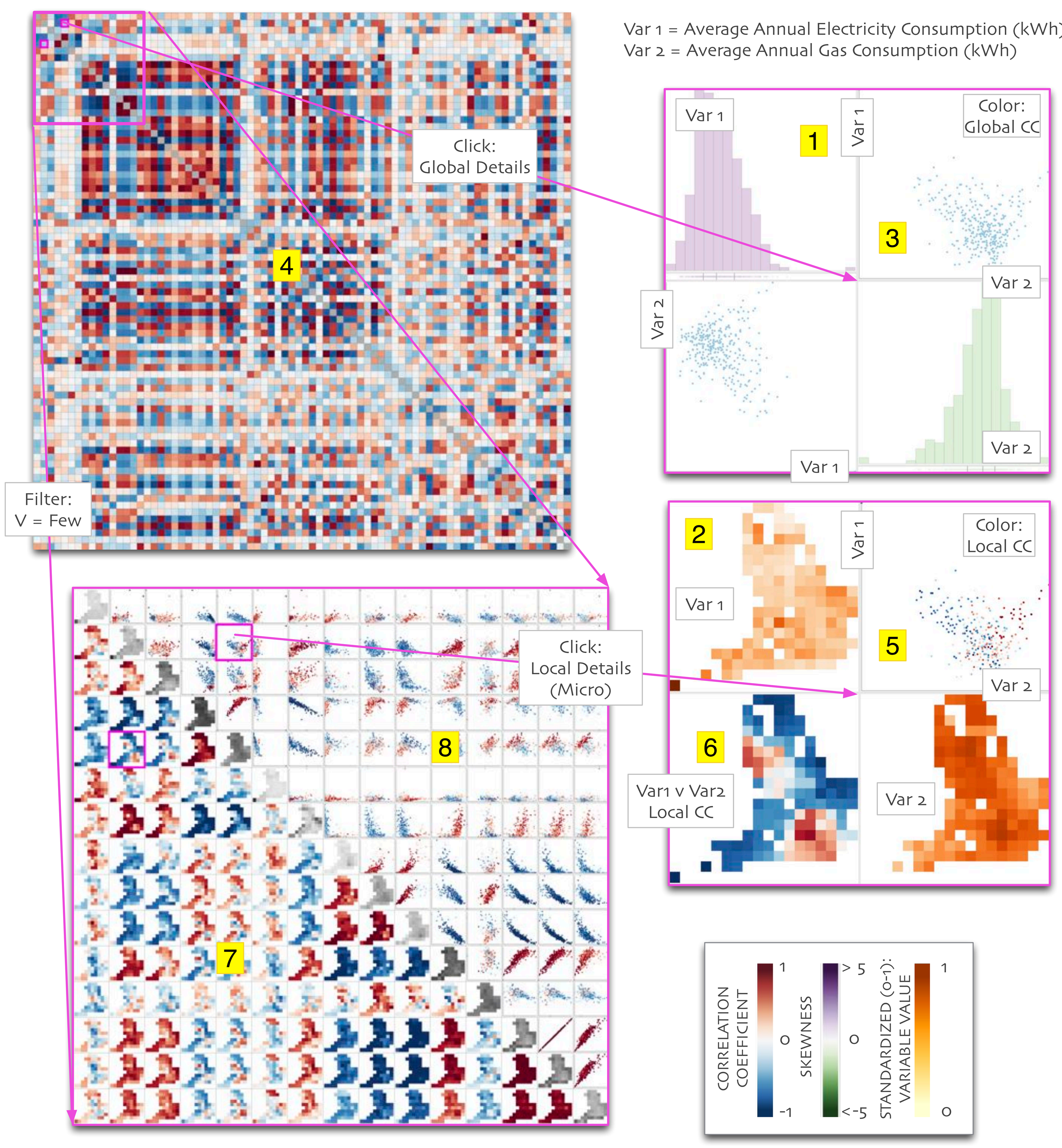
Visualizing the effects of Scale and Geography in Multivariate Comparison

