Exploratory spatial analysis of English place names

Mike Coombes[1] and Colin Wymer[2]

CURDS (Centre for Urban & Regional Development Studies, Newcastle University)

January 27, 2019

Summary

This paper is an initial exploration of the potential for spatial analyses of large place name datasets which are becoming more readily available. It uses the Ordnance Survey gazetteer of Britain that has now been made open data. The example analyses explore how far even simple spatial analyses may help build the evidence for, and visualisation of, the way in which early settlement patterns are identifiable in the geography of place names.

KEYWORDS: place names Ordnance Survey surfacing early settlement England

1. Introduction

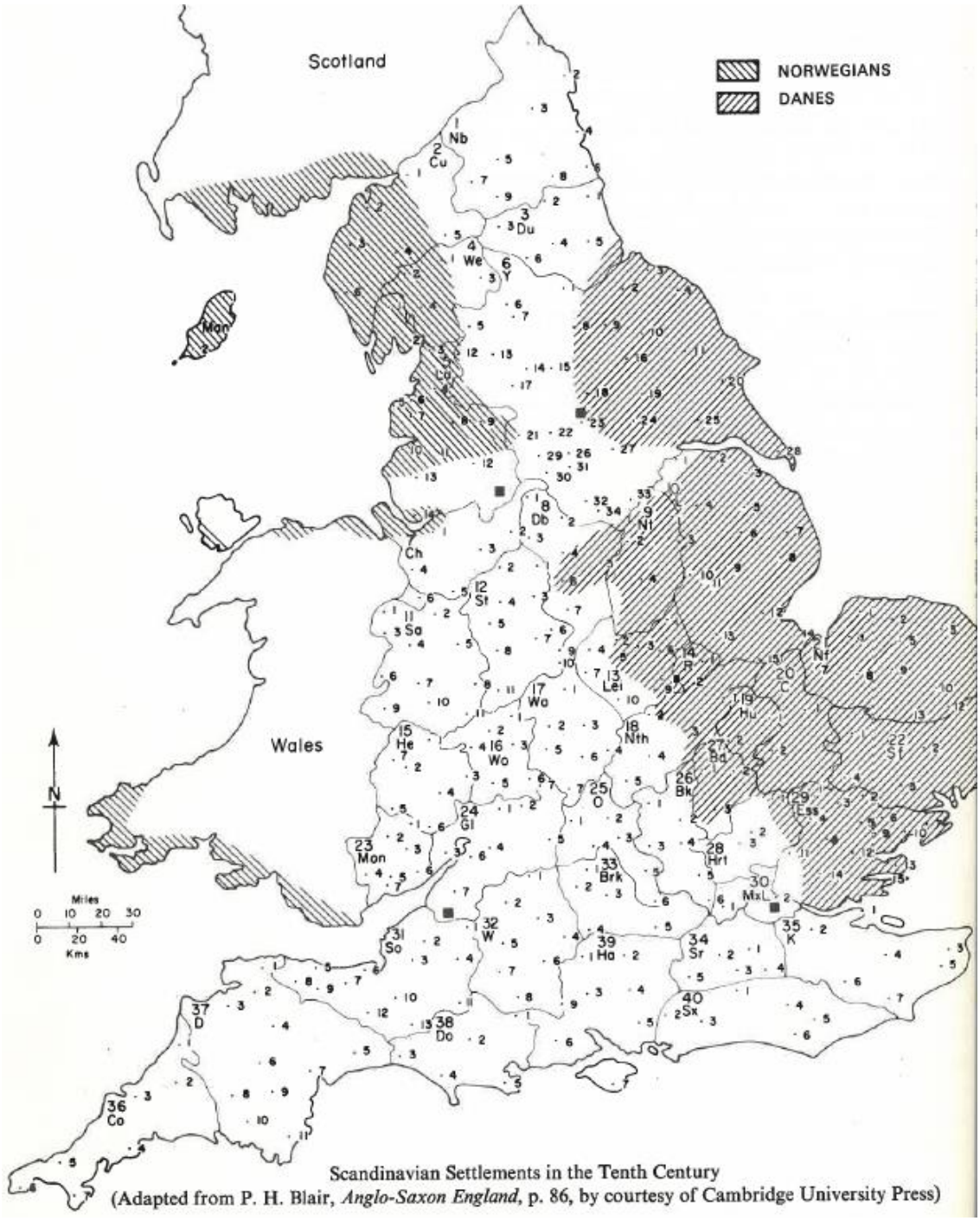
In his renowned agenda-setting article on "GIScience" Goodchild (2004) suggested that it was leading "to increased interest in digital gazetteers and to the processes of naming places" (p.712). In practice the subsequent growth of such work is difficult to detect: the field of toponymy (pace name study) has not seen a burgeoning of the geo-processing of large datasets, but instead of more "critical" case study work (Rose-Redwood et al 2010). This paper is an initial exploration of the potential for the spatial analyses of place name datasets which are becoming more readily available. It uses the potentially 'definitive' name dataset covering Britain, Ordnance Survey's ***Open Names*** (nb. N. Ireland is not covered by the Ordnance Survey but by the Ordnance Survey of Northern Ireland, which is cartographically linked with the Republic's Ordnance Survey Ireland). The example analyses are motivated by the understanding that early settlement patterns are identifiable in the geography of place names (Gelling & Cole 2000).

