Local trend statistics for directional data—A moving window approach

Chris Brunsdon [a,*], Martin Charlton [b]

[a] Department of Geography, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom
[b] National Centre for Geocomputation, John Hume Building, National University of Ireland, Maynooth, Co. Kildare, Ireland

Abstract

The ideas of directional distributions for random data are reviewed, in particular focussing on descriptive directional statistics to summarise these distributions. Consideration is then given to spatial variations in directional distributions; for example how does the directional distribution of wind direction vary across geographical space, and how may this be analysed? To investigate this issue, an approach to moving window-based smoothing of directional data is proposed, based on the application of a geographical kernel-based weighting scheme to find localised mean directions (and related statistics) to directions represented as complex numbers of magnitude one. Consideration is also given to the visualisation of the outputs of an analysis such as this. The paper concludes with two applications of the techniques proposed; an analysis of wind speeds across Europe drawn from NOAA observations, and an analysis of US inter-county net migration counts between 1985 and 1990.
© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

Directions occur in many kinds of geographical situations, but surprisingly little attention is given to the analysis of directional data. As their name suggests, these are data relating to direction (usually in two-dimensional space). Typically, such data will take

---

* Corresponding author.
  E-mail address: cb179@leicester.ac.uk (C. Brunsdon).

the form of angles in the range 0–2π rad (0–360°). Data of this kind has been analysed over a long time span and in a wide variety of disciplines—see for example Mitchell (1767) who provides an interesting astronomical study.

In this paper we discuss the application of localised statistics to this kind of data and the visualisation of directional trends using streamlines. We show how descriptive statistics can be interpreted in terms of complex numbers, and coordinate geometry, and how these interpretations extend naturally to weighted statistics. We then show how these may in turn be used to provide moving window filters to examine geographical non-stationarity in distributions of circular data. We also consider the visualisation of localised directional statistics, and consider how this may be viewed in a modelling context as well as an exploratory approach. In this paper angles are measured in radians, and zero radians corresponds to the direction due east. This is mainly for mathematical convenience—any of the observations made here could be translated to *bearings*, where angles are in degrees and zero is due north. We begin with a brief introduction to circular descriptive statistics.

2. Circular descriptive statistics—a brief introduction

Directional data may be thought of as navigational bearings or as the directional component of a two-dimensional vector. They are cyclic and have no well-defined minimum and maximum value—directions just below 2π are considered to be close to values just above zero. This implies that ordinary summary statistics are sometimes unhelpful. For example, the directional data set in Table 1 has an arithmetic mean of 3.144 rad, although this is almost as far removed as possible from each of the individual directions—see Fig. 1.

The problem here is that although directions should be represented on a cyclic scale, the descriptive statistics are intended for use with a linear scale. An alternative way of representing directions is as two-dimensional vectors, or complex numbers. Here, the magnitude of the vector is fixed (usually taking the value one), and the direction of the vector is that being represented. Equivalently, the direction can be represented as a complex number whose magnitude is one. In this case, we may denote the direction as the complex number z where $z = \exp(i\theta)$, and θ is the direction.

The advantage of this representation is that it overcomes the discontinuous step in value seen as directions straddle due east. For example, if directions 1 and 2 are specified by Table 1 as $\theta_1 = 6.23$ and $\theta_2 = 0.05$ then the directions differ by 6.18 (around 354°). However using the complex number representations we have $z_1 = 0.999 - 0.050i$ and $z_2 = 0.999 + 0.050i$. These differ by 0.1i, which is relatively small, given that all directions lie on the unit circle, and the largest possible difference between two z-values is 2, when the directions are diametrically opposed.

Following this, an alternative definition of a mean may be given for directional data, using the complex number representations of directions. We denote this mean by M_z, and define it as

Table 1
A small data set of directions (in radians)

0.233	0.144	6.179	6.111	5.989	0.208

**Random Direction Sample
(Angles in radians)**
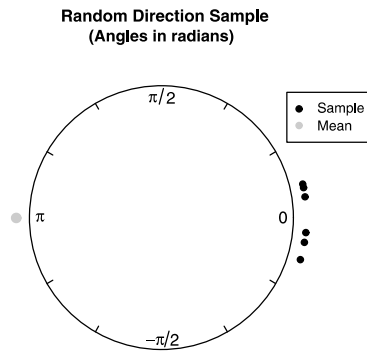
Fig. 1. Relationship of simple mean to directions in Table 1.
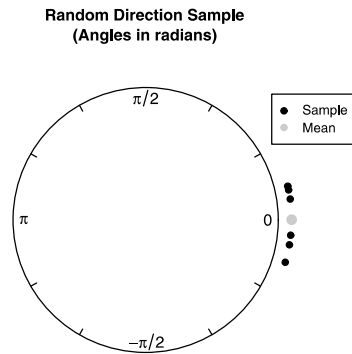
**Random Direction Sample
(Angles in radians)**

