Mapping Landscapes in Transformation
Multidisciplinary Methods for Historical Analysis

Edited by
Thomas Coomans, Bieke Cattoor, and Krista De Jonge

Leuven University Press

12.

Landscape Appreciation in the English Lake District

A GIS Approach

Ian Gregory (Lancaster University)

Christopher Donaldson (Lancaster University)

Joanna E. Taylor (The University of Manchester)

There is a well-established tradition of historical geographers using geographical information systems (GIS) to study historic landscapes and, particularly, landscape transformation (see Cunfer 2005; Donahue 2007; Walford 2018). Such studies tend to be based on quantitative sources and social science paradigms. This tendency is unsurprising, as the GIS data model is well suited to structuring analyses of quantitative sources with attribute data in tabular form linked to spatial data representing precisely located points, lines, or polygons. This chapter presents a new approach to the study of how landscapes were perceived in the past and how this changed over time. Implementing this approach requires a very different integration of GIS from the one used in previous studies because the sources involved are not quantitative: they are texts containing qualitative descriptions and, as such, need to be analysed using a mix of approaches that combine spatial analysis with close reading.

To explore the potential of using GIS to research changing perceptions of historic landscapes we draw on the Corpus of Lake District Writing (CLDW). This is a collection of eighty digitised texts written about the English Lake District between 1622 and 1900. The Lake District is one of the most written about rural landscapes in the world (Fig. 1). Most famously, it was the home of William Wordsworth and his contemporary Lake Poets; however, it has also been described by many lesser-known writers, as well as being the subject of many guidebooks, journals, and travelogues. The CLDW holds a collection of both famous and less-canonical

Fig. 1. Coledale Fells, seen from Crosthwaite, English Lake District (photo Christopher Donaldson).

works.[1] The challenge this chapter explores is two-fold: first, how to convert the texts in the CLDW into a form suitable for importing into GIS; second, how to analyse these texts in an appropriate manner for literary and geographical study. This chapter reports on an approach that meets these challenges through a case study that examines three words commonly associated with the landscapes of the Lake District: beautiful, majestic, and sublime.

From text to GIS database

The CLDW contains approximately 1.5 million words of text marked up with a light XML (eXtensible Markup Language) encoding to identify features such as paragraph breaks and to incorporate basic metadata. The challenge is how to convert this unstructured textual format into a data model suitable for importing into a vector GIS, which requires data in tabular form linked to coordinate-based locations. Performing this conversion involves a two-stage process known as

---

1. For a full list of titles in the corpus: see http://www.lancaster.ac.uk/fass/projects/spatialhum. wordpress/?page_id=64

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

geoparsing (Grover et al. 2010). In the first stage, automated named entity recognition (NER) techniques are used to identify words suspected of being place-names. In the second, these words are matched to a place-name gazetteer, and, where a match exists, the coordinate data the gazetteer provides are assigned to the relevant word. These coordinate data are allocated through additional XML tags that are added to the texts. Conducting this process in an entirely automated manner was found not to be satisfactory for the complex place-names found in the CLDW (see Butler et al. 2017). The process was enhanced using concordance geoparsing — where a small subset of the text is geoparsed, the results are checked, and any corrections fed into processing subsequent subsets (Rupp et al. 2014) — and a considerable amount of manual checking.

Geoparsing the CLDW led to the identification of nearly 40,000 place-names: 2.6% of the total word tokens in the corpus. Linguists count items of punctuation as tokens in addition to words. Of these place-names, 96% are in the UK and 88% are in and around the Lake District. The remainder are frequently places in mountainous parts of North America or Europe that texts in the CLDW compare to locations in the Lakes, or places mentioned in a similarly associative manner. Interestingly, only 60% of the place-names identified are found within the modern Lake District National Park. This suggests that today's National Park is smaller than the area historically associated with the Lake District.

Map 1 offers a visual summary of the results of geoparsing the CLDW. Here, all of the place-names found in and around the Lake District have been located using the coordinates allocated to them in the geoparsed XML. Rather than use a dot-map, which is difficult for the human eye to interpret (particularly when there may be multiple dots at the same location), the map uses a technique called density smoothing to simplify the pattern. The map shows the density of place names near to each location. 'Near' is defined objectively (Fotheringham et al. 2000: 149): to allow comparisons, density on all of the maps in this chapter has been normalised using z-scores. A z-score of 0.0 is the mean density for the dataset; 1.0 is one standard deviation above the mean, and so on. Map 1 shows that there is a very clear and clustered pattern to the place-names mentioned in the CLDW. There are major clusters of places mentioned in and around the larger settlements, including Lancaster, Kendal, Penrith, Bowness, Ambleside, Grasmere, and Keswick. There are also pronounced clusters near the major mountains and lakes. In sum, Map 1 demonstrates that the place-names mentioned most often in the CLDW are located around the principal settlements and landmarks in the centre and east of the region. It is striking, though, that significant areas of this landscape are far

less frequently mentioned by the texts in our corpus. However, representing these clusters is somewhat problematic since large areas such as lakes are represented using a point; the point on the east bank of Windermere, for instance, gives little indication of precisely to where on this large body of water the text is referring.

Map 1 demonstrates that geoparsing can be used to convert a text into a GIS; it also affirms that GIS-generated maps can be used to summarise the patterns of place-naming in a text corpus. That said, **Map 1** tells us nothing about what the texts in the corpus say about these locations. To explore this aspect of the corpus we used a concept taken from corpus linguistics: co-occurrence. Linguists argue that two words appearing near each other in a text are likely to be related (Adolphs 2006: 56-59; McEnery and Hardie 2012: 122-123). To operationalise this principle with place-names we have developed the concept of the place-name co-occurrence (PNC). A PNC occurs when a place-name is found within a set number of word tokens of a particular search-term. In this case, we used a measure of ten word tokens. Take, for example, the following extracts from the CLDW:

> …on the pinnacle of 'the mighty **Helvellyn**'. How <u>sublime</u> an elevation! How glorious a panorama! (Baines, 1829)
> 'The scenery about **Buttermere** is truly <u>sublime</u> and august. On a promontory to the east of the lake…' (Anon, 1804)

In each case, the place-name is highlighted in bold and the search-term, 'sublime', is underlined. In instances like these, the fact that the place-name is within ten tokens of the search-term is taken as indicating that they are associated with one another. Thus, Helvellyn and Buttermere are both (correctly) identified as being described as sublime. Both, in other words, are classed as PNCs of sublime.

Consider, however, the following further extract:

> '–still presenting the same appearance of <u>sublime</u> confusion – of the mountains of **Blackcomb**, **Scaw fell**, Coniston Old Man, Helvellyn, Fairfield, Hill Bell…' (MacKay 1846)

This third example illustrates two ways place-name co-occurrence relationships can be more complicated than the first two examples suggest. Blackcombe and Scawfell[2] are both place-names that lie within ten word tokens of 'sublime'. As a

---

2. Note the non-conventional spelling of these formations in the quotation.

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

result, the clause quoted is the source of not one, but two PNCs: one for Blackcombe and one for Scawfell. The remaining place-names in the quotation (Coniston Old Man, Helvellyn, Fairfield, and Hill Bell) lie outside the ten-word-token span; they are thus erroneously not considered to be associated with 'sublime' and so do not create PNCs. One needs to be mindful of the possibility of such errors when analysing PNCs. Still, in spite of such potential limitations, employing the concept of PNC does allow us to associate place-names with search-terms and thus investigate what places are being associated with particular themes of interest.