2. Early settlement patterns in England and topographical names

The large and prolonged Anglo-Saxon settlement of England which began with the end of Roman control produced village names with endings such as "~ingham" that can be found across the country. A later wave of settlers from Scandinavia occupied parts of eastern and northwestern England and left place names with endings such as "~by" whose spatial distribution is part of the evidence for the mapping by Blair which is reproduced by Orton & Wright (1974), as shown in **Figure 1**.

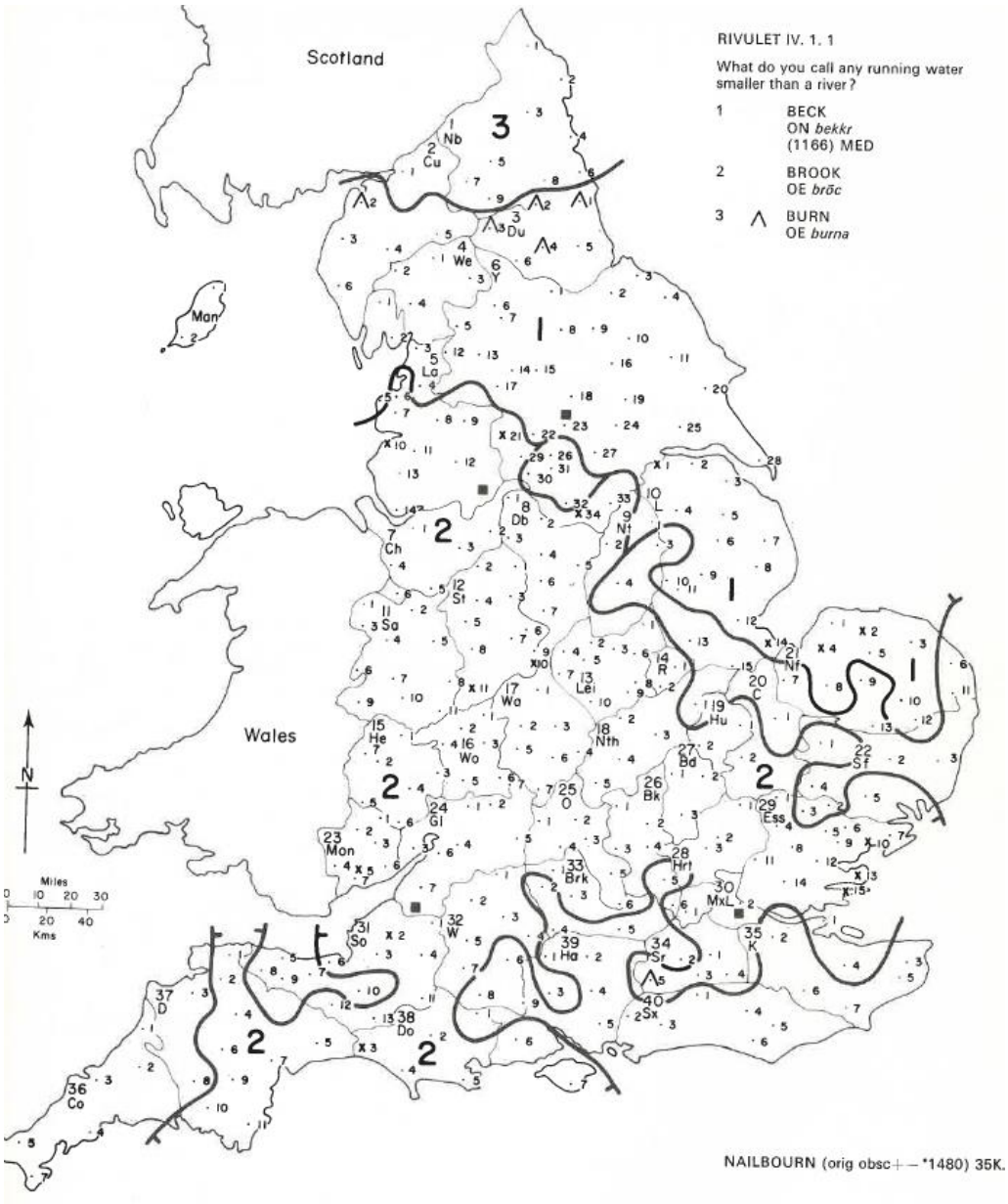
[1] mike.coombes@ncl.ac.uk
[2] colin.wymer@ncl.ac.uk