Fig. 2. Relationship of directional mean to directions in Table 1.

$$\frac{\sum_{i=1}^{n} z_i}{\left| \sum_{i=1}^{n} z_i \right|} \tag{1}$$

this is simply the mean of all of the directions in complex number form, normalised to have a magnitude of 1.

There is a geometrical rationale for defining the mean in this way. It is the complex number lying on the unit circle that minimises the mean squared straight line distances between itself and each of the directions in the sample. For the directions listed in Table 1, it is shown in Fig. 2. As can be seen, the location is now more plausible, when compared to that of Fig. 1.

3. Geometrical interpretation of circular mean

A geometric interpretation for the above definition of M_z can also be given. This is illustrated in Fig. 3. As with Fig. 2, directions are plotted as points on the circumference of the unit circle—and the points are considered as points in two-dimensional space (or on the Argand diagram). Here, the lighter grey point represents the geometrical centroid of the directions of the points in two-dimensional space. The circular mean can then be found by extending a line from the centre of the unit circle, through this centroid, until it meets the circumference of the circle (the meeting shown as the dark grey point).
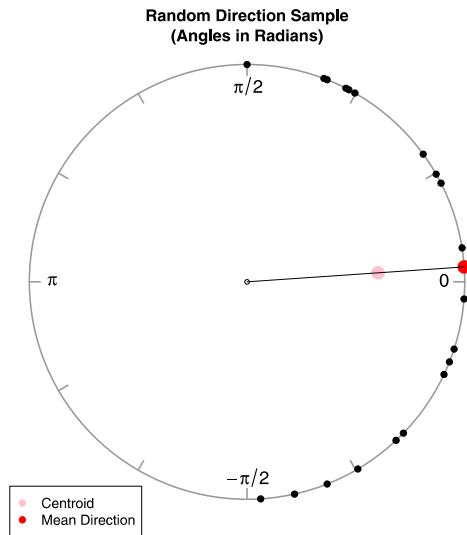
Fig. 3. Geometrical interpretation of the circular mean.

We can extend this idea a further stage, and demonstrate a measure of spread from this geometrical viewpoint. In Fig. 4, two further samples of directions are shown. That on the left shows very little angular dispersion, whereas that on the right shows a much larger amount of dispersion. In both cases, the centroid of the points is shown.

As can be seen, the centroid is much closer to the centre of the unit circle for the sample with the high degree of variation. For the sample with little variation, the centroid is very close to the circumference. In fact, the two extremes are easy to visualise: if all of the
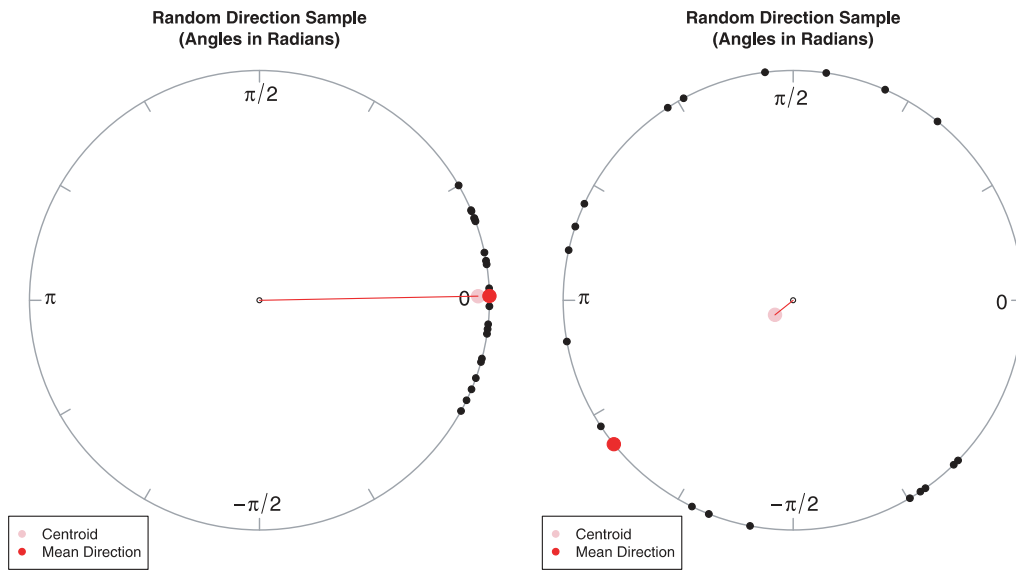

