Spatial Variations in School Performance: a Local Analysis Using Geographically Weighted Regression

A. STEWART FOTHERINGHAM, MARTIN E. CHARLTON & CHRIS BRUNSDON

ABSTRACT   *In Britain, the performance of all state primary schools is assessed by students' attainment levels in a set of standardized tests administered to pupils at the ages of 7 and 11 (the so-called Key Stages 1 and 2, respectively). These data are analysed for 3687 schools in northern England. In particular, school performance is linked to the number of students taking the test at each school and to various socio-economic indicators of the estimated school catchment areas. The latter are based on a geographical weighting function that links census data, an areal coverage, to school locations, a point coverage. Following a traditional global regression analysis, spatial variations in the relationships are examined with geographically weighted regression (GWR) to reveal some interesting geographical variations in the results.*

Introduction

The last ten years have seen a radical shift in parental power in education within Britain with parents now being encouraged to 'shop around' for the best schools for their children. It is now possible in theory for parents to select the state school which they would like their children to attend and the government encourages this selection process by publishing league tables of scholastic achievement by pupils at each state-run school. In practice, however, this inevitably collapses to parents being able to express a preference for a particular school with no guarantee that that school will be able to provide a place because the more popular schools quickly reach their capacity. In fact most schools, particularly primary ones, faced with capacity limitations, give preference to children living close to the school. The net result is that school catchments are still strongly geographically based around each school. However, as a result of the theoretical situation in which parents actually select a state primary school for their children, as opposed to the children just going to the nearest one, state schools are evaluated on a standard set of attainment criteria

A. Stewart Fotheringham, Professor of Quantitative Geography, Department of Geography, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK. E-mail: Stewart.Fotheringham@ncl.ac.uk
Martin E. Charlton, Lecturer in GIS, Department of Geography, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK.
Chris Brunsdon, Reader in Spatial Analysis, Department of Geography, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK.

across the country. These attainment levels are then used to produce 'league tables' of school performance, which can be used by parents as a basis for comparing schools. Obviously, the data on school performance exhibit some very interesting geographical variations and it is a politically 'hot' issue as to what the causes of such variations might be. For instance, does school size affect performance? To what extent is scholastic achievement a product of environment? Are there areas where school performance is consistently below average, and if so, what socio-economic characteristics do such areas have? Is it possible to identify schools that are performing well relatively to their intake of pupils? Clearly these are all very sensitive issues for a government whose stated aim is the elimination of inequalities (see, for example, *The Times* 07/12/98). It is a particularly sensitive issue if it appears that schoolchildren aged between 4 and 11 are being disadvantaged so early in their scholastic careers because of their environment.

With this as background, this paper examines the relationship between school performance and the socio-economic characteristics of school catchment areas. Such an analysis is not new and examples of attempts to discover relationships between school performance and catchment area characteristics can be found in, *inter alia*, Brown *et al*. (1998), Conduit *et al*. (1996), Coombes and Raybould (1997) and McCallum (1996). However, the emphasis here is not so much to determine whether or not relationships between school performance and catchment area characteristics exist, but to determine if there are any interesting spatial variations in these relationships. That is, perhaps some attributes of school catchment areas have an effect on school performance in some areas and not in others and such variations are masked in global results. We undertake a local analysis using the recently developed statistical technique of geographically weighted regression (GWR) (Brunsdon *et al*., 1996; Fotheringham *et al*., 1997, 1998). We begin by defining the data. The difficult issue of estimating school catchment areas is then described. This involves linking areal-based census attributes to the point-based locations of schools. Following this, we describe the results of a global regression model relating school performance to socio-economic characteristics of the estimated school catchment areas and to the numbers of pupils in each school taking the tests. After pointing out the possible flaws in this procedure, we describe the results of a local analysis using GWR which indicates the spatial variations in the determinants of school performance hidden by the global model.

The School Performance Data

Figure 1 shows the spatial distribution of the 3687 primary schools in northern England for which data on pupils' attainment levels were made available.[1] County boundaries and names are also displayed. The varying density of the schools obviously reflects the spatial distribution of population with heavy concentrations of schools in the metropolitan areas of Newcastle and Teesside in the northeast, Leeds/ Bradford in the south, and Manchester in the southwest.

Three performance indicators are available from standardized tests conducted on the same day across the country in May 1997. For each school: the percentage of pupils reaching or exceeding a pre-defined level of attainment in Maths, English and Science is reported. While there are strong correlations between these three scores (pupils who do well in one subject, tend to do well in the other two), there are also some differences and for this reason we decided not to aggregate the scores. Of the three scores, we have chosen to analyse that of Maths because it exhibits the greatest
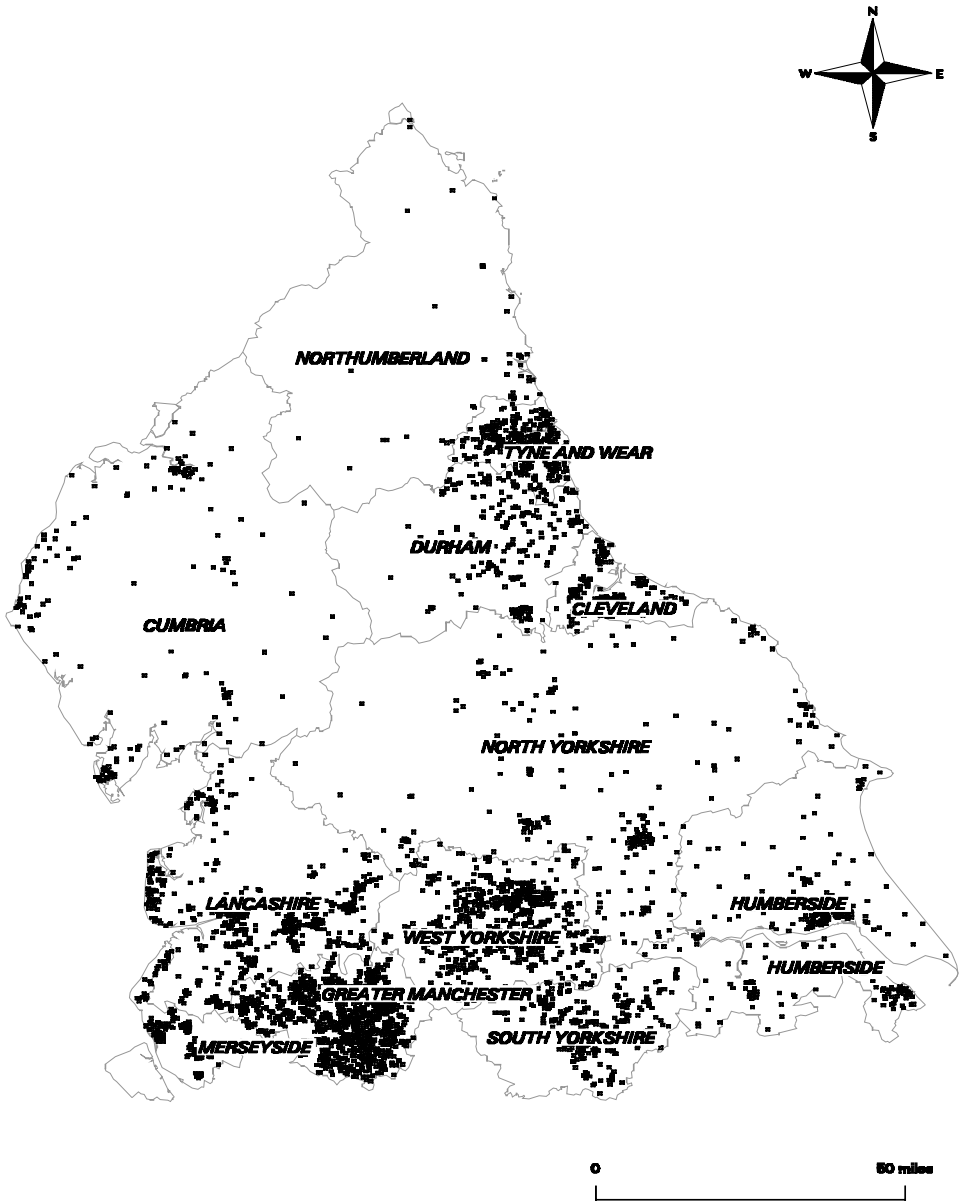
Figure 1. Administrative boundaries and school locations.

variability and it is the one about which there is perhaps the greatest educational concern. The proportion of students reaching the pre-defined level of attainment in Maths is very low in some schools (see Table 1 below).

As mentioned, the scores are the percentages of students eligible to take the test in each school who achieve or exceed the required level of attainment in a particular subject. This is not an ideal indicator of school performance for obvious reasons. It fails to differentiate pupils who are able from those who are very bright and there could, in theory, be quite big differences in pupils' abilities between two schools having the same percentage of students achieving the required level of attainment.

Also, it should be noted that pupils who are absent on the day of the test are counted as having failed and that 'statemented' pupils, those with special educational needs, are also counted in the denominator. While these latter two issues cause some undesirable variability in the scores, they do ensure a certain degree of reliability in the testing procedure with schools unable to 'hide' their poorer students. However, despite these caveats on the school performance data, the data do separate schools where children are performing well from those where children are performing poorly. In addition, the data are available for every state-supported school in the country.

Since the attainment score used for each school is the percentage of students attaining a certain level in Maths, it may be reasonably modelled by a binomial distribution (students either attain the desired level or they do not). Consequently, given the variance of the binomial distribution is $p(1 − p)/n$, with p being the proportion of students at a school who achieve the prescribed standard, the variance of the attainment score will decrease as it approaches 0 or 1. This makes the use of regression analysis, which assumes the error terms to have constant variance, highly questionable. To remove this problem, the attainment scores can be transformed in the following way to produce a variance-stabilized Maths score for each school which can be regressed on a set of independent attributes:

$$TM_i = \arcsin[\text{sqrt}(M_i)] \tag{1}$$

where TM_i is the transformed Maths score and M_i is the raw score (Bartlett, 1936). The effect of the transformation is shown in Figure 2 which shows that the tails of the transformed distribution are stretched and that values of the transformed variable greater than 1.0 are possible.

