Spatial dynamics of household energy consumption and local drivers in Randstad, Netherlands

Mashhoodi, Bardia

DOI
10.1016/j.apgeog.2018.01.003

Publication date
2018

Document Version
Accepted author manuscript

Published in
Applied Geography

Citation (APA)
Mashhoodi, B. (2018). Spatial dynamics of household energy consumption and local drivers in Randstad, Netherlands. *Applied Geography*, *91*, 123-130. https://doi.org/10.1016/j.apgeog.2018.01.003

Spatial Dynamics of Household Energy Consumption and Local Drivers in Randstad, Netherlands

Abstract

This study is an attempt to bridge an eminent knowledge gap in the empirical studies on Household Energy Consumption (HEC): the previous studies implicitly presumed that the relationships between HEC and the geographic drivers is uniform in different locations of a given study-area, and thus have tried to disclose such everywhere-true relationships. However, the possible spatially varying relationships between the two remain unexplored. By studying the performance of a conventional OLS model and a GWR model -adjusted R^2, randomness of distribution of residual (tested by Moran's I), AIC and spatial stationary index of the geographic drivers, ANOVA test of residuals- this study demonstrates that the GWR model substantially provides a better understanding of HEC in the Randstad. In this respect, the core conclusion of this study is: the relationships between HEC and geographic drivers are spatially varying and therefore needed to be studied by means of geographically weighted models. Additionally, this study shows that considering spatially varying relationships between HEC and geographic drivers, by application of hierarchical clustering, the areas of the Randstad can be classified in four clusters: building age and income impact areas, building density impact areas, population density and built-up impact areas, household size and income impact areas.

Highlights

- The geographic drivers of household energy consumption are spatially varying
- Household energy consumption has to be studied by geographically weighted models
- Policies regarding household energy consumption need to be location-specific

Key words

Household Energy Consumption, Geographically weighted regression, Randstad, Netherlands

Spatial Dynamics of Household Energy Consumption and Local Drivers in Randstad, Netherlands

1. Introduction

Curbing level of energy consumption has been matter of policy makers' interest since 1970s subsequent to geopolitical turmoil in 1973 and 1979. The interest has been widened into the environmental impact of energy consumption, particularly greenhouses gases (GHG) emission and global warming, following United Nations Framework Convention on Climate Change (UNFCC) in 1992, and preparation of Kyoto treaty in 1997, and United Nations Climate Change Conference held in Paris, 2015. However, despite the effort spend on international treaties, between 1990 and 2012, final energy consumption in EEA countries (the European Economic Area) increased by 6.5% (European Environment Agency, 2015a). In EU-15 countries between 1990 and 2011, the GHG emission decreased for 14.9% (European Environment Agency, 2013), which is still short of the target set by 2020 climate & energy package: 20% cut from 1990 level (Climate Action 2020 European commission, 2009). The share of Households energy consumption (HEC) in total energy use is substantial. In EU-27 countries in 2010, HEC accounts for some 27% of the total final energy consumption (European Environment Agency, 2015b) and creates 25% of GHG emissions (European Environment Agency, 2012). In the Netherlands, in order to reduce HEC, Third National Energy Efficiency Action Plan for the Netherlands (Ministry of Economic Affairs, 2014) introduces set of incentives and regulations, applicable for all the locations of the country, which mainly aim for improving quality of buildings e.g. low interest loans for building insulation, low-interest loans for building renovation, stricter energy standards for new construction, and compulsory measures to ensure efficiency of buildings' heating and ventilation appliances.

Many previous studies explored the impact of variety of geographic drivers on the HEC. Plenty of the previous studies have established links between level of the income of the inhabitants and the level of HEC (for instance Yun & Steemers, 2011; Druckman & Jackson, 2008; Joyeux & Ripple, 2007). Several previous studies found associations between family type and HEC, mainly concluding that consumption per head drops as the size of family grow (for instance Fong et al., 2007; Lenzen et al., 2006; Tso & Yau, 2003). The age of the inhabitants is also introduced as one of the significant drivers of HEC, particularly the portion of children and senior citizens from total population (Yun & Steemers, 2011; York, 2007; Yust et al., 2002). Moreover, the higher percentage of economically inactive inhabitants –for instance inhabitants with disability or retired- has been seen as sources of higher HEC (for instance Fong et al., 2007). The HEC of the inhabitants of different housing tenure also found to be meaningfully different due to varying level of investment in insulation and different methods of payment for energy cost (for instance Druckman & Jackson, 2008; Tso & Yau, 2003; Aydinalp et al., 2004). Several studies highlighted significant variation of HEC between different types of dwellings, for instance between single-family and multi-family houses, and also between dwellings of different age (for instance Yun & Steemers, 2011; Druckman & Jackson, 2008; Aydinalp et al., 2004). Moreover, land-cover has been found to be effective on HEC due to its links with formation of urban heat islands (for instance Madlener & Sunak, 2011; Georgakis & Santamouris, 2006; Hui, 2001). Wind intensity is found to impact HEC by affecting the thermal exchange between buildings and outside space by affecting infiltration and exfiltration of the buildings (for instance Sanaiean et al., 2014; van Moeseke et al., 2005). Ewing and Rong (2008) suggest that higher building density could decrease the energy used for heating, and increase that for cooling. Several studies suggest that the surface-to-volume ratio of the building affects the heat loss of buildings and HEC (for instance Steemers and Yun, 2009; Druckman and Jackson, 2008; Lenzen et al., 2006). Population density is also considered as an effective determinant of HEC (for instance York, 2007; Lenzen et al., 2006).

A knowledge gap is eminent in the current body of literature on HEC: all of previous studies implicitly presumed that geographic drivers have an unvarying impact on HEC across a given area, and therefore attempted to disclose such everywhere-true impacts. Consequently, the policies-recommendation brought forward by previous study are uniform and generic for all areas in question instead of location-specific and spatially varying. The core objective of this research is to tackle such knowledge gap chasing answers to the following questions: (a) Are the relationships between HEC and the geographic drivers spatially varying across the areas of the Randstad region, the Netherlands? (b) If yes, how such relationships differ across the areas of the Randstad region?