Figure 1 Settlement pattern of Scandinavians in the 10th Century.



The research of Orton & Wright (1974) centred on dialect and used evidence from a very extensive survey of elderly dialect speakers in 313 places across England, the observation points – mostly villages – numbered in **Figure 1**. One possible case of Scandinavian inheritance in dialect is referring to a stream as a ***beck*** [derived from the Old Norse *bekkr*] rather than a ***brook*** [Old English *brōc*] or perhaps ***burn*** [Old English *burna*]. **Figure 2** shows the spatial distribution of the use of the Scandinavian origin word ***beck*** as mapped by Orton & Wright (1974). As would be expected, the underlying pattern is complex, with no 'clean' boundary between areas in which the different words are predominant. The aim of this paper is to test how far the drawing of such boundaries, as also found in Upton et al (1987), can benefit from the application of spatial analysis techniques to gazetteer data.

Figure 2 Use by dialect speakers of different words for a stream.



3. Analysing street names: a training dataset

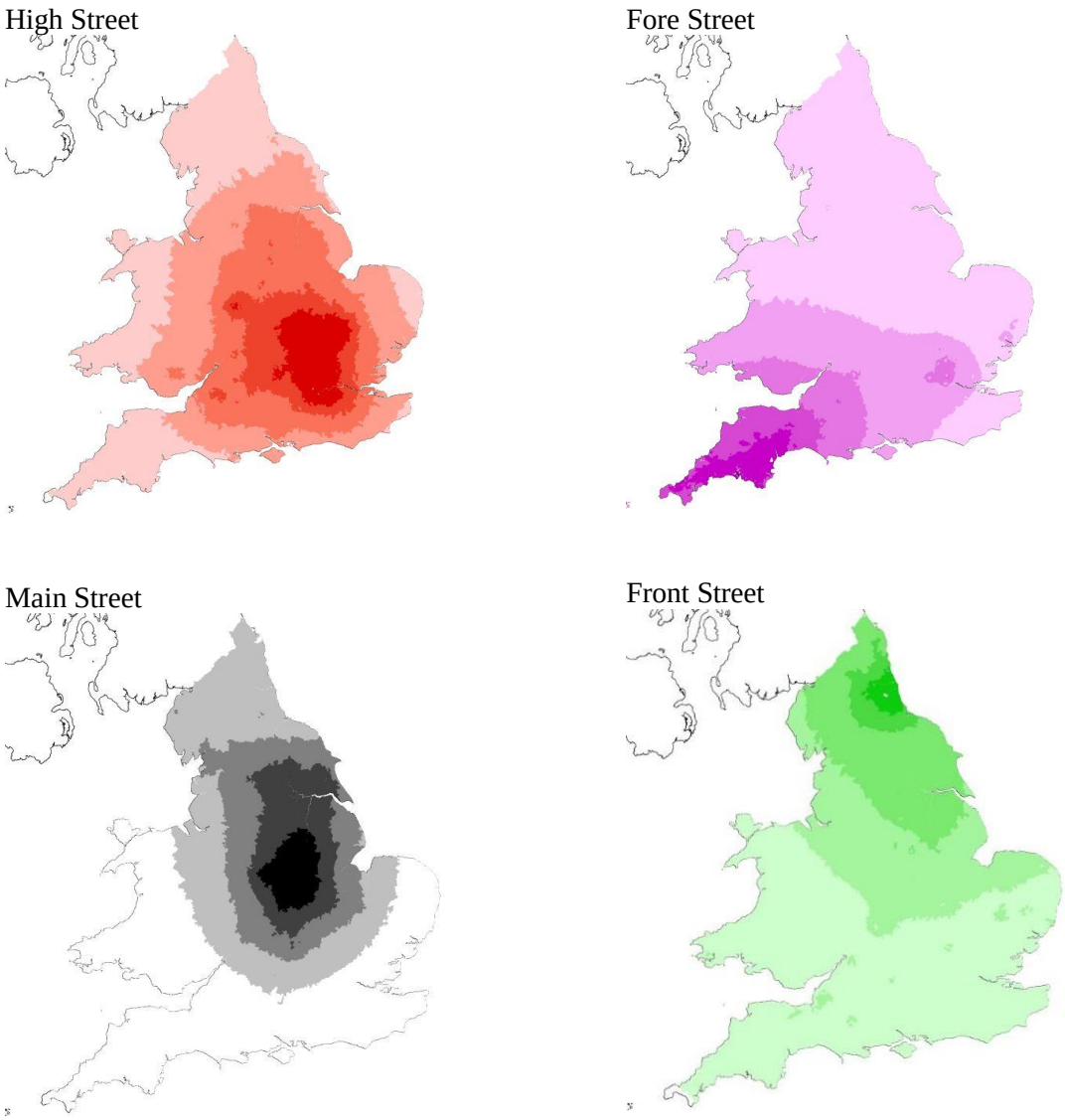
The comprehensive gazetteer dataset OPEN NAMES (Ordnance Survey 2015) includes all street names and these provide a suitable 'training dataset' for the exploratory analysis here seeking to automate the identification of boundaries between areas where the different names used for the same topographical feature suggests the continuing influence of the separate settlement by people from different linguistic origins many centuries ago. The basic steps are as follows: identify sets of names which are cognate words but tend to appear in different parts of the country

