Article:
Taylor, CC and Zempeni, A (2004) Chain Plot: A Tool for Exploiting Bivariate Temporal Structures. Computational Statistics and Data Analysis, 46 (1). 141 - 153 . ISSN 0167-9473

https://doi.org/10.1016/S0167-9473(03)00120-8

Chain Plot: A Tool for Exploiting Bivariate Temporal Structures

C.C. Taylor

Dept. of Statistics,
University of Leeds,
Leeds LS2 9JT, UK

A. Zempléni

Dept. of Probability Theory & Statistics,
Eötvös Loránd University,
Pázmány sétány 1/C,
Budapest, H-1117

Abstract

In this paper we present a graphical tool useful for visualizing the cyclic behaviour of bivariate time series. We investigate its properties and link it to the asymmetry of the two variables concerned. We also suggest adding approximate confidence bounds to the points on the plot and investigate the effect of lagging to the chain plot. We conclude our paper by some standard Fourier analysis, relating and comparing this to the chain plot.

Keywords: Asymmetry of Variable Levels; Environmental; Exploratory Data Analysis; Pollution; Symmetry; Periodic Time Series; Visualization.

1 Introduction

The idea we investigate in this paper has emerged during a relatively simple-looking problem in data analysis. We were given a data set from an automatic measurement station located at Szeged, Southeastern Hungary. Environmental (climate and pollution) measurements were collected with readings every half an hour over a 4-year period. For a detailed description and alternative analysis of the data set see Makra *et al*. (2001). The method we describe in this paper was found to be very useful for the data given. It is generally applicable to the analysis of bivariate time series with cyclic, or seasonal, components.

We suggest the following plot as a visualization of the joint behaviour of the daily pattern of certain pollutants. Let us suppose the time series $(X_t)_{t=1}^T$ and $(Y_t)_{t=1}^T$ have a periodic component with length N (in our case X_t and Y_t are half-hourly readings for two pollutants at the time point t, so $N = 48$). Let

$$\overline{X}_k = \frac{N}{T}\sum_{i=1}^{T/N} X_{k+(i-1)N} \qquad \overline{Y}_k = \frac{N}{T}\sum_{i=1}^{T/N} Y_{k+(i-1)N} \qquad k = 1, \ldots, N. \tag{1}$$

Figure 1 is a scatter plot of these values on the x, y axis, labelled by k, together with the usual separate time series plot of the two components. In this example one of the series has a bimodal structure, whereas the other series is roughly unimodal though asymmetric. The chain-like pattern gave us some ideas for further investigation which we present in the following sections.

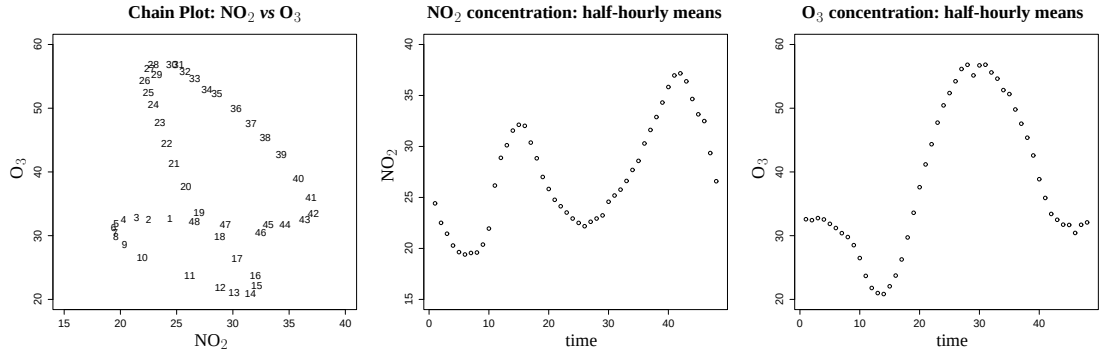

Fig. 1. Half-hourly means for two pollutants: simple chain plot and marginal plots

We investigate behaviour of the chain plot for deterministic functions in Section 2, and bootstrap methodology for inference in Section 3. Some statistical applications are in section 4, and a brief discussion concludes the paper.

2 Chain plot for deterministic functions

In this section we consider continuous, deterministic functions rather than random variables, as this allows us to prove simple results, which have obvious applications to the original setup as well. We suppose our functions $x(t)$ and $y(t)$ to be bounded, continuous and defined for $t \in [0, 1]$. We note here that the time-scale transformations have no effect on the suggested chain plot. Let us consider the simpler of the two, say x, as the reference function. As we want to get chain plots rather than open-ended line plots, we confine ourselves to the cases $x(0) = x(1)$, $y(0) = y(1)$ (which automatically holds for the motivating example of periodic time series). In this setup, the formal definition of the chain plot is $\mathcal{A} := \{(x(t), y(t)) : t \in [0, 1]\}$. This is of course a closed curve, and we shall investigate its properties below, which are relevant to the statistical problem under consideration.

