PSYCHOMETRIKA—VOL. 75, NO. 2, 351–372
JUNE 2010
DOI: 10.1007/S11336-009-9146-8

EXPLORING THE DYNAMICS OF DYADIC INTERACTIONS VIA HIERARCHICAL SEGMENTATION

FUSHING HSIEH, EMILIO FERRER, AND SHU-CHUN CHEN

UNIVERSITY OF CALIFORNIA, DAVIS

SY-MIIN CHOW

UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL

In this article we present an exploratory tool for extracting systematic patterns from multivariate data. The technique, hierarchical segmentation (HS), can be used to group multivariate time series into segments with similar discrete-state recurrence patterns and it is not restricted by the stationarity assumption. We use a simulation study to describe the steps and properties of HS. We then use empirical data on daily affect from one couple to illustrate the use of HS for describing the affective dynamics of the dyad. First, we partition the data into three periods that represent different affective states and show different dynamics between both individuals' affect. We then examine the synchrony between both individuals' affective states and identify different patterns of coherence across the periods. Finally, we discuss the possibilities of using results from HS to construct confirmatory dynamic models with multiple change points or regime-specific dynamics.

Key words: dynamic systems, multivariate analysis, exploratory data analysis, dyadic interactions.

The study of dyadic interactions concerns the understanding of the interrelations between two people. In these interactions (e.g., teacher-student, mother-child, husband-wife), the two elements of the dyad form an intertwined system in which the expression (i.e., behavior, emotion, cognition) of one member is dependent upon the expression of itself and the other member's. If dyadic interactions are to be fully understood, they need to be studied as dynamic systems and the principles underlying such interactions must be expressed in dynamic terms (West, 1985). An area in which the influence of dynamic systems-based concepts has been particularly prominent is the study of marriage and personal relationships (e.g., Boker & Laurenceau, 2006; Gottman, Murray, Swanson, Tyson, & Swanson 2002; Levenson & Gottman, 1983). In all instances, capturing dyadic interactions requires not only an intense set of measurements, but also techniques that can identify the underlying dynamics.

It is well known that the interpersonal dynamics within a dyad can change in critical ways even during a brief episode of interaction (e.g., Boker & Rotondo, 2003; Newtson, 1993). Ideally, methodologies for studying dyadic interactions should be capable of revealing existing patterns of non-stationarity by identifying possible segments from the observed data that correspond to different states of the psychological process. In this article, we present a nonparametric approach, termed hierarchical segmentation (HS), for organizing the dynamic patterns (including nonlinear and nonstationary) of dyadic interactions into a series of discrete "states" and segments of observations with similar recurrence patterns.

R code for the HS algorithm is available upon request.

Requests for reprints should be sent to Emilio Ferrer, Department of Psychology, University of California, One Shields Ave., Davis, CA 95616-8686, USA. E-mail: eferrer@ucdavis.edu

© 2010 The Psychometric Society. This article is published with open access at Springerlink.com

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

Techniques for Studying Dyadic Interactions

A variety of techniques have been employed to study dyadic interactions. One initial and influential approach was the sequential methods developed in the 1970s (Bobitt, Gourevitch, Miller, & Jensen 1969; Castellan, 1979; Goodman, 1970; Gottman, 1979; Sackett, 1979). Some newer techniques have now been developed to deal with discrete dyadic sequential data, such as multinomial logit models (de Rooij & Kroonenberg, 2003). Hazard regression (i.e., survival analysis, event history analysis) has also been used to study dyadic interactions at both the individual dyad level (Gardner & Griffin, 1989) and for groups of dyads (Gardner, 1993). Similarly, various forms of hierarchical linear models have been applied to dyadic interaction data of both cross-sectional (Campbell & Kashy, 2002; Kashy & Kenny, 2000) and longitudinal type (Newsom, 2002; Raudenbush, Brennan, & Barnett 1995). More recently, other techniques have been used that include time-lagged relations in the form of difference and differential equations (e.g., Boker & Laurenceau, 2006; Felmlee & Greenberg, 1999; Ferrer & Nesselroade, 2003; Granic & Hollenstein, 2003; Gottman et al., 2002).

Although highly informative of various aspects of the data, most of these techniques rely on important assumptions to achieve some form of stationarity. Stationarity is a broad and important concept in time series models. Broadly speaking, stationarity means that the statistical properties of a time series do not change over time. Weak stationarity refers to conditions when the mean and variance of a time series remain invariant across different time points, and the lagged covariance between two scores depends only on their distance in time. Strict stationarity, a stronger form, requires invariance of all distributional characteristics of the time series over time. Generally, weak stationarity is sufficient for most statistical procedures of time series (for more detailed definitions, see Shumway & Stoffer, 2006).