To do so, this study aim to conduct geographically weighted regression (GWR) for studying HEC. The method has been successfully deployed in several geographic studies of different disciplines such as afforestation (Clement et al., 2009), regional wealth and land cover (Ogneva-Himmelberger et al., 2009), urban landscape fragmentation (Gao & Li, 2011), agriculture and urbanization (Su et al., 2012), land use and water quality (Tu, 2011), residential land price (Hu et al., 2016), late-stage prostate cancer diagnosis (Goovaerts et al., 2015), urban heat island (Ivajnšič et al., 2014), and fire density (Oliveira et al., 2014). However, surprisingly, HEC studies are lagging behind in application of GWR. To bridge this gap, this study investigates the location-specific effect of variety of socioeconomic, housing, urban morphology, solar radiation and wind-intensity related indicators on HEC in the neighborhoods of the Randstad region, the Netherlands.

2. Material and Methods

2.1 Case study

The study-area is consisted of *buurten*, a spatial division defined by the Dutch central bureau of statistics (CBS), roughly could be translated as neighborhoods, in the Randstad region in 2013 (account for 2413 neighborhoods). The Randstad is a conglomeration of highly urbanized areas located in the south west of the Netherlands comprising the four major Dutch cities of Amsterdam, Rotterdam, the Hague and Utrecht, as well as the relatively less urbanized areas between them – the so-called "green heart". In order to avoid the boundary-effect problem in GWR models, we also defined "analysis areas" which is consist of the study-area plus a 20 km buffer around it (3514 neighborhoods in total). All the calculations are conducted on the analysis area, however at the end only the results obtained for areas within the study-areas are reported (Figure 1).

2.2 Data collection and processing

2.2.1 Dependent variable

The dependent variable of this study is average annual energy expenditure per head within the dwellings on gas and electricity, in 2013 (Figure 1). The data on consumption of gas and electricity are extracted from *wijk-en-buurtkaart 2013* (Centraal Bureau voor de Statistiek, 2013). As the available data does not indicate the neighborhoods with solar energy supply or district heating, the abnormal values of gas and electricity use needed to be filtered out thus univariate outliers of gas and electricity use (incidents with z-value <= -2.5 or z-value >= +2.5) are identified as outlier and excluded. The average cost of gas and electricity for domestic consumption in 2013 in Netherlands, is taken from Eurostat (Eurostat, 2015).

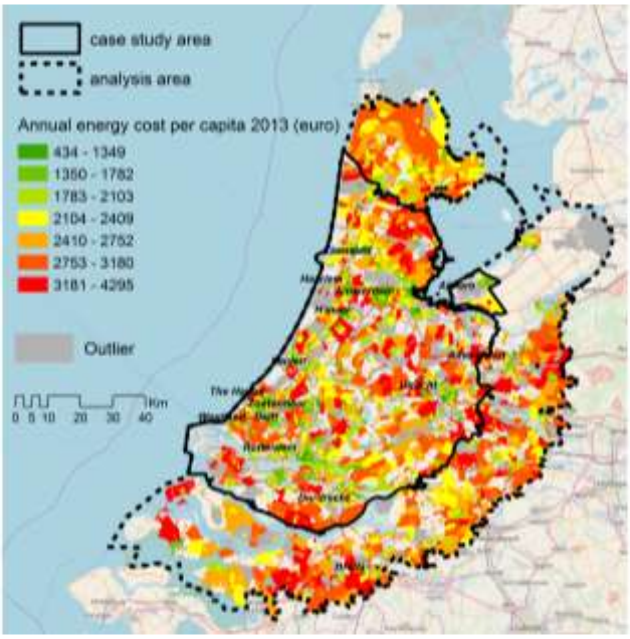

Figure 1: Annual energy expenditure per capita (dependent variable of this study), the study-area and the analysis area.

2.2.2 Independent variables

This study is conducted on 21 independent variables (Table 1). The first two variables indicate the portion of the population aged 14 or younger and aged 65 or older. One variable show population density per square kilometer. One variables specify the household structure by demonstrating average household size. Three variables show economic status of the residents: average annual disposable income per head (in euros), Percentage of population aged 15–64 receiving disability benefits, and Percentage of population aged 15–64 receiving unemployment benefits. Four variables are deployed in order to describe the status of housing tenure in the areas: Property-value (WOZ in Dutch), shows the average value of residential real estate in the areas; percentage of housing tenure owned by public associations (not necessarily social housing); median age of residential buildings; and percentage of residential floor area constructed after the introduction of building energy-efficiency standards in 1988. Land-cover of the areas is further explained by means of two variables including the portion of built-up areas, semi built-up areas and portion of green land covers (consisted of recreational, agricultural and natural areas).

The status of urban morphology (properties related to geometrical distribution of the building masses within space) is described using five variables: floor area ratio (FAR); building coverage ratio (BCR); buildings' surface to volume ratio; frontal area index (λ_f) - the ratio of total building walls facing wind flow to neighborhoods' total area; and rugosity, the variation of building height which, adopted from Adolph (2001), is calculated as the standard deviation of height values of Digitally Elevated Model (DEM) of the neighborhoods. As a proxy for wind speed, aerodynamic roughness length (ARL), the height in which the effective wind speed is theoretically zero, is used. Higher values of ARL correspond with lower wind intensity (Landsberg, 1981). The morphometric model introduced by Macdonald et al. (1998), one of the most comprehensive models according to a review by Grimmond and Oke (1999), is used:

$$\frac{Z_d}{Z_H} = 1 + \alpha^{-BCR}(BCR - 1) \quad \text{(equation 1)}$$

$$\frac{Z_0}{Z_H} = \left(1 - \frac{Z_d}{Z_H}\right) \exp\left(-\left(\frac{0.5\beta C_D \lambda_f}{k^2}\left(1 - \frac{Z_d}{Z_H}\right)\right)^{-0.5}\right) \quad \text{(equation 2)}$$

Where Z_0 is aerodynamic roughness length for momentum, Z_d is zero-plane displacement height, Z_H is height of roughness element (m), BCR is building coverage ratio, λ_f frontal area index, $\alpha = 4.43$, $\beta = 1.0$, $k = 0.4$, and $CD \cong 1$.

Deploying the Arcgis 10.2 solar radiation toolbox, status of solar radiation is described by two variables: solar radiation per square meters of neighborhoods' surface (solar radiation on neighborhood (WH/m2)) and per cubic meters of the buildings (solar radiation per building volume (WH/m3)). Each of the values show the average solar radiation on the longest (21 June) and shortest (21 December) day of 2013.

