cases show possibly high co-benefits. Environ. Res. Lett. 7 (4), 044042.
Creutzig, F., Baiocchi, G., Bierkandt, R., Pichler, P.-P., Seto, K.S., 2015. Global typology of urban energy use and potentials for an urbanization mitigation wedge. Proc. Natl. Acad. Sci. 112 (20), 6283–6288.
DECC, 2011. Department of Energy & Climate Change, http://www.decc.gov.uk/.
DEFRA, 2007. Guidelines to Defra's GHG conversion factors for company reporting: annexes updated June 2007, http://archive.defra.gov.uk/environment/business/reporting/pdf/conversion-factors.pdf.
Fragkias, M., Lobo, J., Strumsky, D., Seto, K.C., 2013. Does size matter? Scaling of CO_2 emissions and US urban areas. PLOS ONE 8 (6), e64727.
Frank, L.D., Pivo, G., 1994. Impacts of mixed use and density on utilization of three modes of travel: single occupant vehicle, transit, and walking. Transport. Res. Rec.: J. Transport. Res. Board 1466, 44–52.
Fujita, M., 1989. Urban Economic Theory: Land Use and City Size. Cambridge University Press.
Glaeser, E.L., Kahn, M.E., 2010. The greenness of cities: carbon dioxide emissions and urban development. J. Urban Econ. 67 (3), 404–418.
Grubler, A., Bai, X., Buettner, T., Dhakal, S., Fisk, D.J., Ichinose, T., Keirstead, J.E., Sammer, G., Satterthwaite, D., Schulz, N.B., Shah, N., Steinberger, J., Weisz, H., 2012. Urban energy systemsIn: Global Energy Assessment – Toward A Sustainable Future, pp. 1307–1400 (Chapter 18).
Gujarati, D.N., 2004. Basic Econometrics, 4th ed. McGraw-Hill.
Gurney, K.R., Razlivanov, I., Song, Y., Zhou, Y., Benes, B., Abdul-Massih, M., 2012. Quantification of fossil fuel CO_2 emissions on the building/street scale for a large U.S. city. Environ. Sci. Technol. 46 (21), 12194–12202.
Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64 (2), 413–430.
Hansen, B.E., 2000. Sample splitting and threshold estimation. Econometrica 68 (3), 575–604.
Harris, R., Sleight, P., Webber, R., 2005. Geodemographics, GIS and Neighbourhood Targeting. John Wiley & Sons, Chichester.
Hillman, T., Ramaswami, A., 2011. Greenhouse gas emission footprints and energy use benchmarks for eight U.S. cities. Environ. Sci. Technol. 44 (6), 1902–1910.
Jones, C., Kammen, D.M., 2014. Spatial distribution of U.S. household carbon footprints reveals suburbanization undermines greenhouse gas benefits of urban population density. Environ. Sci. Technol. 48 (2), 895–902.
Karathodorou, N., Graham, D.J., Noland, R.B., 2010. Estimating the effect of urban density on fuel demand. Energy Econ. 32 (1), 86–92.
Kennedy, C., Steinberger, J., Gasson, B., Hansen, Y., Hillman, T., Havránek, M., Pataki, D., Phdungsilp, A., Ramaswami, A., Mendez, G.V., 2009. Greenhouse gas emissions from global cities. Environ. Sci. Technol. 43 (19), 7297–7302.
Kim, H., Loh, W.-Y., Shih, Y.-S., Chaudhuri, P., 2007. Visualizable and interpretable regression models with good prediction power. IIE Trans. 39 (6) .
Lenzen, M., Dey, C., Foran, B., 2004. Energy requirements of Sydney households. Ecol. Econ. 49 (3), 375–399.
Loh, W.-Y., 2002. Regression trees with unbiased variable selection and interaction detection. Stat. Sin. 12, 361–386.
Loh, W.-Y., 2009. Improving the precision of classification trees. Ann. Appl. Stat. 3 (4), 1710–1737.
Longley, P.A., 2012. Geodemographics and the practices of geographic information science. Int. J. Geogr. Inform. Sci. 26 (12), 2227–2237.
Louf, R., Barthelemy, M., 2014. How congestion shapes cities: from mobility patterns to scaling. Sci. Rep. 4 .
Makido, Y., Dhakal, S., Yamagata, Y., 2012. Relationship between urban form and CO_2 emissions: Evidence from fifty Japanese cities. Urban Clim. 2 (0), 55–67.
Met Office, 2011. Met Office: UKCP09: Gridded Observation Data Sets, http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/.
Min, J., Hausfather, Z., Lin, Q.F., 2010. A high-resolution statistical model of residential energy end use characteristics for the united states. J. Ind. Ecol. 14 (5), 791–807.
Minx, J., Baiocchi, G., Wiedmann, T., Barrett, J., Creutzig, F., Feng, K., Förster, M., Pichler, P.-P., Weisz, H., Hubacek, K., 2013. Carbon footprints of cities and other human settlements in the UK. Environ. Res. Lett. 8 (3), 035039.
Minx, J.C., Creutzig, F., Medinger, V., Ziegler, T., Owen, A., Baiocchi, G., 2011. Developing a pragmatic approach to assess urban metabolism in Europe. A report to the European Environment Agency. ClimatEcon Working Papers, 1.
Newman, P., Kenworthy, J.R., 1989. Cities and Automobile Dependence. Gower, Aldersh.
Newman, P.W., Kenworthy, J.R., 2006. Urban design to reduce automobile dependence. Opolis 2 (1), 35–52.
ONS, 2001. National statistics 2001 area classification of super output areas and data zones: distance from centroids, http://data.gov.uk/dataset/national_statistics_2001_area_classification_of_super_output_areas_and_data_zones_-_distance_from_ce.
ONS, 2011. Office for National Statistics: Neighbourhood Statistics, http://www.neighbourhood.statistics.gov.uk.
Parshall, L., Gurney, K., Hammer, S.A., Mendoza, D., Zhou, Y., Geethakumar, S., 2010. Modeling energy consumption and CO_2 emissions at the urban scale: Methodological challenges and insights from the United States. Energy Policy 38 (9), 4765–4782.
Rickwood, P., Glazebrook, G., Searle, G., 2008. Urban structure and Energy – a review. Urban Policy Res. 26 (1), 57–81.
Sassen, S., 2001. The Global City: New York, London, Tokyo. Princeton University Press.
Schelling, T.C., 1969. Models of segregation. Am. Econ. Rev. 59 (2), 488–493.
Seto, K.C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G.C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J.E., Mueller, D.B., Murakami, J., Nagendra, H., Ramaswami, A., 2014. Human settlements, infrastructure and spatial planning. In: Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., Kriemann, B., Savolainen, J., Schlömer, S., von Stechow, C., Zwickel, T., Minx, J.C. (Eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom/New York, NY, USA.
Tan, C.M., 2010. No one true path: uncovering the interplay between geography, institutions, and fractionalization in economic development. J. Appl. Econom. 25 (7), 1100–1127.
UNESCO-UNEP, 1983. Glossary of Environmental Education Terms. UNESCO, UNEP International Environmental Education Programme.
Viguie, V., Hallegatte, S., 2012. Trade-offs and synergies in urban climate policies. Nat. Clim. Change 2, 334–337.
Von Thünen, J.H., 1875. Der isolirte staat in beziehung auf landwirtschaft und nationalökonomie, vol. 1. Hempel & Parey, Wiegant.
Weisz, H., Steinberger, J.K., 2010. Reducing energy and material flows in cities. Curr. Opin. Environ. Sustain. 2 (3), 185–192.