As a motivation of our results, we show another example of a chain plot in Figure 2 where we observe an almost one-dimensional behaviour. This is substantially different from Figure 1. What is the main reason behind these differences?

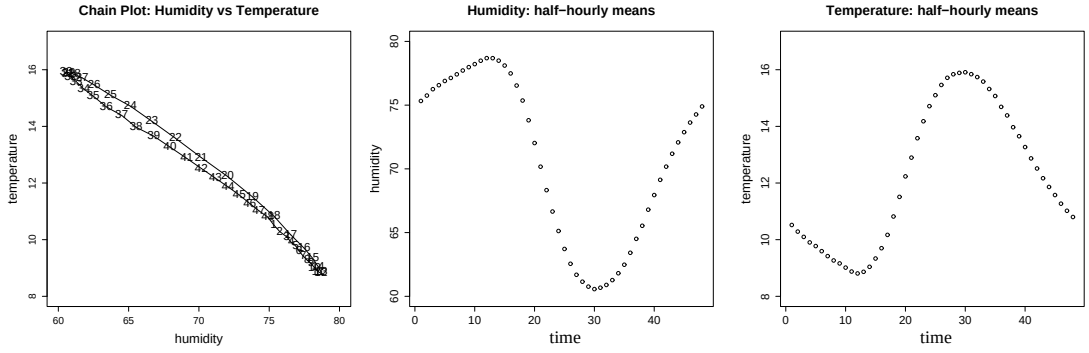

Fig. 2. Half-hourly means for climate variables: connected chain plot and marginal plots

Let us define $A(x, y)$ as the total area within the closed curve, where

we might omit the arguments if it does not cause confusion.

2.1 The simplest case

In this section we introduce the main notions of this paper in a setup which allows an easy interpretation.

Definition 1 *We say our reference function x is* simple *if it is strictly monotonically increasing in $[0, t_0]$ and strictly monotonically decreasing in $[t_0, 1]$.*

Remark 1 *We note that we have not claimed the derivative of x to exist at all the points.*

We supposed that $x(0) = \min\{x(t) : t \in [0, 1]\}$, but it is not a real condition, as the endpoints of the interval can be chosen arbitrarily. We have not posed any conditions for the set $\{x(t) : t \in [0, 1]\}$, but in order to make the chain plot area for different pairs of functions (x, y) to be comparable, it is advised to normalize all the variables.

There are lots of real-life cases, where one of the components of the bivariate time series can be considered as a simple function (temperature over a day being the most obvious example, see also Figure 2). The conditions of the following lemma are not at all unrealistic in real-life examples.

Let x be a simple function and define $x_1 : [0, t_0] \to \mathbb{R}$ as $x|_{[0,t_0]}$ and $x_2 : [t_0, 1] \to \mathbb{R}$ as $x|_{[t_0,1]}$.

Lemma 1 *Let x be simple and suppose that $\mathcal{A}$ consists of a single chain (*i.e. *it is homeomorphic to a circle). Then*

$$A = \left| \int_{x(0)}^{x(t_0)} y(x_1^{-1}(u))du - \int_{x(0)}^{x(t_0)} y(x_2^{-1}(u))du \right|. \qquad (2)$$

Proof We just have to observe that Definition 1 and the remarks afterwards imply that the chain can uniquely be cut into two parts, where the cutting points are its unique minimal and maximal values

along the x-axis:

$$\{(x(t), y(t)) : t \in [0,1]\} = \{(u, y(x_1^{-1}(u))) : u \in [x(0), x(t_0)]\} \cup \{(u, y(x_2^{-1}(u))) : u \in [x(0), x(t_0)]\} \quad (3)$$

As the two parts defined in (3) have no intersection points in their interior, the assertion (2) is a simple consequence of (3).

In the remainder of this section we use a transformation, which is again the easiest to be introduced for simple reference functions. It is similar to the probability-integral transformation — see Embrechts *et al.* (1999) for example — used to transform marginal distributions of bivariate random variables to uniform ones. We use the transformation for one coordinate only, in order not to change the value of A. In our actual deterministic world, the role of the uniform distribution is of course played by the function $x(t) = ct$ or $x(t) = c(1-t)$. Let

$$(\tilde{x}(t), \tilde{y}(t)) := \begin{cases} (2tx(t_0), y(x_1^{-1}(2tx(t_0)))) & \text{if } 0 \le t \le 0.5 \\ (2(1-t)x(t_0), y(x_2^{-1}(2(1-t)x(t_0)))) & \text{if } 0.5 < t \le 1 \end{cases} \quad (4)$$

Remark 2 *It is obvious that transformation (4) does not affect the chain plot.* $\tilde{x}$ *is a simple function, too.*

We now reformulate — and at the same time generalize — our previous result (Lemma 1). This makes it easier to understand the meaning of the area A we investigate and even more importantly it allows further generalizations.