Mathematically, a time series $X_1, X_2, \ldots, X_N$ is nonstationary if, for some m, the joint probability distribution of the m-vector $X_{i+1}, X_{i+2}, \ldots, X_{i+m}$ is dependent on the time index i (Priestley, 1988). From a practical standpoint, an observed time series $x_1, x_2, \ldots, x_N$ of length N could be deemed nonstationary if, for some m ($\ll N$), there are significant variations in the estimated joint distribution of $x_{i+1}, x_{i+2}, \ldots, x_{i+m}$ (e.g., Casdagli, 1997).[1]

Under stationarity assumptions, any deviation in statistical and distributional properties over time is undesirable and considered as noise (Manuca & Savit, 1996). Other researchers have argued that nonstationarity is inherent in most processes and should indeed be the focus of the analysis. To this end, exploratory techniques have been proposed for identifying the presence of nonstationarities in, e.g., the cross-correlation (Boker, Xu, Rotondo, & King, 2002) and spectral properties of empirical data (Weber, Molenaar, & Van der Molen, 1992). In contrast, confirmatory approaches serve to capture the ways in which, and the point(s) where, a system deviates from a preconceived model of change. Developments along this line include regime switching models (Kim & Nelson, 1999; Dolan, Schmittmann, Lubke, & Neale, 2005), change point analysis (Hinkley, 1970, 1971), Bayesian change point detection models (Carlin, Gelfand, & Smith, 1992; Stephens, 1994), and state-space models with parameter change (e.g., Chow, Ferrer, & Nesselroade, 2007; Kitagawa, 1981; Molenaar & Newell, 2003).

In cases where a time series undergoes several discrete but unobserved phases—or so-called regimes—during which the time series is characterized by distinct change patterns, regime switching models[2] or change point analysis are two common ways of modeling the shifts in dynamics across regimes and change points (i.e., the points at which the deviations in dynamics take

[1] There might be some instances where changes in the joint distribution can be observed but the implied process may still be stationary. Examples of this could include regime-switching models (see Hamilton, 1988).

[2] Some interesting properties of regime switching models are worth noting here. Whereas these models allow researchers to model changes in a system's mean and variance over different regimes, processes that conform to this kind of changes are not necessarily nonstationarity. As long as the switches between regimes do not depend on time itself, then at any given occasion t, some of the realizations of the ensemble will be in one regime, while some realizations will

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

place). In many empirical applications, however, the nature of the distinctions across regimes or change points may not be clearly understood and exploratory tools can provide valuable insights into the underlying dynamics of a system, thereby facilitating the development of confirmatory models of change in the long-run.

The Current Study

The HS approach proposed in the present article is exploratory, does not depend on stationarity conditions, and can be contrasted with some of these other existing techniques. It has some functional similarities to cross recurrence plots (see e.g., Eckmann, Kamphorst, & Ruell, 1987; Marwan & Kurths, 2002), which can be used to extract patterns of recurrence in bivariate time series without imposing any assumption of stationarity and linearity. Our approach differs, however, in that the HS is used to discern the recurrence of discrete, as opposed to continuous states. Finally, one unique strength of the proposed approach is its ability to extract exploratory information concerning the number of regimes or distinct phases in the data, as well as the dynamics within regimes. This information can then be used to generate confirmatory change-point or regime switching models, from a data driven perspective.

The remainder of the article is organized as following. First, we describe the steps and functions of the HS algorithm. Here, we define each step and its associated goal. Second, we present results from auto- and cross-correlation analyses using illustrative empirical data on affect. These techniques assume stationarity in the data, and thus provide a basis against which results from HS can be compared. Third, we apply HS to the empirical data and describe the results at each step. Here, we also elaborate on how and in what ways HS adds to results from the earlier auto- and cross-correlation analysis. Fourth, we outline a chi-square test for evaluating concurrent and lagged coherence between dyad members using information from HS. In all steps, we use the empirical data set to illustrate the application of the method.

