20 Turbulence in the River Severn: a Dynamic Systems Approach

LEONARD A. SMITH
Mathematical Institute, University of Oxford, Oxford, UK

20.1 INTRODUCTION

It is well known that the analysis of complex geophysical time series is a difficult task and that recent developments in non-linear dynamic system theory have suggested a variety of new approaches including prediction techniques (Casdagli, 1989; Farmer and Sidorowich, 1988). In addition to their obvious applications, these methods may be used to test implicit assumptions made in alternative techniques (e.g. that the series is a realization of a linear stochastic process) through the method of surrogate data (Smith, 1992; Theiler *et al.*, 1992). The comparison with surrogate series provides a quantitative measure of the quality of different models of the data and, in turn, an indication of whether the underlying dynamics is linear or non-linear, deterministic or stochastic.

This chapter illustrates non-linear prediction by way of two-dimensional interpolation in a toy model of particulate transport, and then applies it to the prediction of observational data. The particulate transport model used, one of the first to show the effects of chaotic advection, is used for spatial and temporal predictions. A preliminary analysis of velocity data collected from the River Severn concludes the chapter. The results of this analysis reject the hypothesis that this signal is linear red noise, yet no evidence of low dimensional determinism is obtained. Several suggestions for the design of future data collection are proposed.

20.2 STOMMEL FLOW

In this section we introduce non-linear prediction with a simple two-dimensional time-dependent flow which can exhibit chaotic advection (Smith, 1984; 1987; Smith and Spiegel, 1985). The flow was previously considered as a model for the motion of particles suspended in a time-dependent flow, a situation of interest in many areas of geophysics (Huppert, 1984; Maxey and Corrsin, 1985). By introducing periodic time dependence, we generalize the laminar flow originally considered by Stommel (1949) to allow chaotic advection. In the chaotic regime, particles may take much longer to traverse a series of cells than the retention time indicated by dimensional calculations. Related flows have been investigated experimentally by Tooby, Wick and Isaacs (1977), Gollub and Solomon (1987) and Chaiken *et al.* (1987).

Mixing and Transport in the Environment. Edited by K.J. Beven, P.C. Chatwin and J.H. Millbank

20.2.1 Steady flow

Intrigued by the observation that the yield of plankton tows taken along the direction of the wind were more variable than those taken perpendicular to the wind, Stommel (1949) considered the trajectories of negatively buoyant bodies immersed in steady fluid rolls. If the fluid motion is described by a stream function $\Psi(x,y,t)$, then, when particle inertia is negligible, the particle trajectories also allow a stream function. That is, the trajectories are solutions of

$$v_x = \frac{d\psi(x,y,t)}{dy} \tag{20.1}$$

$$v_y = -\frac{d\psi(x,y,t)}{dx} \tag{20.2}$$

where the particle stream function is

$$\psi(x,y,t) = \Psi(x,y,t) + v_s x \tag{20.3}$$

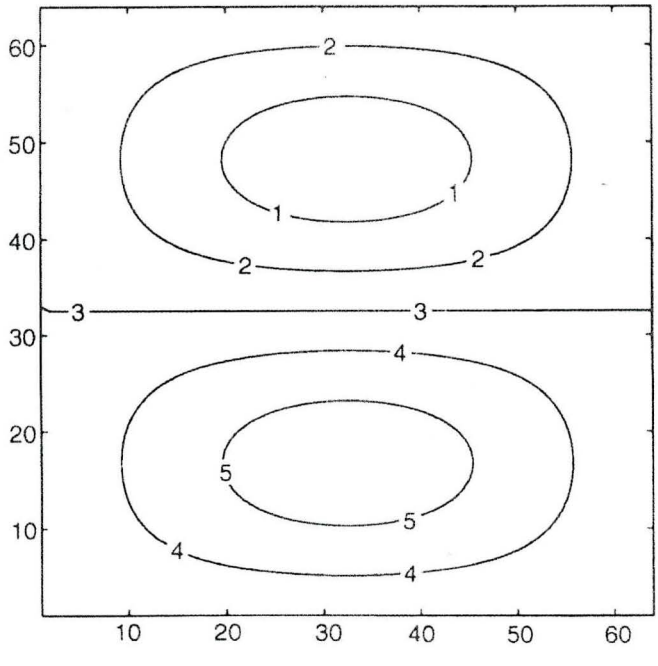
and v_s is the Stokes velocity (the terminal velocity of a particle in free fall through a quiescent, viscous fluid). Stommel considered a two-dimensional, vertical cross-section and adopted the stream function

$$\Psi(x,y) = A\sin x \sin y \tag{20.4}$$

Streamlines of this flow are shown in Figure 20.1. The particle stream function is then

$$\psi(x,y) = A\sin x \sin y + v_s x \tag{20.5}$$

where x represents the horizontal direction and y the vertical direction. Particle stream lines



CONTOUR KEY	
1	-0.800
2	-0.400
3	0.000
4	0.400
5	0.800

Figure 20.1 Fluid streamlines of the steady flow. Here x is the horizontal axis and y is the vertical axis; the figure shows a single cell (i.e. $0 \leq x \leq \pi$, $0 \leq y \leq 2\pi$)

for $v_s = 0.25$ are shown in Figure 20.2. Stommel (1949) concluded that the flow divided into two classes: one which swept through vertically stacked cells and another, trapped within the closed contours, which formed a region of retention in each cell. This model may be interpreted in a downstream guise (where x is the cross-stream direction and y the downstream direction) with the Langmuir cells corresponding to dead zones due to boundary conditions not explicitly modeled; v_s then corresponds to a superimposed downstream velocity. In the steady flow case, downstream motion is trivial. Dead zones really are dead as there is no transport across the stream lines, and the streamlines themselves have the simple structure shown in the figures. We could consider more spatially complicated flows by superimposing additional structure periodic with half the wavelength of this flow. Repeating the process at still smaller length scales would yield a flow similar to the β model (Frisch, Sulem and Nelkin, 1978). As shown in the following, the particle trajectories from Equation 20.5 are already fairly intricate when the flow is time dependent.

20.2.2 Time dependent flow

When ψ is independent of time, topological constraints prevent trajectories from displaying chaos. A time-dependent stream function, however, corresponds to a non-autonomous Hamiltonian system; in this instance the system has a three-dimensional phase space and admits qualitatively different trajectories (see, for example, Mackay and Meiss, 1987). To investigate this case, consider

$$A(t) = A_0[1 + \epsilon\sin(\omega t)] \tag{20.6}$$

where ϵ quantifies the strength of the time dependent element of the flow. For $\epsilon \neq 0$, this has the immediate effect of detaching the dead zone from the wall; for small ϵ an isolated