The spatial distribution of the Maths scores transformed in this way over the study region is shown in Figure 3 where the darker shading indicates the higher transformed score. Clearly, there are variations in this surface with schools in some parts of the country, such as North Yorkshire, generally performing well and schools
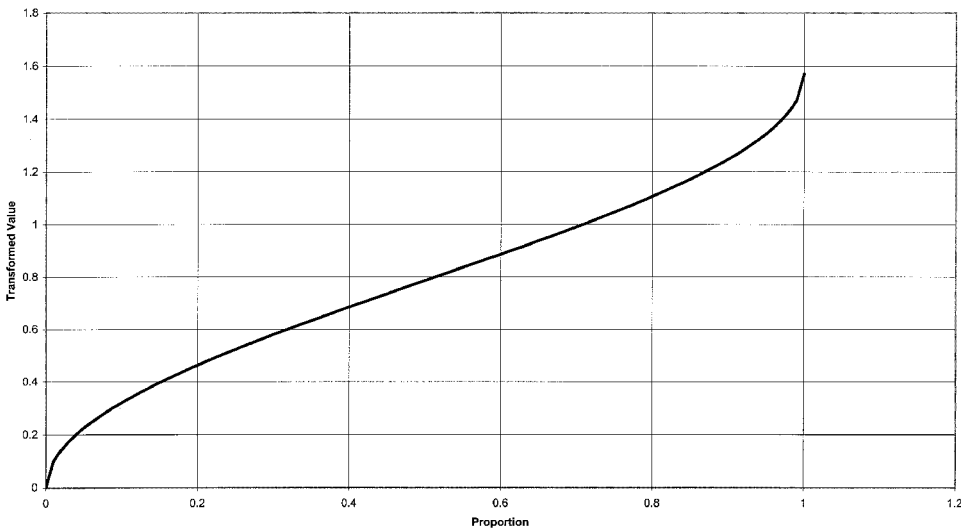

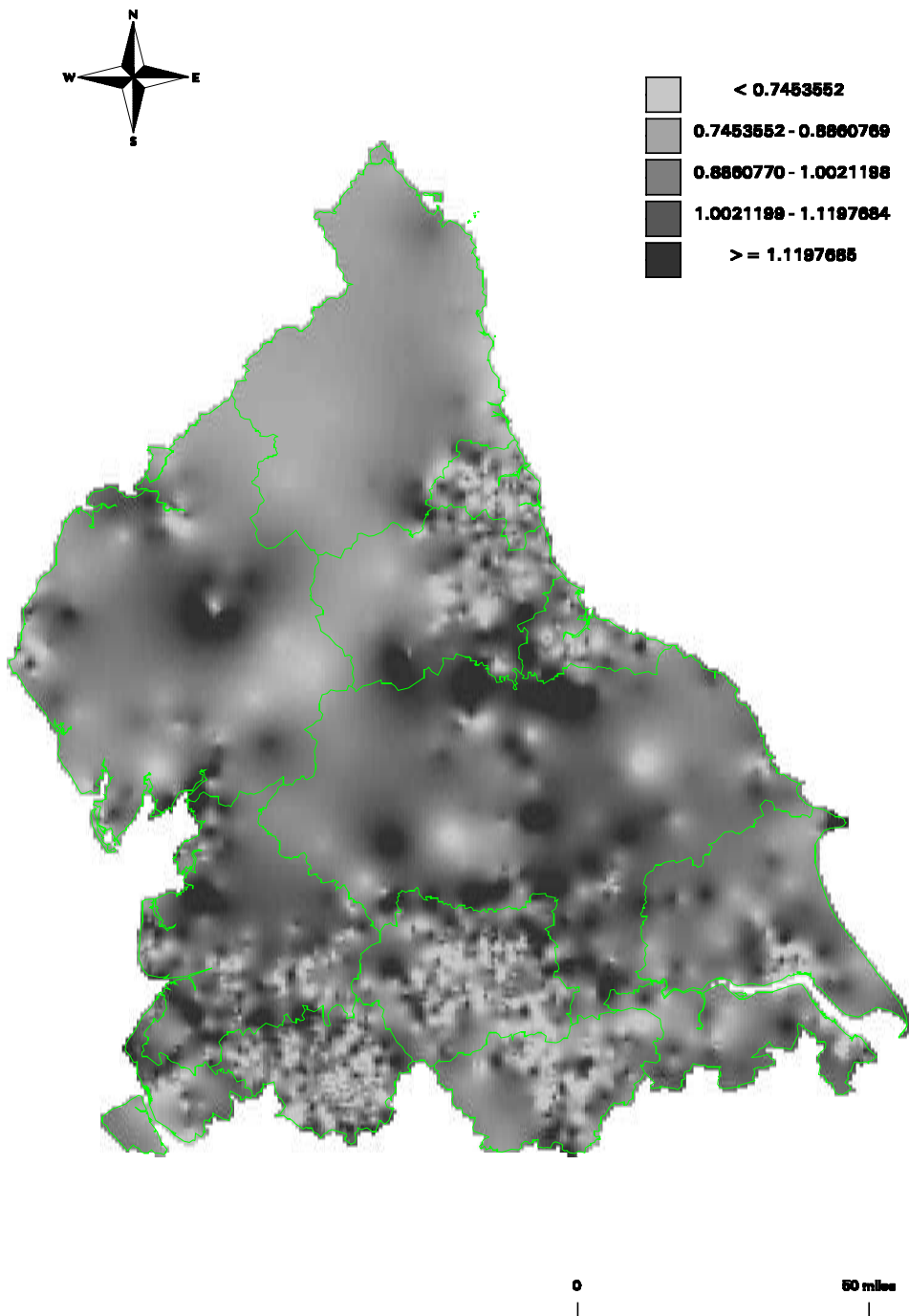**Figure 2.** Arcsin[sqrt(x)] transformation.

Figure 3. Spatial distribution of the transformed Maths score.

in other parts of the country, performing not so well. There tends to be a complicated mosaic of performance in the major metropolitan areas representing perhaps the diversity of population within such areas. It is the task of the remainder of this paper to try to uncover any regularities in the determinants of the spatial distribution of the Maths scores mapped in Figure 3.

The Socio-economic and School Roll Data

The main aim of this paper is to examine the role of various attributes in determining differences in school performance in Maths. One of these attributes is a school size variable (termed the school 'roll') supplied with the school performance data. This is the total number of pupils from each school eligible to take the standardised test.[2] The other attributes used to examine variations in school performance are various socio-economic characteristics of each school catchment area. As the school catchments themselves are unknown (most primary schools probably do not have discrete catchment boundaries), and as the socio-economic data are obtained from the census at the level of enumeration districts, the catchment area data for each school have to be estimated. One way to do this, which is perhaps overly simplistic, is to draw Thiessen polygons around each school and assume that all pupils within a school's polygon attend that school and all pupils outside it attend some other school. Another, more accurate, way of estimating a school's catchment area characteristics is shown in Figure 4. Given that there is likely to be strong distance-decay in primary school attendance (most parents would want their children to travel relatively short distances to primary school and most schools will give preference in admittance to pupils who live nearby), it is reasonable to assume that a school will attract most of its pupils from nearby enumeration districts. Schools are less likely to attract pupils from enumeration districts that are long distances away. However, distance is not the sole criterion by which pupil's choose schools or by which schools define catchment areas and there is a growing proportion of students who, for one reason or another, do not attend the nearest school to them. Therefore, it seems reasonable to estimate
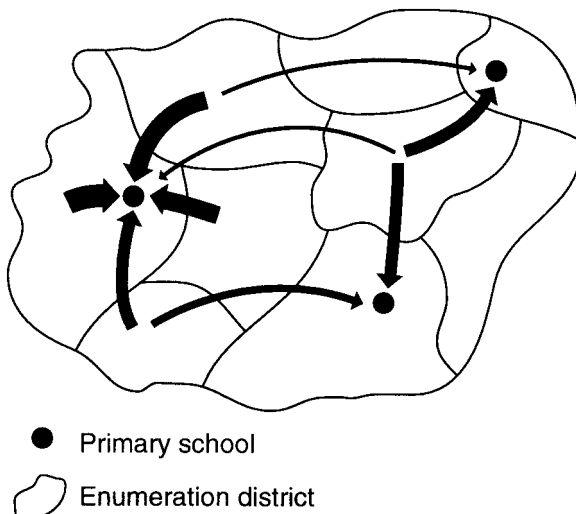

● Primary school
◇ Enumeration district

Figure 4. Allocation of enumeration district characteristics to schools.

the characteristics of a school's catchment area from a distance-weighted function of enumeration district characteristics in which the characteristics of enumeration districts closest to a school are given greatest weight in determining the catchment area characteristics. In this instance, the function used to derive the distance-weighted catchment area characteristic of school i, CA_i, is

$$CA_i = \sum_{j}^{N} w_{ij} ED_j \qquad (2)$$

where ED_j is the value of the socio-economic characteristic for enumeration district j, N is the number of enumeration districts and w_{ij} is a spatial weighting function.[3] In this case, the number of enumeration districts is 28 858 and the weighting function is defined by

$$w_{ij} = \exp[-(d_{ij}/h_i)^2] \qquad (3)$$

where d_{ij} represents the distance between school i and enumeration district j and h_i is the bandwidth of the spatial weighting function for school i. As the bandwidth decreases, the school catchment areas become more localized. One problem in applying this ED-to-catchment area transformation is to select an appropriate bandwidth. If the bandwidth is too large, the catchment areas become too similar (in the limit as the bandwidth tended to infinity, all schools would have the same catchment profile). Conversely, if the bandwidth is too small, the catchments would be too localized and reflect only the conditions in the immediate vicinity of each school. Reflecting the current situation in the UK where the majority of primary children still attend their nearest school but where an increasing proportion do not, we selected a bandwidth range between 0.5 and 2 km. In an attempt to reflect reality further, the bandwidth was made a function of school size with larger schools having larger bandwidths, simulating the situation in which larger schools draw pupils from larger catchment areas. In practice, this could be criticized as some small rural schools have large catchment areas. However, this is not really a problem as the schools in rural areas will tend to be separated by large distances and their predicted catchment areas will still be much larger than those for urban schools. The bandwidth selection has more of an effect on urban schools which are located much closer together.