Beautiful, majestic, and sublime places in the Lake District

Using PNCs enables us to produce maps such as **Map 2a**, which shows the locations of the PNCs for all instances of the search-term 'beautiful' in the CLDW. The word beautiful occurs 1,233 times in the corpus, creating 485 PNCs. This amounts to 39.4 PNCs per 100 instances, suggesting that 'beautiful' is strongly associated with place-names. There is a clear geographical pattern to these PNCs, with clusters around Windermere, Ambleside, and Grasmere in the south, and Keswick and Derwentwater in the north. This pattern is potentially somewhat problematic, however. **Map 1** showed that there is a clear pattern to place-naming in the corpus; consequently, the pattern shown in **Map 2a** could simply identify places that are associated with 'beautiful' because those places are often mentioned in the CLDW overall. To determine whether this is the case, we implemented advanced spatial analysis techniques, particularly Martin Kulldorff's Spatial Scan Statistic (Kulldorff 1997). Kulldorff's Statistic allowed us to identify places that were more closely associated with the search-term than would be expected given the background geography of all the place-names mentioned in the corpus.

Map 2b displays the results of using Kulldorff's Spatial Scan Statistic to identify areas that have more PNCs with 'beautiful' than would be expected based on the underlying place-name geography shown in **Map 1**. These clusters are termed hotspots. Kulldorff's Spatial Scan Statistic also identifies those areas which have fewer PNCs than would be expected; these are termed cold spots. Thus, **Map 2b** shows that the area around Grasmere is a hotspot as well as being a density cluster in **Map 2a**. This suggests that Grasmere is both a frequently mentioned location and, moreover, one that is described as 'beautiful' more often than would be expected. The area around Keswick is also a hotspot in **Map 2b**, but it is noticeable that this is further west than the density cluster shown in **Map 2a**. A third hotspot is found in the southern part of the Lake District, in the area covering the Cartmel Peninsula and High Furness. This is an area that has a low density of

place-name instances [**Map 1**] and of 'beautiful' PNCs [**Map 2a**], which suggests that, although it is not an area mentioned frequently in the CLDW, when it is mentioned it is often described as 'beautiful'. **Map 2b** also shows two cold spots. The large one in the north-east corner of the map can probably be ignored, as this is an area well outside what would normally be thought of as the Lake District. The smaller one in the west, however, is the Western Fells, which include Scawfell Pike, the highest mountain in the Lake District. This finding suggests that this area is not often described as beautiful in the corpus; rather, this region inspired somewhat grander language.

Indeed, the area around Scawfell is more associated with the word 'sublime', as **Map 3** shows. 'Sublime' occurs 270 times in the corpus and has 87 PNCs: 32.2 PNCs per 100 instances. This percentage suggests that sublime is a little less strongly associated with named places than the other two terms. The density pattern in **Map 3a** shows a strongly clustered pattern with clusters around Keswick, Ullswater, Windermere, Coniston, and the Western Fells (including Scawfell). Most of these are major tourist sites. The Western Fells, however, were less frequently visited and are mentioned less often in the corpus. The fact that the only hotspot for 'sublime' is located at these fells [**Map 3b**] indicates the emotional effect the mountains had on the writers in the CLDW. As with **Map 2b**, the cold spot to the east of the map can be ignored as it is beyond the Lake District.

Map 4 shows two similar maps for the term 'majestic'. 'Majestic' is a far less common word than 'beautiful' or 'sublime' in the CLDW, wherein it occurs only 173 times. However, it forms 79 PNCs (45.7 PNCs per 100 instances), which makes it more commonly associated with named places than 'beautiful'. The patterns shown in **Map 4** are very different from those in **Maps 2** and **3**. These patterns are concentrated in two major density clusters, one around the mountain Skiddaw in the north, and a less pronounced one around the point representing the River Duddon in the south. These are both hotspots as well as density clusters; there are no cold spots.

Map 5 draws these three patterns together on two single maps. **Map 5a** shows the density clusters with z-scores above 2.58 for all three search-terms. What is clear from the pattern is how different the geographies associated with the three search-terms are. The only place where all three terms overlap, a tiny sliver between the Keswick cluster and the Skiddaw cluster, can be dismissed as an edge effect. The only major overlaps are between 'beautiful' and 'sublime', which cluster together in Keswick and Windermere. There is also a more limited

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

overlap between 'beautiful' and 'majestic' centred on Skiddaw, but there is no overlap between 'majestic' and 'sublime'. **Map 5b** shows that for hotspots this pattern of no overlap is even more pronounced: none of the hotspots overlap at all. However, the Western Fells provide a hotspot for 'sublime' and — although not shown — a cold spot for 'beautiful', suggesting a marked contrast between the two terms.

Figure 2 shows that, as well as spatial variation in how the words are used, there is also temporal variation. The corpus can be divided into four period groupings: the Early Modern period, the 'long' eighteenth century, the Romantic period and the Victorian era. Figure 2a shows the percentage of instances of the three search-terms that occur in each grouping. These percentages are influenced by the fact that the corpus contains more texts from the Romantic period and the Victorian era than the other two groupings. Accordingly, Figure 2b compares each of the three terms to what would be expected if its use were proportional to the number of words in each period grouping. The Early Modern period has been left off this graph as it represents a much smaller portion of the corpus. This chart shows not only that 'majestic' was most used in the Romantic period texts in the corpus, but also that the word is much less seldom used either before or afterwards. 'Sublime' was also most used in works from the Romantic period, but the use of the term appears to have decreased during the Victorian era. The use of 'beautiful', by contrast, appears to have remained broadly constant across all three periods.

Exploring transformation in landscape representation involves evaluating change over space and time simultaneously. In order to do this, we have used the three period groupings from Figure 2: the 'long' eighteenth century, the Romantic period, and the Victorian era. We can then apply a technique called Spatial Segregation Analysis (Rowlingson 2015; Diggle et al. 2005) to assess how patterns of instances vary between periods. Spatial segregation analysis is a technique that compares point patterns, in this case PNCs from different periods, and identifies parts of the map that have statistically significantly more points from one set of points than the other. It also provides a global summary statistic that identifies whether the two patterns are statistically significantly different from each other across the entire map area.