To illustrate our approach, we use a data set throughout all the different steps in our analysis. These data consist of daily reports on affect from a husband–wife dyad. Details of these data are reported elsewhere (Ferrer & Nesselroade, 2003) and only crucial information is presented here. The variables from the data represent self-reports of emotional experiences using the Positive and Negative Affective Schedule (PANAS; Watson, Lee, & Tellegen, 1988) reported by both individuals daily for 182 days. The PANAS contains 20 items describing positive affect (PA) and negative affect (NA). Participants are asked to mark the extent to which they experienced each of the items on a 5-point Likert-type scale ranging from 1 (very slightly or not at all) to 5 (extremely). For reasons of parsimony, and in line with standard use of the PANAS, we created unit-weighted composites of positive and negative affect for both individuals. Figure 1 presents plots of the positive and negative affect composites for both individuals. These plots reveal differences in positive affect between both people and an apparent coupled fluctuation in the negative affect. Both plots indicate the lack of a manifest trend in the data. All the analyses described in this article were conducted using these four dimensions.

1. Hierarchical Segmentation

HS is a technique originally developed for representing nonparametric dynamics in animal behavior (Hsieh, Hwang, Lee, Lan, & Horng, 2006) and has also been applied to affect data (Ferrer, Chen, Chow, & Hsieh, 2010). HS does not require prior assumptions concerning the nature of

be in another regime. Because this will be the same for any other occasion, the average across the realizations will be the same at each occasion, such that the process is stationary. Thus, a Markov-switching autoregressive model, for example, can be described as stationary if this particular property is fulfilled (p. 389; Hamilton, 1988).

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

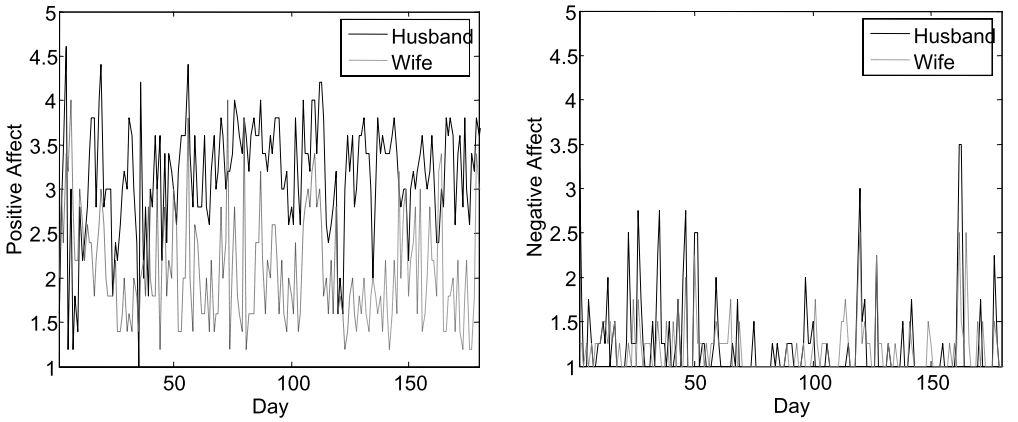


FIGURE 1.
Time series of daily ratings of positive and negative affect from each dyad member.

Hierarchy	*Coding*	*Interpretation*	*Sequence*
Level 0		Raw data	
Level 1	C_1	Extreme expression of the dimensions (0_1)	
Level 2	C_2	Long consecutive segments of low expressions (0_2)	Sub-routine (*SR*)
Level 3	C_3	High frequent segments of extreme expression (0_3)	Main program (*MP*)

FIGURE 2.
Conceptual representation of hierarchical segmentation and sequence coding.

the change process of interest, and does not rely on many of the common statistical assumptions, one important example of which is stationarity. The goal of HS is to identify segments that could be governed by different phases of that process.

HS starts with multivariate time series data of q dimensions (or variables). To each dimension, a coding schema is then applied. In our example, $q = 4$, representing PA and NA from each person in the dyad. HS involves recoding each person's time series for each variable into three layers—or levels of hierarchy—each serving a specific purpose for extracting information about dynamics. These layers are denoted as C_1, C_2, and C_3. Each level represents a summary statistic of data reduction. Figure 2 provides a conceptual description of the different levels of HS. We explain this process next.