Proposition 1 *Let* x *be simple. Then*

$$A = \int_0^{0.5} |\tilde{y}(u) - \tilde{y}(1-u)| du. \quad (5)$$

Proof By Remark 2 we know that A can be calculated using the functions $(\tilde{x}, \tilde{y})$. Since x is simple A is a union of the areas of simple chains, defined by the intersection points: $\{u \in [0,1] : \tilde{y}(u) = \tilde{y}(1-u)\}$. For each simple chain, the area can be calculated by Lemma

1. It only has to be observed that the continuity of x and y implies that on the whole domain of integration either $\tilde{y}(u) > \tilde{y}(1-u)$ or $\tilde{y}(u) < \tilde{y}(1-u)$, so we can move the absolute value into the integral.

Formula (5) turns out to be important for calculating the area of a chain plot for observed data.

Definition 2 *The* asymmetry index *of y with respect to a simple x is defined as $AI(y,x) := A$.*

Analogues of this definition can be found in the literature: for the dependence function of bivariate extremes, an analogous definition was given in Villa-Diharce (2001). Our definition is easily motivated, see the first of the following properties:

(i) $AI(y,x) = 0 \iff \tilde{y}(u) = \tilde{y}(1-u)$ holds for all $u \in [0,1]$. This means that the behaviour of y during the period of which x increases is exactly symmetric to its behaviour during the decrease of x.

(ii) If $0 \leq y \leq 1$, then $0 \leq AI(y,x) < 1$, where the latter inequality can be changed into "smaller or equal" if we allow for noncontinuous y: $y = \delta_{[t_0,1]}$ has $AI(y,x) = 1$.

(iii) The visible asymmetry in Definition 2 can be easily resolved if y is itself simple, since then we can choose y as the reference function and thus we are allowed to define $AI(x,y) = A$ as well.

Neither co-ordinates of Figure 1 are simple functions (even after shifting the time scale to ensure $x(0) = \min\{x(t)\}$), so we must introduce necessary modifications in order to cover this and similar cases.

2.2 Some generalizations

Our aim is by no means to find the possible boundaries of the generalizations, since we are mostly interested in questions arising from statistical analysis.

So we loosen the conditions imposed for our reference function x to such an extent only, which allows the investigations of practically all real-life applications.

Definition 3 *We say our reference function x is* normal *if there are points $0 = s_1 < t_1 < s_2 < t_2 < s_3 < \cdots < t_m < s_{m+1} = 1$ such that x is strictly monotonically increasing over $[s_i, t_i]$ and x is strictly monotonically decreasing over $[t_i, s_{i+1}]$ or it is constant over the whole interval $[s_i, s_{i+1}]$ ($i = 1, \ldots, m$) and $\min\{x(t) : t \in [0, 1]\} = x(0)$.*

Remark 3 *The simple reference functions are exactly those normal ones, which are non-constant and for which $m = 1$.*

Now the definition of the area of the chain plot is far from straightforward, as in this more complicated case several inner loops and unusual configurations might arise (we do not think them to be very common in real applications). We choose one possible definition, which is just an iteration of cases defined in the previous section.

Let $\mathcal{A}$ be a chain plot which intersects itself by finitely many points. Then we define its area recursively as $A = A_1 + A^*$, where A_1 is the area of the first subchain (which is obtained simply by drawing the plot until it is first closed). We then omit this part from the chain ($\mathcal{A}^*$ is the remainder), and define its area similarly (this is denoted by A^*). More formally:

Step 1 Let $u_1 := \min\{u : \exists v < u : x(u) = x(v), y(u) = y(v)\}$ (let v_1 be the corresponding v) and the chain $A_1 := \{(x(t), y(t)) : v_1 \leq t \leq u_1\}$. (It should be noted that extreme cases, where one segment of the chain exactly coincides with another segment, are excluded from this definition — but one can get rid of such parts by moving one of its components slightly, for example).

Step 2 The remaining part of the chain is then $\{(x(t), y(t)) : 0 \leq t \leq v_1\} \cup \{(x(t), y(t)) : u_1 \leq t \leq 1\}$ and go back to Step 1 (to find the next sub-chain etc.)

Remark 4 *Our definition of the area makes it possible to calculate a*

certain area twice (see Figure 3), but it seems to be logical, as these parts play a multiple role in the asymmetry.

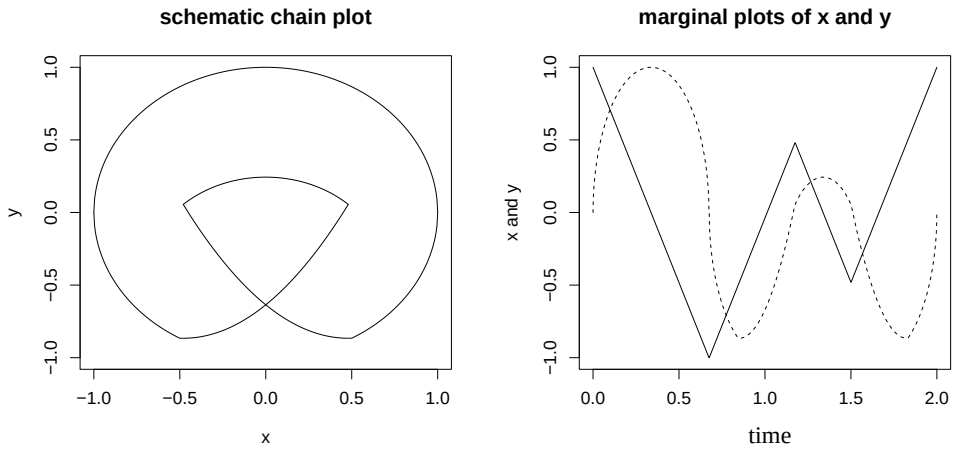

Fig. 3. Left: A schematic chain plot, in which there is a region which is counted twice when the total area is calculated. Right: the two functions $x(t)$ (line) and $y(t)$ (dashed).