- locate all occurrences of each name (nb. implementing this involves deciding what is to count as an 'occurrence')
- measure the prevalence of each name in the vicinity of each small area (*c.*30,000 2011 Census LSOAs/DZs)
- devise boundaries separating groups of areas for which different names were measured as more prevalent (nb. implementing this involves deciding what is to count as 'prevalence').

This approach is applied first to street names which indicate the principal thoroughfare of a settlement which is often its commercial core. High Street is the most frequently occurring such name, and the most widespread alternative is Main Street: very few settlements have both. To further test the approach it is useful to also include alternative names with few occurrences but which are spatially concentrated and so may achieve local prevalence; these are Fore Street (mainly in the west of England), and its equivalent in the northeast, Front Street. These names are not chosen for distinct linguistic backgrounds, but simply as suitable data for the development of the proposed form of analysis.

The initial results simply show the output of a basic analysis similar to the population potential method used for measuring relative accessibility. It produces a basic form of surfacing which can be interpreted as answering the question 'are there many such names near here?' **Figure 3** shows a separate map for each street name, with the darkest colour showing the areas where that name occurs most frequently. (The colour gradients here are fairly arbitrary, essentially aiming only to emphasise peak occurrence.)

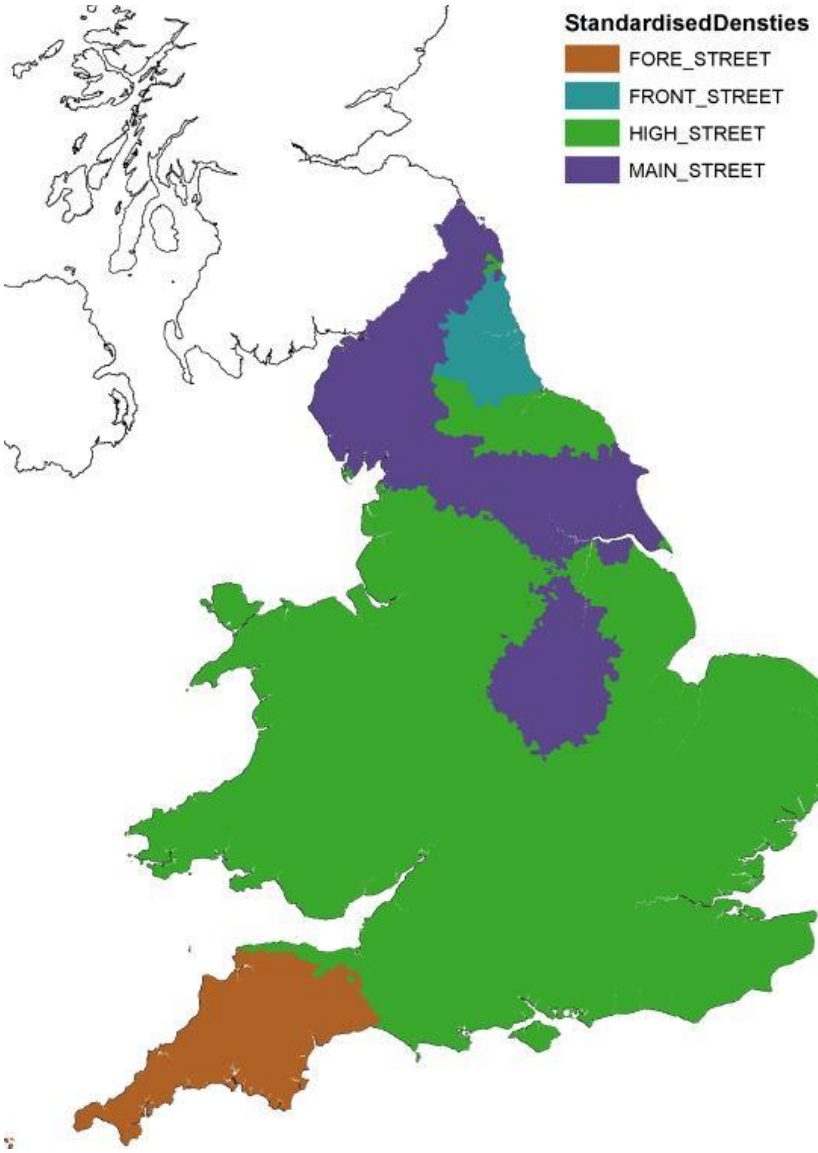
Figure 3 Areas where a specific name for a principal thoroughfare occurs more frequently.



The next step is to identify the name with the highest prevalence score for each individual LSOA. In this dataset, High Street is not only by far the most frequently occurring of these four in England & Wales, it is also widely dispersed; as a result it has the highest value for this type of prevalence

score of the four names here in every single LSOA. The parameterisation of the prevalence analyses means the 'standard' settings used here could be changed to produce somewhat different results. The alternative approach trialled here is to index each name's set of results, so that the original LSOA values are all expressed as proportions of the value of the LSOA with the highest value for that name. Thus the highest point on the 'surface' of each of the above maps has the same value: 100%. This step virtually ensures that each of the names will have at least some LSOAs where its value was the highest of the four (here too there are analytical choices which can be altered so that the sensitivity of the results can be assessed). **Figure 4** shows the result of this step.

Figure 4 Results of analysis to find the principal thoroughfare name which is locally most salient



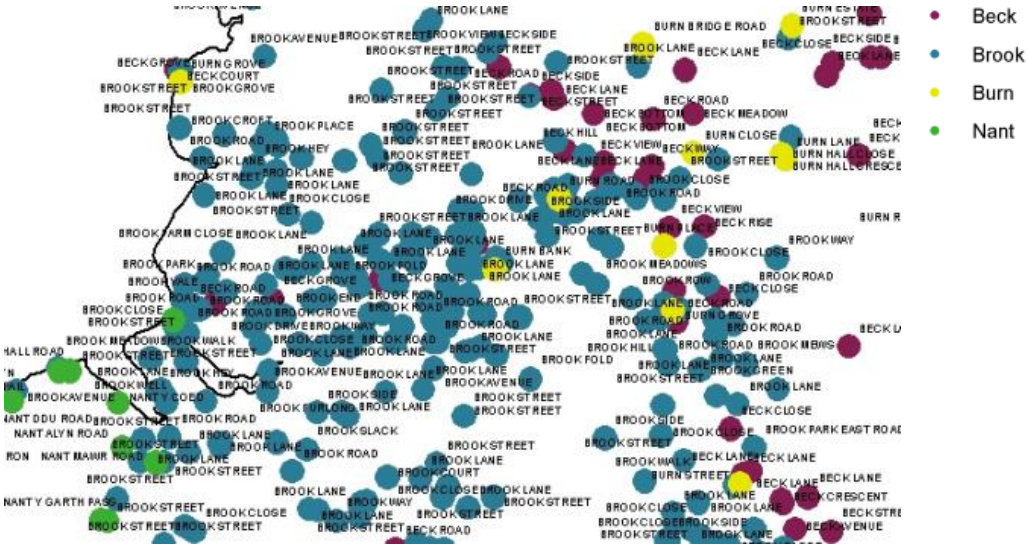
4. Alternative manes for streams: the effect of early settlement patterns?