Fig. 4. Geometric illustration of a measure of circular dispersion.

directions in a sample were identical (so that there was no variation at all in the sample) the centroid would lie on the circumference—since it would just take the value of the unique direction exhibited. Alternatively, if directions were uniformly spread around the unit circle (for example in the locations of vertices of a regular polygon)—the centroid would lie in the centre of the unit circle. The length of the grey radial line in Fig. 4 measures the nearness of the centroid to the centre of the unit circle, and hence gives an indication of sample spread. This could be used directly to measure spread, although it has the counter-intuitive property that greater lengths are associated with smaller degrees of sample variability. For this reason, the length of the line is subtracted from one to give the measure of spread.

Translating this geometrical entity into a formula using complex numbers, we obtain

$$v = 1 - \left| \frac{1}{n} \sum_{i=1}^{n} z_i \right| \tag{2}$$

as a measure of spread.

This geometric interpretation also identifies an interesting pathological case. If the directions in a sample are evenly positioned around a circle (and in some other cases) then their centroid is at the centre of the unit circle. In this case, the spread is 1, but the mean is undefined as the two points needed to define the line in Fig. 4 are coincident. This is perhaps not unreasonable—if directions are evenly spread then no direction would provide a 'typical' direction in the sense that a mean direction would if some directions were favoured over others. Although this observation may seem esoteric, it is important to be aware that for some circular distributions, there is no mean direction.

4. Circular probability distributions

In addition to considering circular descriptive statistics, one can also think of the probability distributions used to model the variations in directional data. Perhaps the most commonly used distribution for circular data is the *von Mises distribution*:

$$f(\theta) = [2\pi I_0(\kappa)]^{-1} \exp\left[\kappa \cos(\theta - \mu)\right] \tag{3}$$

where θ is a random direction, $I_0(\kappa)$ is the modified Bessel function of order zero, and κ and μ are distributional parameters. In a number of ways, the von Mises distribution can be thought of as the circular equivalent of the normal distribution—see for example Mardia (1972, 1975).

As well as being defined for samples, circular means and directional variances are also defined for distributions—with expected values being substituted for sample means in Eqs. (1) and (2). For the von Mises distribution we have

$$\text{Circular mean} = \mu \tag{4}$$

and

$$\text{Circular variance} = 1 - A_1(\kappa) \tag{5}$$

where $A_1(\kappa)$ is the ratio of the modified Bessel function of order one to the modified Bessel function of order zero. Note that sample estimates of these two quantities can be obtained using these relationships. μ can be estimated by M_z and κ estimated by solving (5) with v substituted for the circular variance. Standard errors of these estimates can be obtained using bootstrap methods (Efron, 1982).

5. Localising directional statistics

Although it is unusual to consider directions as complex numbers or from a geometrical viewpoint, the definitions above are nothing new. Directional means may be considered in a number of ways—see for example Mardia and Jupp (1999) or Fisher (1993). Similarly, Eq. (4) defines the *sample circular variance*—Fisher (1993). However, it is useful to define the sample circular mean and variance in the ways above if one wishes to extend this idea by considering weighted mean directions, and in particular *geographically* weighted mean directions. A weighted mean direction can easily be defined in an intuitive way:

$$M_z = \frac{\sum_{i=1}^{n} w_i z_i}{\left| \sum_{i=1}^{n} w_i z_i \right|} \qquad (6)$$

where w_i is the weight associated with direction z_i. Similarly, a weighted circular variance can be defined as

$$v = 1 - \left| \frac{\sum_{i=1}^{n} w_i z_i}{\sum_{i=1}^{n} w_i} \right| \qquad (7)$$