Using the spatial weighting function defined in equations (2) and (3), various socio-economic characteristics of each school catchment area were estimated. The attributes selected from the 1991 census were as follows:

- *SC1*: the percentage of families where the head of household has a professional or managerial occupation. Given the 1991 Census of Population did not ask any income information, this variable is often used as a proxy for income. High levels of this variable indicate a school having a relatively rich catchment area.
- *CH*: the percentage of households living in state-provided council housing.
- *UN*: the percentage of unemployment.
- *LP*: the percentage of families headed by a single parent.
- *B*: the percentage of black residents.
- *I*: the percentage of Indian residents.
- *C*: the percentage of Chinese residents.

Information on the distributions of these variables across the school catchment

Table 1. Five number summary of the attributes

	Min.	Q1	Median	Q3	Max.
Transformed Maths score	0.22	0.79	0.94	1.08	1.57
School roll	11	25	34	51	181
Social class 1	0	2.3	3.9	6.5	45.0
Council house	0	10.2	18.9	32.3	84.0
Unemployment rate	0	6.4	9.7	14.4	38.1
Lone parents	0	0.9	1.5	2.2	9.3
Black	0	0.1	0.2	0.4	26.0
Indian	0	0.1	0.3	0.9	62.0
Chinese	0	0.2	0.3	0.6	7.7

Refer to text for full definitions of variables.

areas is given in Table 1 which displays a typical five-number summary for each variable showing the minimum and maximum values across the 3687 schools, the median and the lower and upper quartiles. The variation in the transformed Maths score indicates some large differences in school performance across the region. The value ranges from a low of 0.23 to a high of 1.57 and a median of 0.94. Looking at the transformation in Figure 2, these figures relate to proportions of schoolchildren achieving the set standard ranging from about 0.05 to 1 with a median of approximately 0.7. School size varies from 11 to 181 with the median size being 34 which in most schools will represent a single class (maximum school size in many school districts is 34/35). The percentage of social class 1 residents, the unemployment rate, and the percentage of lone parent families also vary considerably across the catchment areas. The percentage of each of the three ethnic groups (black, Indian and Chinese residents) is generally very low across the whole study areas with just a very small number of school catchment areas having a sizeable ethnic composition.

There is an issue that the school performance data were recorded in 1997 and we are using 1991 census data to describe the schools' catchment areas. However, given the general stability in socio-economic patterns of population in the UK, this is probably of very little consequence: what were relatively deprived areas in 1991 will most definitely have remained relatively deprived areas in 1997.

Global Regression Results

In order to investigate whether there are any environmental factors which might explain, in part, the spatial variation in transformed Maths scores shown in Figure 3, the following regression model was calibrated by weighted least squares regression with data on all 3687 schools.

$$TM_i = \alpha_0 + \alpha_1 SR_i + \alpha_2 SC1_i + \alpha_3 CH_i + \alpha_4 UN_i + \alpha_5 LP_i + \alpha_6 B_i + \alpha_7 I_i + \alpha_8 C_i \quad (4)$$

where the values of α are parameters to be estimated and SR_i represents the school roll (the number of pupils in each school within the age band for which the test is designed). The other variables are defined above. Weighted least squares regression is used here to calibrate the global model with weights equal to $1/SR_i$ because the variance of the scores is likely to be an inverse function of the school roll: the scores for smaller schools will be more sensitive to absentees and special needs students. As

Table 2. Global regression results

Variable	Parameter estimate	t
Intercept	1.11	94.7
SR	− 0.0013	− 6.5
SC1	0.0060	6.9
CH	− 0.0029	− 7.4
UN	− 0.0035	− 2.9
LP	− 0.0234	− 4.5
B	0.0013	0.4
I	− 0.0074	− 7.7
C	0.0133	1.3

mentioned earlier, the denominator of the school performance indicator includes both students with special needs and who were absent on the day of the test. Calibration of this model by weighted least squares yields the results shown in Table 2. It appears that schools with good performances in Maths are characterized by: low numbers of pupils; catchments with high percentages of people in professional and managerial occupations; low percentages in council housing; low percentages of Indian residents; low unemployment rates; and low percentages of lone parent households.

However, the model explains only 24% of the variance in the transformed Maths scores and there are clearly many other determinants of school performance not accounted for in the model. The more obvious variables missing from the model are those related to the school such as the average class size, the degree of parental involvement, resources, the number of special needs pupils, and the level of truancy. Also missing are attributes which are difficult to quantify such as the quality of the teaching. To some extent, the absence of these variables from the model is exacerbated by the fact that the model being calibrated is a global one that is assumed to apply equally to *all* parts of the region. In fact, it may not apply to *any* part of the region. The estimated parameters represent global averages of processes that might exhibit a substantial degree of spatial variation. We now examine the data with GWR which allows us to calibrate the model given in equation (4) locally rather than globally so that we can investigate the nature of any spatial non-stationarity in relationships (Brunsdon *et al.*, 1996; Fotheringham *et al.*, 1997, 1998). This also provides us with a great deal more information on the determinants of school performance.

Geographically Weighted Regression

Consider a general form of linear regression model given by

$$y_i = \alpha_0 + \Sigma_k a_k x_{ik} + \varepsilon_i \tag{5}$$

of which the school performance model in equation (4) is a particular example. In the calibration of this model, one parameter is estimated for the relationship between each independent variable and the dependent variable and this relationship is assumed to be constant across the study region. GWR is a relatively simple technique that extends the traditional regression framework by allowing local rather than global parameters to be estimated. The philosophies behind local and global

modelling are discussed by Fotheringham and Brunsdon (1999) and more details and examples of local modelling approaches are given by Fotheringham (1997) and Fotheringham and Brunsdon (1999). The GWR model equivalent to equation (5) is

$$y_i = a_{0i} + \Sigma_k a_{ki} x_{ik} + \varepsilon_i \tag{6}$$

where a_{ki} represents the value of a_k at point i (Brunsdon *et al.*, 1996; Fotheringham *et al.*, 1998).

In the calibration of the GWR model it is assumed that observed data near to point i have more of an influence in the estimation of the values of a_{ki} than do data located farther from i. In essence, the equation measures the relationships inherent in the model *around each point i*. Hence, in GWR an observation is weighted in accordance with its proximity to point i so that the weighting of an observation varies with i. Data from observations close to i are weighted more than data from observations farther away. The GWR estimator is

$$\hat{\mathbf{a}}_i = (\mathbf{X}^t \mathbf{W}_i \mathbf{X})^{-1} \mathbf{X}^t \mathbf{W}_i \mathbf{y} \tag{7}$$

where the bold type denotes a matrix, $\hat{\mathbf{a}}_i$ represents an estimate of $\mathbf{a}_i$, the place-specific parameters and $\mathbf{W}_i$ is an n by n matrix whose off-diagonal elements are zero and whose diagonal elements denote the geographical weighting of observed data *for point i*. That is

$$\mathbf{W}_i = \begin{matrix} w_{i1} & 0 & 0 & \dots & 0 \\ 0 & w_{i2} & 0 & \dots & 0 \\ 0 & 0 & w_{i3} & \dots & 0 \\ . & . & . & \dots & . \\ 0 & 0 & 0 & \dots & w_{in} \end{matrix} \tag{8}$$

where w_{in} denotes the weight of the data from school n on the calibration of the model around point i (in this case the point i will be a school but it need not be a point at which data are collected). The weights are defined as continuous functions of distance so that the closer a data point is to the calibration point, the greater is its weight in the estimation of the parameters for that calibration point. An alternative weighting procedure is a discrete one in which all data points within a prescribed distance from i are given a weight of 1 and all points beyond this distance are given a weight of 0. However, this seems an unrealistic representation of most spatial processes that are usually continuous. The selection of a particular continuous weighting function does not appear to be very important (Fotheringham *et al.*, 1998) and in this instance the weighting function selected is a Gaussian one so that

$$w_{ij} = \exp(d_{ij}^2 / h^2) \tag{9}$$

where d_{ij} is the distance between school i and the data point j and h is a bandwidth that affects the distance-decay of the weighting function. Too large a bandwidth will produce a flat surface with little spatial variation and too small a bandwidth will result in estimation problems with some of the local regressions. Brunsdon *et al.* (1996) and Fotheringham *et al.* (1997, 1998) describe how the bandwidth can be calibrated.

It is perhaps useful to emphasize the fact that the application of GWR to the global model means that the data are in fact weighted twice. The local version of the global model is also weighted by the inverse of the school roll to account for the increased variance of scores from small schools. This WLS model is then geographically weighted by GWR to allow for the local spatial effects to be measured.