Map 6a offers an example of the application of this method. Here, we compare the PNCs of 'beautiful' from the 'long' eighteenth century with those from the Romantic period. The global statistic for this is $p=0.07$, which suggests that overall
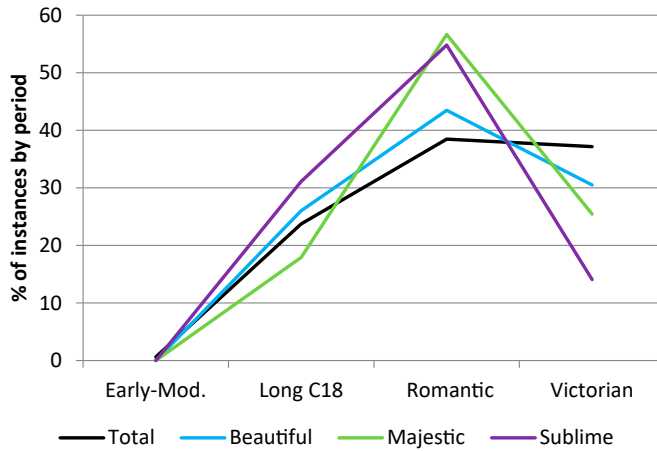
Fig. 2a. Word use by period showing overall percentages of instances occurring in each period.
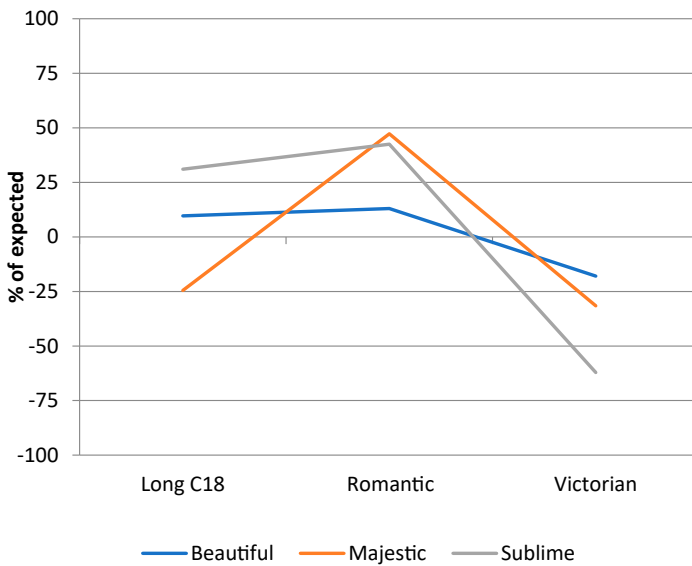
Fig. 2b. Word use by period showing overall percentages normalised by overall word count in each period.

In general these diagrams show that the usage of all three terms declines in the Victorian era although this is least pronounced for 'beautiful'. 'Majestic' was relatively uncommon in the 'long' eighteenth century, but became popular in the Romantic era before dropping out of use again. 'Sublime' was popular until the Victorian era.

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

the differences between the distributions in the two groupings are not significantly different. **Map 6a** implies that this is a slight over-simplification as the area around Grasmere has more PNCs of 'beautiful' in the Romantic era than it did in the 'long' eighteenth century. Further to the west, the area around Ravenglass and Eskdale also has significantly more beautiful PNCs in the Romantic period, although this is largely because there are none in the 'long' eighteenth century. **Map 6b**, which compares the Romantic period and Victorian era, suggests that this development was part of a growing trend. This map shows significantly more 'beautiful' PNCs in the Victorian era, particularly in the area west of Coniston Water. This suggests that this previously neglected area was increasingly described as 'beautiful' as time went on. At the same time, areas around Keswick and Derwent Water and Ullswater have significantly more 'beautiful' PNCs in the Romantic era, suggesting that these traditionally popular landscapes became less associated with 'beautiful' as time went on. The global statistic for the comparison between the Romantic and Victorian eras is p=0.01, meaning that these patterns are significantly different from each other and suggesting that the rate of change in the use of this word increases over time.

Map 7 shows the same technique applied to 'sublime' PNCs. Interestingly, the two patterns are similar to those for 'beautiful' shown in [**Map 6**]. The contrast between the 'long' eighteenth century and the Romantic era is not statistically significant globally (p=0.09) while the comparison between the Romantic period and Victorian era is significant (p=0.01). Locally, the pattern again shows a move towards the southern and western Lake District over time although the emphasis on Grasmere disappears. The analysis for 'majestic' has not been mapped because the pattern is too concentrated on Skiddaw and the Duddon Valley, with the Duddon emerging as a PNC only really in the Victorian era.

Overall, the spatial segregation analysis suggests that there is an increasing tendency over time to describe landscapes in the south and west of the Lake District as both 'beautiful' and 'sublime'. Following up this finding further using Kulldorff analysis, however, does not suggest that this is because there appear to be new clusters of PNCs in these areas when compared to the background geographies. Instead, it appears to be the case that, over time, the south and west of the Lake District received more attention within the corpus. This trend probably reflects the fact that these localities began to be more frequently visited, especially following the completion of railways to locations such as Windermere and Coniston during the Victorian era. Such developments can be seen to have worked in

Place-centred reading

The above analyses all present abstract summaries from a large corpus of historic texts. These summaries identify and describe aggregate geographical and temporal patterns within those texts. The patterns found should, in turn, be thought of as a question to the researcher: e.g. why is this pattern as it is? The ability to perform such descriptive analyses is a feature of most maps and diagrams produced by GIS analyses. The use of qualitative textual sources, though, enables us to address those questions with reference to the material from which the patterns arise. By focusing on the text around any given PNC we can begin to understand more precisely the nature of the relationship that underlies the observed search-term and place-name co-occurrences. In other words, we can move from macro (or 'distant') reading approaches (Jockers 2013; Moretti 2013) to more conventional — in literary studies, at least — close readings of the parts of the text that the macro-approaches identified as significant.

As noted above, 'majestic' presents perhaps the most striking pattern of our three search-terms, as its PNCs are highly clustered in two distinct areas, Skiddaw and the Duddon. Identifying places in the text where these place-names are found within ten word tokens of the word 'majestic' can either be done by looking at the PNCs themselves or by using corpus software such as AntConc (Anthony 2005) or CQPweb (Hardie 2012), which provides more contextual text. This investigation reveals that Skiddaw, a mountain that towers above the tourist town of Keswick, is most commonly described as 'majestic' in guidebooks from the Romantic era. It is difficult to determine if there was a particular text that established this trend; it appears to be common throughout the CLDW, and the word 'majestic' is used as a descriptor in many contexts. The Duddon PNC presents an interesting contrast. In his *The River Duddon: A Series of Sonnets*, William Wordsworth (1820: 32) writes 'now expands majestic **Duddon**, over smooth flat sands'. This is the first time that the Duddon is described as 'majestic' in the corpus, and five of the other six texts that describe the Duddon as 'majestic' use exactly this quotation.

The major hotspot for 'sublime' is in the Western Fells [**Map 4**], an area that can be seen from a distance in much of the Lake District but that was, and remains, relatively inaccessible. In stark contrast to 'majestic', close reading reveals that 'sublime' tends to be used to describe these mountains as a group; for example:

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

'on the west the mountains of *Red Pike*, **High Stile**, and **High Crag** rise precipitously to a <u>sublime</u> elevation' (Baines 1829); and 'that <u>sublime</u> group of mountains, **Yewbarrow**, **Pillar**, **Great Gable**, **Kirkfell**, *Lingmell*, and *Scawfell*' (Waugh 1861).[3] Thus, 'sublime' seems to be used quite specifically to refer to groups of mountains in the distance, rather than individual summits. This application of the word was maintained during the Victorian era, even as the use of 'sublime' appears to have decreased. For example, Mackay (1846) refers to 'the <u>sublime</u> chain extending from **Coniston** to **Ennerdale**', a phrase that is repeated by a number of subsequent writers.