Now we have reduced our task of calculating the area of a chain plot, to the calculation of A_i, which is homeomorphic to a circle — but of course not always simple in the sense of our Definition 1. Practical calculations of such an area can easily be performed by numerical methods. But in order to prepare the statistical procedures of section 3 we sketch an iterative algorithm for the reduction of the normal case to the sum of simple ones as follows.

For a non-intersecting chain plot with a normal reference function, we can calculate its area iteratively: Let us use the notation of definition 3 and let $\varepsilon > 0$ be so small that $\exists\ u_0 \in [0, t_1] : x(u_0) = x(s_2) - \varepsilon$. Then the line segment

$$I := \{(x(s_2) - t\varepsilon, y(s_2) - (y(s_2) - y(u_0))t) : t \in [0, 1]$$

is entirely within the closed curve. Now the chain

$$\{(x(u), y(u)) : u_0 \leq u \leq s_2\} \cup I \tag{6}$$

is a chain with a simple reference function and the x coordinate of the remainder chain

$$\{(x(u), y(u)) : 0 \leq u \leq u_0\} \cup I \cup \{(x(u), y(u)) : s_2 \leq u \leq 1\} \tag{7}$$

has exactly one less maximum point, so the iteration is finished in just a finite number of steps.

As a final remark, we mention that the total area of the chain plot is the sum of local asymmetries for subchains, corresponding to the constructed simple parts of the normal reference function.

3 Statistical applications

3.1 Asymmetry index-calculation for observed data

Using the notions introduced in the previous section, one can calculate the asymmetry index for the observed data. In order to do so, one only needs to interpolate the values between the observation points $(\overline{X}_t, \overline{Y}_t)$ and $(\overline{X}_{t+1}, \overline{Y}_{t+1})$. The simplest method is the linear interpolation, which we preferred in the current paper.

Carrying this out for the data under investigation, we got the following results. After normalizing both variables so that the minimum is 0 and the maximum is 1 — which is preferred here to the more usual normalization based on the standard deviation, since that would allow arbitrarily large values for A in spite of the bounded variance — we get area 0.453 for the chain plot in Figure 1 (NO_2 *vs.* O_3) and 0.053 for the chain plot in Figure 2 (humidity *vs.* temperature). For a chain plot which forms a circle, the (normalized) area is $\pi/4 = 0.785$, and a square gives area 1 which is the maximal value for plots which only include areas once. For the plot in Figure 3 the asymmetry index is 0.879.

3.2 Alignment and Lagging

The structure of the chain plot (its area and number of intersecting points) is associated with the general alignment of the two time series. To illustrate this further, and to indicate another possible use of chain plots, we plot *lagged* versions in Figure 4. In this case we can see that

the first part of the series is “aligned” (though possibly with opposite sign) when the y variable is lagged $+2$ whilst the second part of the day is closely aligned when the y variable is lagged -4. This means that by finding those lags which provide the smallest area of the chain plot, we can find those shifts where the two variables show the most symmetrical behaviour (this lag can be interpreted as the lag for the effect of one variable on another).