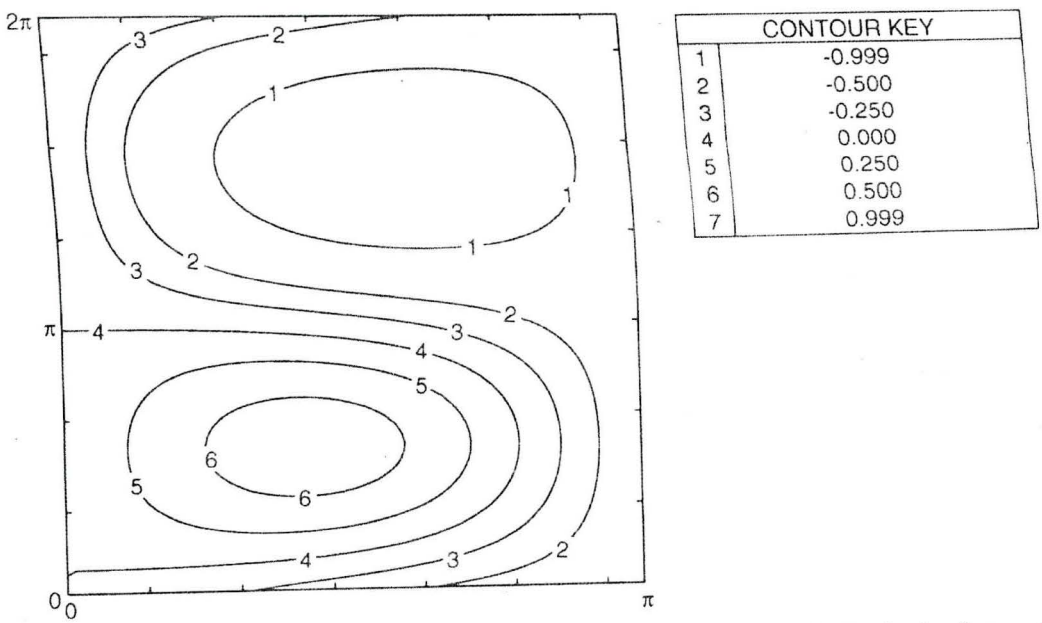
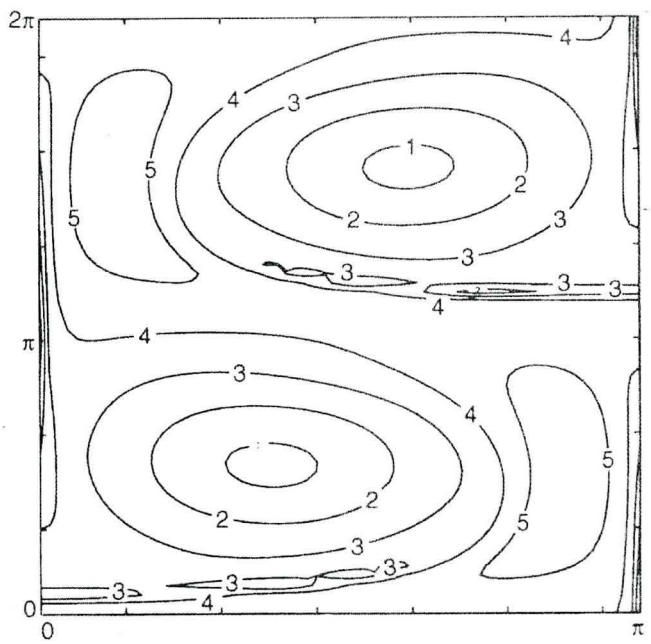


Figure 20.2 Particle streamlines in the steady flow with $v_s = 0.25$, $A = 1$. Again the figure shows a single cell which may either be considered as one of a series of vertically stacked cells, or to have periodic boundary conditions identifying the bottom with the top of the cell

dead zone(s) persists, but the boundaries can be very complicated; whether or not a particle will be retained for another period of the flow depends sensitively on its location. We will consider the case $A_0 = 1.0$, $\nu_s = 0.25$, $\omega = 2\pi/4.5$ and apply periodic boundary conditions in y, so that particles passing through the bottom of the cell are reintroduced at the top.

This complexity is reflected in Figure 20.3 in the total displacement, δ, of a particle from its initial position after one period [$\delta = |x(2\pi/\omega) - x(0)|$]. The concentric contours 1, 2 and 3 within the dead zone reflect stable, roughly circular motion, whereas the oblong contour 5 represents particles which fall out of the cell. Similar contours for longer evolution times maintain the relatively simple structure within the dead zone, but reveal complex structures near its boundary as some particles fall repeatedly through cells while their near neighbours become entrained (temporarily) near the dead zones of other cells. Examining the variation in residence time of a particle near the dead zone boundary reveals the sensitive dependence of residence time on initial position.

In the next section, we shall consider the question of interpolation in this field, specifically: given the initial position of an ensemble of particles and their final position after one period of the fluid forcing, how can we approximate the final position of some other initial condition (if the underlying equations are unknown)? First, we consider the results of a long run shown in Figure 20.4. Here we have plotted the location of a single initial condition ($\nu_s = 0.25$) once per cycle of the background forcing and applied periodic boundary conditions identifying the top and bottom of the cell; this stroboscopic graph is equivalent to a Poincaré section. The figure sketches the region accessible to migratory particles. Note that within each cell the dead zone is divided into several disconnected regions; this is typical of Hamiltonian systems of this form. If a line of particles is introduced (Smith and Spiegel, 1985), it is quickly contorted into a complex shape that appears to be self-similar; as experimental techniques of flow visualization typically rely on tracers (see, for example, Corrsin, 1950), and as material lines



CONTOUR KEY	
1	0.065
2	0.188
3	0.311
4	0.435
5	0.558

Figure 20.3 Contours of the displacement after one period as a function of initial position for the case $A_0 = 1.0$, $\nu_s = 0.25$, $\omega = 2\pi/4.5$. The fine structure already visible near the centre and bottom of the cell (contours 2,3,4) becomes much more intricate as the number of periods increases

in a turbulent flow are expected to develop into fractal structures, it is interesting to note that such structures can occur in a periodic laminar flow, leading to the formation of self-similar distributions. There is nothing turbulent here; the flow is smooth but chaotic. Pollutant dispersal may be more widespread and more concentrated than expected from laminar flows without requiring turbulent mixing. Additional discussion of this and related flows is given in Smith and Spiegel (1985), Pasmanter (1987), Smith (1987), Broomhead and Ryrie (1988), Crisanti *et al.* (1990; in press), Yu, Grebogi and Ott (1990) and Huppert (1991).

20.2.3 Interpolation of vertical displacement in two dimensions

Consider the problem of interpolating a function $s(\boldsymbol{x})$ which determined the contours shown in Figure 20.3. We shall initially consider a global predictor (or map), $F(\boldsymbol{x}): R^2 \rightarrow R^1$, which estimates s for any $\boldsymbol{x}$. First, choose n_c base points or centers in the two-dimensional space

$$c_j, \quad j = 1,2,\ldots,n_c; \quad c_j \in R^2 \tag{20.7}$$

We will consider $F(\boldsymbol{x})$ of the form

$$F(\boldsymbol{x}) = \sum_{j=1}^{n_c} \lambda_j \phi(\|\mathbf{x} - \mathbf{c}_j\|) \tag{20.8}$$

where $\phi(r)$ are radial basis functions (Powell, 1985). We will consider $\phi(r) = r^3$ and $\phi(r) = e^{-r^2/c^2}$ where the constant is based on a multiple of the average distance between data