Weighting is useful for a number of reasons—for example one may have a set of directions measured with varying reliability, and wish to reflect this when computing a mean. Also, one may wish to attach greater weight to more recent observations. The idea is important here as a basis for geographical weighting of directional descriptive statistics. In this approach descriptive statistics are 'localised' around a given geographical point by weighting observations using a kernel centred on that point. Allowing the point (and consequently the kernel) to move around the map one can build up a continuous field of localised statistics which may be mapped to investigate local trends (Brunsdon, Fotheringham, & Charlton, 2002). For any given localisation point (x,y) we obtain a unique set of w_i's depending on the distance between each point in the data set and (x,y). Typically, if d_i is the distance between the location associated with the directional observation z_i and (x,y) a Gaussian weighting scheme is chosen:

$$w_i = \exp\left(-d_i^2 / 2b^2\right)$$

where b is the *bandwidth* of the kernel. This controls the smoothness of the localised statistic, if it is viewed as a continuous function of (x,y). Applying this approach to directional data allows geographical patterns in directional phenomena to be graphically explored.

Finally, geographically weighted directional means also have a theoretical interpretation. If directions at some geographical point (x,y) have a von Mises distribution as set out in Section 4, but with mean $\mu(x,y)$ then the geographically weighted mean direction centered on (x,y) is a maximum localised likelihood estimator (Tibshirani & Hastie, 1987) of $\mu(x,y)$. This likelihood interpretation can also be used in a method for selecting an 'optimal' bandwidth. Suppose we have an observation θ_i at location (x_i,y_i), then leaving each observation i out of the data set, it is possible to estimate $\mu(x_i,y_i)$ from the remaining observations using the localised weighting procedure described above for a given bandwidth. Next, one can compute a cross-validation likelihood for each observation by substituting the estimated $\mu(x_i,y_i)$ into Eq. (3). Summing the logarithms of these gives a cross-validation likelihood score associated with the bandwidth. Thus we can consider

this score as a function of the bandwidth. Choosing the bandwidth that maximises this score is a useful 'automatic' approach to bandwidth selection.

6. Two examples

6.1. Wind direction

In this section two examples of the localised directional analysis described above are given. The first of these is an analysis of METAR data obtained from the US National Oceanographic and Aeronautical Administration (NOAA). 'Cyclic Files' contain weather information from a large number of locations throughout the world. Among the recorded data are the wind speed and direction. Here, a set of data for a number of stations in western Europe for midnight, 18th January 2003, are analysed. Local directional mean wind directions are computed as set out above, using a moving kernel with a bandwidth equivalent to 4° of arc on a great circle. Here double weighting occurs—firstly there is a geographical weighting as set out in earlier parts of this paper, but secondly weighting occurs to allow for the *magnitude* of the wind speed as well as the direction. Thus each w_i is the product of the geographical weight multiplied by the wind speed at location i. Thus, the numerator in Eq. (6) is the geographically weighted resultant wind speed vector at a given spatial location i, and dividing by the denominator yields the unit vector giving the directional component of this quantity.

A key question here is how such localised statistics may be visualised. One simple way of depicting directional data is the use of *glyphs*. These are symbols—such as arrows—which may be depicted pointing in any direction. A glyph depiction of the raw data is given in Fig. 5.

An initial approach was to show the geographically weighted mean directions as glyphs on a regular grid. The problem here is that visual attention is drawn to the lattice arrangement of the glyph locations regardless of the direction in which they point. In a sense there
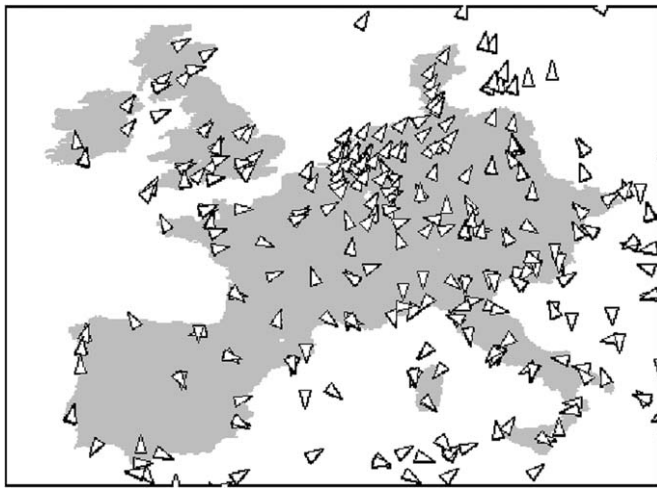