The usage of 'beautiful' is more difficult to pin down. The most obvious question driven by **Map 2b** is why the Southern Lakes are shown to be a 'beautiful' hotspot. Referring back to the corpus reveals that this finding is perhaps misleading. In 1860, Edwin Waugh published a guidebook entitled *Over the Sands to the Lakes*, which frequently uses the word 'beautiful' to describe places in this part of the Lake District. In 1861, Waugh went on to publish a collection entitled *Rambles in the Lake Country*, wherein he reproduced his earlier text as a chapter. The hotspot displayed in **Map 2b** is thus a result of these two texts, combined with the sparsity of other writing about this area. Beyond this, however, it seems clear that 'beautiful' is often used to describe well-visited parts of the Lakes at relatively low elevations, particularly near Grasmere and Keswick, and that this use of the word persists over time. By contrast, 'beautiful' is not used to describe high, remote, 'sublime' landscapes such as the Western Fells.

Conclusions

In this chapter we have shown how three words that are historically used to describe the Lake District — majestic, sublime, and beautiful — are actually employed in distinctive ways to describe different places and different types of landscapes. We are also able to explore if and how the use of these words in the corpus changed over time. 'Majestic' is used for a small number of places in quite specific ways, and in one case (e.g. the River Duddon) we are able to show that Wordsworth's use of this word resulted in many subsequent writers describing it in this way. 'Sublime' tends to be used to describe mountain ranges as lists of peaks. It was particularly associated with the Western Fells, but as the term became less popular it also

---

3. Italics refer to words that are place-names that lie outside the ten-word span and are thus (wrongly) excluded from defining the PNC.

began to be increasingly used in other similar areas. 'Beautiful' is the most widely used word but seems to have a fairly specific association with low-lying, popular parts of the Lake District.

The implications of this chapter are broader than the examples we have used here. They demonstrate that, in addition to traditional quantitative and cartographic sources, we can now use GIS-led methodologies to explore how landscapes and environments are represented. More than this, we can discover how the perceptions of those landscapes, and the language used to describe them, transformed over time.

Acknowledgements

Bibliography

Adolphs S. (2006). *Introducing Electronic Text Analysis: A Practical Guide for Language and Literary Studies*. London: Routledge.

Anon (1804). *Observations, Chiefly Lithological, Made in a Five Weeks' Tour.* London: T. Ostell.

Anthony L. (2005). AntConc: Design and development of a freeware corpus analysis toolkit for the technical writing classroom. In: *Proceedings of the International Professional Communication Conference*. p. 729–737.

Baines E. (1829). *Companion to the Lakes*. Hurst, London: Chance & Co.

Butler J.O., Donaldson Ch.E., Taylor J.E. and Gregory I.N. (2017). Alts, Abbreviations, and AKAs: Historical onomastic variation and automated named entity recognition. *Journal of Map and Geography Libraries* 13, p. 58–81

Cunfer G. (2005). *On the Great Plains: Agriculture and Environment.* College Station: Texas A&M University Press.

Diggle P., Zheng P. and Durr P. (2005). Nonparametric estimation of spatial segregation in a multivariate point process: Bovine tuberculosis in Cornwall, UK. *Journal of the Royal Statistical Society Series A* 54, p. 645–658.

Donahue B. (2007). *The Great Meadow: Farmers and the Land in Colonial Concord.* New Haven: Yale University Press.

Fotheringham A.S., Brunsdon C. and Charlton M. (2000). *Quantitative Geography: Perspectives on Spatial Data Analysis*. London: Sage.

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

Grover C., Tobin R., Byrne K., Woollard M., Reid J., Dunn S. and Ball J. (2010). Use of the Edinburgh geoparser for georeferencing digitized historical collections. *Philosophical Transactions of the Royal Society A*, 368, p. 3875–3889.

Hardie A. (2012). CQPweb – Combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics* 17, p. 380–409.

Jockers M.L. (2013). *Macroanalysis: Digital Methods and Literary History.* Urbana-Champagne: University of Illinois Press.

Kulldorff M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods* 26, p. 1481–1496.

Mackay C. (1864). *The Scenery and Poetry of the English Lakes.* London: Longman & Co.

McEnery T. and Hardie A. (2012). *Corpus Linguistics: Method, Theory and Practice.* Cambridge: Cambridge University Press.

Moretti F. (2013) *Distant Reading.* London: Verso.

Rowlingson B. (2015). *Spatial Segregation Analysis.* Personal Communication.

Rupp C.J., Rayson P., Gregory I., Hardie A., Joulain A. and Hartmann D. (2014). Dealing with heterogeneous big data when geoparsing historical corpora. In: *Proceedings of the 2014 IEEE Conference on Big Data.* p. 80–83

Walford N. (2018). Using GIS to transition from contemporary to historical geographical research: exploring rural land use change in southern England in the twentieth century. In: Gregory, I.N., DeBats D. and Lafreniere D. (eds), *The Routledge Companion to Spatial History.* Abingdon: Routledge, p. 394–413.

Waugh E. (1860). *Over the Sands to the Lakes.* Manchester: Alexander Ireland & Co.

Waugh E. (1861). *Rambles in the Lake Country and Its Borders.* London: Whittaker & Co.

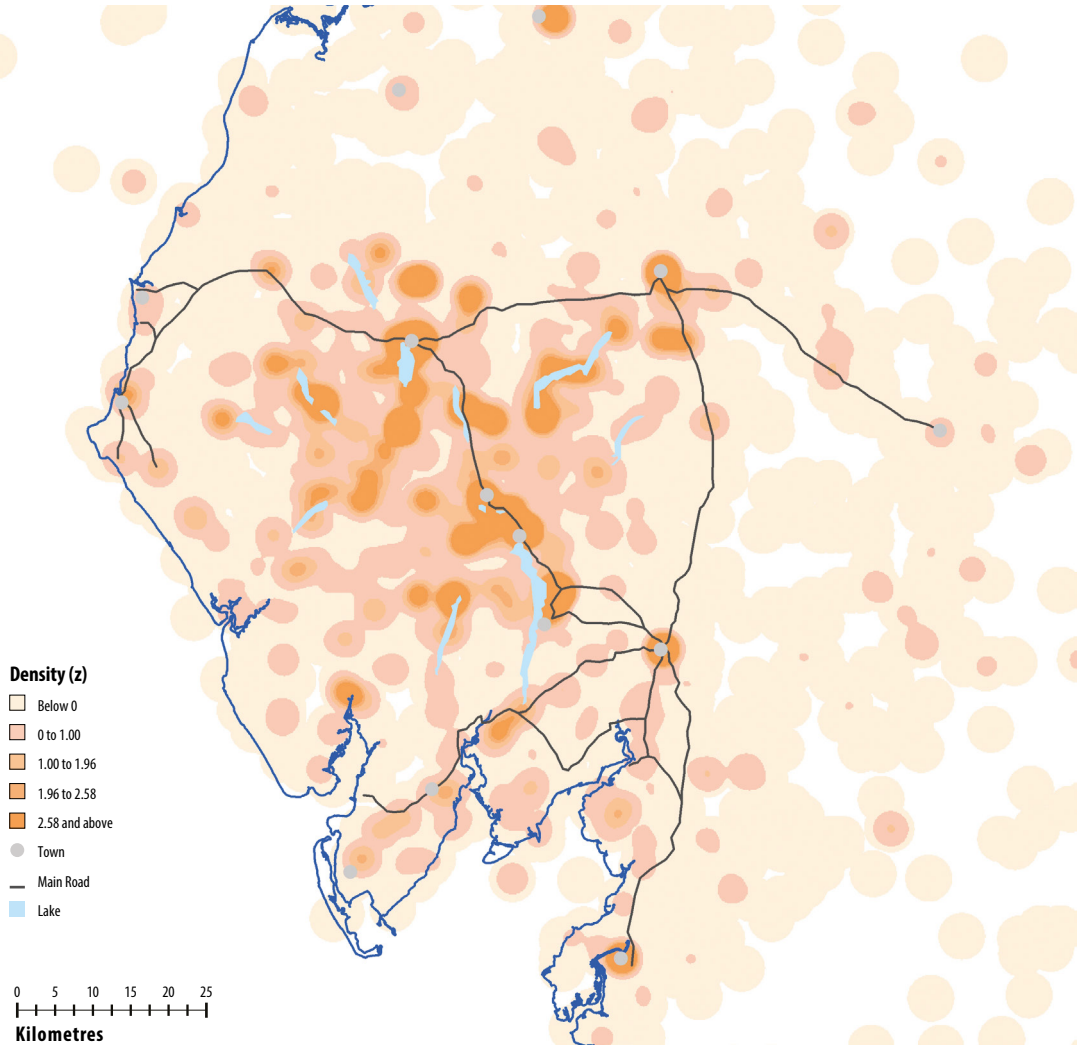Wordsworth W. (1820). *The River Duddon, Vaudracour and Julia, and Other Poems.* London.