The data on the first socioeconomic are provided by *wijk-en-buurtkaart 2013* (Centraal Bureau voor de Statistiek, 2013). The data on land-cover are extracted from *Bodemgebruik* database. 2012(Bodemgebruik, 2012). The DEM used to prepare the urban morphology and wind and solar variables, is prepared based on the building height database in the Netherlands, the so-called as 3D BAG (Esri Netherlands, 2016).

2.2.3 Factor analysis of the independent variables

To avoid the potential misleading results caused by multicollinearity between the 21 independent variables, factor analysis, with extraction method of principal component analysis and rotation method of Oblimin with Kaiser Normalization, is deployed. As result, the effect of the variables is compressed in five factors (Table 1). The five factors account for almost 75% of the total variance of the variables. The first factor, *FAC1 Population density & built-up areas*, is positively loaded onto built up coverage (%), BCR, λ_f, population density and FAR, and negatively on green-coverage (%). The second component, *FAC2 Income & private tenure*, is positively loaded onto income per capita and property value, and negatively loaded onto disability (%), unemployment (%) and public rental (%). *FAC3 Household size & population younger than 14*, is positively loaded onto population ages 0–14 (%) and household-size, and negatively loaded onto population ages 65+ (%). *FAC4 Building age*, is positively loaded onto building median age, and negatively onto floor area after 1988 (%). *FAC5 Building density*, is and positively onto FAR, rugosity and ARL and negatively onto solar radiation per building volume (WH/m3) and solar radiation on neighborhood (WH/m2).

Variables	Factors				
	FAC1 population density & built-up areas	FAC2 Income & private tenure	FAC3 Household size & population younger than 14	FAC4 Building age	FAC5 Building density
built-up coverage (%)	**,977**	-,089	-,091	-,177	-,067
building coverage ratio (BCR)	**,905**	,075	,005	,177	-,005
green-coverage (%)	**-,891**	,086	,075	,216	-,065
frontal area index	**,750**	,021	,064	,201	,291
population-density	**,621**	-,165	,231	,125	,270
income per capita	,126	**,892**	-,304	-,113	,121
public-rent (%)	,050	**-,780**	-,070	-,047	,183
property-value	-,276	**,739**	-,058	,020	-,085
disability (%)	-,147	**-,631**	-,266	-,024	,088
unemployment (%)	,221	**-,481**	-,056	-,040	-,014
population ages 65+(%)	,019	,037	**-,891**	-,067	-,064
population ages 0-14 (%)	-,020	,002	**,748**	-,343	-,125
household-size	-,167	,218	**,478**	-,338	-,380
building median age	-,061	,110	,046	**,855**	,119
floor area after 1988 (%)	-,013	,205	,283	**-,674**	,267
solar radiation per building volume	,028	,089	-,055	,002	**-,919**
rugosity	,288	-,021	,026	,139	**,751**
solar radiation on neighbourhood	-,260	-,031	-,066	-,273	**-,741**
aerodynamic roughness length (ARL)	,175	-,168	-,001	-,143	**,721**
floor area ratio (FAR)	**,484**	,099	,067	,306	**,532**
Buildings' surface to volume ratio	,067	-,005	,191	,138	-,379

Table 1: Independent variables of the study and pattern matrix showing the loading of factors on independent variables. Coefficients with absolute value greater than 0,400 are marked bold.

2.3 Geographically weighted regression

The first session of the method is consisted of a conventional linear regression model, (see equation 3), which assess the generalizable influence of geographic drivers on HEC:

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \quad \text{(equation 3)}$$

Where y_i represent the estimated value of HEC in the location i, β_0 show the intercept of the estimation, β_k denote the coefficient slope of the factor k, x_{ik} represents its value of factor in location i

. ε_i accounts for the random error term in location i. The second session, GWR model, (see equation 4), is deployed on the same dataset:

$$y_i = \beta_0(\mu_i, \nu_i) + \sum_k \beta_k(\mu_i, \nu_i) x_{ik} + \varepsilon_i \qquad \text{(equation 4)}$$

Where (μ_i, ν_i) express the geographic coordination of location i. $\beta_k(\mu_i, \nu_i)$ and $\beta_0(\mu_i, \nu_i)$ are the local coefficient and intercept of factor *k* estimated specific to location i. The local estimates are obtained by weighting the instances around location i (equation 5):

$$\hat{\beta}(\mu, \nu) = (X^T W(\mu, \nu) X)^{-1} X^T W(\mu, \nu) y \qquad \text{(equation 5)}$$

Where $\hat{\beta}(\mu, \nu)$ denote the unbiased estimate of β, $W(\mu, \nu)$ is weighting matrix obtained by means of adaptive Gaussian function (equation 6):

$$\begin{aligned} W_{ij} &= \exp(-d_{ij}^2 / \theta_{i(k)}^2), \text{ if } d_{ij} < \theta_{i(k)} \\ W_{ij} &= 0, \qquad \text{otherwise} \end{aligned} \qquad \text{(equation 6)}$$

Where W_{ij} denote the weight of instance observed at location *j* for estimating the coefficient at location i, d_{ij} is the bird-fly metric distance between i and j, and $\theta_{i(k)}$ is an adaptive bandwidth defined as the distance from the *k*th nearest neighbor distance. In this study, using ArcGIS (version 10.2), the bandwidth is specified as 108 neighbors, in order to minimize the Akaike Information Criterion (AIC) of the GWR model.

The performance of the OLS and GWR model are compared by means of five test: improvement of adjusted R^2; reduction of AICc (for at least three points as previously established by other authors e.g. Hu et al., 2016; Gao & Li, 2011); the randomness of the spatial distribution of the residual of the two models (assessed by Moran's I); ANOVA test of improvement of residual in GWR model; and spatial stationary index - the ratio of interquartile ranges of the standard error of coefficients in GWR model to twice of standard error of the coefficients in OLS model (Charlton et al, 2003).

At the last session, is cluster analysis of GWR results. The advantages of GWR models is provision of an extensive number of local coefficients. However, such an advantage is also a challenge where the summarization and interpretation of the results for the end users –e.g. policy makers- could be challenging (Mennis, 2013, Matthews & Yang, 2012). In this respect, in order to summarize the results of GWR in an interpretable format, *hierarchical clustering technique*, with Ward's method and squared Euclidean distance, on the local standardized coefficients of GWR model is conducted (insignificant coefficients are considered equal to zero). The study areas are subsequently clustered into two, three and four groups (see dendrogram in Fig 4a). The clusters are compared by one-way ANOVA test of the local coefficients and named after the effects which differentiate them the most from one another.