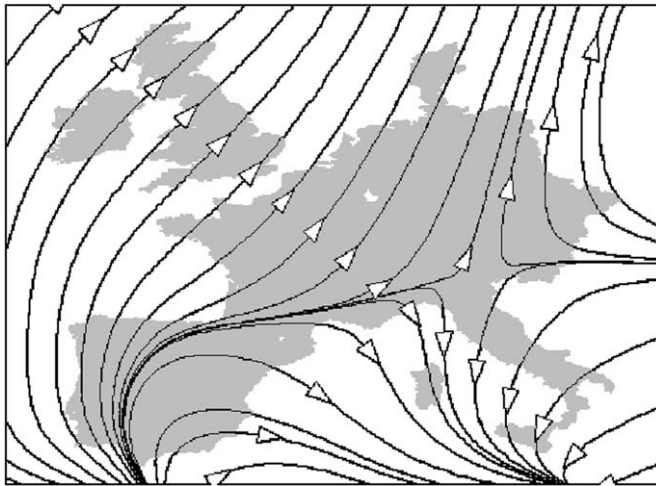Fig. 5. Raw NOAA data depicted with glyphs.

Fig. 6. Streamlines showing wind direction for NOAA data.

were two patterns: the pattern of interest (the geographical trend in wind direction) and a 'nuisance pattern' (the regular grid on which the glyphs are located).

A more helpful approach uses *streamlines*. Here a location is chosen (the initiator point) and the local mean direction is computed. The next sample point is then chosen by moving a small amount in the mean direction. This process continues and selects a series of sample points following the direction of flow. These points are joined to create a streamline. Similarly, starting again at the initiator point, points may be drawn in the opposite direction to the mean direction giving streamlines leading *into* the initiator point. Finally a single glyph is drawn at the initiator point to indicate the direction of the stream line. A number of judiciously chosen initiator points give a clear picture of trends in mean direction—see Fig. 6.

In Fig. 7, the local circular variance is plotted. This is derived as in Eq. (2)—and can be thought of as a local variability in wind direction. As before, wind speeds are included in the weighting for this calculation. Note that this is not defined as variability over time— since all measurements are essentially a 'snapshot' taken at a single point in time. Rather, this is a measure of spatial variability—a high value of this quantity suggests that near to the sampling point, winds were recorded in a wide range of directions. Here there is a general trend for increasing spatial divergence of wind direction as one heads from north to south, with a number of local maxima. Most notable of these is a maximum circular dispersion in the north of Italy. This is consistent with the smoothed direction map, where it appears that wind streamlines bifurcate.

6.2. US migration data

An interesting way of considering migration data is as directional flow information. Every person moving from place A to place B can be thought of as a unit of population flow in the direction of the vector AB. This notion of modelling human flows as a continuous phenomenon has a long history—see for example Beckmann (1952). This idea has
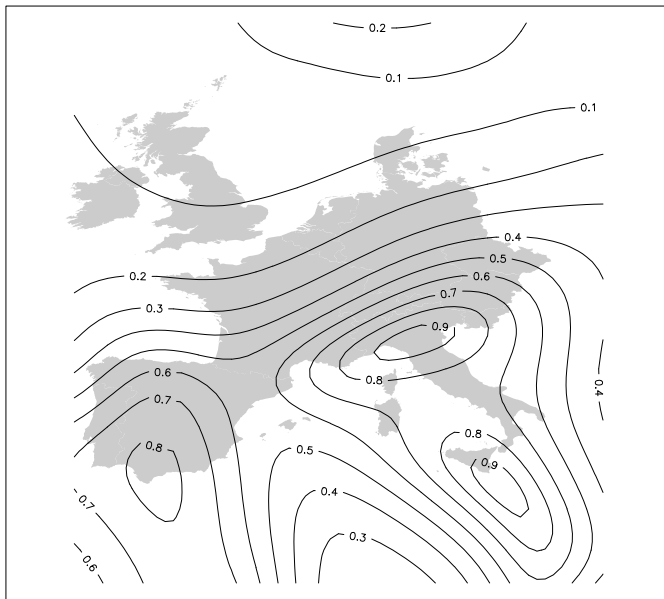
Fig. 7. Contour lines showing local circular variance.

been exploited by Tobler, who derives a model of population flow on the basis of 'attractiveness potentials' (Tobler, 1978, 1981). The idea here is that each location in space has an attractiveness potential at any point in time, and that population at any location flows in the direction of greatest increase in potential. Thus, if population flow is regarded as a vector field, then this field is proportional to the gradient vector of the population potential surface. Tobler demonstrates how discrete approximations may be used to estimate the potential field using tabulated migration counts. Since population flows here are considered as a vector field, one can obtain a field of directions varying continuously over geographical space, and represent these using streamlines similar to those provided in Fig. 6—indeed this is done in Tobler's 1981 paper.