For further reading

Donaldson Ch., Gregory I.N. and Murretia-Flores P. (2015). Mapping Wordsworthshire: A GIS Study of Literary Tourism in Victorian Lakeland. *Journal of Victorian Culture* 20 (3), p. 287–307.

Donaldson Ch., Gregory I.N. and Taylor J.E. (2017). Locating the Beautiful, Picturesque, Sublime and Majestic: Spatial Analysing the Application of Aesthetic Terminology in Descriptions of the English Lake District. *Journal of Historical Geography* 56, p. 43–60.

Rayson P., Reinhold A., Butler J., Donaldson Ch., Gregory I. and Taylor J. (2017). A Deeply Annotated testbed for Geographical Text Analysis: The Corpus of Lake District Writing. In: *GeoHumanities'17. Proceedings of the 1st ACM SIGSPATIAL Workshop on Geospatial Humanities.* Redondo Beach, CA, p. 9–15.

Maps

Map 1: *Place-names Mentioned in the Corpus of Lake District Writings in the North-west of England.*

The map started as a point map of the locations of all of the place-names within the Corpus of Lake District writing. These points were density smoothed to highlight the places that are named most frequently in the corpus (shown in dark shading) and contrast these with areas that are rarely mentioned (shown in yellow). The map identifies major clusters in and around the larger settlements, including Lancaster, Kendal, Penrith, Bowness, Ambleside, Grasmere, and Keswick. There are also pronounced clusters near the major mountains (particularly Skiddaw, Helvellyn, Scawfell, the Old Man of Coniston) and lakes (Windermere, Coniston Water, Ullswater, and Derwent Water).

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

Density (z)
- Below 0
- 0 to 1.00
- 1.00 to 1.96
- 1.96 to 2.58
- 2.58 and above
- Town
- Main Road
- Lake

0 5 10 15 20 25
Kilometres

Map 1: Place-names Mentioned in the Corpus of Lake District Writings in the North-west of England.

Map 2a: *Density Smoothed Map of PNCs for 'Beautiful'.*
Map 2b: *Kulldorff Spatial Scan Statistic Map of PNCs for 'Beautiful'.*

Place-name co-occurrences (PNCs) are found when a place-name appears within 10 words of the search-term, in this case the word 'beautiful'. The resulting pattern is then density smoothed to create the pattern shown in Map 2a. This shows that the places described as 'beautiful' are clustered in and around Keswick and Derwent Water in the north-west, Grasmere and Ambleside on the road south from Keswick, and Windermere further south again. This pattern could, however, be a result of the fact that people are writing about these places more than other places. To test for this, Map 2b uses Kulldorff's Spatial Scan Statistic to identify places where there are more 'beautiful' PNCs than would be expected given the overall distribution of place-names shown in Map 1 (termed hot spots), and fewer than would be expected (cold spots). This shows that Grasmere is a hotspot and so is the area to the west of Keswick and Derwent Water. The Grizedale Forest area, in the south of the Lake District, is also a hotspot, suggesting that although this area is not discussed often, when it is it is frequently described as 'beautiful.' The area around the Western Fells is a cold spot, meaning that it is not described as 'beautiful' as often as would be expected.

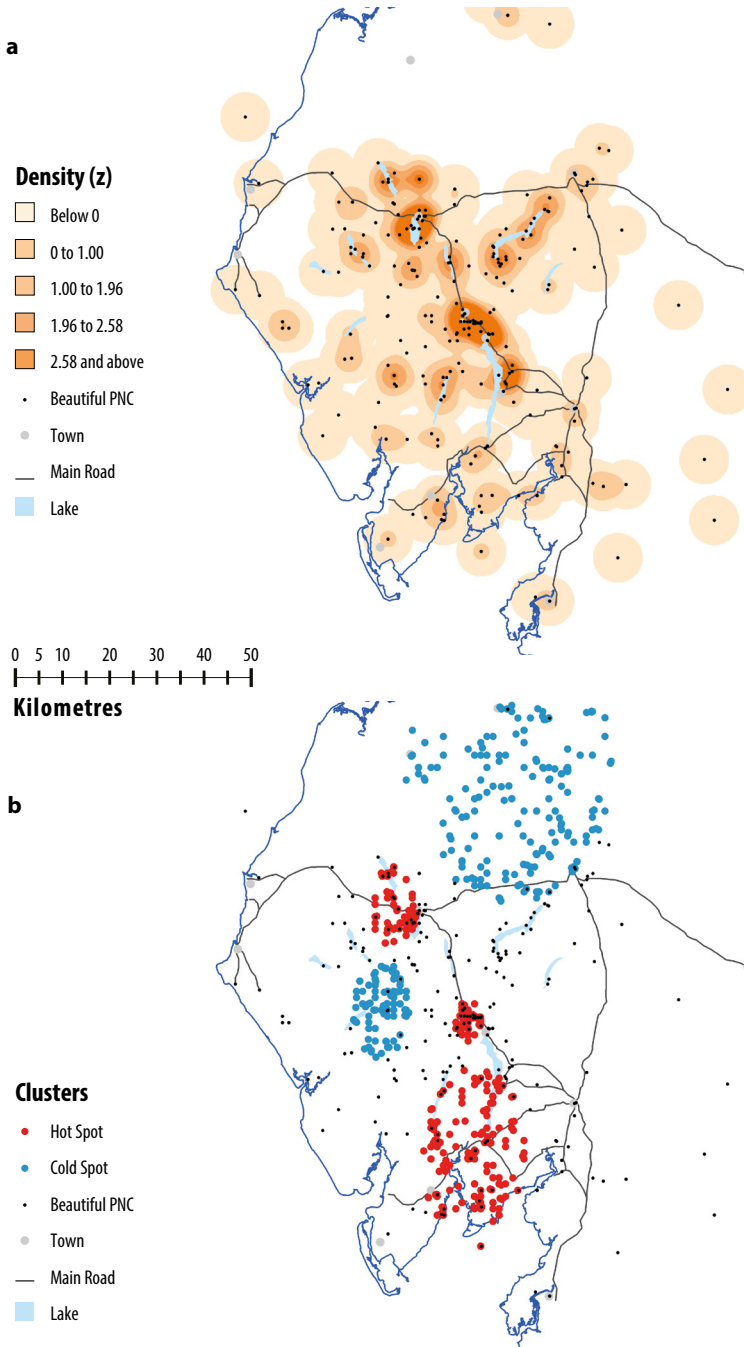**Map 3a:** *Density Smoothed Map of PNCs for 'Sublime'.*
Map 3b: *Kulldorff Spatial Scan Statistic Map of PNCs for 'Sublime'.*