3. Results

3.1. Comparison between performance of OLS and GWR models

Comparison between adjusted R^2 of the two models (see Table 2) show some 10% improvement of the estimation by deploying geographically weighted model (0.796 in GWR model compare to 0.691 of OLS). The spatial variation of the adjusted R^2 is demonstrated in **Fig. 2**. The range of the local adjusted R^2 (0.48 to 0.91) show that the goodness-of-fit of the estimation of some 76% of the studied areas is higher than that of OLS model. The geographic pattern of the values show a concentration of higher values of R^2 around The Hague, Haarlemmermeer, Amsterdam west and Zoetermeer, Utrecht west and Barendrecht. in contrary, the goodness-of-fit in the areas of central Rotterdam, central Utrecht, Leiden and Dordrecht are the lowest values within Randstad. Presumably, the latent variables affecting HEC, such as detailed information on dwellings quality as well as individual habits, have a stronger impact in these areas.

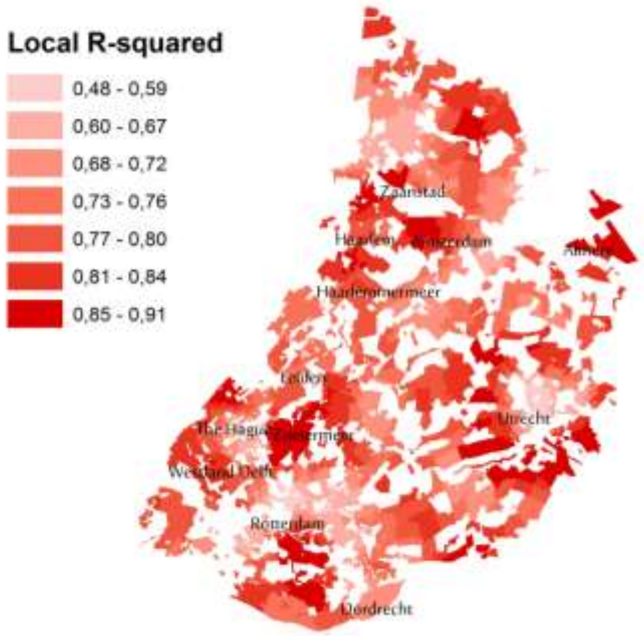

Figure 2 : Local adjusted R-squared of GWR estimation of HEC in the Randstad

AICc of the GWR model is substantially smaller than that of OLS (4780 of GWR compare to 5882 in case of OLS model), indicating remarkable better performance in this respect. In case of this study, Moran's I of the GWR model is substantially closer to 0, implying higher randomness of distribution of its residual compare to that of OLS model , -0.008 in case of GWR compare to 0.272 of OLS. As all the spatial stationary indices are greater than one, the results demonstrate that the impact of all the geographic factors are spatial non-stationary and therefore need to be locally studied. (Table 2).

Variable	GWR results					OLS results
	β Mean	β Min	β Max	β SD	Stationary index	β
Intercept	-0,004	-0,537	0,507	0,199		0,000*
FAC1 population density & built-up areas	-0,178	-0,729	0,215	0,118	1,125	-0,199*
FAC2 Income & private tenure	0,459	-0,065	0,793	0,123	1,772	0,420*
FAC3 Household size & population younger than 14	-0,453	-0,848	0,062	0,124	1,142	-0,482*
FAC4 Building age	0,432	-0,141	0,861	0,143	2,263	0,361*
FAC5 Building density	-0,261	-1,069	0,247	0,183	2,622	-0,321*
R-squared	0,830					0,692*
Adjusted R-squared	0,796					0,691*
AICc	4780,15					5852,00
Residuals Moran's I	-0,0078					0,2715
Neighbours	108,000					

β: standardized regression coefficient

* *p-value* <0,05

Table 2. Estimated parameters and diagnostic statistics in the OLS and GWR models.

The ANOVA test of the residuals of GWR and OLS model show the significant improvement in case of the former (Table 3).

	Df	Sum Sq	Mean Sq	F value
OLS Residuals	6.000	1083.53		
GWR Improvement	92.037	272.63	296.213	
GWR Residuals	3.415.963	810.90	0.23739	12.478*

**p-value* < 0,001

Table 3. ANOVA test of residuals of GWR and OLS models

Local coefficient of the *FAC1 population density & built-up areas* ranges from -0.729 to 0.215 where the global coefficient of the factor, obtained from OLS model, is -0,199 (Table 2). Study of the significance level of the local coefficient at *p<0.05* level reveals that merely some 58% of the local coefficients of the *FAC1* are significant. Almost all of the significant local coefficients are negative. In other words, in almost three fifth of the areas the higher values of the factor are associated with lower levels of HEC. The highest negative elasticity between *FAC1 population density & built-up areas* and HEC is observed in some areas of city of Utrecht. Some dispersed pockets of high negative elasticity are also identified in the so-called green heart areas (Fig 3a).

Local coefficients of *FAC2 Income & private tenure* range from -0.065 to 0.793 compare to 0.420 of the global model (Table 2). Some 99% of the local coefficients are found significant at the *p<0.05* level, which are all positively associated with HEC. The elasticity between *FAC2 Income & private tenure* and HEC reaches its maximum in Haarlemmermeer and Harlem. The magnitude of the positive elasticity roughly resembles in case of Amsterdam, Utrecht and The Hague. Whereas, in case of Rotterdam either the coefficient estimate is not significant or the its magnitude is marginal (Fig 3b).

Although local coefficients of the *FAC3 Household size & population younger than 14* range from -0.848 to 0.062 (compare to -0.482 in global model), however all of the significant coefficients, account for some 97% of the areas, are positive. Relatively high elasticity between *FAC3 Household size & population younger than 14* and HEC is estimated in case of city center of Amsterdam and Leidn. No significant elasticity between the factor and HEC is estimated in city the centers of Utrecht. Though the estimated coefficient in case of Rotterdam is significant, however the magnitude is relatively modest (Fig 3c).