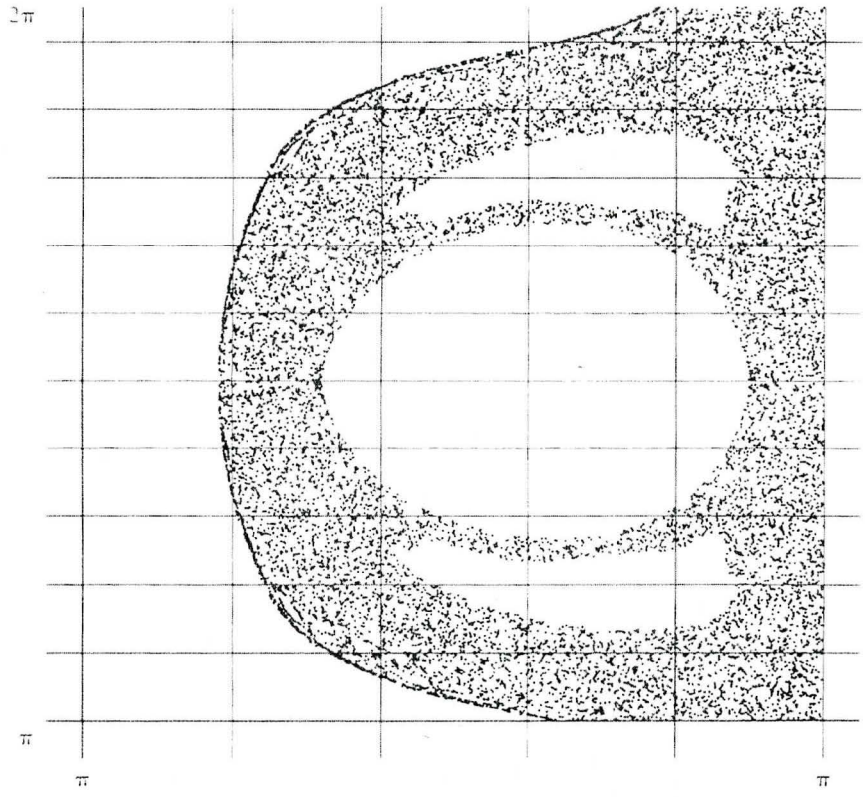


Figure 20.4 Stroboscopic view of a single particle falling for 2^{12} periods with periodic boundary conditions. Only the upper half of the cell is shown

points, d_{nn}. The λ_j are constants which are determined by observations

$$F(\boldsymbol{x}_i) \approx s_i, \quad i = 1,2,\ldots,n_L \tag{20.9}$$

where the $\boldsymbol{x}_i$ are the initial conditions of the n observations and the s_i are their observed displacements. We shall call the data used in determining the λ_j the *learning set*.

Determining the λ_j corresponds to the solution of the (linear) problem

$$\boldsymbol{b} = \boldsymbol{A}\lambda \tag{20.10}$$

where λ is a vector of length n_c whose jth component is λ_j and $\boldsymbol{A}$ and $\boldsymbol{b}$ are given by

$$A_{ij} = \omega_i \phi(\|\boldsymbol{x}_i - \boldsymbol{c}_j\|) \tag{20.11}$$

and

$$b_i = \omega_i s_i \tag{20.12}$$

where $i = 1,\ldots n_L$ and $j = 1,\ldots,n_c$. Traditionally, the weights ω_i reflect the confidence associated with the ith observation, we shall restrict attention to the case where all ω_i are equal, assuming that the errors are independent with equal variance (but note the discussion in Section 3.2 of Smith, 1992). When there are fewer centres than data points, $\boldsymbol{A}$ is not square and we are left with a standard least-squares problem of finding a λ which minimizes $\chi^2 = \|\boldsymbol{b} - \boldsymbol{A}\lambda\|^2$. The solution which also minimizes $\|\lambda\|^2$ corresponds to

$$\lambda = \boldsymbol{A}^+\boldsymbol{b} \tag{20.13}$$

where $\boldsymbol{A}^+$ is the Moore–Penrose pseudo-inverse of $\boldsymbol{A}$. Efficient methods to calculate $\boldsymbol{A}^+$ are given in Press *et al.* (1987). Additional discussion of the details of the construction of this type of predictor is provided by Broomhead and Lowe (1988), Casdagli (1989), Farmer and Sidorowich (1988) and Smith (1992).

A function F constructed in this way will be called a radial basis function (RBF) predictor. We shall also consider local linear (and quadratic) interpolation. In all instances the data in the learning set used to construct the predictor are kept distinct from the test data set on which it is evaluated. Thus the time series results represent out of sample prediction. Out of sample statistics are crucial if we are to establish that an RBF predictor is robust in the presence of noise; RBF predictors can minimize in-sample error by overfitting the noise in the learning set.

20.2.3.1 Global and local prediction

We now contrast the results of constructing one global predictor for the entire cell with those from constructing a local predictor for each point of interest based on that point's near neighbours. This simple two-dimensional example illustrates qualitative features of global reconstructions less easily visualized in higher dimensional constructions. Examining contours of the predictor error as a function of location in the cell shows that the largest prediction errors occur near the corners where the gradient is greatest; when the data density is low, global predictors may maintain a smoothness over the cell lost in local predictors. In this instance, using a grid of 32 × 32 sample points as the learning set, the 32 neighbour local RBF predictor does a superior job of resolving the details of the flow. When the data density is very high, local linear predictors should be most accurate; at this data density, however, local quadratics were found to give the smallest mean absolute error.

The predictions of one local RBF with $\phi(r) = r^3$, one local linear, and one global predictor

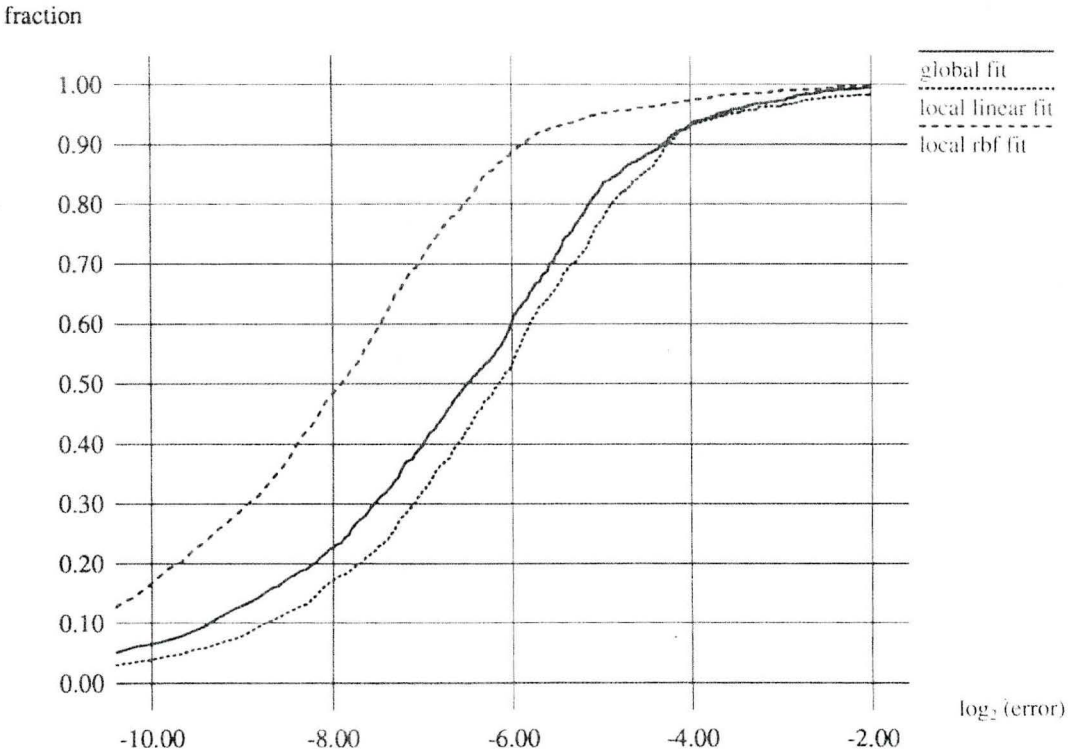


Figure 20.5 Comparison of prediction error profile for (a) global and two 32 neighbour local predictors. The horizontal axis is the $\log_2$ of the error balance. The local predictors are (b) local linear and (c) RBF with $\phi(r) = r^3$ with $n_c = 16$

are summarized in the predictor error profiles of Figure 20.5. Each curve shows the fraction of the out of sample points which can be predicted to within a given (absolute) accuracy; for long test sets this should converge to the cumulative probability distribution of the prediction errors. The local RBF predictor is clearly the best; for example, 50% of its predictions have an error of less than $2^{-8} \approx 0.004$, almost a factor of four less than the corresponding level for the local linear predictor. Evaluating a predictor with a single measure of an 'average' error in the prediction may be misleading when a few per cent of the predictions have very large errors. In extreme cases this may result in a predictor which is more accurate 90% of the time having a greater average absolute error. As long as the prediction error profiles are well separated there is no confusion, but if they cross (as in Figure 20.8), it is important to ascertain which properties of the predictions are considered the more important when evaluating predictors; in this instance, different definitions of the error can result in different 'optimal' predictors.