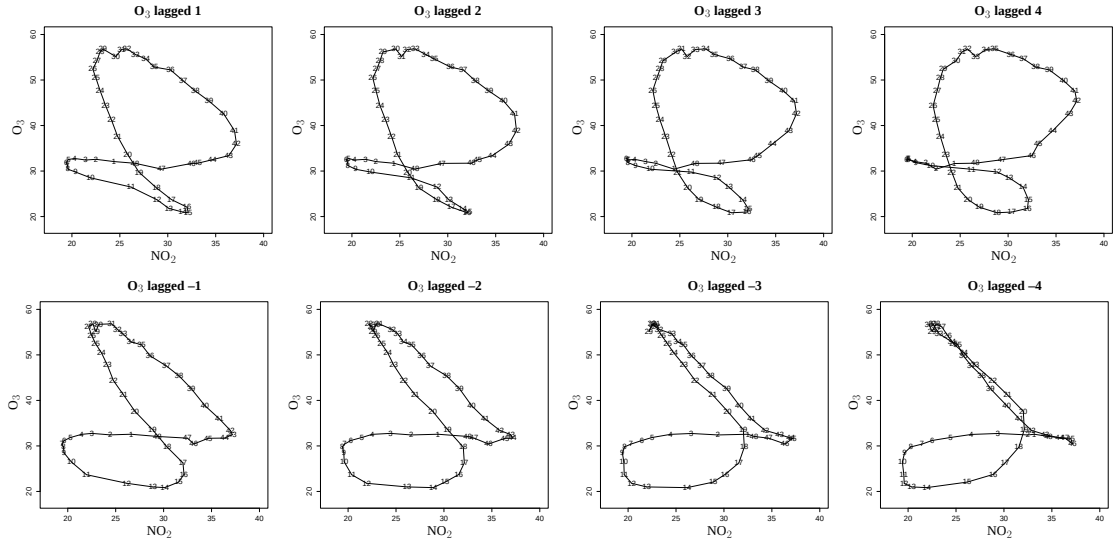

Fig. 4. Connected chain plots for lagged series. Top row: positive lags; Bottom row: negative lags

3.3 A method for testing symmetry

Our aim in this subsection is to give a procedure for testing symmetry between the variables. It is worth mentioning that in some cases, the physical processes behind the observed data suggest time lags (see Figure 4). In order to cope with this phenomenon, we first minimize the asymmetry index with respect of possible lags:

$$A_{\min} := \min_m AI(\overline{X}_t, \overline{Y}_{t-m}) \tag{8}$$

From now on let us suppose that the Y has been shifted so that the question is the symmetry of $(\overline{X}_t, \overline{Y}_t)$ is the most symmetric data set for which the symmetrical behaviour with respect to the reference $\overline{X}$ is investigated. The question we consider is the following: can the

symmetry be accepted, based on an observed asymmetry index A?

In order to tackle this question, we might either use parametric modelling with a symmetric model:

$$\overline{Y}_t = \overline{X}_t + \varepsilon_t \tag{9}$$

where in the simplest case ε_t can be considered as an i.i.d. sequence of 0-mean random variables. For such models more or less straightforward methods for statistical inference are available. An obvious disadvantage of such an approach is that the cause of the possible rejection is not clear: it might well happen that there is just an acceptable level of discrepancy from symmetry, and rejection is only based on the poor fit of the model. We thus suggest an alternative nonparametric method, focusing only on the symmetry.

Let us first suppose that the reference function constructed by interpolation is simple. If it is only normal, then the procedure of (6–7) can be used to cut the curve into simple parts and then the symmetry of these parts can be tested independently. Let us use the transformation $\sim$ defined in the previous section in (4). This is of course based on the observed averages $\overline{X}$ and $\overline{Y}$, but from now on we omit the overline, when using $\sim$.

Our method relies on the fact that the area for the normalized AI can be approximated by

$$\frac{1}{n}\sum_{k=1}^{n/2-1} |\tilde{Y}(k/n) - \tilde{Y}(1-k/n)| \tag{10}$$

(see (5), we supposed n to be an even number). Equation (10) shows that it is much more useful to base our procedures on the estimated values $\tilde{Y}((k/n))$ rather than the observed $\tilde{Y}_t$, for which the contribution very much depends on the difference $\tilde{X}_{t+1} - \tilde{X}_t$. Let us introduce the notation $D(k/n) := \tilde{Y}((k/n)) - \tilde{Y}((1-k/n))$ and consider the increments of the process $D(k/n)$: $\delta_k := D(k/n) - D((k-1)/n)$ for $k = 1, 2, \ldots, n/2$. It is obvious that $\sum_{k=1}^{n/2} \delta_k = D(1/2) - D(0) = 0$ and that part of Equation (10) where there is no sign change (*i.e.* the absolute value is not needed) can be expressed as $\sum_{k=1}^{n/2}(n/2-k)\delta_k$,

so we get the maximal value if the δ values are ordered.

Now we can describe the procedure we suggest: if we generate a random permutation of the set $\{1, \ldots, n/2\}$ and permute the vector $(\delta_1, \ldots, \delta_{n/2})$ then we get another closed curve. If there was no asymmetry between the variables, then we could suppose that the original permutation was just a typical one, so the area of the original curve would be near to this generated one.

If we repeat the permutation procedure M times and calculate (10) for all, then we get M different possible values of the symmetry index. If the observed AI is larger than the 95 (99 etc)% quantile of this observed distribution, then the symmetry can be rejected.

There is an open question how to choose n in Equation (10). We suggest to use the same number of points as for the original observations. This could be investigated, but for an X with constant derivative we use just the original observations. Another point against a too refined grid is the extensive computing time and that in such cases almost always there is strong dependence among neighbouring values, which indicates that there is little gain in using all of them.

The suggested approach shows similarities to the methods presented in Schmid and Trede (1995), where the classical two sample problem is tested by a method, based on the area between two curves. The asymptotic distribution of their test statistic is the integral of the Brownian bridge (see Shepp, 1982 for its tabulated distribution). We also get this limit distribution for the scaled and randomized sequence if certain mixing conditions can be supposed for the sequence δ, but as in our problems we have a definite, not too large number as n, we do not exploit this idea here.