Local Regression Results

Instead of producing a single global average parameter estimate for each relationship, GWR produces a set of local parameter estimates that can be mapped. For instance, with a bandwidth of 15 km, the GWR surfaces for the school roll and social class parameters are shown in Figures 5 and 6, respectively. The global parameter estimate for the relationship between school parameter and school size (roll) is significantly negative and most of the local parameter estimates are negative. However, there are parts of the region, particularly in the northwest, where there appears to be a positive relationship between school size and performance (the darker shaded parts of the map in Figure 5). Similarly, although the global relationship between school performance and the percentage of professionals in the school catchment area (*SC1*) is significantly positive, the GWR surface in Figure 6 indicates that there appear to be some interesting spatial variations in this relationship. The darker shaded areas of this map indicate parts of the region where school performance is particularly sensitive to variations in *SC1*.

However, although the spatial distributions of the raw parameter estimates are interesting, the values represented on these surfaces do not take into account the standard errors of the estimates. Consequently, it is often more illuminating to view surfaces of *t*-statistics by dividing each local estimate by the corresponding local standard error of the estimate. These *t*-surfaces are useful, not in a formal sense, but in a purely exploratory role, to highlight parts of the map where interesting relationships appear to be occurring.

The *t*-surfaces for each of the local parameters in equation (4) are shown in Figures 7–15. Although these surfaces generally depict the relationship shown in the global model, there are some interesting spatial variations that would be missed completely if we relied solely on the global analysis. For instance, the *t*-surface for the relationship between school performance and school roll is generally negative with a positive region in the northwest. This was also noted in the interpretation of the raw local parameter surface above. However, the *t*-surface suggests that there is something interesting in the relationship between school score and school roll in two metropolitan areas: Newcastle in the northeast and Leeds/Bradford in the south. In both these areas the *t*-values are strongly negative and there appears to be a strong inverse relationship between school performance and size. The GWR analysis has therefore uncovered something that was hidden in the global analysis and has raised the interesting question of why the relationship is strongly negative in these two areas and nowhere else in the region.

Similarly, the *t*-surface for the relationship between school performance and the percentage of people in the school catchment area employed in professional and managerial occupations (*SC1*) is generally weakly positive with one or two areas exhibiting a weakly negative relationship. However, there is a strong positive relationship between these two variables in the south of the region corresponding almost exactly to the Manchester metropolitan area. Again, the local analysis highlights a facet of the relationship completely hidden in the global results and forces us to ask
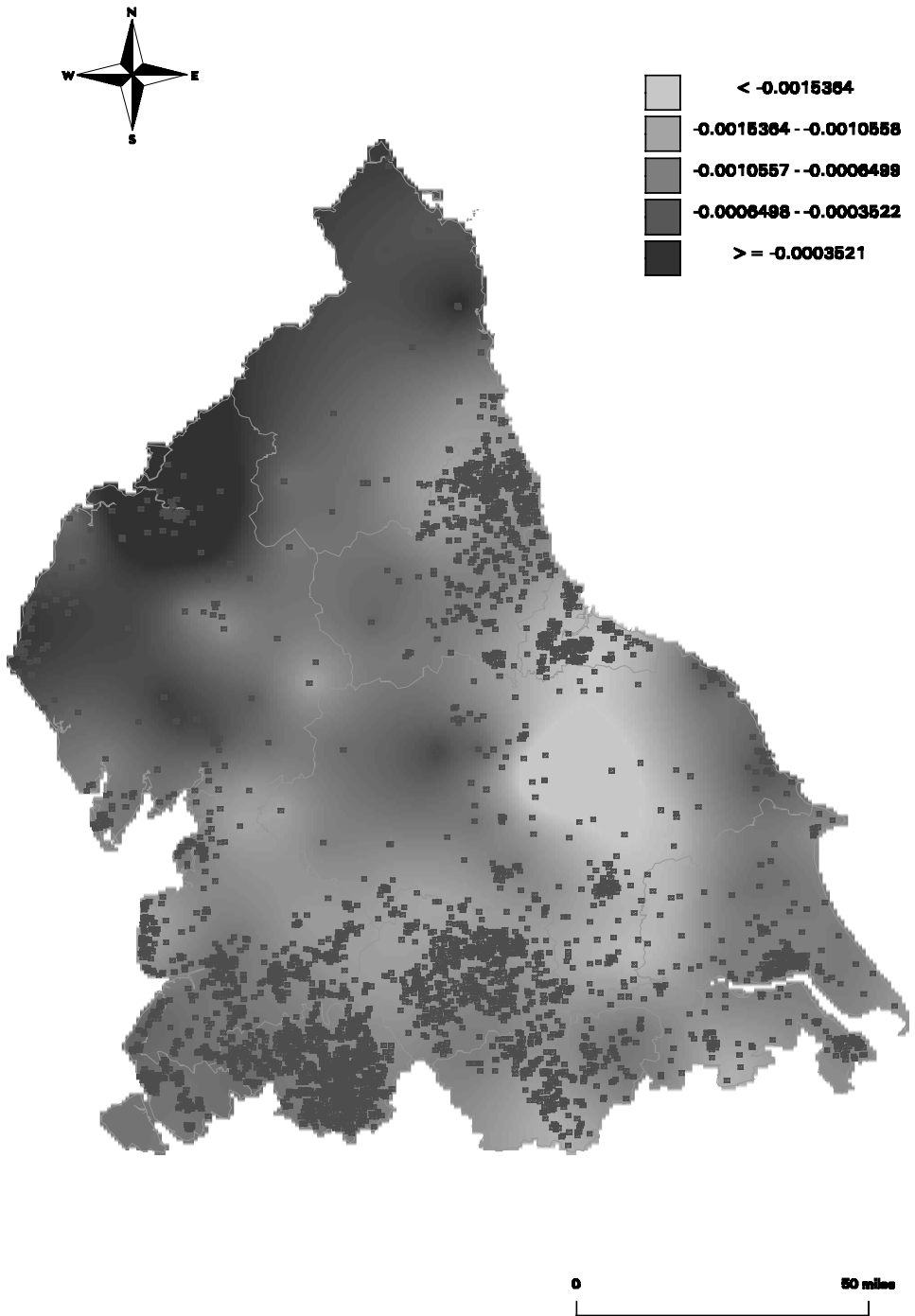
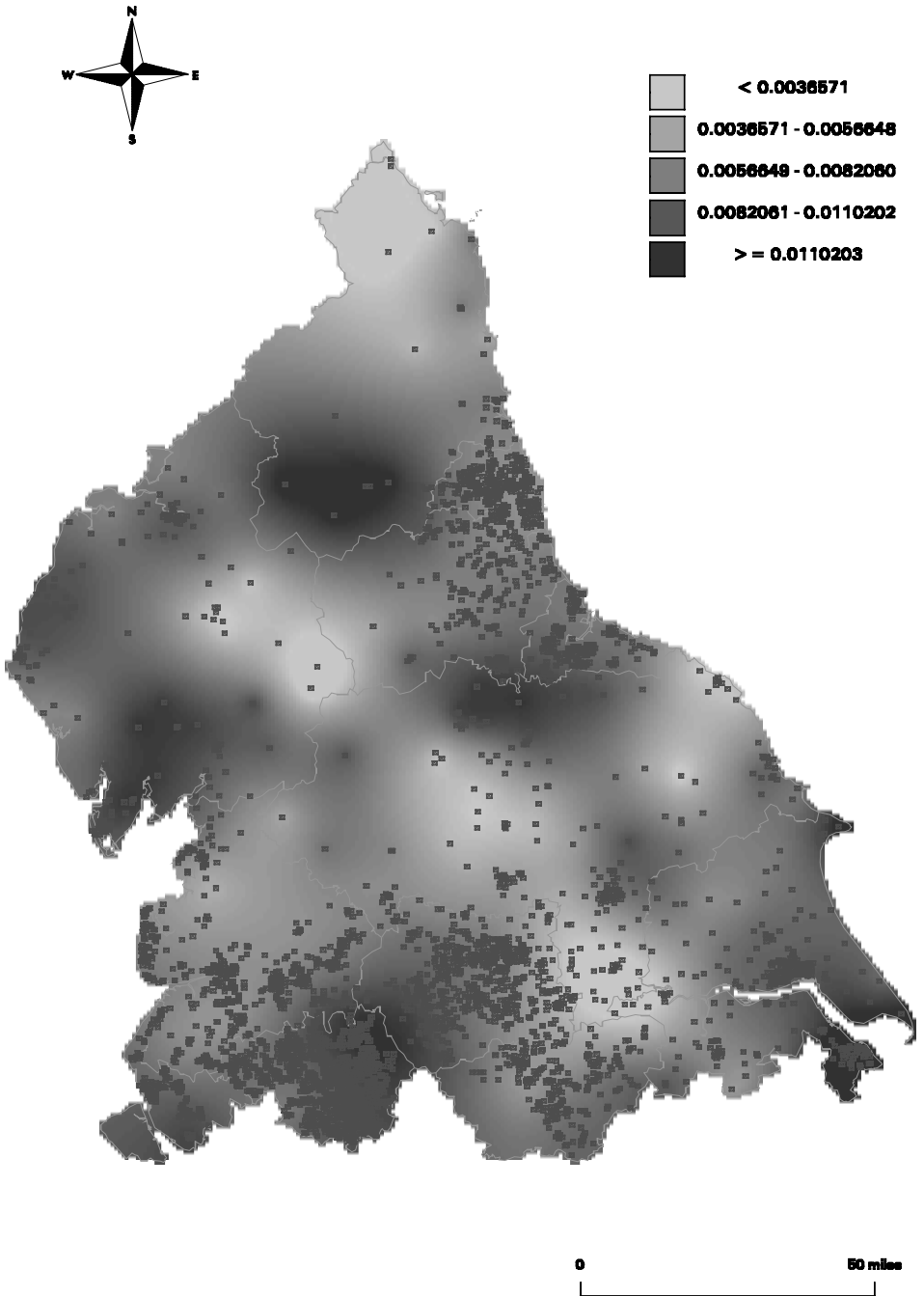**Figure 5.** Spatial distribution of the local school roll parameter.