Here we consider the problem from a slightly different standpoint—that of a local likelihood statistical model. Rather than using a potential field model to derive the direction of population movement, we work from a probabilistic model, where population at any point has a probability distribution associated with its direction of movement. Thus, given an individual who lives at a location (x, y) who decides to migrate, one may attach probabilities to each possible direction of migration and using a locally weighted mean of observed directions one can estimate the expected direction for any point (x, y). Streamlines may then be derived from these directions streamlines as in Fig. 6. In this case the streamlines indicate trends in population flow. Working with US inter-state migration data for the period 1985–1990 gives the result shown in Fig. 8.

Clearly the data supplied does not specify the exact origin and destination as latitude and longitude for each migrant in the five year period. For this reason all 'from' and 'to' locations for migrations are allocated to county centroids. For each state, there are observations of the numbers of migrations to each other state for which migrations occur (intra-state migrations are not considered here). Each non-zero state-to-state migration arc
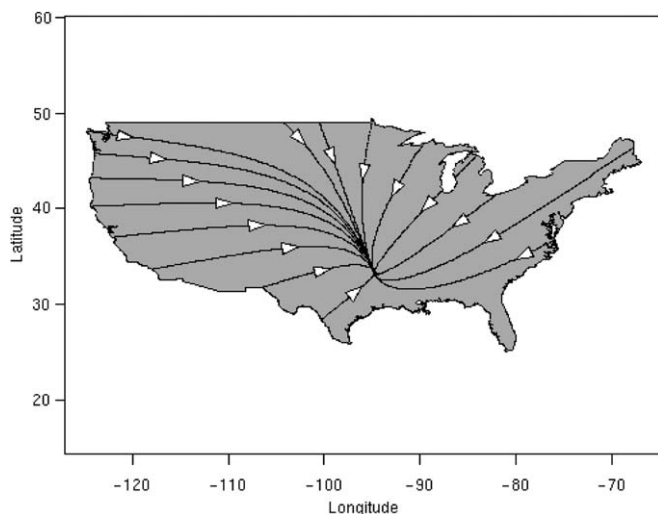
Fig. 8. Streamlines showing the directional components of trends in US inter-state migration (1985–1990).

is then counted as a directional observation. Unfortunately this does reduce the resolution of the analysis, so that more subtle sub-state trends in migration cannot be identified here. To compute spatial trends, geographically weighted directional means are computed—as before 'double weighting' occurs. In this case the non-geographical weight applied to each arc is the number of people associated with that arc. Note that the directions being considered here relate to ex-migration—so that at a given states, the angles refer to the directions to all counties to which people have migrated *to*. Note also that the angles refer to directions on the surface of a sphere.

A number of methods of depicting population movement have been considered in the past—Tobler (1987) gives a comprehensive review of possibilities. A more recent consideration is Holland and Plane (2001). The continuous nature of the smoothing approach taken here suggests the use of a flow-based method of visualisation, similar to that used by Tobler in his 1981 paper. It is worth noting that this approach typically associates just one flow direction to any point in geographical space—with the exception of a small number of bifurcation points, and therefore cannot depict complex one-to-many flow situations, as discussed for example in Gou (1993) and Marble, Gou, Liu, and Saunder (1997). However, the aim here is to provide a visual synopsis of the data and identify trends, and some degree of simplification is justified.