Applying the above method to the bivariate dataset of sample means illustrated in Figure and 2 we get an estimated p-value (based on 10 000 bootstrap samples) of 0.0268, and a few of the simulated chain plots, corresponding to various critical values, are shown in Figure 5. These plots are constructed with the same end points as for the raw data. This p-value may be surprisingly small, given that the marginal

plots have a very similar structure, and the asymmetry index itself was small. However, this sample size of $n = 1460$ means that the chain plot is very smooth and so there is little variability in the resulting δ_k.

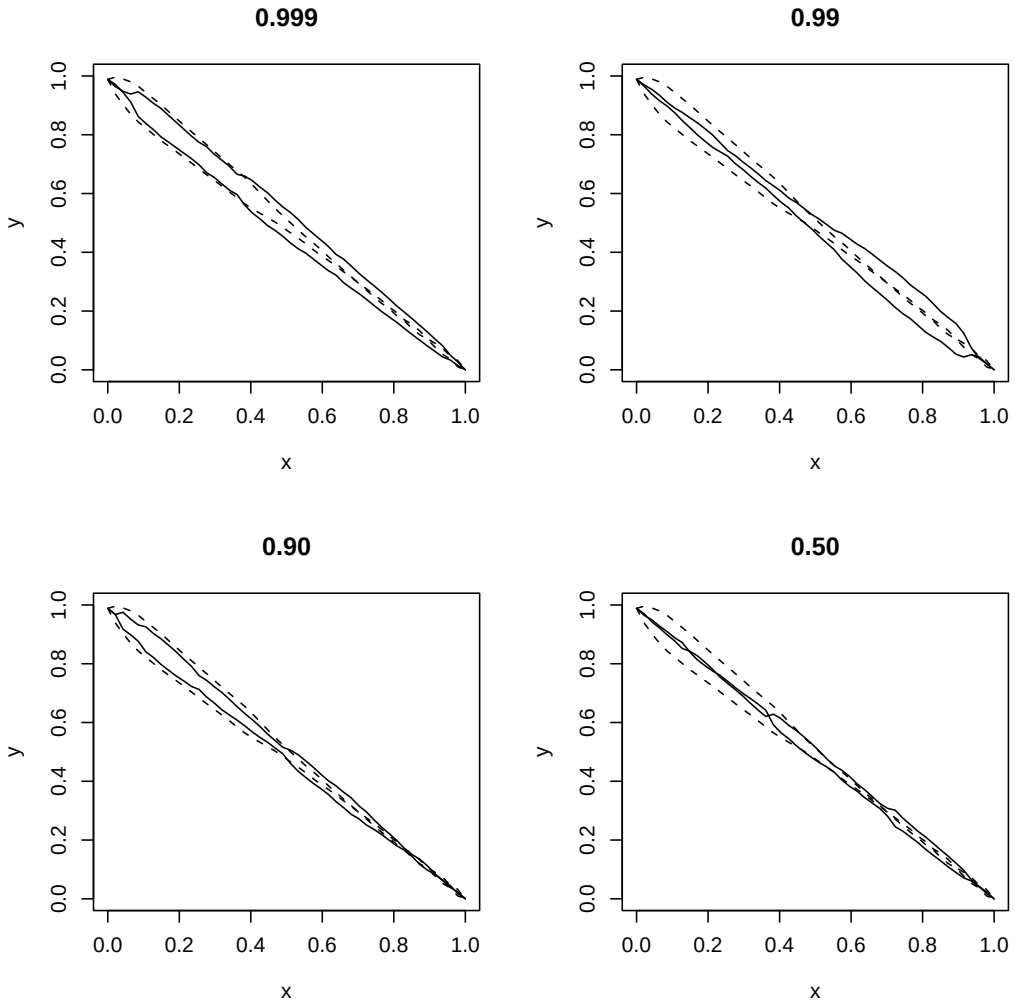

Fig. 5. Simulated chain plots (continuous line) corresponding to the 0.999, 0.99, 0.9 and 0.5 quantiles of the empirical distribution of the resampled asymmetry index, based on the (rescaled) chain plot of the data (dashed line) shown in Figure 2.

4 Extensions

4.1 Correlations

The cross-covariance between $(X_t)_{t=1}^T$ and $(Y_t)_{t=1}^T$ at lag j is estimated by

$$c_{xy}(j) = \frac{1}{T}\sum_{t=1}^{T-j}(X_t - \overline{X})(Y_{t+j} - \overline{Y}) \qquad j = \ldots, -1, 0, 1, \ldots$$

and the cross-correlation at lag j is estimated by

$$r_{xy}(j) = \frac{c_{xy}}{s_x s_y}$$

where $s_x^2 = T^{-1}\sum(X_t - \overline{X})^2$. For the type of data considered here most of the cross-correlation can be attributed to the fact that both X_t and Y_t have very pronounced periodic components. For example, for the data considered in Figures 1 and 2 the cross-correlations are given in Figure 6. In such cases, we ought to consider the correlation *after*