Although the appropriate sensitivity testing still needs to be undertaken, the basic approach does seem to have promise. The same set of steps was then carried out on the gazetteer data recording stream names. **Figure 5** shows the results from applying the above form of indexed analysis to the three terms for a stream identified earlier.

Figure 5 The locally most salient term used for a stream: evidence from steam names

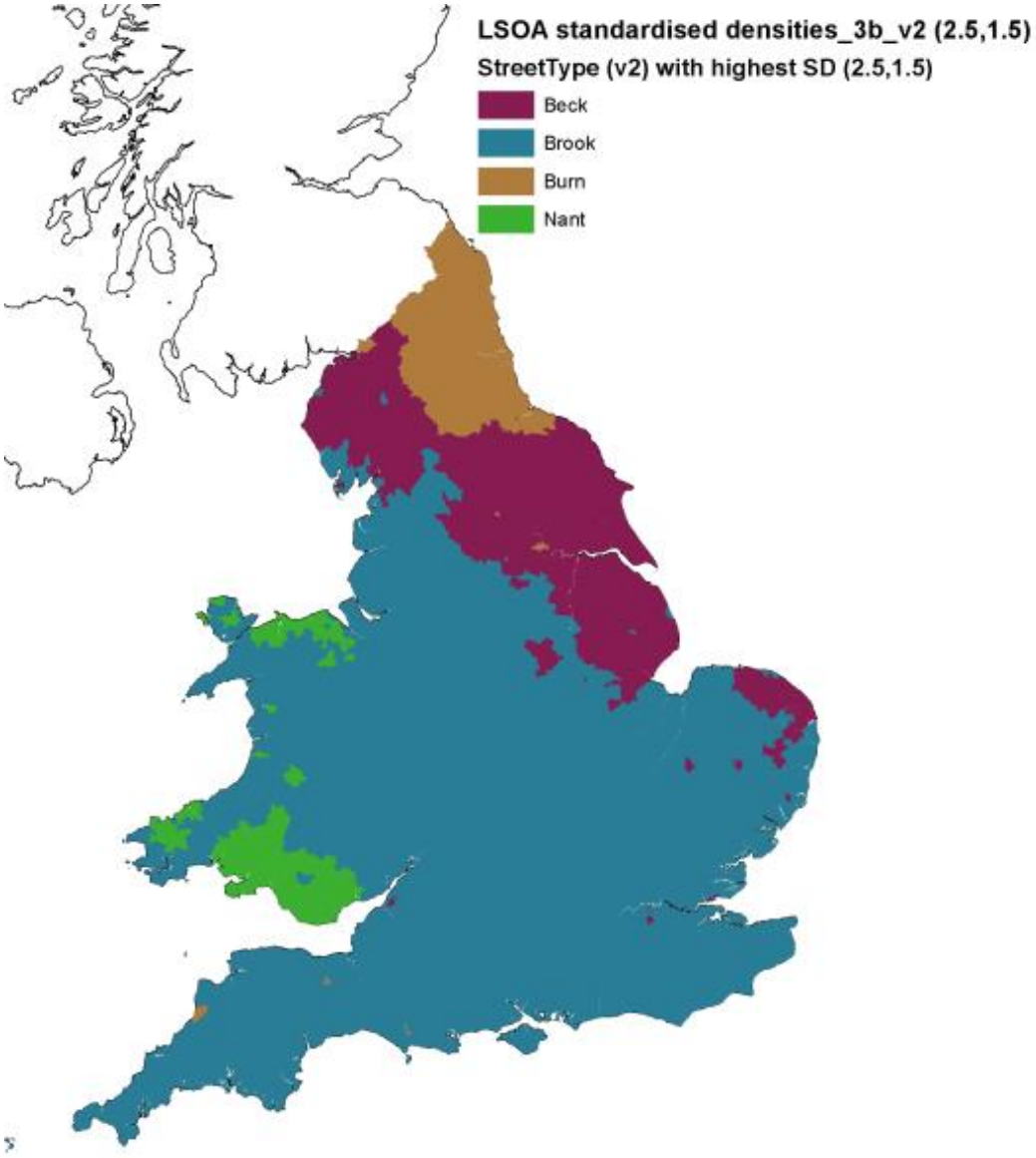
The simplicity of this pattern may be partly due to there being relatively few streams whose names can provide the input data. The alternative is to search the street name data for occurrences of the three terms for stream (although of course stream names are likely to be far more long-established and perhaps 'significant' than are many street names). **Figure 6 s**hows the Pennine region in the north of England the names of all the individual streets which include one of the three stream names, or of ***nant*** which is an equivalent name found in or near Wales (Kitson, 1996).