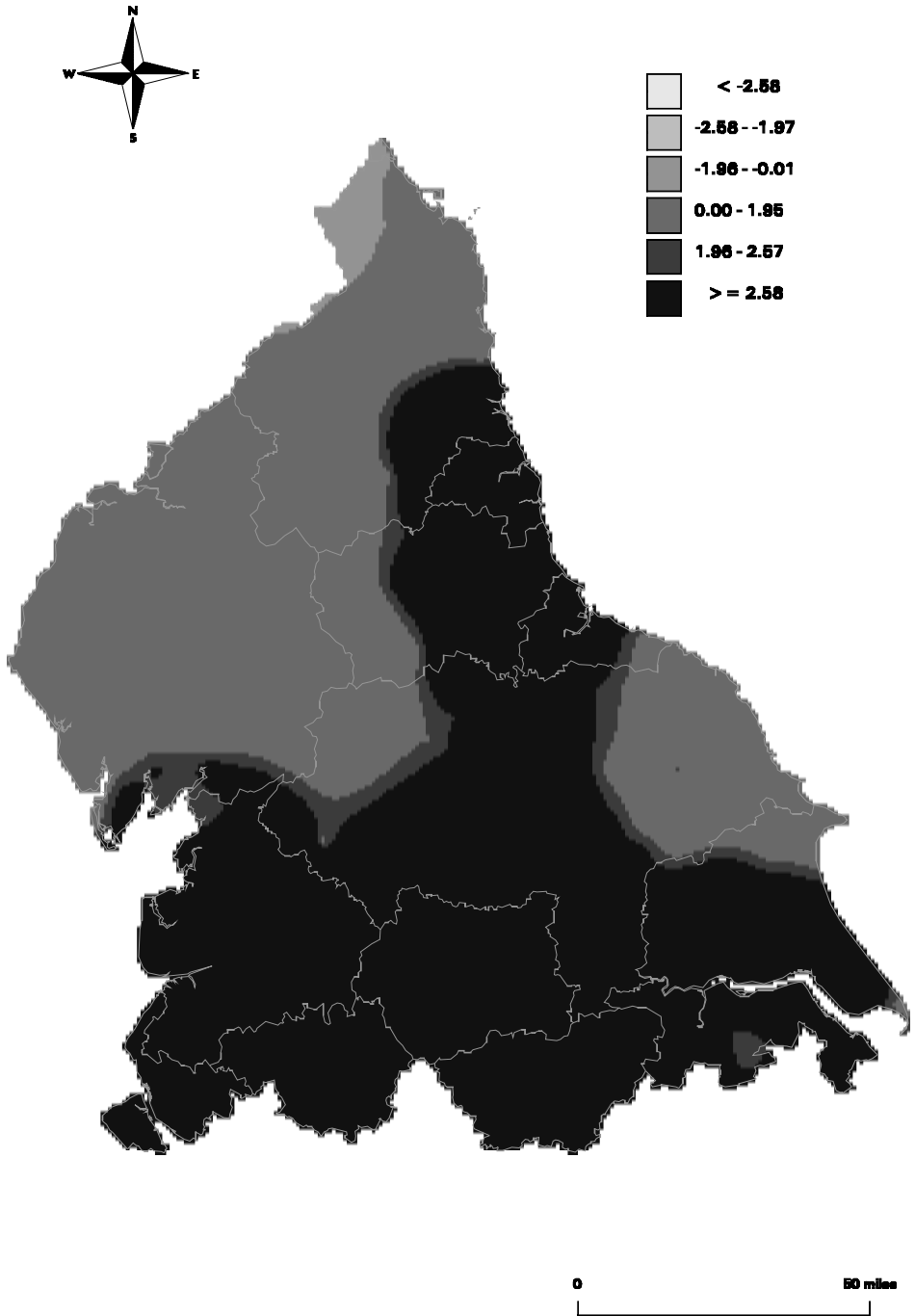**Figure 6.** Spatial distribution of the local social class parameter.
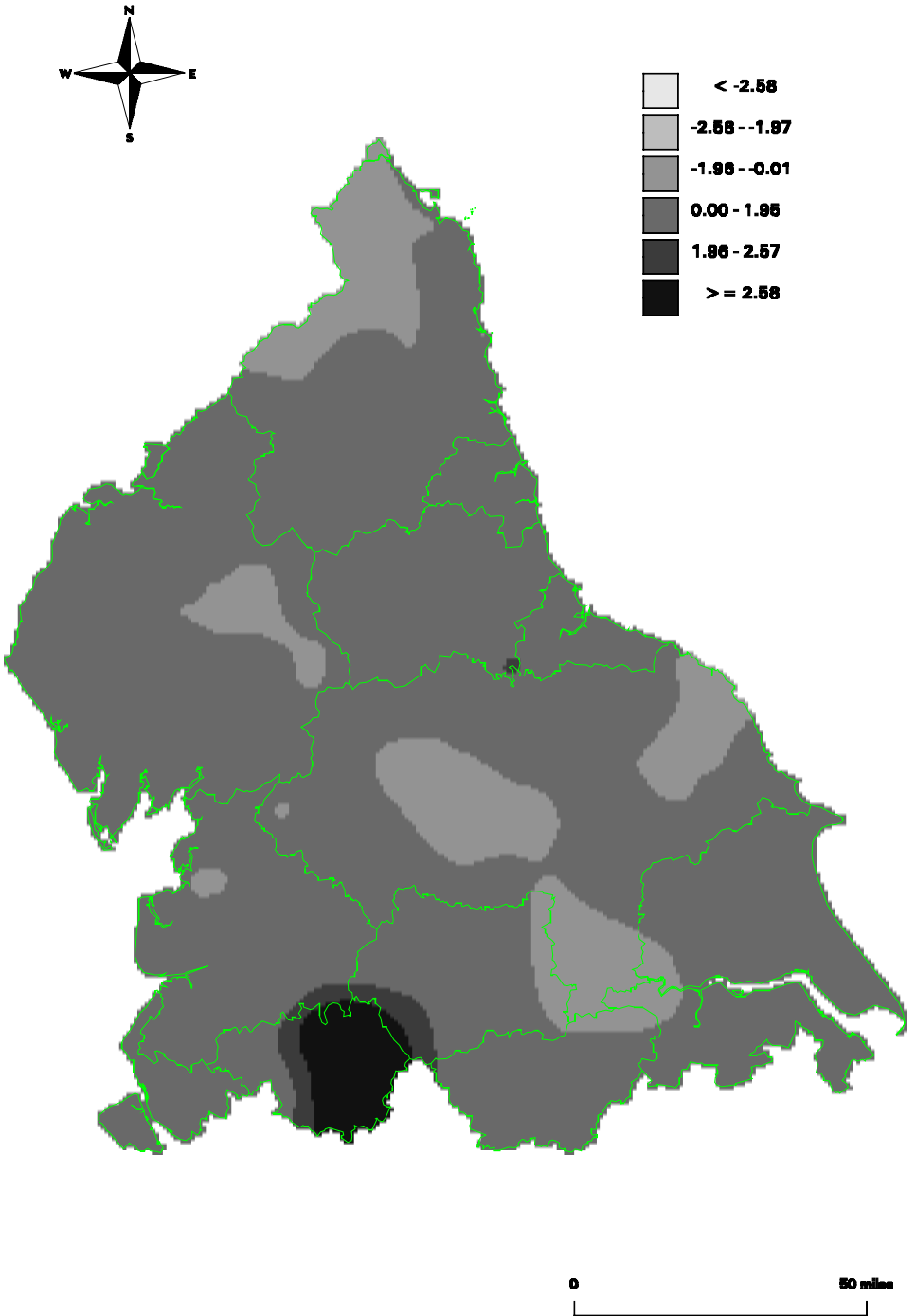
Figure 7. *t*-Surface for the intercept.