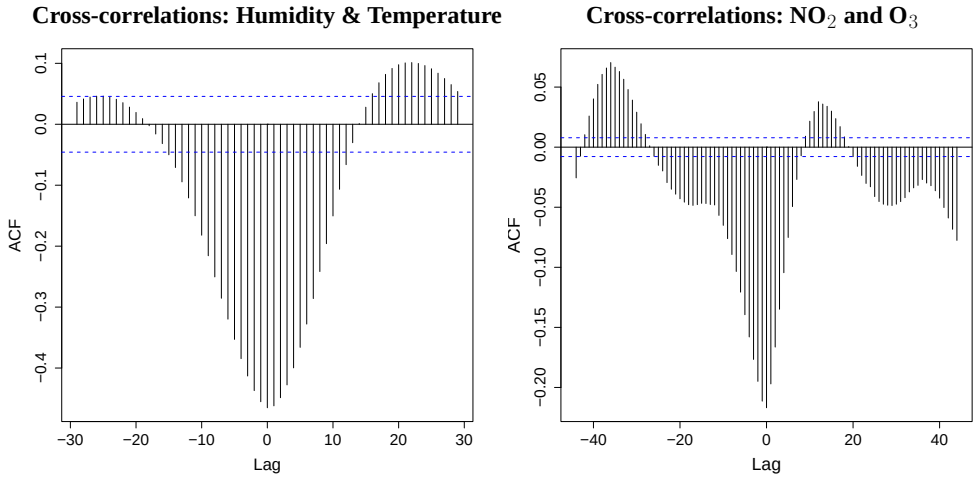

Fig. 6. Cross-correlations of bivariate data

allowing for the periodic effects. This can be done by considering the cross-covariance at lag j for points at seasonal time k, which is given by

$$c_{xy}(k, j) = \frac{N}{T}\sum_{t=1}^{T/N-k}(X_{k+(t-1)N} - \overline{X}_k)(Y_{k+j+(t-1)N} - \overline{Y}_{k+j}) \quad (11)$$

for $k = 1, \ldots, N,\ j = \ldots, -1, 0, 1, \ldots, N - j$ where $\overline{X}_k$ and $\overline{Y}_{k+j}$ are given in Equation (1). The adjusted sample cross-correlation can be estimated in an analogous manner. Incorporating all of this information on the chain plot is difficult, since we now have N cross-correlation plots to display. However, if we initially consider $j = 0$ and use a normal approximation to describe the joint distribution of $\overline{X}_k$ and $\overline{Y}_k$, then we can include $100(1 - \alpha)\%$ confidence limits around each point in the chain plot. An example is shown in Figure 7 which shows how the variability in the two variables changes over time. An efficient way to construct such contours is to use polar co-ordinates as follows.

Let θ be a sequence of length l angles in $[0, 2\pi]$, and form the $l \times 2$ matrix $M = [\cos\theta, \sin\theta]$. Given k, j form the covariance matrix for the relevant data, say

$$\Sigma = \begin{pmatrix} c_{xx}(k,j)\ c_{xy}(k,j) \\ c_{yx}(k,j)\ c_{yy}(k,j) \end{pmatrix}$$

Now solve $v = \Sigma^{-1} M^T$ and then plot polar co-ordinates $r(\theta), \theta$ around $(\overline{X}_k, \overline{Y}_k)$ where $r(\theta)$ is a vector of length l determined by

$$r^2(\theta) = \frac{\chi^2_2(1-\alpha)}{v^T M u}$$

where $u = (1, 1)^T$ and $\chi^2_2(1 - \alpha)$ is the $100(1 - \alpha)\%$ point from a χ^2 distribution with 2 degrees of freedom.

It is evident in Figure 7, except around the middle of the day (times 20–30), that there is a small negative correlation between O_3 and NO_2. Such a plot also indicates times of greater joint change in the daily means — in this example between time period 10 and 11 (5:00 and 5:30 a.m.), and from time periods 16 to 21 (8:00 to 10:30 a.m.). See Makra *et al.* (2001) for a more detailed analysis of the data.