20.3 TIME SERIES APPLICATION

20.3.1 Residence times: transition to a Lagrangian frame

In the time-dependent Stommel flow, the velocity components at a fixed position each show a regular periodic oscillation which is straightforward to predict. We therefore adopt a Lagrangian reference frame, moving with a particle, to show the practical problems involved with the prediction of chaotic systems, and consider the series of residence times in each

cell (defined as the time interval between passages through the lower periodic boundary of the cell). At first glance, this problem appears far removed from the interpolation problem presented earlier. On reflection, however, they are seen to be very similar; Casdagli (1989) and Farmer and Sidorowich (1987; 1988) were the first to make this application to chaotic time series.

Figures 20.6 and 20.7 are derived from two segments of the trajectory which generated Figure 20.4; here the residence time in consecutive cells is given. This spiky appearance, with regions of activity separated by quiescent periods, is common to many observed series in geophysics and is difficult to predict. This series is extremely varied (more so than can be shown in these figures); there are stretches of time during which the particle will fall through dozens of cells without being re-entrained; alternatively, there are cells in which a particle becomes trapped for a large fraction of the total integration time. Although there can be no attractors in this Hamiltonian system, the probability density in these 'reef' regions where the particle is temporarily trapped can become fairly large (Meiss, 1986) and may appear as a strange accumulator (Smith and Spiegel, 1987).

20.3.2 Method of delays

The first step in applying this method of prediction is to *reconstruct* the time series into a geometrical framework. [An introduction to this method is given by Broomhead and Jones (1989).] In many instances, $s(t)$ would be a physical variable measured at regular intervals

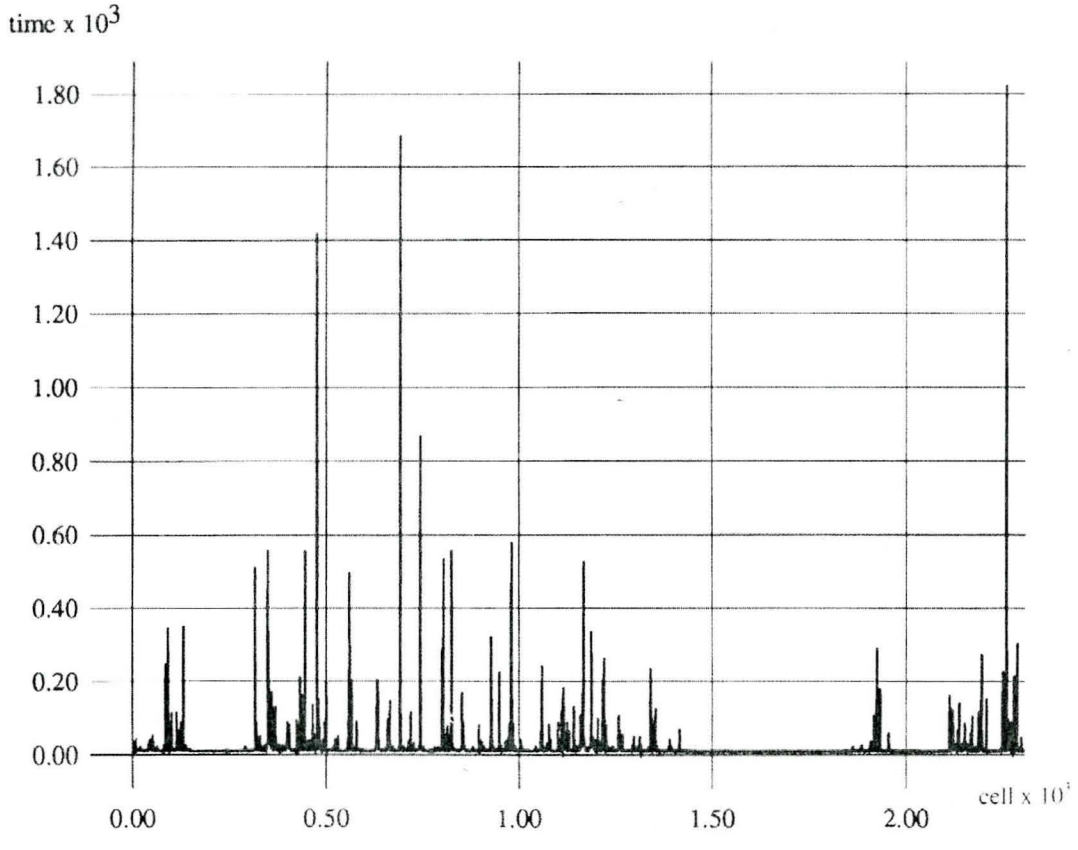


Figure 20.6 Segment of the series of time a particle remains in a given cell; the data corresponds to the same trajectory as the stroboscopic section shown in Figure 20.4. The x axis corresponds to the ith cell; the residence times in a few thousand cells are shown

(τ_s, the sampling time)

$$s_i = s(i\tau_s) \quad i = 1,2,\ldots,n_s \tag{20.14}$$

In our case a 'sampling time' is exactly what we are trying to predict, and the s_i are taken as consecutive observations.

A trajectory, $\boldsymbol{x}(t)$, of this system is reconstructed in M dimensions from a time series of a single observable, $s(t)$, by the method of delays to yield a series of vectors

$$\boldsymbol{x}_i = (s_i, s_{i-j}, \ldots, s_{i-j(M-1)}) \tag{20.15}$$

where j (or $j\tau_s$) is called the delay time, τ_d. (As noted by a referee, the phrase 'delay time' may be misleading in this case, as the subscript i relates to the time spent in the ith cell.) For a deterministic system with phase space dimension M_s and a generic observable, this reconstruction preserves many of the characteristics of the original system for sufficiently large M (see, for example, Casdagli, 1992; Packard *et al.*, 1980; Sauer, Yorke and Casdagli, 1991; Takens, 1981). As shown in the following, multivariate series can also be considered, often with significantly shorter time series in terms of the total duration of the experiment. This is easily understood as multivariate probes can distinguish well separated states in phase space which appear similar to univariate probes due to projection effects. For example, combining the phase of the background flow when the particle entered the cell with the residence time series could improve the predictions *significantly* more than doubling the length of the residence time series. It is the information content, not the length of the data set, which is more important.