@sgeoviz
Sarah.Goodwin.1@city.ac.uk

Sarah Goodwin, Jason Dykes and Aidan Slingsby - giCentre, City University London, UK

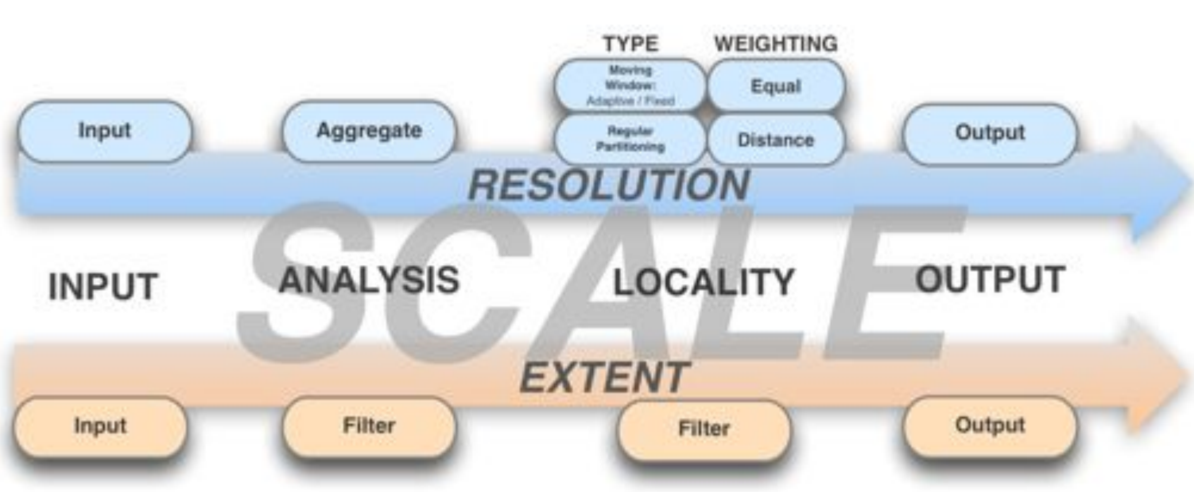
The comparison of geographically varying phenomena is both position and scale dependent. We propose a framework to compare multivariate data across multiple scales whilst allowing local geographical variations and the effects of their parameters to be explored. The statistical (top row) and geographical (bottom row) visualization possibilities adapt as the number of variables (V) and/or the number of local summary statistics (L) increase. An interactive visualization prototype has been built to demonstrate the framework, with aspects (1 - 8) highlighted. Matrices of scatterplots [1], correlation maps or color encoded statistical values (i.e correlation coefficient (CC) and skewness) are utilized to allow as many V as possible to be shown in one view:



Visualizing the FRAMEWORK:

	DISTRIBUTION	CORRELATION		
	V = 1	V = 2	V = FEW	V = MANY
GLOBAL L = 1	1 Histogram or Dot/Box Plot; Map (Choropleth) of Raw Values 2	3 Scatterplot; Pair of Maps or Difference Map	Scatterplot Matrix; Series of Maps or Difference Maps	4 Color Encoding; Color Encoding
MACRO-LOCAL L = FEW	Series of Histograms or Dot/Box Plots; Map (Choropleth) of L Values	Scatterplot colored by L Values; Pair of Maps or Correlation Map	Scatterplot Matrix colored by L Values; Correlation Map Matrix	Less Data Items (Aggregate: Vis or SR/ Filter: SE); High Data Density (Saliency Threat); More Pixels
MICRO-LOCAL L = MANY	Dot/Box Plots or Color Encoding; Map (Choropleth) of L Values	Scatterplot colored by L Values 5; Pair of Maps or 6 Correlation Map	Less Data Items (Aggregate: Vis or SR/ Filter: SE) 7; High Data Density (Saliency Threat) 8; More Pixels	