These maps were created in the same way as the maps in Map 2. They show density clusters around Keswick, Ullswater, Windermere, Coniston, and Scawfell. The Kulldorff map shows that only the Western Fells around Scawfell are a hotspot.
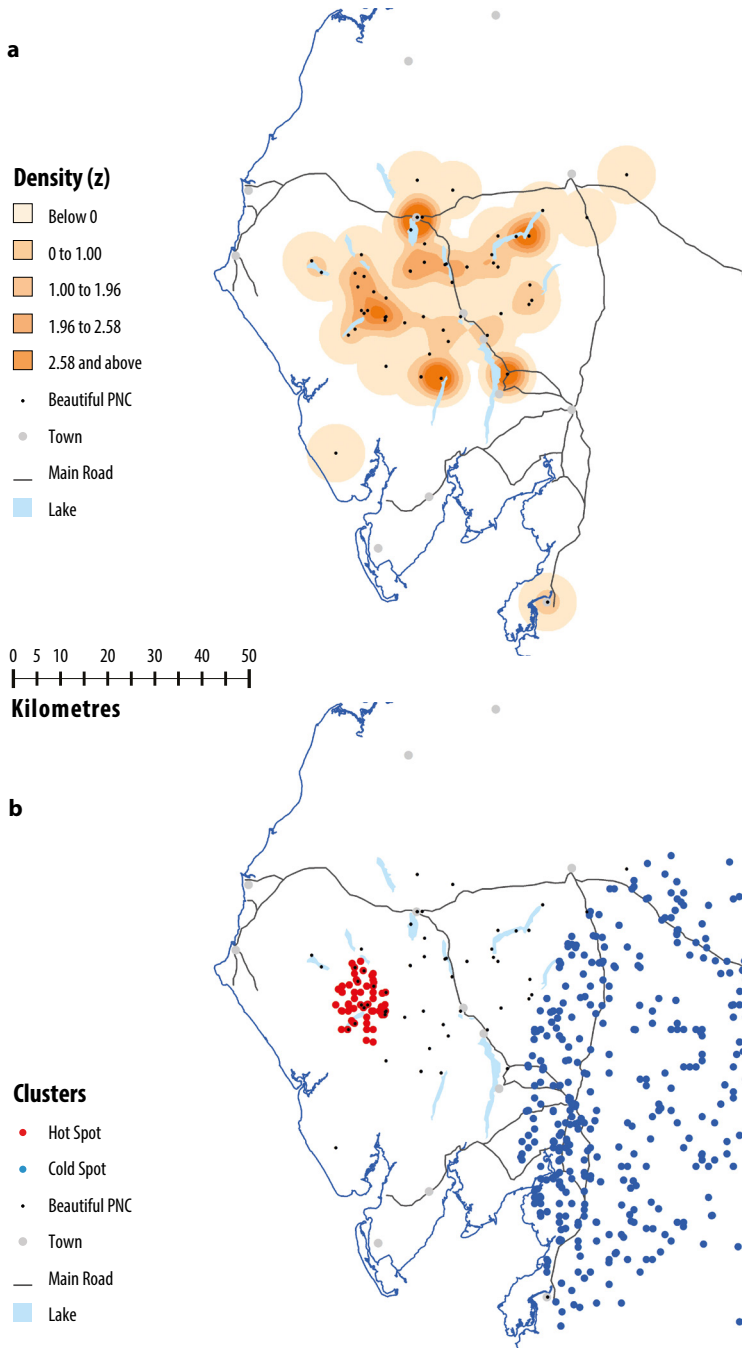
Map 4a: *Density Smoothed Map of PNCs for 'Majestic'.*
Map 4b: *Kulldorff Spatial Scan Statistic Map of PNCs for 'Majestic'.*
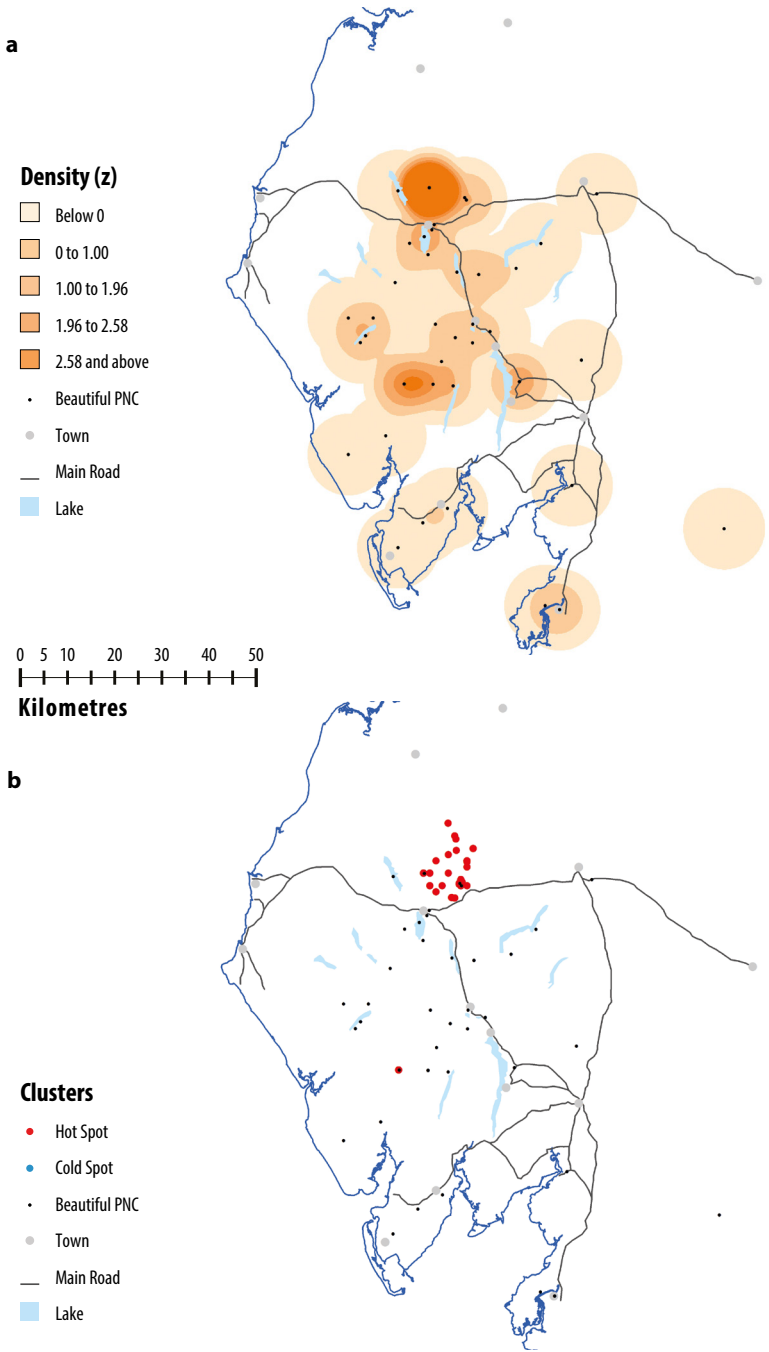
In this case, the density clusters and Kulldorff hot spots are found in two similar locations, namely Skiddaw in the north and the Duddon Valley south of there.