Consider the problem of predicting a fixed distance $\tau_p = k$ steps into the future. Once we have the set of observed initial conditions $\boldsymbol{x}_i$ and later observations s_{i+k}, the same

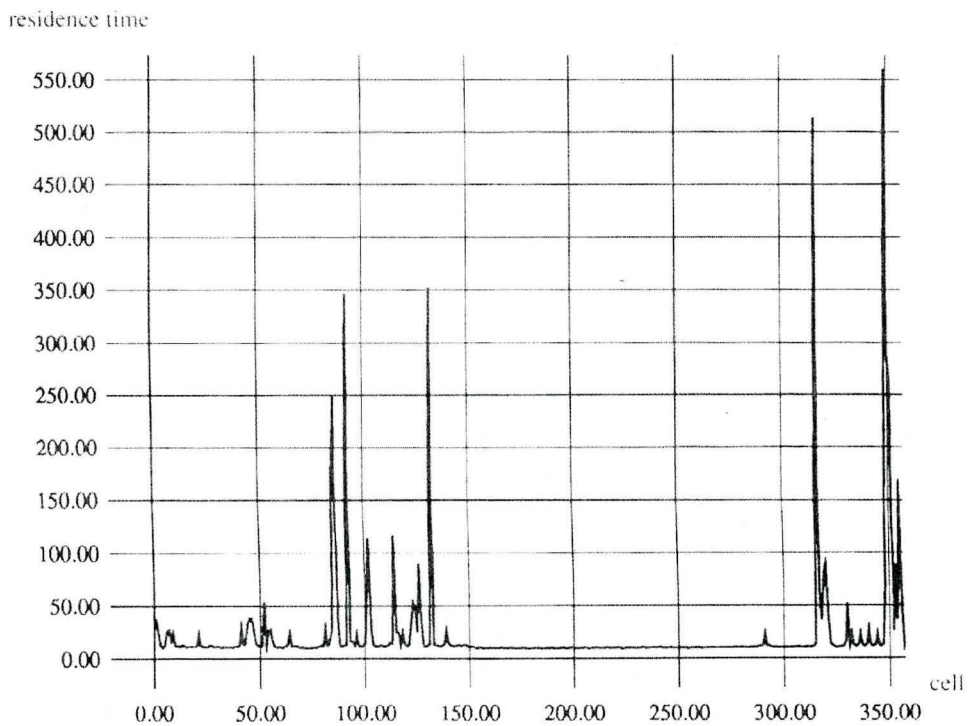


Figure 20.7 Initial segment of the series in the previous figure; only the first 350 residence times are shown here. Note the change in the scale of the vertical axis

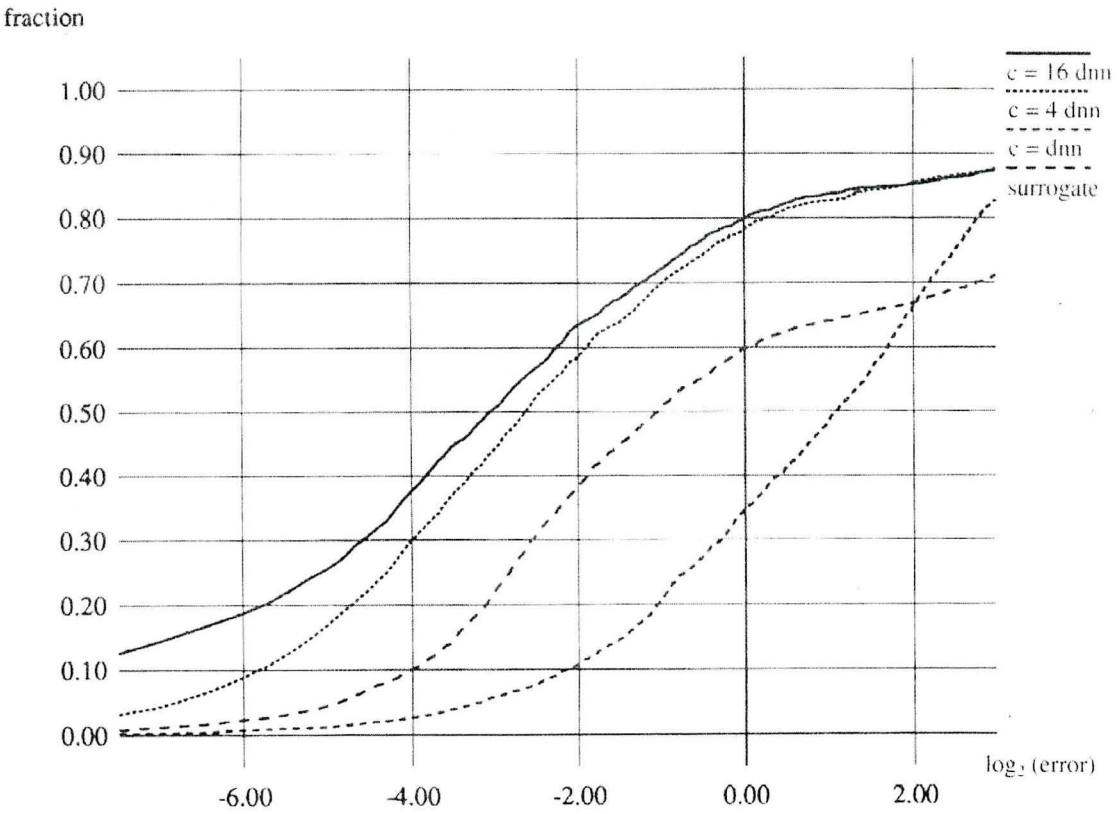


Figure 20.8 Cumulative predictor error used to optimize basis function parameters and show the significance with respect to surrogate data for the prediction of particle residence times. RBF predictors using $\phi(r) = e^{-r^2/c^2}$ for $c = 1$, 4 and 16 times each *local* average nearest neighbour distance. Local linear prediction is of similar quality as the best RBF predictor

machinery set up to solve the two-dimensional example may be used to predict the value of s, k steps ahead in this M-dimensional case. The most immediate difficulties to arise are those of data density in higher dimensional spaces and, of course, visualization of the result.

We will use a learning set consisting of the first 2000 points of a 6000 point series and evaluate the predictors by making one step (i.e. cell) ahead predictions on each of the remaining points. Figure 20.8 provides the prediction error profiles for this example. Three of the curves correspond to different choices for the constant c in the basis function $\phi(r) = e^{-r^2/c^2}$, illustrating how the profile may be used to optimize free parameters. In this instance we knew there was an underlying deterministic system; how would we evaluate the quality of this result if we did not?

20.3.3 Null hypothesis testing

The significance of a result is determined through the consideration of surrogate data and surrogate predictors; other algorithms developed from non-linear dynamical systems theory can also be evaluated in this manner (Smith, 1992; Theiler *et al.*, 1992). In brief, the procedure is to construct a stochastic data series with statistical characteristics similar to the original data and then to determine whether a given algorithm can distinguish the two series. An ensemble of surrogates may be analysed and the probability of a surrogate series producing the observed result may be estimated. Alternatively, surrogate predictors quantify the errors

that should be expected by even the most naïve forecasts, the simplest forecast being either the persistence of the last observed value, or a random choice from the observed distribution.

20.3.4 Construction of surrogate data

The method of surrogate data can be used to quantify the significance of a claim for non-linearity, chaos, or even periodicity. There are a variety of methods for constructing surrogate series; the appropriate method in a specific example will depend on the statistic evaluated and null hypothesis to be tested. In the following we will consider examples which reproduce the distributions observed in, or the autocorrelation function of, the data set. In general, any method of simulation (including, of course, bootstrap and parametric bootstrap methods well known in statistics; Tong, 1990) may be used; however, the correspondence between an algorithm for generating surrogates and a well posed null hypothesis may be obscure.

The basic idea is to take surrogate series from a process which does not, for example, reproduce the sensitivity to initial value found in chaotic series; discovering whether a given algorithm can detect this difference (by distinguishing the chaotic series within an ensemble of surrogates) is the key point. In short, we want the surrogates to reproduce the statistics of the series, but not the physics of the system.

20.3.4.1 Shuffling

A typical method for generating surrogates for iterated systems such as the retention times is to draw from the observed distribution at random; the resulting surrogate series are independent and identically distributed (IID) random variables with the same distribution as the observations. (This is similar to shuffling the data, but allows the construction of arbitrarily long series.) The true retention time signal is easily distinguished from surrogate signals of this type. This should be the case whenever there is some memory in the system — that is, when the expected value of the next observation is conditioned on the current value. A simple method to simulate this conditional probability distribution is given in the following.