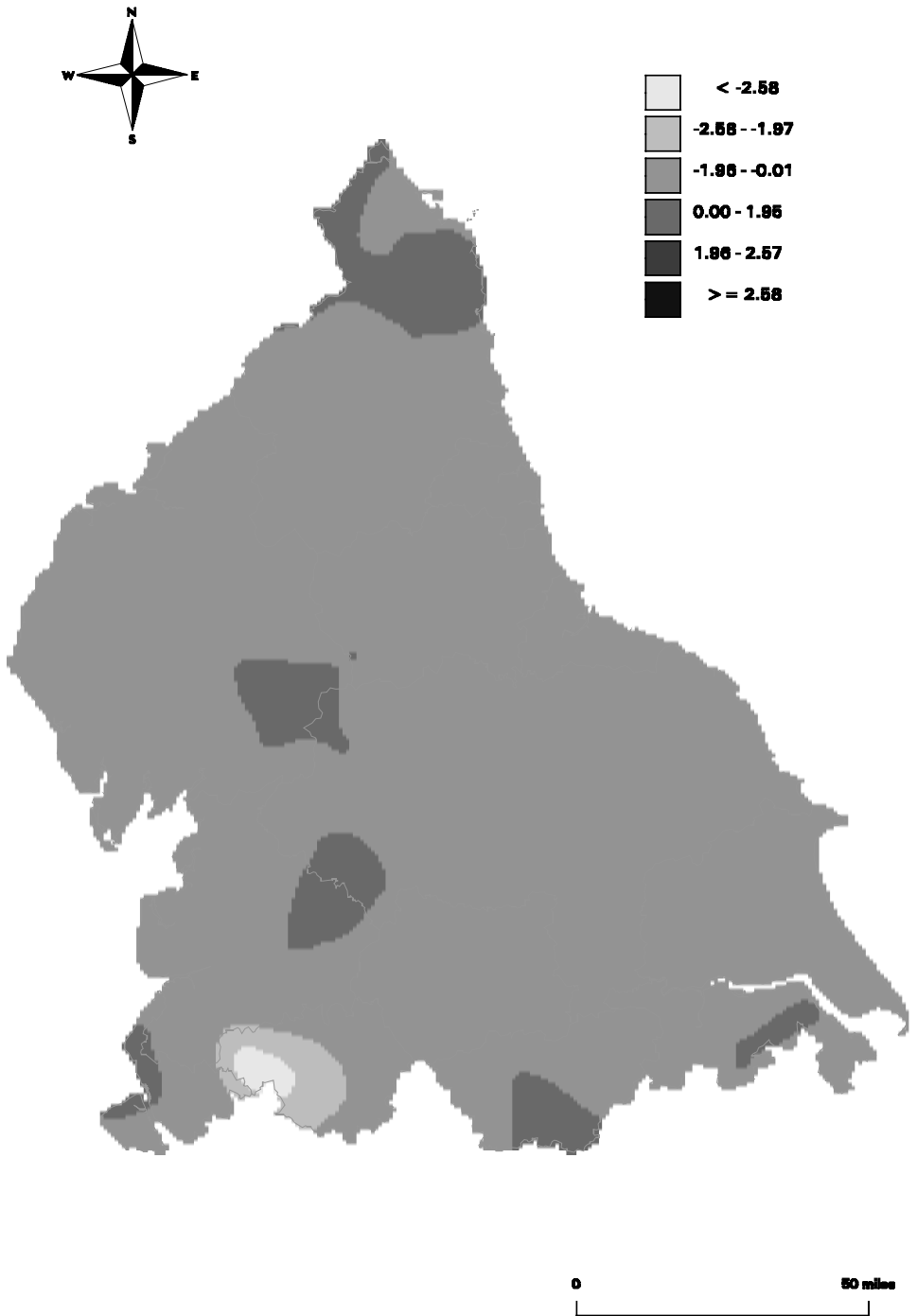**Figure 8.** *t*-Surface for the social class parameter.

Figure 9. *t*-Surface for the council house residents parameter.
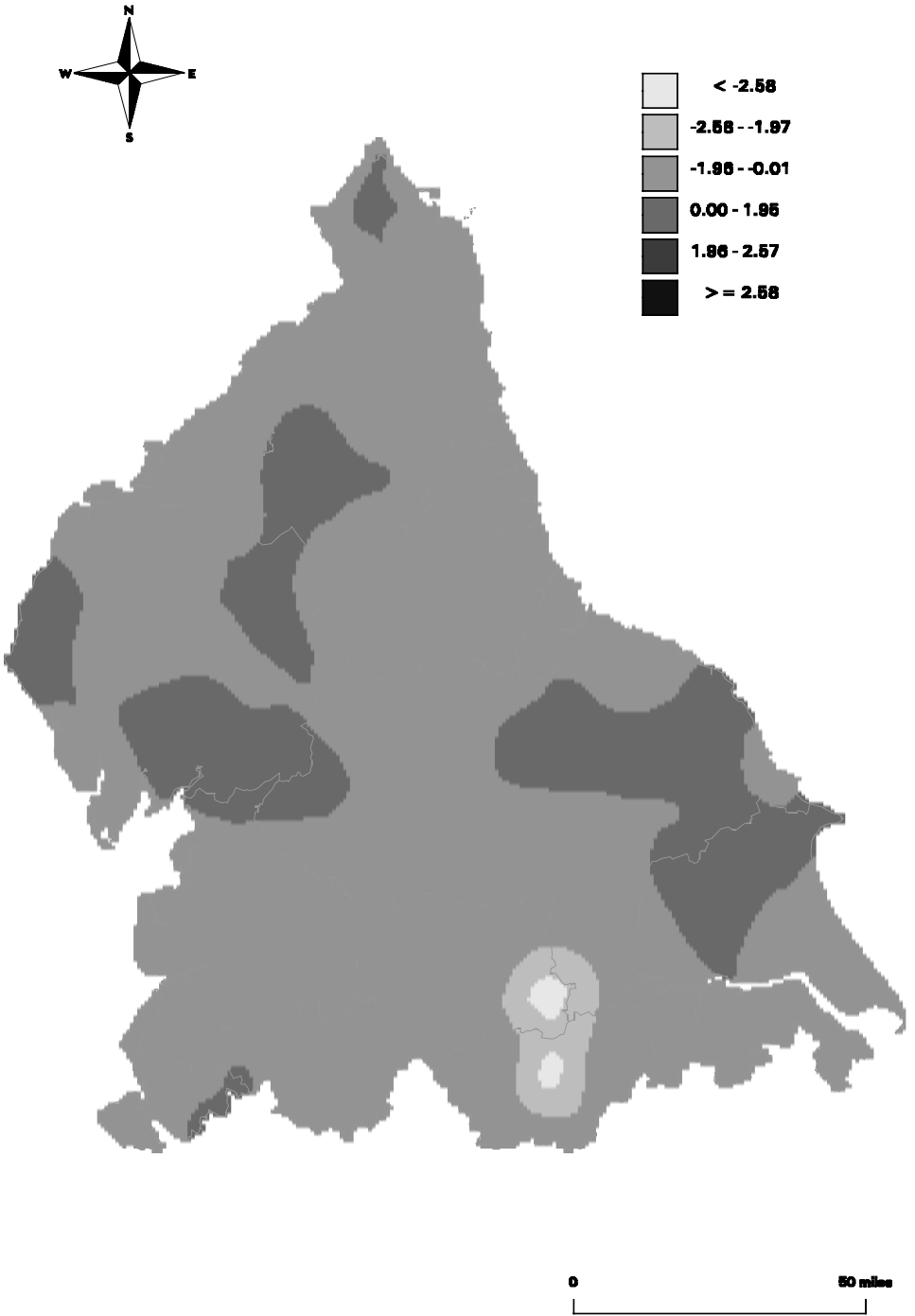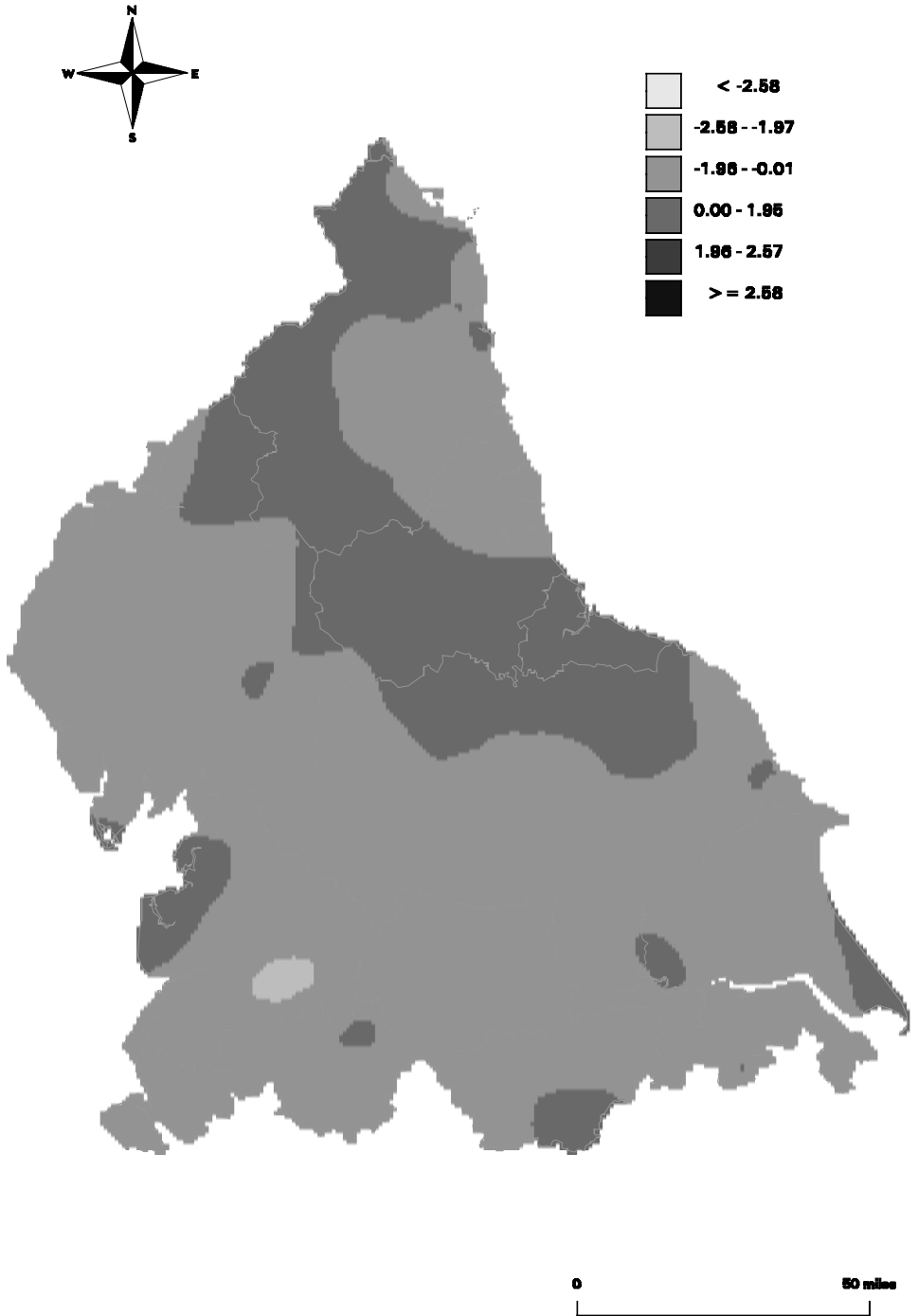
Figure 10. *t*-Surface for the unemployment parameter.