Local coefficient of the *FAC4 Building age* ranges from -0.141 to 0.861 (compare to 0.361 in global model). Some 95% of the estimated coefficients values are significant (at *p-value*< 0.05 level) which all are all positive. The largest elasticity between HEC and *FAC4 Building age* is estimated in some areas of the so-called green heart particularly in vicinity of Zoetermeer. *FAC4 Building age* is estimated to substantially increase level of HEC in vicinity of Zandaan and Dordrecht (Fig 3d).

In case of local coefficient of the *FAC5 Building density*, although the values are ranged from -1.069 to 0.247 (compare to -0.321 of the OLS model), however almost all the significant coefficients, observed in some 62% of the study areas, are negative. The concentration of the high values of estimated coefficient is central areas of Utrecht and Rotterdam. Also, high elasticity are estimated for areas north of Amsterdam and around port of Rotterdam (Fig 3e).

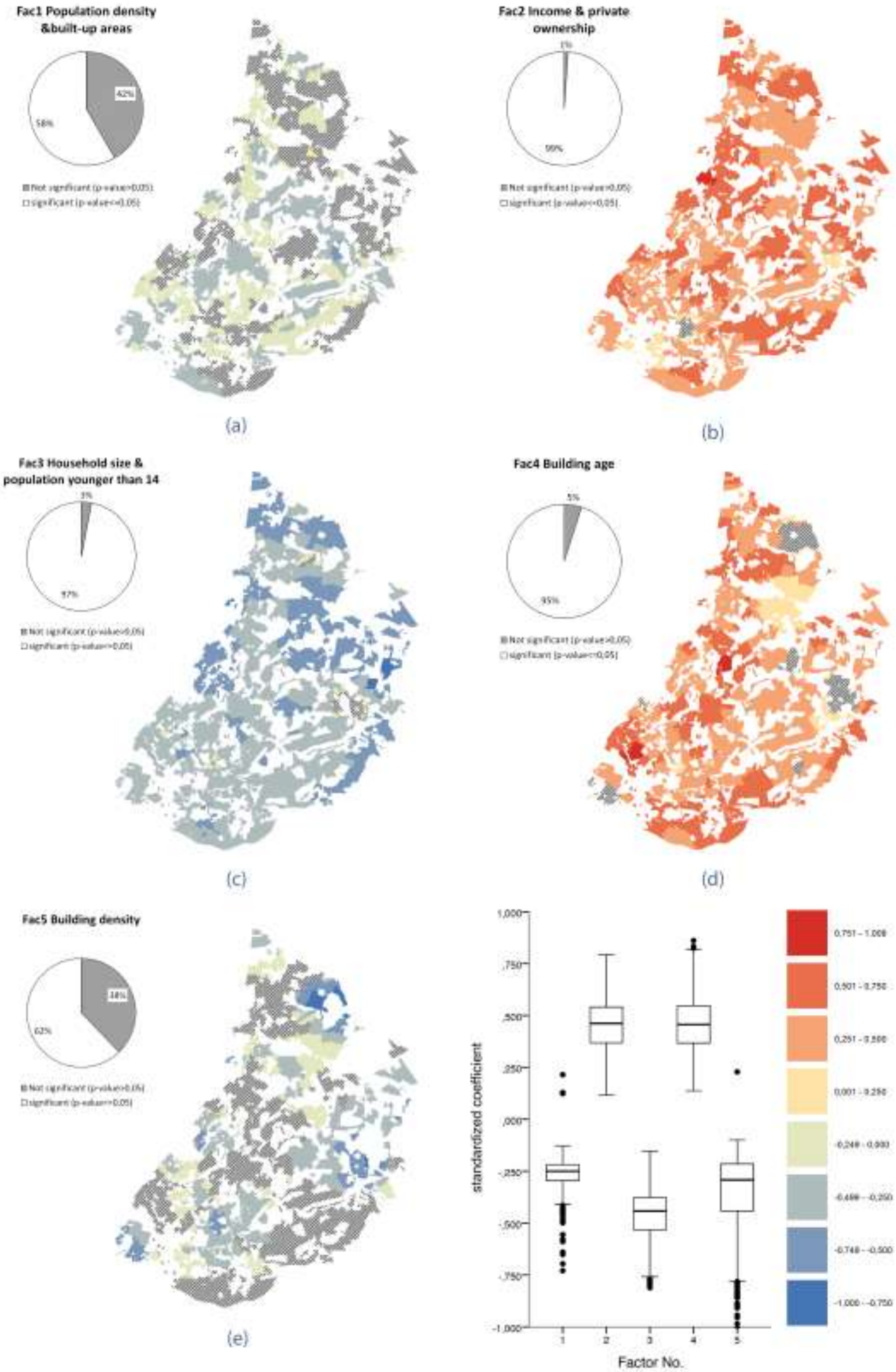

Figure 3: Local standardized coefficient of the independent factors and their level of significance. The box plot illustrates the variability of the significant coefficients.

3.2. The typologies of local geographic impacts on HEC

As result of hierarchical clustering of the local standardized coefficients of the independent factors, insignificant coefficients are considered to be equal to zero, four clusters are identified. ANOVA table show that all the clusters are significantly differentiated based on the mean value of local standardized coefficients (Fig 4).

The first impact-type, differentiated at the first stage of clustering (see dendrogram at Fig 4a) labelled "Cluster1 building age and income", accounts for some 39,9% of the studied-areas. The areas of the type are differentiated from those of the other impact-types according to substantial positive coefficients of *FAC4 Building age* and *FAC2 Income & private tenure*. The impact of *FAC1 population density & built-up areas* and *FAC3 Household size & population younger than 14* are roughly at the average level of local coefficients in the Randstad. The impact of *FAC5 Building density* on HEC in the areas of this type is marginal (Fig 4c).

The areas of the second impact-type, differentiated in the second stage of clustering, account for 11,1% of the areas, are identified as "Cluster2 building density" as *FAC5 Building density* show the largest negative coefficient value. The impact of *FAC1 population density & built-up areas* is roughly at the average level of local coefficients in the Randstad. That of *FAC2 Income & private tenure*, *FAC3 Household size & population younger than 14* and *FAC4 Building age* are lower than other clusters.

Two clusters are identified in the third stage of clustering. In the areas of the third impact-type, labelled as "Cluster3 population density and built-up area", accounting for 23% of the study areas, merely one factors have remarkable impact on HEC: *FAC1 population density & built-up areas*. Whereas, the impact of other factors is almost at the average of the Randstad areas.