20.3.4.2 Malicious shuffling

A method of shuffling which retains the general association between consecutive observations provides more realistic surrogates. For the residence time series, short residence times often follow very short residence times, whereas long residence times tend to follow those of intermediate length. We may produce a surrogate series with this quality from a series of N observations with the following algorithm:

(1) Sort the first $(N - 1)$ members of the series into increasing order, recording the value which immediately followed each member in the unsorted series (i.e. its image).
(2) Divide the sorted list into K subgroups.
(3) Pick a member from the entire series at random; this is the first data point of the surrogate series.
(4) Determine which of the K subgroups the *image* of this value falls in, and pick an element from this subgroup at random as the next surrogate data point.
(5) Repeat step 4 until a series of the desired length is achieved.

These surrogates cannot only preserve the distribution of the original data set, but can also

approximate the (one-step) conditional probability distribution [i.e. $P(s_i = c|s_{i-1} = b)$] depending on the value chosen for K and are, in essence, Markov chains. For $K = 1$, the surrogates are IID, alternatively, for $K \approx N$ segments of the original are reproduced. [Assuming that the observations are distinct (i.e. $s_i \neq s_j$ for $i \neq j$) thereby avoiding an ambiguity often introduced by quantization effects. In this case, the final image used must be grouped with an earlier observation, hence $K < N - 1$.] The choice of a particular value for K depends on the structure of the distribution. Although surrogate generators will provide the advertised results in the long run, it should always be verified that, for the desired length of surrogate series, the *observed* distributions of surrogate data are not sensitive to either the choice of K or that of the location of the boundaries between subgroups.

This variant of shuffling was originally envisaged for series which displayed distinct 'modes' of behaviour; modes which, in turn, were reflected by the magnitude of the observation. The subgroups are (individually) IID and there is a fixed probability for remaining in a given subgroup. As in the following example, *K was taken to be much less than N*. This appears more natural for iterated systems than for those which evolve smoothly in time: the one-step conditional probability distribution is often blind to 'trends' in stationary series. [Exceptions exist in processes whose conditional probability distributions change under time-reversal (i.e. $P(s_i = c|s_{i-1} = b) \neq P(s_{i-1} = b|s_i = c)$), like a saw-tooth series.] It should also be noted that it is straightforward to extend this method of shuffling to maintain correlations over periods longer than one step [e.g. $P(s_i = c|s_{i-1} = b, s_{i-3} = a)$].

Not surprisingly, the expected prediction error of surrogate series with $K = 5$ is lower than that for the IID surrogates produced by straight shuffling (equivalent to $K = 1$); none the less, the observed series can be distinguished from surrogates of this type, as illustrated in Figure 20.8.

20.3.4.3 Fourier transforms

Surrogates which test whether 'good' predictions result from the autocorrelation function alone may be generated with Fourier transforms (Osborne *et al.*, 1986; Smith, 1992; Theiler *et al.*, 1992). A surrogate is generated through the inverse Fourier transform of the observed Fourier amplitudes and random values for the phase of each frequency component. As noted by Theiler *et al.* (1992), this is equivalent to testing whether the signal is distinguishable from a linear stochastic noise. This method is used for the Severn data in the next section.

20.4 DYNAMIC RECONSTRUCTIONS OF THE RIVER SEVERN

We now turn to the analysis of velocity time series taken on the River Severn. The data consists of simultaneous, three-component velocity measurements taken in 1989 at the Leighton New Upstream Pool Site; details of the collection and previous analysis of this and related data sets are given elsewhere in this volume (Heslop *et al.*, 1994) and in Beven and Carling (1992) and Heslop and Allen (1990). It should be noted that, due to noise contamination, the analogue signal was redigitized and low-pass filtered (Holland, 1991); it is possible this would have a negative impact on the current analysis. Renormalization would not affect reconstructions of the individual components, but would affect the analysis of the total velocity series.

20.4.1 'Predicting' the total velocity from observed components

As stressed earlier, the non-linear models applied here are constructed completely from the data; no underlying model of the system is required (although additional information can be used). This makes the method sensitive to variations in the density of data points in the reconstruction space; in particular, when the system is in a region of reconstruction space not explored during the learning set, these predictions correspond to extrapolations (rather than interpolations) and their uncertainty is high even when the other assumptions of the technique are satisfied. To stress the importance of this effect, and the sensitivity to details of the reconstruction, we will pose a simple problem: Given the observations v_x, v_y, v_z, what is the total velocity? This test, and its intermittent failure, also serves to stress that no underlying model of the system is assumed.

Figure 20.9 shows the results of a local linear 'prediction' of the current total velocity given its three component parts (i.e. $\tau_p = 0$). Of course, this computation can be performed exactly as we know the underlying relationship; this information is not available to the non-

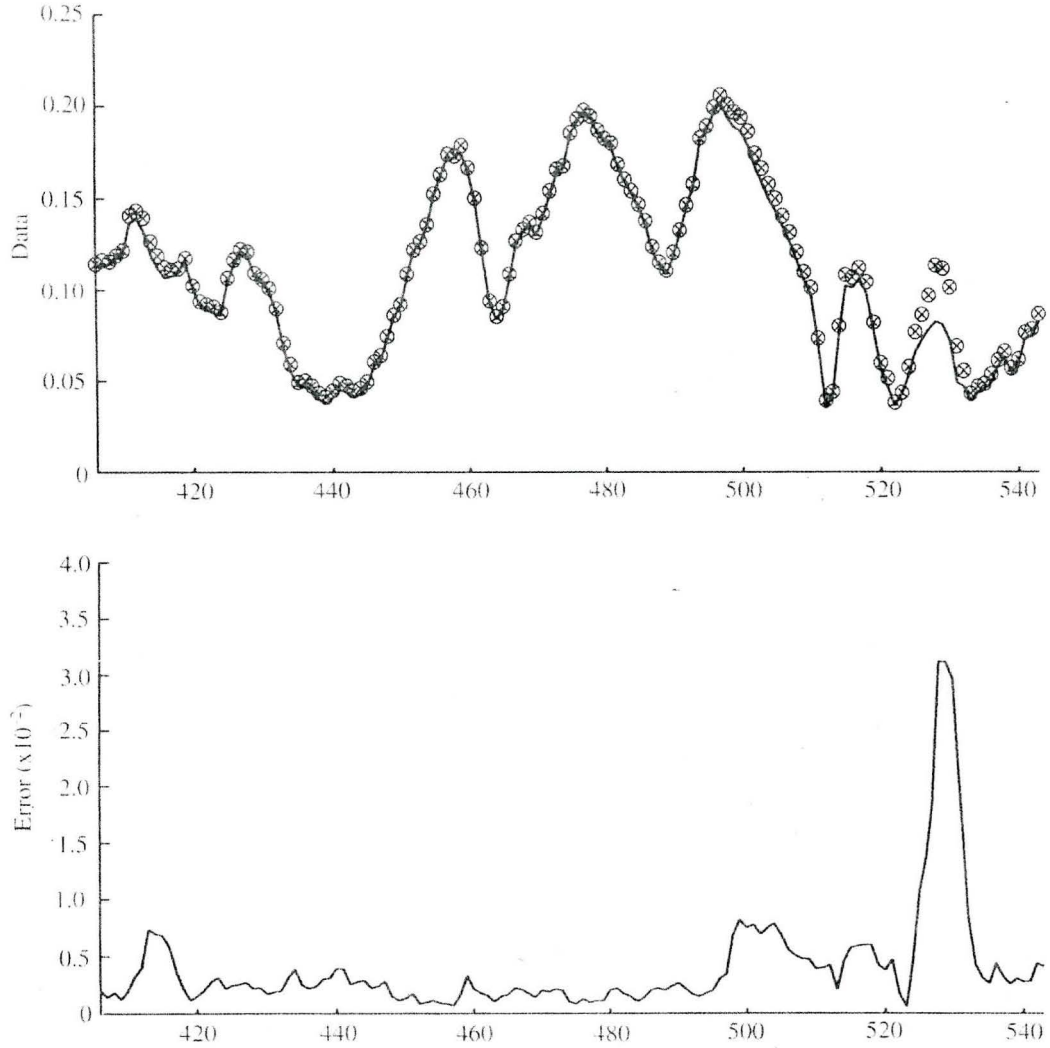


Figure 20.9 Observed (solid) and 'predicted' (symbol) values of the total velocity from a local linear predictor using 32 nearest neighbours (upper). Simultaneous time series of the prediction error (lower)