Figure 6 Street names in the Pennine region including one of four terms for a stream



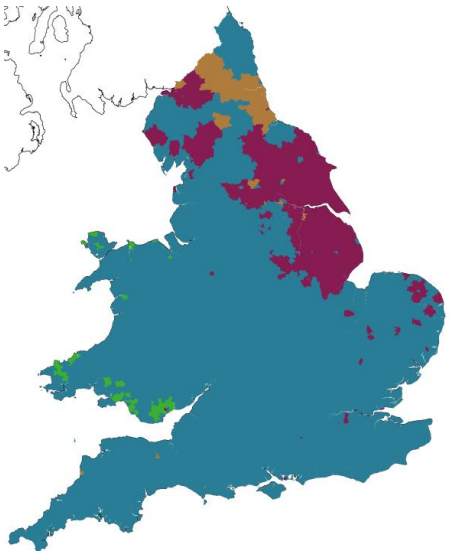
Applying the analysis developed with the training dataset on Main Street and equivalent names produces the results in **Figure 7**.

Figure 7 The locally most salient term used for a stream: evidence from street names



The patterning in **Figure 7** does offer an intriguing 'echo' of the settlement pattern in **Figure 1**. Scandinavian settlement areas of a millennium earlier are found to be partly reflected in the areas where the Scandinavian-origin word for a stream ***beck*** is often used in contemporary street names, even though many of those names will have been created quite recently (and perhaps by house building firms without any regional identity). Such evidence can at best provide systematic background information to the in-depth toponymic studies typically involving hugely intensive research into historical sources which reveal how individual place names evolved over centuries. For the present focus on the potential of spatial analysis of gazetteer data, an immediate next step can be to explore the sensitivity of the results obtained here to the specific analytical decisions made in these analyses, and as a first step **Figure 8** shows the results of changing one of the parameters in the analysis which produced the results in **Figure 7**.

Figure 8 The locally most salient term used for a stream: results from an alternative analysis



Acknowledgements

Acknowledgement is due to Prof Diana Whaley for specialist guidance in relation to place names.

References

Gelling M and Cole A (2000). *The landscape of place-names*. Shaun Tyas, Stamford.

Goodchild M (2004). GIScience, geography, form and process. *Annals of the Association of American Geographers,* 94(4), 709-714

Kitson P (1996). British and European river-names. *Transactions of the Philological Society,* 94(2), 73-118.

Mills A (1991). *Dictionary of English place-names.* Oxford University Press, Oxford.

Ordnance Survey (2015) *OS open names: product guide and technical specification* Ordnance Survey, Southampton https://www.ordnancesurvey.co.uk/docs/user-guides/os-open-names-user-guide.pdf

Orton H and Wright N (1974). *A word geography of England.* Seminar Press, London.

Rose-Redwood R, Alderman D and Azaryahu M (2010). Geographies of toponymic inscription: new direction in critical place-name studies. *Progress in Human Geography,* 34(4), 453-470.

Upton C, Sanderson S and Widdowson J (1987). *Word maps: a dialect atlas of England.* Croom Helm, London.

Biographies

Mike Coombes is now Emeritus Professor of Geographic Information, having joined CURDS in the 1980s. His aim is to analyse social and economic statistics to maximise the 'intelligence' they provide policy-makers and others. Mike has for 40 years led studies to define TTWAs, the only official boundaries defined by academics.

Dr Colin Wymer is now a CURDS Visiting Fellow and has worked with CURDS on key research projects for over 25 years. He has specialist skills in programming, especially for analysing large sparse matrices (eg. to define TTWAs). He is joint author of numerous journal articles and chapters.