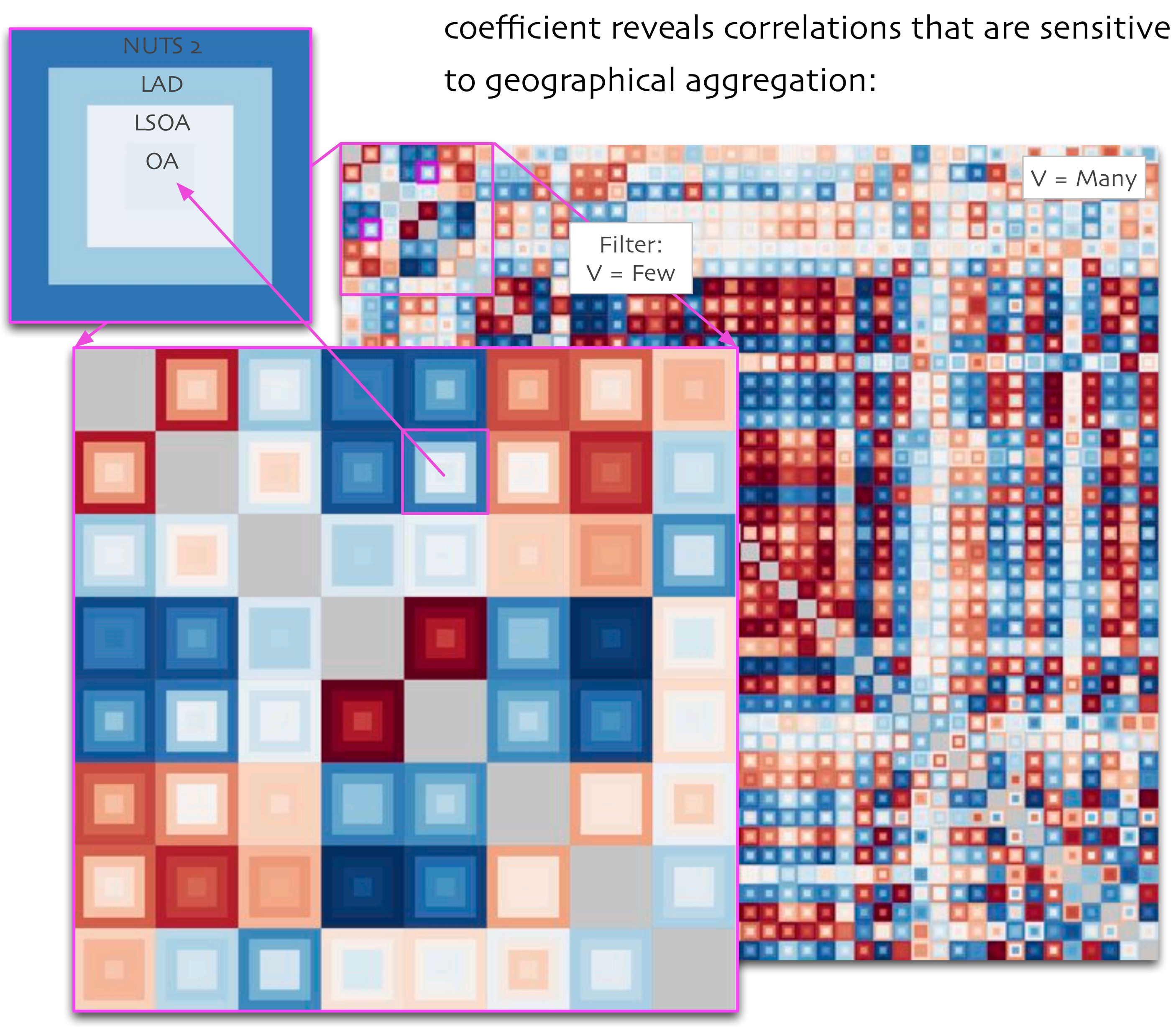
Varying SCALE RESOLUTION:



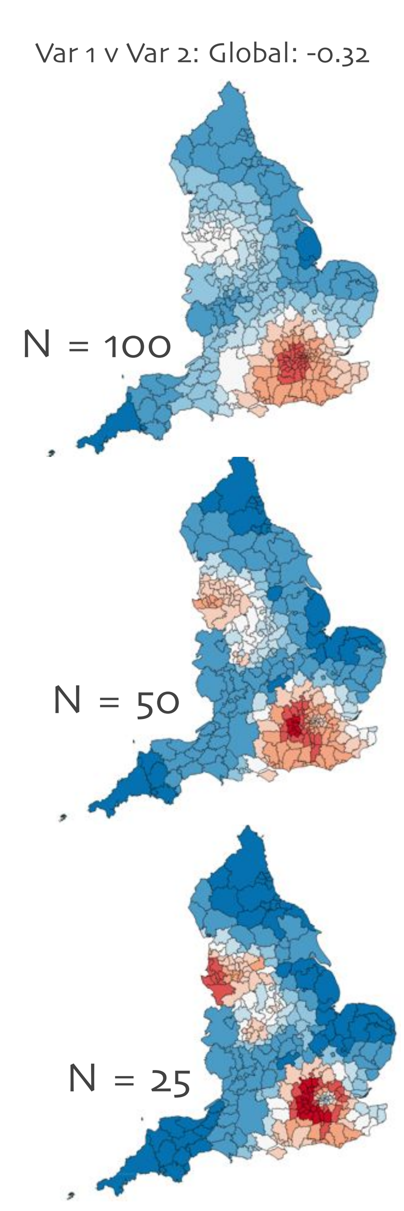
The framework is investigated in the context of 'variable selection' for energy-based geodemographic classification [2,3]. Four stages (Input, Analysis, Locality and Output) are identified in which Scale Resolution (SR) and Extent (SE) [4,5] can be varied. 78 variables are compared over 4 geographical aggregations (SR) covering a SE of England: NUTS2 European Regions (30), Local Authority Districts (326 LADs), Lower Super Output Areas (32,844 LSOAs) and Output Areas (171,372 OAs). The sensitivities and complexities of varying SR are investigated through visualization:

AGGREGATION

Visualizing the color encoded global correlation coefficient reveals correlations that are sensitive to geographical aggregation:



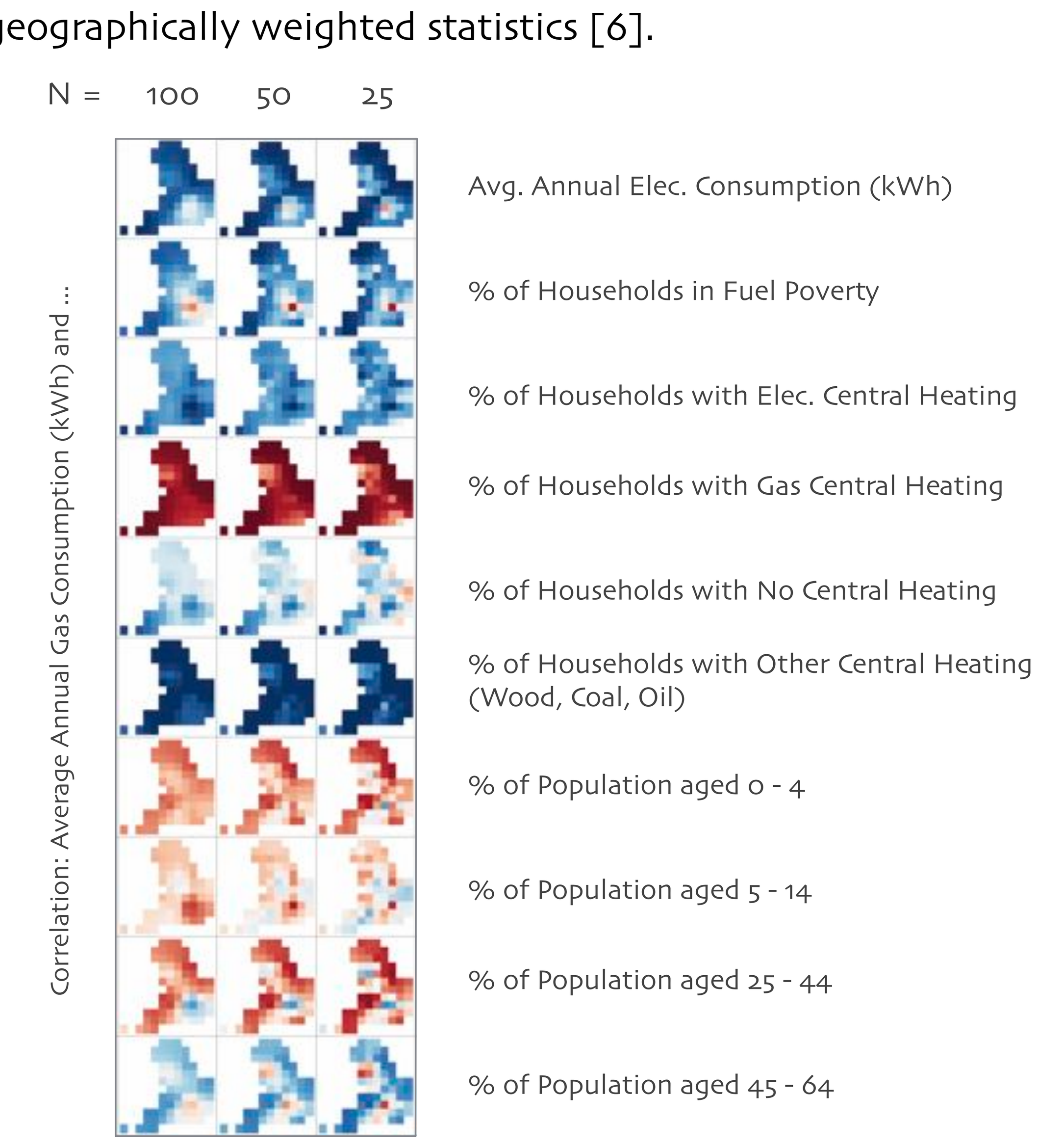
LOCALITY



.... allows for the calculation of local summary statistics (macro or micro). TYPEs are based on geographically weighted statistics [6].

Here, an adaptive moving window is used to calculate local correlation coefficients.

The number (N) of nearest neighbours is adjusted to reveal more detailed local variations when comparing variable pairs (left) or multiple (right):



SUMMARY

This framework and prototype visualization enables the visual comparison of multivariate distributions and correlations across geographical scales and allows for local variations to be explored. The visual representations used in this case-study can be adapted to compare the effects of scale resolution and scale extent that occur when we aggregate and filter by time or attribute as well as geography in our analysis.

[1] M. Monmonier. Geographic Brushing: Enhancing exploratory analysis of the scatterplot matrix. Geographical Analysis, 21(1): 81-84, 1989
[2] S. Goodwin and J. Dykes, Visualizing Variations in Household Energy Consumption. In IEEE Conference on VAST 2012, pages 217-218, Oct 2012
[3] R. Harris, P. Sleight, and R. Weber, Geodemographics: GIS and Neighbourhood Targeting. Wiley-Blackwell, 2005.
[4] N. Lam and D. A. Quattrochi. On the Issues of Scale, Resolution and Fractal Analysis in the Mapping Sciences. The Professional Geographer, 44(1):88–98, 1992.
[5] C. Turkay, A. Slingsby, H. Hauser, J. Wood, and J. Dykes. Attribute signatures: Dynamic Visual Summaries for Analyzing Multivariate Geographical Data. IEEE TVCG, Dec 2014
[6] P.Harris, C. Brunsdon, and M. Charlton. Geographically Weighted Principal Components Analysis. Int. Journal of Geographical Information Science, 25(10): 1717-1736, 2011.

CITY UNIVERSITY LONDON

gicentre.org

Conference Paper: openaccess.city.ac.uk/3876