The fourth impact-type, account for 26% of the areas, is identified as "Cluster4 household size and income" are differentiated by substantial impact of two factors: *FAC2 Income & private tenure*, *FAC3 Household size & population younger than 14*. The impact of *FAC1 population density & built-up areas* on the areas of this cluster is almost zero, and that of *FAC4 Building age* and *FAC5 Building density* stands at average level.

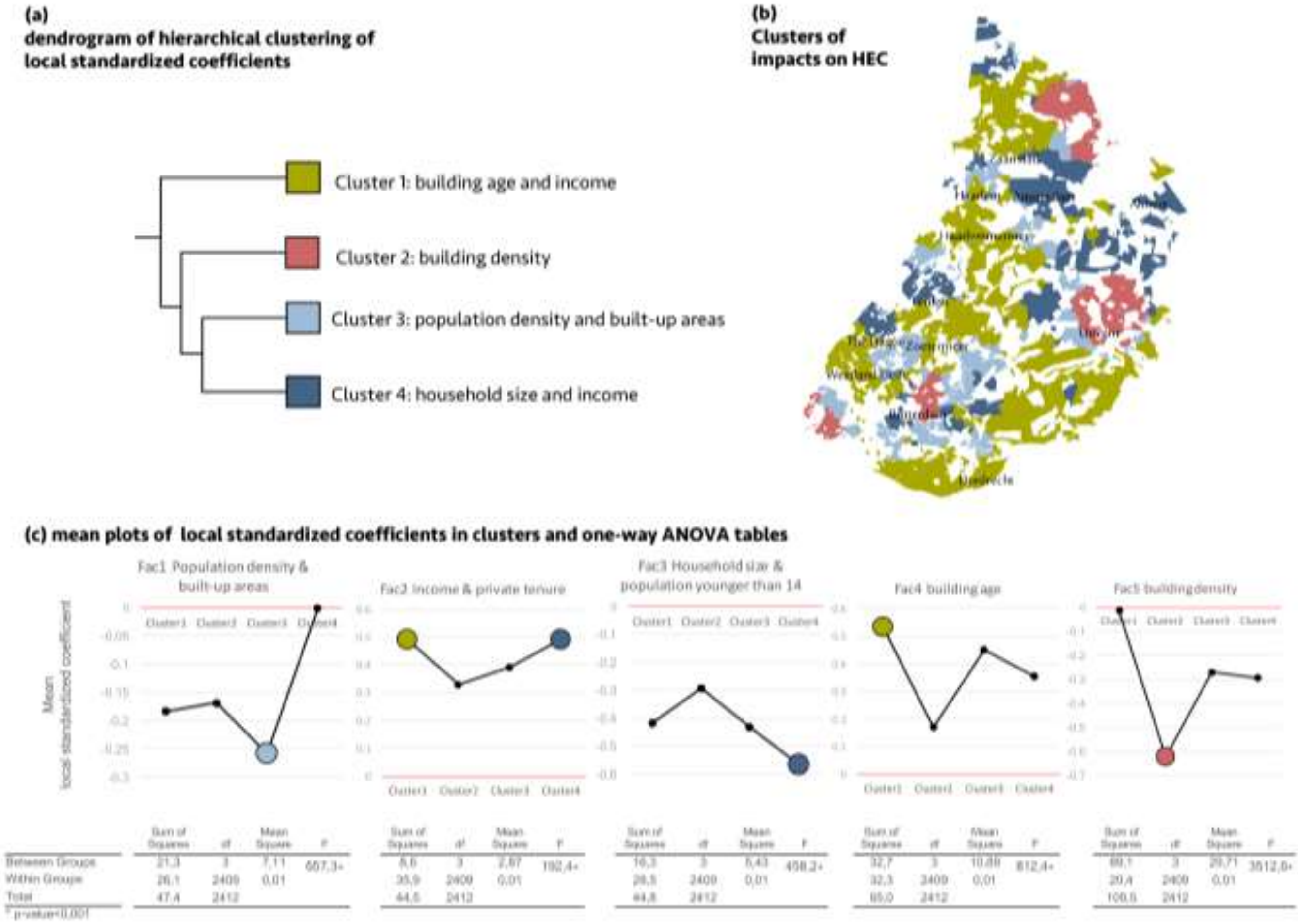

Figure 4: Four types of impact of geographic drivers on HEC obtained by hierarchical clustering of local standardized coefficients

4. Discussion

The core aspect of this exploration was whether the impacts of geographic drivers on HEC are spatially non-stationary or not, and whether GWR models provide a better understanding of HEC rather than conventional OLS. As illustrated by the comparison between conventional OLS model and GWR model on HEC, the latter model significantly improves our understanding of HEC's drivers in different aspects: goodness-of-fit of estimate is some 10% higher (measured by R^2); AIC is substantially lower; and the residual of the model is smaller and more randomly distributed (tested by means of ANOVA and Moran's I test on residual). In addition, verified by spatial stationary index, it is demonstrated that the impacts of all the geographic factors on HEC vary over the study areas.

Considering the second research question, how the impacts of geographic drivers on HEC differ across the urban areas of the Randstad urban region, subsequent to application of GWR model, four types of impacts on HEC are identified: building age and income impact, building density impact, population density and built-up area impact, and household size and income impact. However, the output of GWR models is limited into discovering the associations and does not disclose the causal mechanisms. in this section, for sake of opening up new discussion and stimulating further studies, some speculations of the causal mechanisms are presented.

The first type of impact, called as “building age and income”, highlights the areas in which HEC is the most increased by building age, inhabitants’ income and property value. The neighborhoods of this cluster are mainly less urbanized areas of the Randstad. Presumably, considering the higher amount of free standing dwellings, the impact of quality of buildings on HEC is remarkably higher compare to other clusters. In a similar fashion, higher income and private tenure, which presumably is associated with larger dwelling size and possession of more appliances, has a substantial impact on increasing HEC. Observed positive elasticity of income shows that though the more affluent inhabitants can afford better maintenance and insolation for their dwellings, however, due to different life style, ultimately their energy consumption outnumber that of those with lower income.