linear model; regions of poor fit (e.g. near $t = 525$) correspond to areas of relatively low data density.

Local non-linear predictors cope better with this difficulty, in part because they do not assume that the local behaviour is linear and hence can more accurately interpolate low data regions where it is indeed not linear. Non-linear predictors, in turn, can over fit the noise in the data and are sensitive to the additional parameters in their definition (for example, the constant c in $\phi(r) = e^{-r^2/c^2}$). What is needed is a robust method to determine how much of the local structure should be reproduced by the predictor; some progress in this area has been achieved in collaboration with A. Mees (manuscript in preparation). The results for the total velocity 'prediction' are summarized in Figure 20.10, where the optimum RBF predictor reduces the error with which 50% of the predictions are made by more than a factor of 32 over the local linear case illustrated in Figure 20.9.

20.4.2 Predicting short-term fluctuations in velocity

Finally, we consider the 'standard problem' of time series prediction. Given a single, univariate time series of the total velocity, $v(t)$, how well can we predict a future value $v(t + \tau_p)$? For the Severn data with $\tau_p = 2$ and $\tau_p = 7$, predictor error profiles indicate that the best predictors have a delay time $\tau_d = 1$ and either $M = 2$ or $M = 3$. This has implications for future data collection. The short delay time indicates that a shorter sampling rate would be useful, whereas the low embedding dimension also points to data density problems. In addition to requesting a longer, more densely sampled data stream, comparison with Fourier transform surrogate series indicates that, for short times, the observed series ($\tau_p = 2\tau_s$) may be slightly more predictable than the surrogates, but for longer time scales ($\tau_p = 7\tau_s$) the observed series can be distinguished as significantly *less* predictable than the linear stochastic surrogates.

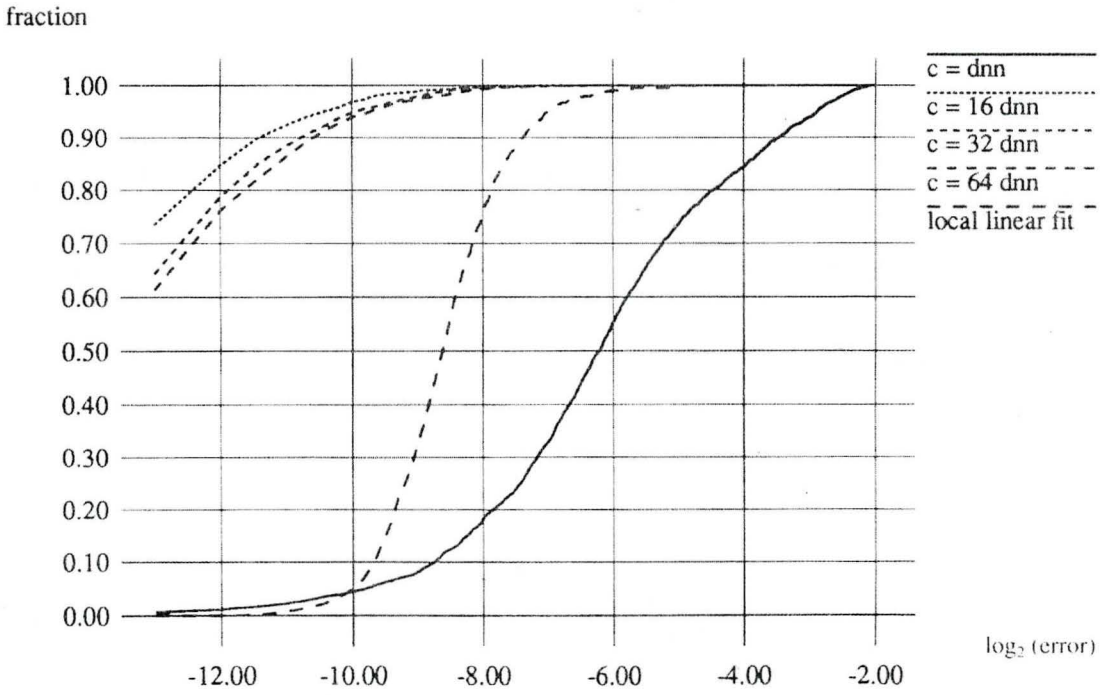


Figure 20.10 Prediction error profile for local linear and RBF predictors using $\phi(r) = e^{-r^2/c^2}$ for $c = 1$, 16, 32 and 64 times the *local* average nearest neighbour distance, d_{nn}

This result indicates that the series contains significant non-linearities which cannot be captured by linear models. In addition, it implies that the data density is too low to fill out the reconstruction space — that is, the data record is too short to allow sufficient recurrence in most regions of the reconstruction space even though the statistics of the distribution of data values may appear stationary. Examination of the series shows that many distinct types of behaviour are indeed observed.

Predictions from a three-dimensional reconstruction based on the three simultaneous velocity components resulted in an improvement in the case $\tau_p = 7\tau_s$. The average absolute errors and predictor profiles of this predictor are also better than those of a persistence predictor. This interesting result indicates that the orientation of the velocity vector is an important diagnostic of short time fluctuations.

20.5 DISCUSSION

We should also address the question of whether these techniques are worth applying. The most robust answer is yes. From the point of view of simplicity, low dimensional chaotic systems are intermediate between periodic (or static) systems and rich turbulent systems; without performing tests such as these, they cannot be dismissed. From a more pragmatic point of view, these techniques can be useful in the event that the dynamics are simple but *not* deterministic. For example, reasonable predictions are obtained when these methods are applied to data originating from non-linear stochastic models (and also for linear ARMA type models). If the system is equivalent to a linear ARMA model, the parameters of which are well estimated, then the techniques discussed here can add little more. If, on the other hand, the observations are non-linear and not easily transformed to linearity (Theiler *et al.*, 1992), then deterministic predictors which are robust in the presence of noise should yield good predictions (in the sense of expected values). The data requirements needed to apply these techniques are currently under investigation; ultimately, it will most likely be their relative speed of convergence (as the amount of data in the learning set increases) which will determine their usefulness.

We have illustrated a non-linear prediction technique and presented the initial results of an analysis of data from the River Severn. The analysis shows (1) that the velocity series contains important non-linearities which cannot be described within a linear model and (2) that the three-dimensional orientation of the velocity vector is relevant to the short-term evolution. The most serious limitations encountered appear to arise from a lack of recurrence (low data densities), implying the need for longer data series. There is little doubt that series such as this, and those of geophysical systems in general, will provide a tough proving ground for these techniques.