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

a

Density (z)

☐ Below 0
☐ 0 to 1.00
☐ 1.00 to 1.96
☐ 1.96 to 2.58
☐ 2.58 and above
· Beautiful PNC
● Town
— Main Road
▨ Lake

0 5 10   20   30   40   50
Kilometres

b

Clusters

● Hot Spot
● Cold Spot
· Beautiful PNC
● Town
— Main Road
▨ Lake

Map 2: Map of PNCs for 'Beautiful'. a: Density Smoothed Map; b: Kulldorff Spatial Scan Statistic Map.

a

Density (z)

□ Below 0

□ 0 to 1.00

□ 1.00 to 1.96

□ 1.96 to 2.58

□ 2.58 and above

· Beautiful PNC

○ Town

— Main Road

▨ Lake

0 5 10   20   30   40   50
├─┼─┼─┼─┼─┼─┼─┼─┼─┼─┤
Kilometres

b

Clusters

● Hot Spot

● Cold Spot

· Beautiful PNC

○ Town

— Main Road

▨ Lake

Map 3: Map of PNCs for 'Sublime'. a: Density Smoothed Map; b: Kulldorff Spatial Scan Statistic Map.

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

a

Density (z)

□ Below 0
□ 0 to 1.00
□ 1.00 to 1.96
□ 1.96 to 2.58
■ 2.58 and above
· Beautiful PNC
● Town
— Main Road
▪ Lake

0 5 10   20   30   40   50

Kilometres

b

Clusters

● Hot Spot
● Cold Spot
· Beautiful PNC
● Town
— Main Road
▪ Lake

Map 4: Map of PNCs for 'Majestic'. a: Density Smoothed Map; b: Kulldorff Spatial Scan Statistic Map.

a

Clusters
- All
- Beautiful, majestic
- Beautiful, sublime
- Majestic, sublime
- Beautiful
- Majestic
- Sublime
- Town
- Main Road
- Lake

0 5 10   20   30   40   50
Kilometres

b

Clusters
- Majestic
- Sublime
- Beautiful
- Town
- Main Road
- Lake
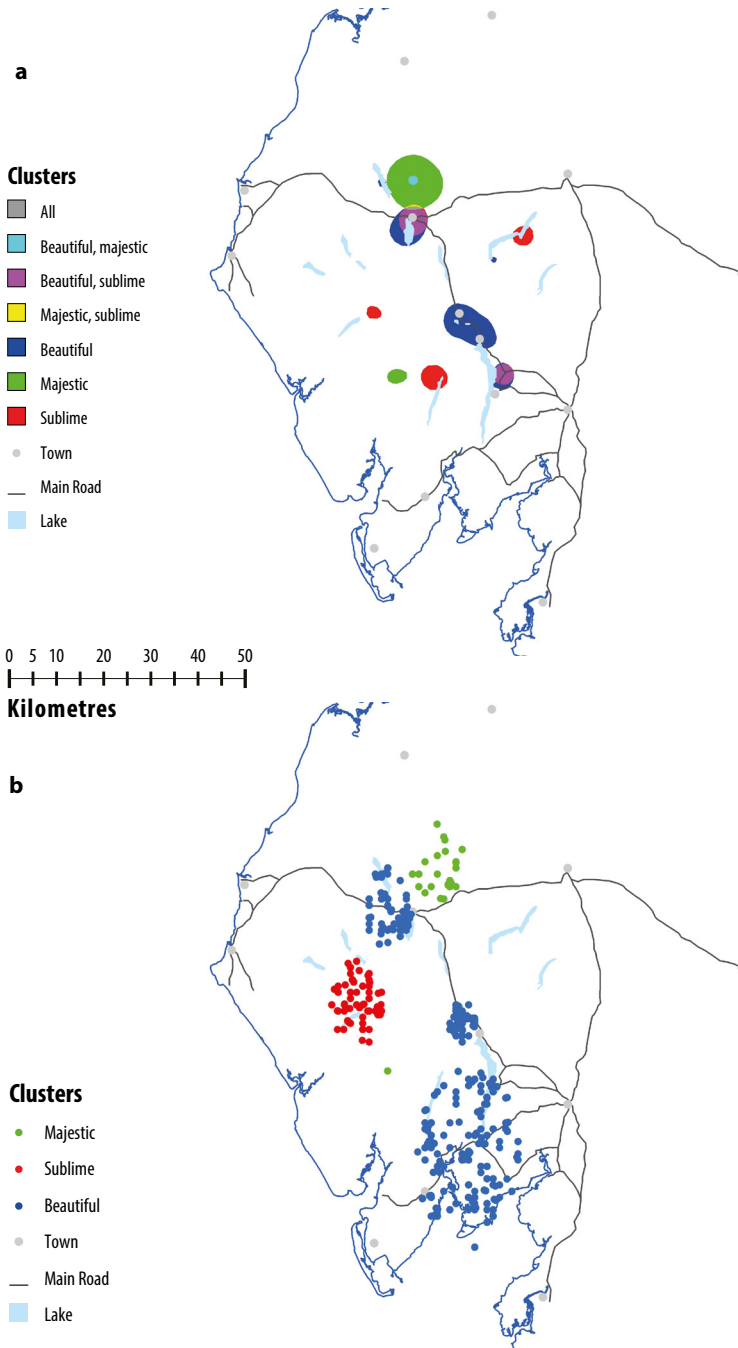
Map 5: Map of PNCs for 'Beautiful', 'Sublime' and 'Majestic'; a: Density Smoothed Map; b: Kulldorff Spatial Scan Statistic Map.