The second type of impact, labelled as “building density”, is mainly identified by remarkable impact of high FAR and low solar radiation and wind intensity (associated with high values of ARL) on decreasing HEC. The areas of this cluster are mainly located in Rotterdam and Utrecht. Presumably the remarkable impact of these indicators in these cities is related to higher variability of building density compare to rest of the neighborhoods. One possible reason for impact of FAR on decreasing HEC is compactness of dwellings and higher heat exchange between them. FAR could be also associated with formation of urban-heat-islands (UHI) which can result in higher air temperature and thus decrease HEC (similar to conclusions drawn by Ewing and Rong, 2008). The association between lower HEC and lower solar radiation and wind intensity could be due to two causal mechanisms. First, higher solar radiation presumably raises electricity use for cooling and ventilation in warm and sultry months, whereas it is supposedly not intense enough to decrease the amount of energy used for warming in cold seasons. Second, presumably high wind intensity increases thermal loss of the buildings due to higher levels of infiltration and exfiltration – which can raise gas use (Sanaiean et al., 2014, van Moeseke et al., 2005). Apparently, such energy loss offsets the thrift gained by better ventilation in windy areas.

The third impact-type, labelled as “population density and built-up areas”, highlights the areas in which HEC is the most affected by population density and presence of built-up areas. The areas of this cluster are mainly located in the fringes of the big cities of the Randstad. Such areas could vastly vary in population density as they include different types of developments ranged from populated modernist developments (as Zoetermeer) to suburban areas with villas (as Vrijenburg located in North of Barendrecht). Higher population density in the fringe areas is presumably associated with more vital urban environment which, according to a study by Heinonen et al. (2013), could increase participation of residents in outdoor activities and thus reduce amount of time spent at dwellings and HEC.

The last impact-type, labeled as “household size and income”, point out the areas in which HEC is remarkably affected by presence of larger households with children and adolescences (negative coefficient) as well as higher income of the residents (positive coefficient). The areas of this cluster are mainly located in highly urbanized areas of Amsterdam, The Hague, Leiden and Almeer. Decrease in level of HEC in response to presence of larger households is supposedly due to economies of scale (similar to the conclusion drawn by O'Neill and Chen, 2002). Presumably, the remarkable impact of household size and younger age groups on HEC is due to distinguished life style of such families from that of retired citizens living in small households.

5. Conclusion

HEC has been a hot topic in the policy-making and scholar circles in the last decades. However, one knowledge gap in the existing body of literature on HEC is eminent: all the previous studies implicitly presumed that the influence of geographic drivers on HEC resemble across the study areas. Therefore, deploying conventional statistical method, merely the average global impact of geographic drivers on HEC has been estimated, where location specific relations has remained unexplored. The main conclusion of this study is: HEC is vastly affected by location specific impacts and thus understanding of such impacts is essential for enhancing further understanding of HEC.

This result of this study has also two policy implication. Policies aimed at the reduction of HEC in the Netherlands, as like The Third National Energy Efficiency Action Plan (Ministry of Economic Affairs, 2014), follow two unwritten presumptions: First, that it is possible to formulate certain policies which are optimally suitable in all the locations of the country. Secondly, that the main way to reduce HEC is to improve energy efficiency of buildings. According to the result of this study both these presumptions could be revisited. First, it is established that the effects of socioeconomic, housing, land cover and morphological indicators on HEC are spatially variant. In this case, a certain set of policy guidelines would not fit the circumstances of all the areas and thus one-size-fits-all type policies need to be completed with location-specific strategies. By proposing location-specific strategies, decision makers could prioritize different incentives and obligations in different areas of the region. Secondly, the results show that the effect of energy efficiency of buildings on reduction of HEC is not necessarily the only effective determinant of HEC in all the areas. Thus, the policies need to add socioeconomic and morphological angles to their approach.

This study also has one major limitations: there are some latent variables which potentially affect HEC such as behavioral habits of the inhabitants or detailed data on building quality. Although obtaining such data on the scale of an urban region in size of the Randstad is practically impossible, however the potential “omitted variable bias” need to be acknowledged. Finally, further study on HEC could chase the possibilities for application of geographically weighted structural models, such as path analysis - which are typically used for studying HEC.

6. Acknowledgment

This study is part of DCSMART project founded in the framework of the joint programming initiative ERA-Net Smart Grids Plus, with support from the European Union’s Horizon 2020 research and innovation programme.

Bibliography

Adolphe, Luc., 2001. A simplified model of urban morphology: application to an analysis of the environmental performance of cities. Environment and planning B: planning and design, 28(2), pp.183-200.

Aydinalp, M., Ugursal, V.I. and Fung, A.S., 2004. Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks. Applied Energy, 79(2), pp.159-178.

Bodemgebruik, 2012. National Georegister, http://www.nationaalgeoregister.nl/geonetwork/srv/dut/search#|09c2856f-8541-44c0-9621-44d496f3990d / Accessed 28.06.2016

Centraal Bureau voor de Statistiek, 2013. Wijk- en buurtkaart 2013, https://www.cbs.nl/nl-nl/achtergrond/2014/08/wijk-en-buurtkaart-2013 / Accessed 28.06.2016

Charlton, M., Fotheringham, S. and Brunsdon, C., 2003. Software for geographically weighted regression. Spatial Analysis Research Group, Department of Geography, University of Newcastle.

Clement, F., Orange, D., Williams, M., Mulley, C. and Epprecht, M., 2009. Drivers of afforestation in Northern Vietnam: assessing local variations using geographically weighted regression. Applied Geography,29(4), pp.561-576.

Climate Action 2020 European commission, 2009. climate and energy package: http://ec.europa.eu/clima/policies/strategies/2020/index_en.htm / Accessed 28.06.2016

Druckman, A. and Jackson, T., 2008. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model.Energy Policy, 36(8), pp.3177-3192.