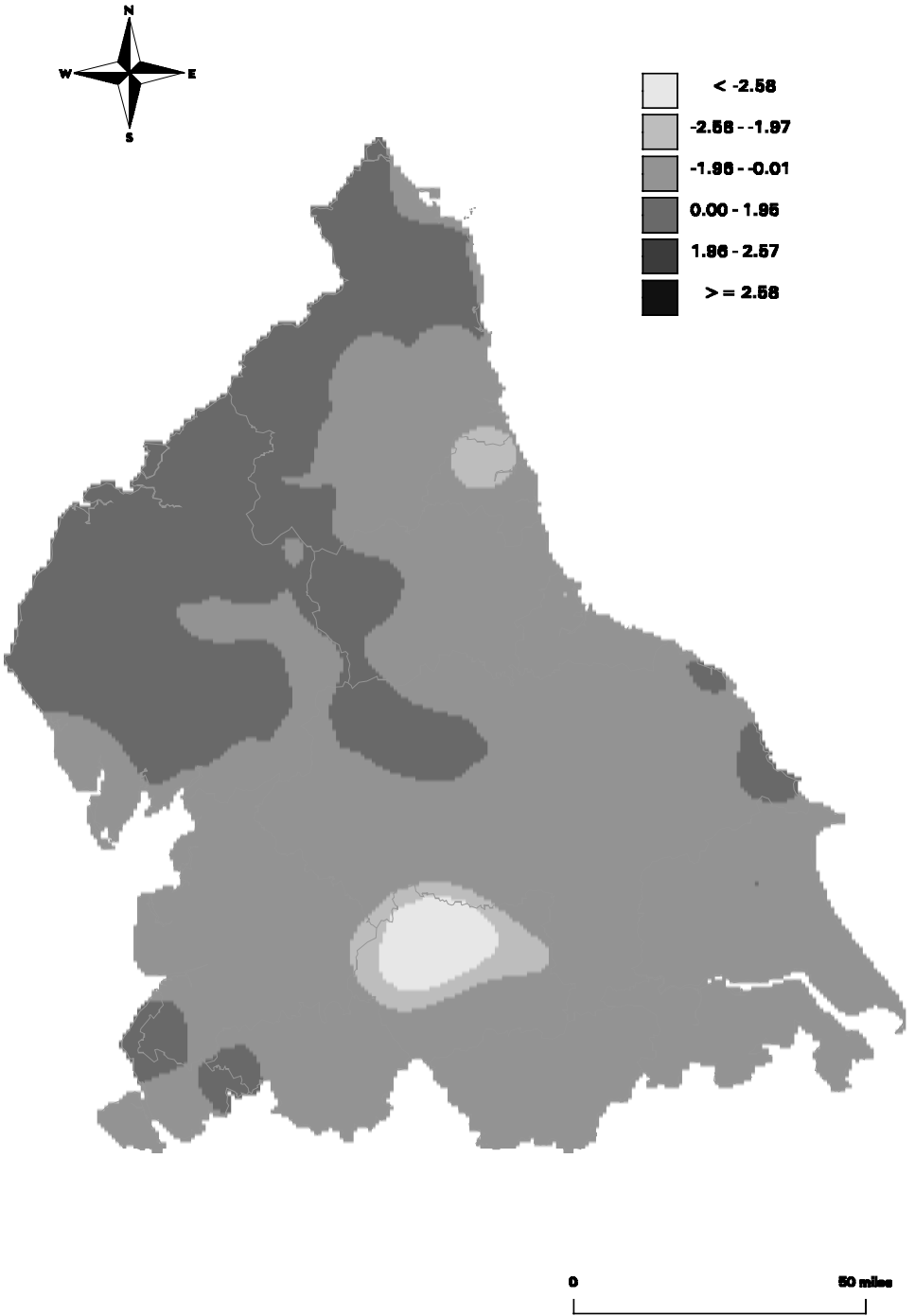**Figure 11.** *t*-Surface for the lone parents parameter.

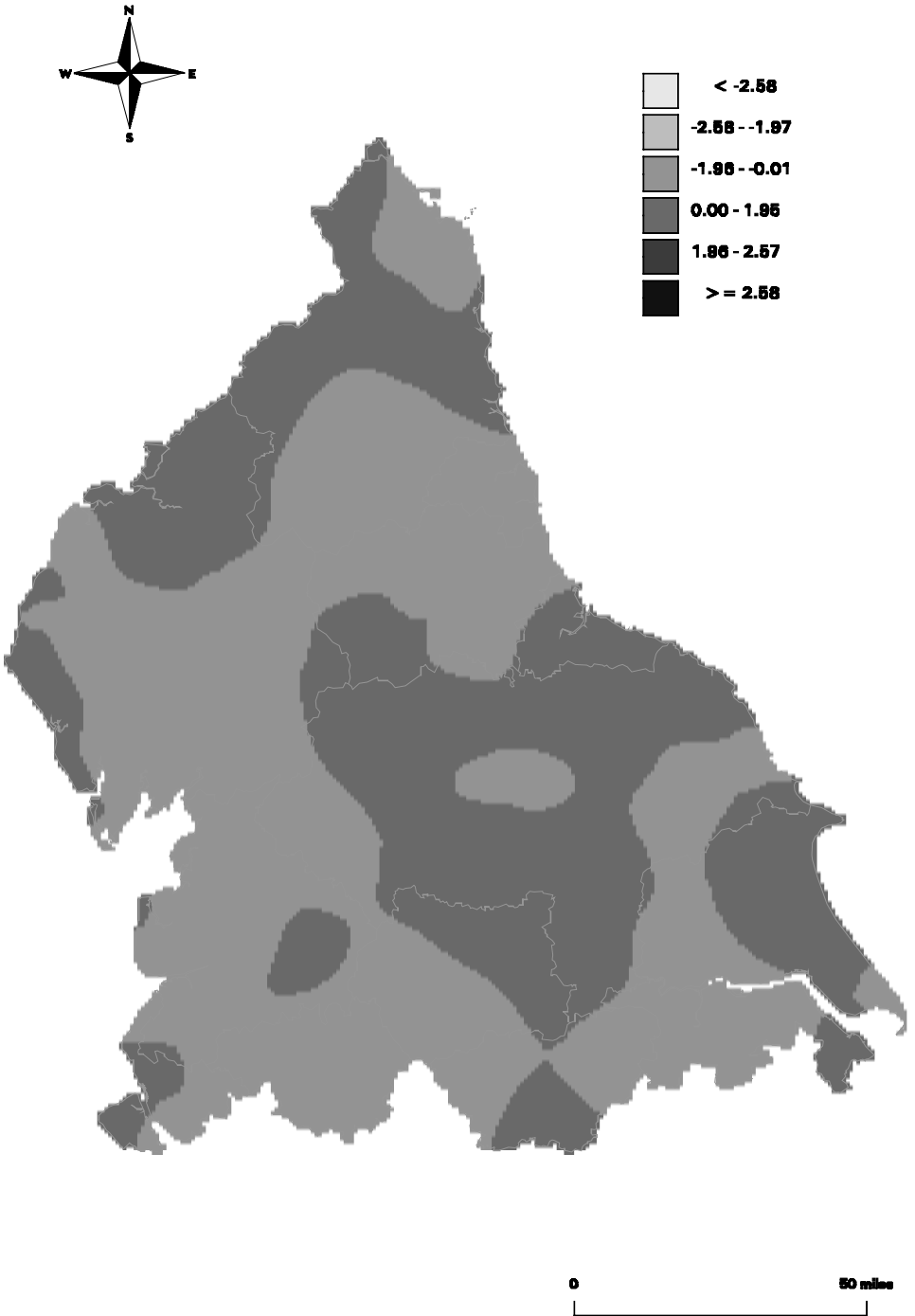**Figure 12.** *t*-Surface for the school roll parameter.