ACKNOWLEDGEMENTS

I thank Keith Beven for introducing me to this problem and for his patience, and Alistair Mees for discussion on a wide range of issues. I also thank Mark Muldoon and Catherine Lanone for a critical reading of the manuscript. I met Cath Allen only once, while visiting Lancaster, and am still struggling to solve several of the many questions she asked after my seminar. I would like to express my admiration of her courage and that of other scientists strong enough to battle both cancer and ignorance simultaneously.

REFERENCES

Beven, K. and Carling, P. (1992) Velocities, roughness and dispersion in the lowland River Severn. In: *Lowland Floodplain Rivers: Geomorphological Perspectives* (Eds P. Carling and G.E. Petts). Wiley, Chichester, 71–93.

Broomhead, D.S. and Jones, R. (1989) Time-series analysis. *Proc. R. Soc. London* **423**, 103–121.

Broomhead, D.S. and Lowe, D. (1988) Multivariable functional interpolation and adaptive networks. *J. Complex Syst* **2**, 321–355.

Broomhead, D.S. and Ryrie, S.C. (1988) Particle paths in wavy vortices. *Nonlinearity* **1**, 409–434.

Casdagli, M. (1989) Nonlinear prediction of chaotic time series. *Physica D* **35**, 335–356.

Casdagli, M. (1992) Chaos and deterministic *versus* stochastic non-linear modeling. *J. R. Statist. Soc. B* **54**, 303–328.

Chaiken, J., Chu, C.K., Tabor, M. and Tan, Q.M. (1987) Lagrangian turbulence and spatial complexity in stokes flow. *Phys. Fluids* **30**, 687–694.

Corrsin, S. (1950) Turbulent flow. *Am. Sci.* **38**, 300–325.

Crisanti, S., Falcioni, Provenzale, A. and Vulpiani (1990) Passive advection of particles denser than the surrounding fluid. *Phys. Lett. A.* **150**, 79.

Crisanti, S., Falcioni, Provenzale, A. and Vulpiani Dynamics of passively advected impurities in simple two-dimensional flow models. *Phys. Fluids A*, in press.

Farmer, J.D. and Sidorowich, J. (1987) Predicting chaotic time series. *Phys. Rev. Lett.* **59**, 8.

Farmer, J.D. and Sidorowich, J. (1988) Exploiting chaos to predict the future and reduce noise. In: *Evolution, Learning and Cognition* (Ed. Y.C. Lee). World Scientific, 277.

Frisch, U., Sulem, P.-L. and Nelkin, M. (1978) A simple dynamical model of intermittent fully developed turbulence. *J. fluid Mech.* **87**, 719–736.

Gollub, J.P. and Solomon, T.H. (1987) Complex particle trajectories and transport in stationary and periodic convective flows. In: *Chaos and Related Nonlinear Phenomena: Where do we go from here?* (Ed. I. Procaccia.) Plenum Press.

Heslop, S.E. and Allen, C.M. (1990) Turbulence and dispersion in larger UK rivers. In: *Proceedings 23rd Congress of the International Association of Hydraulic Research, Ottawa, Canada*, D75–D82.

Heslop, S.E., Holland, M.J. and Allen, C.M. (1994) Turbulence measurements in the River Severn. In: *Mixing and Transport in the Environment* (Eds. K. Beven, P.C. Chatwin and J.H. Millbank). Wiley, Chichester, Ch. 3.

Holland, M.J. (1991) A detailed analysis of turbulence on the River Severn. Preprint. Dept. of Environmental Sciences, University of Lancaster.

Huppert, H.E. (1984) Lectures. In: *Lectures on Geological Fluid Dynamics* (Ed. W. Malcus). *Woods Hole Oceanogr. Inst. Tech. Rep. WHOI-84-44*, WHOI, Woods Hole, 1–71.

Huppert, H.E. (1991) Buoyancy-driven motions in particle-laden fluids. In: *Of Fluid Mechanics and Related Matters* (Eds R. Salmon and D. Betts). Scripps Institution of Oceanography Ref. Ser., San Diego, 141–159.

Mackay, R.S. and Meiss, J.D. (Eds) (1987) *Hamiltonian Dynamical Systems*. Adam Hilger, Bristol.

Maxey, M.R. and Corrsin, S. (1985) Gravitational settling of aerosol particles in randomly oriented cellular flow fields. *J. Atmos. Sci* **43**, 1112–1134.

Meiss, J.D. (1986) Class renormalization: Islands around islands. *Phys. Rev. A* **34**, 2375–2383.

Osborne, A.R., Kirwan, A.D., Provenzale, A. and Bergamasco, L. (1986) A search for chaotic behavior in large and mesoscale motions in the Pacific Ocean. *Physica D* **23**, 75–83.

Packard, N.H., Cruchfield, J:P., Farmer, J.D. and Shaw, R.S. (1980) Geometry from a time series. *Phys. Rev. Lett.* **45**, 712.

Pasmanter, R.A. (1987) Deterministic diffusion. In: *Proceedings of the International Symposium on Physical Processes in Estuaries* (Ed. W. Van Leussen). Springer Verlag, New York.

Powell, M.J.D. (1985) Radial basis functions for multivariate interpolation: a review. In: *IMA Conference on Algorithms for the Approximation of Functions and Data*. RMCS, Shrivenham.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1987) *Numerical Recipies*. Cambridge University Press, Cambridge.

Sauer, T., Yorke, J.A. and Casdagli, M. (1991) Embedology. *J. Stat. Phys.* **65**, 579–616.

Smith, L.A. (1984) Particulate dispersal in a time-dependent flow. In: *Lectures on Geological Fluid Dynamics* (Ed. W. Malcus). Woods Hole Oceanogr. Inst. Tech. Rep. WHOI-84-44, WHOI, Woods Hole, 243–265.

Smith, L.A. (1987) Lacunarity and chaos in nature. *PhD thesis*, Columbia University, New York.
Smith, L.A. (1992) Identification and prediction of low-dimensional dynamics. *Physica D* **58**, 50–76.
Smith, L.A. and Spiegel, E.A. (1985) Pattern formation by particles settling in viscous flows. In: *Macroscopic Modeling of Turbulent Flows*. Vol. 230. *Lecture Notes in Physics*. (Eds. U. Frisch *et al.*). Springer-Verlag, New York.
Smith, L.A. and Spiegel, E.A. (1987) Strange accumulators. In: *Chaotic Phenomena in Astrophysics*. (Eds J.R. Buchler and H. Eichhorn). *Ann. NY Acad. Sci.* **497**, 61–65.
Stommel, H. (1949) Trajectories of small bodies sinking slowly through convection cells. *J. Mar. Res.* **8**, 24–29.
Takens, F. (1981) Detecting strange attractors in fluid turbulence. In: *Dynamical Systems and Turbulence*. Vol. 898. (Eds D. Rand and L.S. Young). Springer-Verlag, New York, 366.
Theiler, J., Eubank, S., Longtin, A., Galdrikan, B. and Farmer, J.D. (1992) Testing for nonlinearity in time series: The method of surrogate data. *Physica D* **58**, 77.
Tong, H. (1990) *Non-Linear Time Series Analysis*. Oxford University Press, Oxford.
Tooby, P.F., Wick, G.L. and Isaacs, J.D. (1977) The motion of a small sphere in a rotating velocity field: A possible mechanism for suspending particles in turbulence. *J. Geophys. Res.* **82**, 2096–2100.
Yu, L., Grebogi, C. and Ott, E. (1990) Fractal structure in physical space in the dispersal of particles in fluid. In: *Non-linear Structure in Physical Systems* (Eds L. Lam and H.C. Morris). Springer-Verlag, New York, 223–231.