Esri Netherlands, 2016. 3D BAG, http://www.esri.nl/nl-NL/news/nieuws/sectoren/nieuw-in-arcgis-voor-leefomgeving , accessed 09/03/2017

European Environment Agency, 2012. End-user GHG emissions from energy, Reallocation of emissions from energy industries to end users 2005–2010, http://www.eea.europa.eu/publications/end-user-ghg-emissions-energy / Accessed 28.06.2016

European Environment Agency, 2013. Why did greenhouse gas emissions decrease in the EU in 2011?, http://www.eea.europa.eu/publications/european-union-greenhouse-gas-inventory-2013/why-did-greenhouse-gas-emissions / Accessed 28.06.2016

European Environment Agency, 2015a. Final energy consumption by sector and fuel, Over the period 1990 and 2012 final energy consumption in EU28 increased by 2, http://www.eea.europa.eu/data-and-maps/indicators/final-energy-consumption-by-sector-9 /Accessed 28.06.2016

European Environment Agency, 2015b. Final energy consumption by sector, Between 1990 and 2010, the final energy consumption in the EU-27 increased by 7, http://www.eea.europa.eu/data-and-maps/indicators/final-energy-consumption-by-sector-5/assessment Accessed 28.06.2016

Eurostat, 2015. Half-yearly electricity and gas prices, second half of year 2012–14 (EUR per kWh), http://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Half-yearly_electricity_and_gas_prices,_second_half_of_year,_2012%E2%80%9314_(EUR_per_kWh)_YB15.png&oldid=238282 /Accessed 28.06.2016

Ewing, R. and Rong, F., 2008. The impact of urban form on US residential energy use. Housing policy debate, 19(1), pp.1-30.

Fong, W.K., Matsumoto, H., Lun, Y.F. and Kimura, R., 2007. Influences of indirect lifestyle aspects and climate on household energy consumption. Journal of Asian Architecture and Building Engineering, 6(2), pp.395-402.

Gao, J. and Li, S., 2011. Detecting spatially non-stationary and scale-dependent relationships between urban landscape fragmentation and related factors using geographically weighted regression. Applied Geography, 31(1), pp.292-302.

Georgakis, C. and Santamouris, M., 2006. Experimental investigation of air flow and temperature distribution in deep urban canyons for natural ventilation purposes. Energy and Buildings, 38(4), pp.367-376.

Goovaerts, P., Xiao, H., Adunlin, G., Ali, A., Tan, F., Gwede, C.K. and Huang, Y., 2015. Geographically-weighted regression analysis of percentage of late-stage prostate cancer diagnosis in Florida. Applied Geography, 62, pp.191-200.

Grimmond, C.S.B. and Oke, T.R., 1999. Aerodynamic properties of urban areas derived from analysis of surface form. Journal of applied meteorology, 38(9), pp.1262-1292.

Heinonen, J., Jalas, M., Juntunen, J.K., Ala-Mantila, S. and Junnila, S., 2013. Situated lifestyles: I. How lifestyles change along with the level of urbanization and what the greenhouse gas implications are—a study of Finland. Environmental Research Letters, 8(2), p.025003

Hu, S., Yang, S., Li, W., Zhang, C. and Xu, F., 2016. Spatially non-stationary relationships between urban residential land price and impact factors in Wuhan city, China. Applied Geography,68, pp.48-56.

Hui, S.C., 2001. Low energy building design in high density urban cities. Renewable energy, 24(3), pp.627-640.

Ivajnšič, D., Kaligarič, M. and Žiberna, I., 2014. Geographically weighted regression of the urban heat island of a small city. Applied Geography,53, pp.341-353.

Joyeux, R. and Ripple, R.D., 2007. Household energy consumption versus income and relative standard of living: a panel approach. Energy Policy,35(1), pp.50-60.

Landsberg, H.E., 1981. The urban climate (Vol. 28). Academic press.

Lenzen, M., Wier, M., Cohen, C., Hayami, H., Pachauri, S. and Schaeffer, R., 2006. A comparative multivariate analysis of household energy requirements in Australia, Brazil, Denmark, India and Japan. Energy, 31(2), pp.181-207.

Macdonald, R.W., Griffiths, R.F. and Hall, D.J., 1998. An improved method for the estimation of surface roughness of obstacle arrays. Atmospheric environment, 32(11), pp.1857-1864.

Madlener, R. and Sunak, Y., 2011. Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management?. Sustainable Cities and Society, 1(1), pp.45-53.

Matthews, S.A. and Yang, T.C., 2012. Mapping the results of local statistics: Using geographically weighted regression. Demographic research, 26, p.151.

Mennis, J., 2013. Mapping the results of geographically weighted regression. The Cartographic Journal.

Ministry of Economic Affairs, 2014. Third National Energy Efficiency Action Plan for the Netherlands, https://ec.europa.eu/energy/sites/ener/files/documents/NEEAP_2014_nl-en.pdf, accessed 31/07/2017

O'Neill, B.C. and Chen, B.S., 2002. Demographic determinants of household energy use in the United States. Population and development review, 28, pp.53-88.

Ogneva-Himmelberger, Y., Pearsall, H. and Rakshit, R., 2009. Concrete evidence & geographically weighted regression: a regional analysis of wealth and the land cover in Massachusetts. Applied Geography, 29(4), pp.478-487.

Oliveira, S., Pereira, J.M., San-Miguel-Ayanz, J. and Lourenço, L., 2014. Exploring the spatial patterns of fire density in southern Europe using geographically weighted regression. Applied Geography, 51, pp.143-157.

Sanaieian, H., Tenpierik, M., van den Linden, K., Seraj, F.M. and Shemrani, S.M.M., 2014. Review of the impact of urban block form on thermal performance, solar access and ventilation. Renewable and Sustainable Energy Reviews, 38, pp.551-560.

Steemers, K. and Yun, G.Y., 2009. Household energy consumption: a study of the role of occupants. Building Research & Information, 37(5-6), pp.625-637.

Su, S., Xiao, R. and Zhang, Y., 2012. Multi-scale analysis of spatially varying relationships between agricultural landscape patterns and urbanization using geographically weighted regression. Applied Geography,32(2), pp.360-375.

Tso, G.K. and Yau, K.K., 2003. A study of domestic energy usage patterns in Hong Kong. Energy, 28(15), pp.1671-1682.

Tu, J., 2011. Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. Applied Geography, 31(1), pp.376-392.

Van Moeseke, G., Gratia, E., Reiter, S. and De Herde, A., 2005. Wind pressure distribution influence on natural ventilation for different incidences and environment densities. Energy and buildings, 37(8), pp.878-889.

York, R., 2007. Demographic trends and energy consumption in European Union Nations, 1960–2025. Social science research, 36(3), pp.855-872.

Yun, G.Y. and Steemers, K., 2011. Behavioural, physical and socio-economic factors in household cooling energy consumption. Applied Energy,88(6), pp.2191-2200.

Yust, B.L., Guerin, D.A. and Coopet, J.G., 2002. Residential energy consumption: 1987 to 1997. Family and Consumer Sciences Research Journal, 30(3), pp.323-349.