Ian Gregory, Christopher Donaldson & Joanna E. Taylor

Map 5a: *Density Smoothed Map of PNCs for 'Beautiful', 'Sublime' and 'Majestic'.*
Map 5b: *Kulldorff Spatial Scan Statistic Map of PNCs for 'Beautiful', 'Sublime' and 'Majestic'.*

Comparing the patterns of all three search-terms using density smoothing and Kulldorff Spatial Scan Statistic. The density smoothed map [Map 5a] defines clusters as areas with a z-score of over 2.58, while the Kulldorff map [Map 5b] uses statistically significant hot spots. The two maps show that there are very few overlaps between the places described using the three search terms. 'Beautiful' and 'sublime' do have density clusters around Keswick/Derwent Water and Windermere, but these are the only major areas of overlap. There is no overlap on the Kulldorff hotspots.

Map 6a: *Spatial Segregation Analysis Comparing PNCs for 'Beautiful' for the 'Long' Eighteenth Century and the Romantic Eras.*
Map 6b: *Spatial Segregation Analysis comparing PNCs for 'Beautiful' for the Romantic and Victorian Eras.*

Spatial Segregation Analysis comparing PNCs for 'beautiful' for the 'long' eighteenth century and the Romantic era [Map 6a], and the Romantic and Victorian eras [Map 6b]. The techniques identify areas that have statistically significantly more PNCs from one period than the other and highlights them using contour lines. The maps suggest that Grasmere became increasingly associated with 'beautiful' in the Romantic era and that there was a more general rise in describing places in the southern and western parts of the Lake District as 'beautiful' over time.

Map 7a: *Spatial Segregation Analysis Comparing PNCs for 'Sublime' for the 'Long' Eighteenth Century and the Romantic Eras.*
Map 7b: *Spatial Segregation Analysis Comparing PNCs for 'Sublime' for the Romantic and Victorian Eras.*

Spatial Segregation Analysis comparing PNCs for 'sublime' for (a) the 'long' eighteenth century and the Romantic era, and (b) the Romantic and Victorian eras. These maps also show a move to the south and west over time.
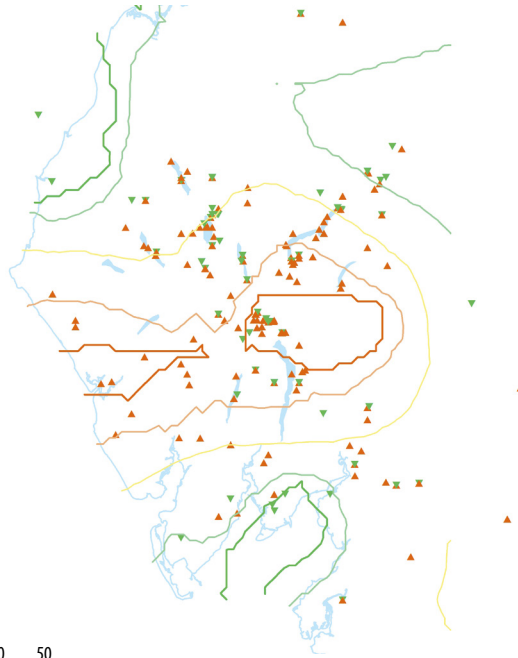
a

Significance

— Long C18 p<.01
— Long C18 p<.05
— Equal
— Romantic p<.05
— Romantic p<.01

PNC

▼ Long C18
▲ Romantic Era
▆ Lake

0 5 10   20   30   40   50

Kilometres
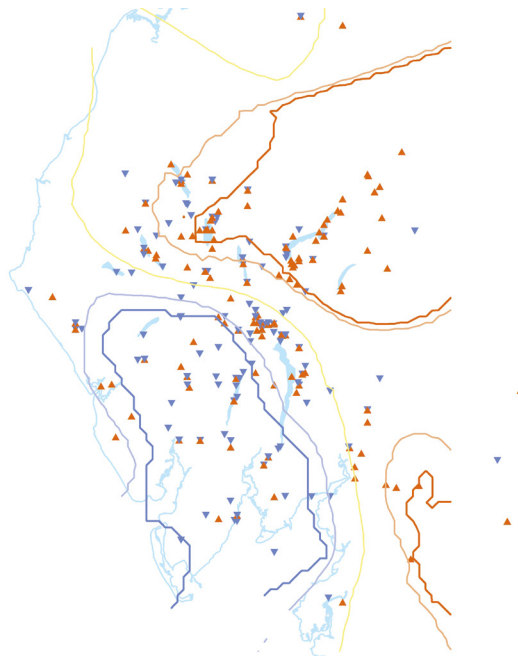
b

Significance

— Romantic p<.01
— Romantic p<.05
— Equal
— Victorian p<.05
— Victorian p<.01

PNC

▲ Romantic Era
▼ Victorian Era
▆ Lake

Map 6: Spatial Segregation Analysis Comparing PNCs for 'Beautiful'. a: For the 'Long' Eighteenth Century and the Romantic Era; b: For the Romantic and Victorian Eras.
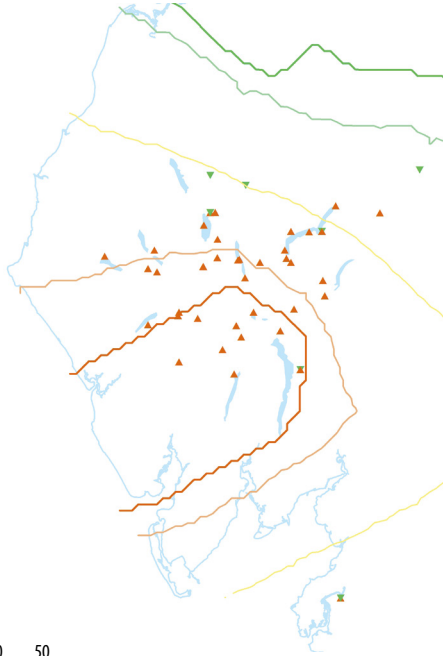
Ian Gregory, Christopher Donaldson & Joanna E. Taylor

a

Significance
— Long C18 p<.01
— Long C18 p<.05
— Equal
— Romantic p<.05
— Romantic p<.01

Period
▼ Long C18
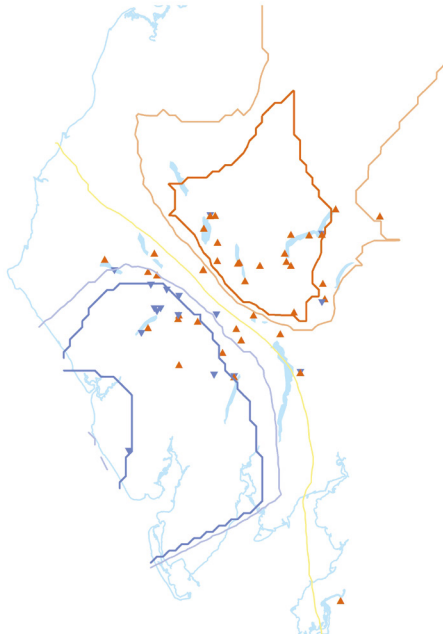▲ Romantic Era
■ Lake

0 5 10 20 30 40 50

Kilometres

b

Significance
— Romantic p<.01
— Romantic p<.05
— Equal
— Victorian p<.05
— Victorian p<.01

Period
▲ Romantic Era
▼ Victorian Era
■ Lake

Map 7: Spatial Segregation Analysis Comparing PNCs for 'Sublime'. a: For the 'Long' Eighteenth Century and the Romantic Eras; b: For the Romantic and Victorian Eeras.