6.3. Degree of smoothing

One way in which this approach differs from other flow-based methods is that it is possible to control the degree of smoothing, and indeed produce a range of flow maps each depicting a 'low-pass filtered' image of the raw directional data, each with a different degree of filtering. The migration process may be thought of as a superposition of patterns of different scales—some local, some intra-state and some spanning the entire country—and the effect of smoothing is to filter out the lower frequency (ie more locally-scaled) patterns. Different levels of smoothing reveal trends at different geographical scales. The
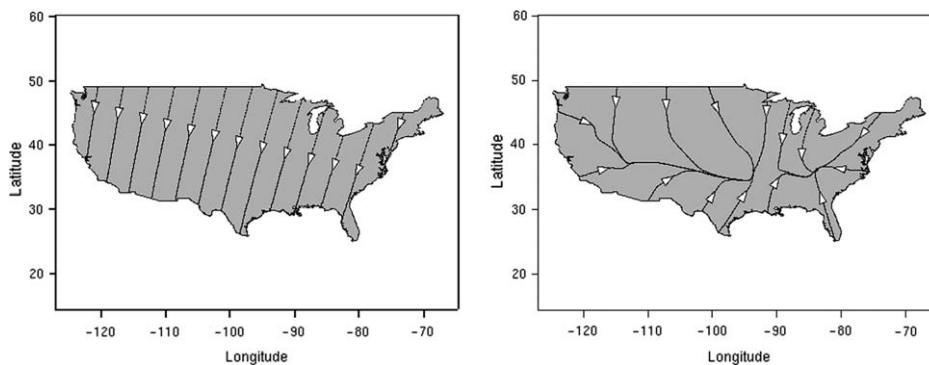
Fig. 9. Effect of greater (LHS) and less (RHS) smoothing than in Fig. 8.

degree of smoothing is governed by the choice of the smoothing kernel bandwidth. As with more usual smoothing operations (such as an 'ordinary' moving window averaging) and as discussed above, smaller bandwidths tend to give 'rougher' trends. In the case of surfaces, the roughness takes the form of large numbers of peaks and troughs. For directional trends roughness can be considered in terms of numbers of *sink holes* or *vortices*—points at which several streamlines meet. Fig. 8 has just one vortex—however the effect of less smoothing and more smoothing relative to this are shown in Fig. 9.

In the left hand panel a very large degree of smoothing is used—all streamlines more or less follow the global mean migration direction. On the right hand panel a low degree of smoothing is used, resulting in two vortices rather than one as in Fig. 8. From these, it is possible to demonstrate patterns at different geographical scale as discussed in the above section. Thus, one viewpoint is that it is good to visualise maps at different levels of smoothing.

However, there may be situations in which it is helpful to depict the situation with just one map. 'Correct' (or at least automatic) choice of degree of smoothing can be achieved using methods such as cross-validation is discussed in Section 5 (or see for example Brunsdon, 1995), although it can be helpful to look at the results of several levels of smoothing and see directional trend patterns at different scales. The degrees of smoothing for the streamlines in Figs. 6–8 were all selected using cross-validation.

7. Concluding discussion

The two examples above suggest there are a number of areas in which localised circular statistics may provide some useful exploratory tools. However, a number of questions are also raised, and a number of areas for future research may be identified. A few of these will be outlined here.

7.1. Working with orientational data

Orientational data is similar to directional data, but is used in situations where there is rotational symmetry through 180°. Thus, a straight road will have an orientation, but the velocity of a car driving along that road has a direction. More generally, data relating to

axes of some description are orientational. For example, Spain, Okello-Oloya, and John (1983) consider the orientation of termite mounds—which tend to be elongated along one axis. There are situations where orientation may vary geographically—and in these situations the methods outlined in this paper could be adapted to describe the mean orientation and spread in orientation. A typical approach for estimating orientational means is to double the angles and treat as circular data, and then halve the computed mean—but as well as these straightforward computational issues, the visualisation of trends in orientation must also be considered.

7.2. Semi-parametric models for circular data

Although the emphasis in this paper has been the use of localised circular statistics as an exploratory method it has been hinted that the method may be used as a semi-parametric model. If we assume that a direction located at (x, y) has a von Mises distribution with a mean $\mu(x, y)$, where $\mu(., .)$ is an unspecified function of x and y then, as discussed in Section 5, the local circular mean is actually the local likelihood-based point estimator of $\mu(x, y)$. Note that the parameter κ must also be estimated—this could either be constant across space, or itself a function of (x, y). This can be thought of as a parallel to a typical semi-parametric regression model for linear data of the form $z = f(x, y) + \varepsilon$ where $z(., .)$ is an unspecified mean function of x and y, and e is a random normal variate with mean zero. In particular, in the latter situation one may choose between ε having a constant variance with regard to space, or having a variance that is itself a function of (x, y). One advantage of considering this from a modelling viewpoint is that some inferential procedures may then be applied. For example, one could use a bootstrap-based technique applicable to local-likelihood methods to estimate confidence intervals for $\mu(x, y)$—see Galindo, Kauermann, Liang, and Carroll (2000).