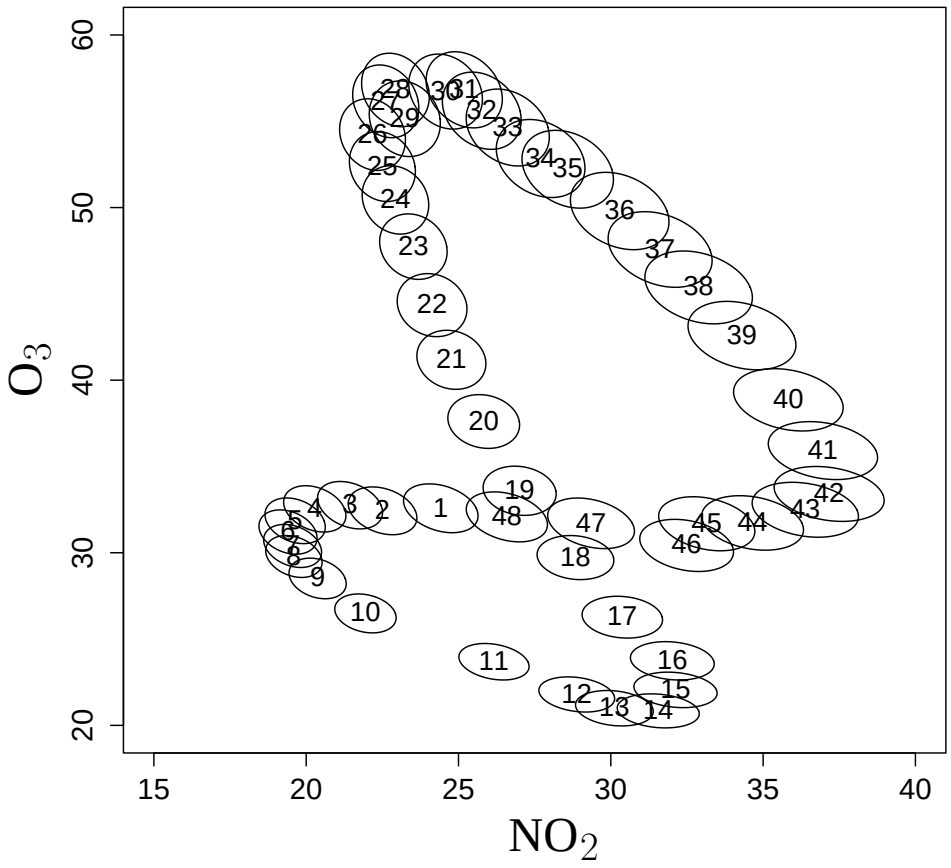

Fig. 7. Modified chain plot with 90% confidence intervals (using a multivariate normal approximation) based on the conditional covariance at each time point

4.2 Fourier Analysis

A straightforward Fourier analysis was made harder by the large number of missing values. So, in order to compare with the above results we fitted a linear model corresponding to only daily and 12 hourly periods (Bloomfield, 2000). That is, for each time series, we fitted the model:

$$X_t = \alpha + R_1 \cos(\omega_1 t + \phi_1) + R_2 \cos(\omega_2 t + \phi_2) + \varepsilon_t$$

where $R_i, \phi_i, i = 1, 2$ are the fitted amplitudes and phases, respectively, corresponding to the frequencies ω_i, with ω_i corresponding to 24-hourly and 12-hourly periods for $i = 1, 2$, respectively. The results for the four time series used as examples in this paper are shown

in Table 1.

	Daily cycle		12-hour cycle	
Time series	Amplitude	Phase	Amplitude	Phase
NO_2	4.0	1.37	7.06	2.18
O_3	16.1	2.31	7.61	-0.94
Humidity	9.1	-0.94	2.12	-1.04
Temperature	3.5	2.24	0.76	-0.66

Table 1
Amplitude and phases corresponding to two fitted periods for each of four time series

For the climate time series (humidity and temperature) the dominating period is the daily cycle, and the two phases are very similar. This implies that the asymmetry index A is small (see Figure 2) and visually we would expect the chain plot to be long and thin. For the two pollutants, the phases for the daily cycle are very different, and the 12-hourly cycle is relatively more important for NO_2 than for O_3. These values are consistent with the chain plot analysis (due to the phase differences, the lagging was necessary for achieving some symmetry), but we think not so immediately informative. In particular, it is not obvious how the two series can be partially aligned using the given phases for the two cycles.

5 Discussion

We end our paper with some comments on the use of the suggested graphical tool. As here we have practically three dimensions to investigate (time+two variables), the traditional sequence of bivariate scatterplots (2 time series plots plus a scatterplot in our case) do not give a clear picture of the evolution of the relation over time.

We estimated the proposed measure of asymmetry for the data under investigation. An objective comparison of asymmetry in different data sets was possible by the proposed nonparametric method.

Acknowledgements: We acknowledge support from the British Council and the Hungarian Ministry of Education under the British-Hungarian Academic Research Programme.

References

Bloomfield, P. 2000. *Fourier Analysis of Time Series: An Introduction (second edition)*. John Wiley, New York.

Embrechts, P., McNeil, A. and Straumann, D. 1999. Correlation and dependency in risk management: properties and pitfalls. Preprint ETH, Zürich. www.math.ethz.ch/∼embrechts.

Makra, L., Horváth, Sz., Taylor, C.C., Zempléni, A., Motika, G. and Sümeghy, Z. 2001. Modelling air pollution in countryside and urban environment, Hungary. *Proc. of the 2nd International Symposium on Air Quality Management at Urban, Regional and Global Scales.* Istanbul, Turkey. pp. 189-196. Eds.: Topcu, S. Yardim, M.F. and Incecik, S.

Schmid, F. and Trede, M. 1995. A distribution free teset for the two sample problem for general alternatives. *Comput. Stat., Data Anal.* **20**, p. 409-419.

Shepp, L.A. 1982. On the integral of the absolute value of the pinned Wiener process - calculation of its probability density by numerical integration. *Ann. Probab.* **10**, p. 240-243.

Villa-Diharce, E. 2001. Estimation of the tail dependence index. Research Report of CIMAT, Guanajuato, Mexico.