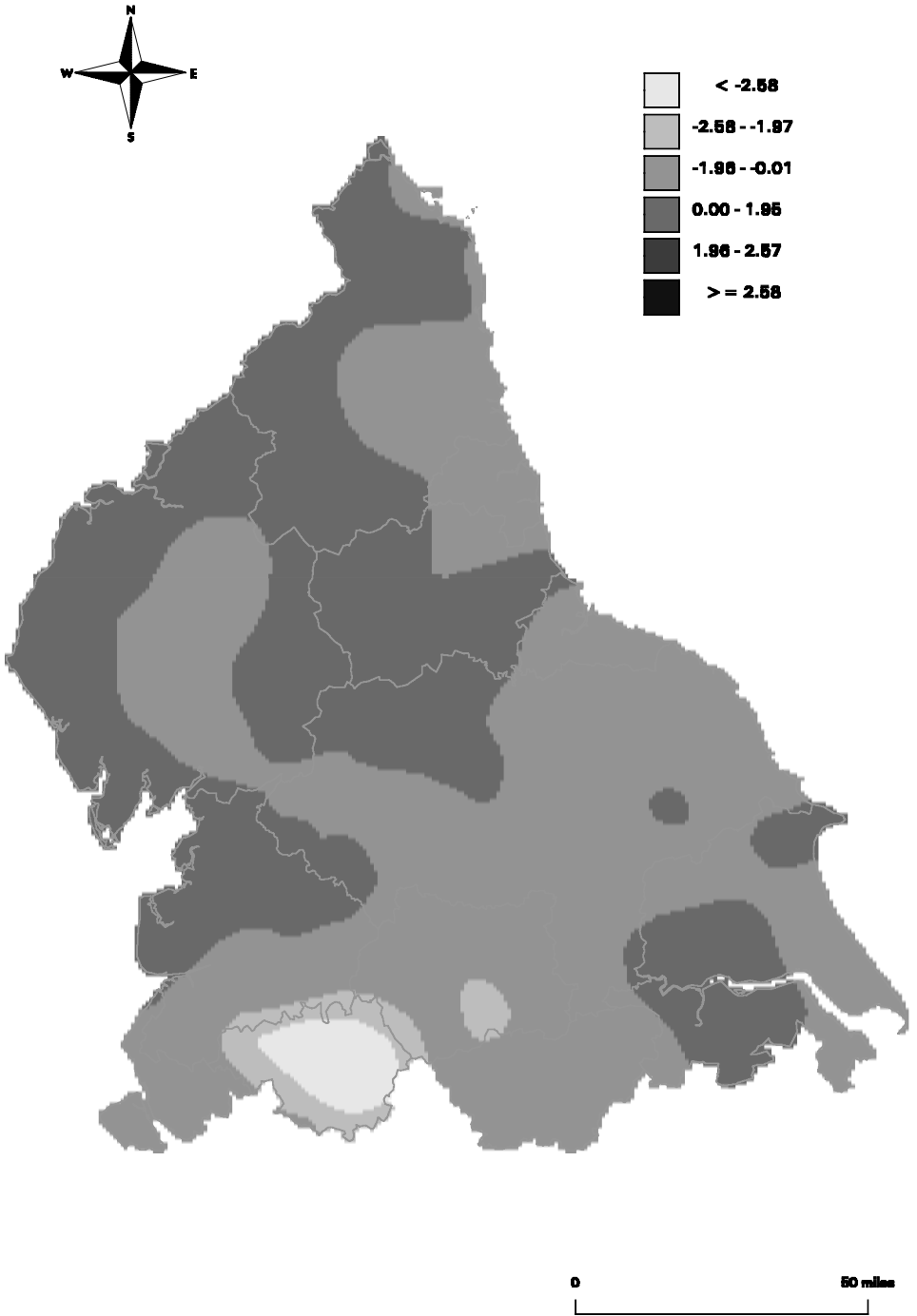**Figure 13.** *t*-Surface for the black residents parameter.

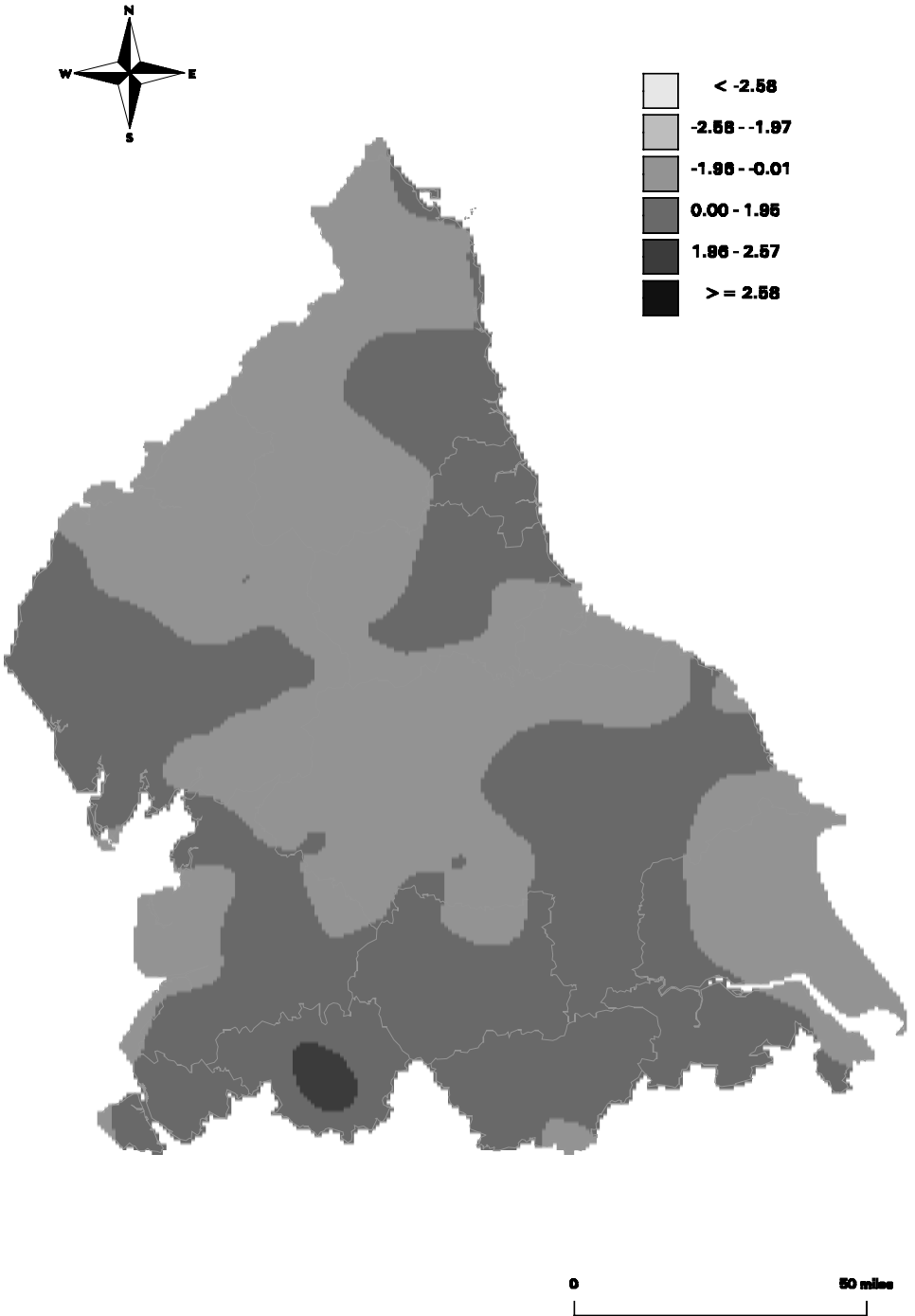**Figure 14.** *t*-Surface for the Indian residents parameter.

Figure 15. *t*-Surface for the Chinese residents parameter.

questions we would otherwise not have asked: why is this relationship so strongly positive in just the Manchester metropolitan area?

Summary

The geography of school performance is clearly an important social and political issue. There appear to be areas in which schoolchildren are disadvantaged in their educational performance because of their environment. A high level of unemployment, council house tenants, and lone parent households in a school's catchment area appear to have an adverse effect on the school's overall performance in standard tests. Conversely, high levels of professionals in a catchment area appear to have a positive effect on school performance. School performance appears to be inversely related to school size. It is impossible, of course, to provide a causal link between school performance and these factors but the circumstantial evidence is certainly there.

There is also a great deal of spatial variation in school performance which is not explained by the global regression framework and the global modelling results are only averages across a very diversified region. It is more illuminating to examine the spatial variations in the relationships as shown in the GWR output. The global model results can hide a lot of interesting spatial variation in relationships that is illuminated in the local analysis. The maps of parameter surfaces, and more usefully, the *t*-surfaces, allow us to get a good feel for the stationarity or non-stationarity of relationships. The results also force us to ask questions about the nature of the relationships being examined that would not have arisen in the global analysis. In this case, for example, there are some interesting local differences in relationships. These could be facets of model misspecification being manifested spatially or they could indicate intrinsically different relationships over space. In either case, we can only profit by exploring the relationships in more detail.

Notes

1. The authors would like to thank Dr Robin Flowerdew for making these data available to us.
2. Although this gives a good idea of the size of the school, it would have been preferable to have the number of children *per class* eligible to take the test rather than the total number of a certain age. Class size, rather than school size, is more likely to have an effect on pupils' performance (smaller classes in theory being better for educating pupils) than school size, for which the school roll variable is a surrogate. It is impossible to know, for example, whether a school roll of 60 represents two classes of 30 or three classes of 20 in a particular school.
3. It should be noted that proportional data are assigned to schools on the basis of raw counts initially along with the denominator as a raw count. Only when all the *ED* values have been assigned to schools in this way is the proportion calculated for each school.

References

Bartlett, J. (1936) The square root transformation in analysis of variance. *Journal of the Royal Statistical Society*, 3, 68–78.

Brown, P., Marsden, J., Batey, P. & Hirschfield, A. (1998) Relationships between pupil performance and social conditions: a GIS-based analysis using geodemographics. Paper presented at the conference 'Investigating Locational Data', Lancaster University, 9–10 July.

Brunsdon, C., Fotheringham, A.S. & Charlton, M.E. (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, 28, 281–298.

Conduit, E., Brooks, R., Bramley, G. & Fletcher, C.L. (1996) The value of school locations. *British Educational Research Journal*, 22, 199–206.

Coombes, M. & Raybould, S. (1997) Modelling the influence of individual and spatial variations in the levels of secondary school examination results. *Environment and Planning A*, 29, 641–658.

Fotheringham, A.S. (1997) Trends in quantitative methods I: stressing the local. *Progress in Human Geography*, 21, 88–96.

Fotheringham, A.S. & Brunsdon, C. (1999) Local forms of spatial analysis. *Geographical Analysis*, 31, 340–358.

Fotheringham, A.S., Brunsdon, C. & Charlton, M.E. (1998) Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30, 1905–1927.

Fotheringham, A.S., Charlton, M.E. & Brunsdon, C. (1997) Measuring spatial variations in relationships with geographically weighted regression. In: M.M. Fischer & A. Getis, Eds, *Recent Developments in Spatial Analysis: Spatial Statistics, Behavioral Modeling and Computational Intelligence*, Chap. 4. Berlin: Springer, pp. 60–80.

McCallum, I. (1996) The chosen ones? *Education*, 187(3), 12–13.

The Times (07/12/1998) Exam results 'not linked to class size', p. 4.