7.3. Alternative approaches to choosing bandwidth

In this paper a cross-validation based approach was used for bandwidth selection. However a number of other options are possible if one adopts the modelling approach suggested above. Cross-validation exploits the likelihood function for a set of observations, but this is also the case for a number of other approaches. In particular, an approach using either the Akaike Information Criterion (AIC) (Akaike, 1973) or the Bayesian (or Schwarz) Information Criterion (BIC) (Schwartz, 1978) could be used to select the optimal bandwidth. Either of these latter approaches are likely to have a lower computational overhead than the cross-validation approach.

7.4. Analysis of time-based data

The key characteristic of circular data is its cyclic nature. However, several other kinds of data have this cyclic structure. In particular, data relating to time of day (or time within a weekly 168 h cycle) have this structure. Thus, the methods here may be useful for modelling events whose time of day and location are recorded. An example of this might be Police data for recorded public order incidents. It would be useful to model the variations in the mean time of day for incidents over geographical space. However, it is likely that

different methods of visualisation than those used to examine directional trends should be used for this kind of data.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proceedings of the second international symposium on information theory* (pp. 267–281).

Beckmann, M. (1952). A continuous model of transportation. *Econometrica, 20*, 643–660.

Brunsdon, C. (1995). Estimating probability surfaces for geographical point data: An adaptive kernel algorithm. *Computers and Geosciences, 21*, 877–894.

Brunsdon, C., Fotheringham, S., & Charlton, M. (2002). Geographically weighted summary statistics—A framework for localised exploratory data analysis. *Computers Environment and Urban Systems, 26*, 501–524.

Efron, B. (1982). *The Jacknife, the bootstrap and other resampling plans.* SIAM (Regional conference series in applied mathematics), Philadelphia.

Fisher, N. (1993). *Statistical analysis of circular data.* Cambridge: Cambridge University Press.

Galindo, C. D., Kauermann, G., Liang, H., & Carroll, R. J. (2000). *Bootstrap confidence intervals for local likelihood, local estimating equations and varying coefficient models.* Available from http://citeseer.nj.nec.com/galindo00bootstrap.html.

Gou, Z. (1993). Scientific visualisation and exploratory data analysis of a large spatial flow data set. Ph.D. Thesis, unpublished. Ohio State University.

Holland, S., & Plane, D. (2001). Methods of mapping migration flow patterns. *Southeastern Geographer, 41*, 89–104.

Marble, D., Gou, Z., Liu, L., & Saunder, J. (1997). Recent advances in the exploratory analysis of interregional flows in space and time. In Z. Kemp (Ed.), *Innovations in GIS 4* (pp. 75–88). London: Taylor and Francis.

Mardia, K. (1972). *Statistics of directional data.* London: Academic Press.

Mardia, K. (1975). Statistics of directional data (with discussion). *Journal of the Royal Statistical Society (series B), 37*, 349–393.

Mardia, K., & Jupp, P. (1999). *Directional statistics.* New York: Wiley.

Mitchell, J. (1767). An inquiry into the probable parallel and magnitude of the fixed stars, from the quantity of light they afford us, and the particular circumstances of their situation. *Philosophical Transactions of the Royal Society, 57*, 422–433.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Spain, A. V., Okello-Oloya, T., & John, R. D. (1983). Orientation of the termitaria of two Species of *Amitermes* (Isoptera: Termitinae) from Northern Queensland. *Australian Journal of Zoology, 31*, 167–177.

Tibshirani, R., & Hastie, T. (1987). Local likelihood estimation. *Journal of the American Statistical Association, 82*, 559–567.

Tobler, W. (1978). Migration fields. In W. Clark, & E. Moore (Eds.), *Population mobility and residential change*, Department of Geography, Studies in Geography, No. 25, Evanston (pp. 215–232).

Tobler, W. (1981). A Model of geographical movement. *Geographical Analysis, 13*, 1–20.

Tobler, W. (1987). Experiments in migration mapping by computer. *The American Cartographer, 14*, 155–163.