1.1. C_1 and 0_1: Identifying Segments with Flags for Extreme Expressions

The code sequence C_1 is the first level of the HS hierarchy. At this level, each of the dimensions of interest is separately transformed into a time series of p categories. We use percentiles

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

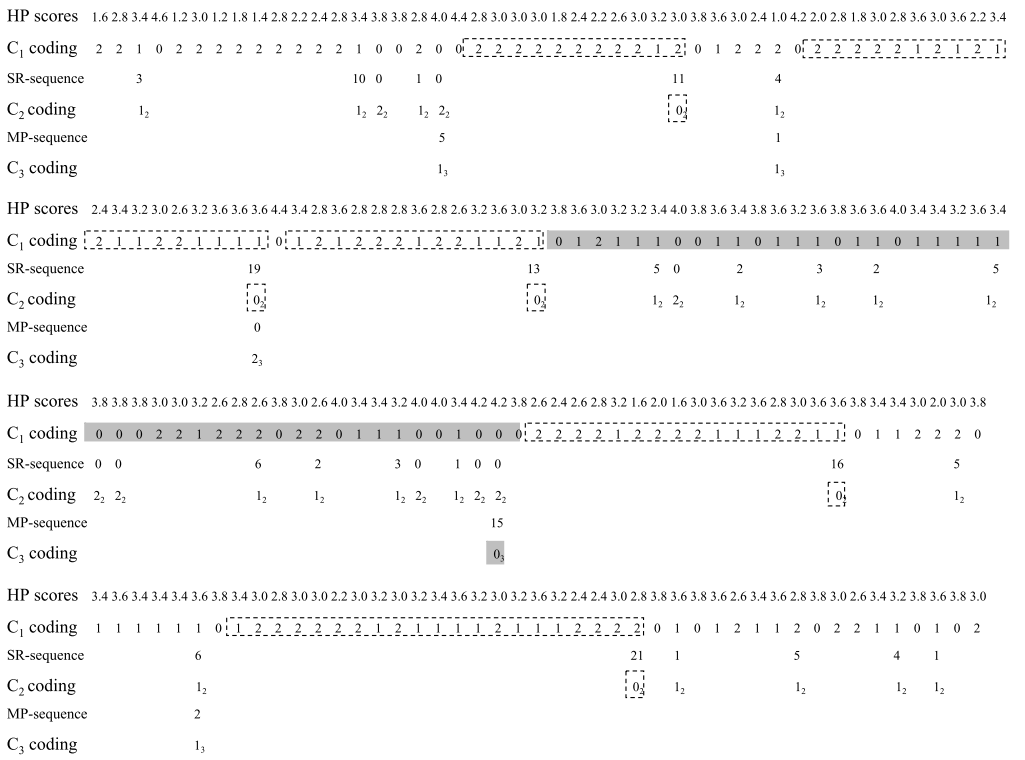


FIGURE 3.
Full sequential coding for a selected data segment.

(computed separately for each individual and each of the q variables) to recode each individual's raw scores into digital strings of base p with values $0, 1, \ldots, p-1$. In our application, for example, we use two cutoff points to form three (ordinal) categories. Thus, $p = 3$ and the resultant recoded series is of base 3. We note that the selection of the cutoff points is arbitrary. In our application, we rely on exploratory information deduced from visually inspecting the frequency distribution of each time series in a histogram. Other criteria (e.g., an externally defined clinical cutoff score) can also be used, as appropriate, to select the cutoff points. After recoding the raw data into values of 0, 1, and 2, the entire sequence of 0, 1, and 2 codes is denoted as the C_1 code sequence. In this particular layer of the HS, the value 0 serves as a flag that marks the presence of extreme high (positive as well as negative) affects.[3]

Figure 3 describes the sequence coding after applying HS to a selected segment of the husband's PA data (HP in Figure 3). Although this figure includes the coding resulting from the full HS algorithm (i.e., from the raw data to the top level of the sequence hierarchy), for now we focus on the first and second lines. The first line shows the husband's PA raw scores. The second line represents the C_1 code sequence, resulting from the first transformation of the raw data. The codes 0, 1, and 2, indicate high, medium, and low expression of positive affect, respectively.

From the standpoint of information theory (Chaitin, 1987), one aspect to consider about this code sequence C_1 is whether or not it is generated via simple random sampling. However,

[3]Whereas the use of 0 as a marker of extreme high affect may seem counterintuitive in our application, the event of interest in other applications may be a low extreme expression or, possibly, an event that is nondirectional in nature (e.g., in the identification of unusual genes in a Deoxyribonucleic Acid (DNA) sequence; Ewens & Grant, 2005). We thus use 0 as a general marker in our formulation.

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

one of the assumptions of simple random sampling is that each measurement occasion is independent of all other measurement occasions. This assumption is hardly tenable in psychological data. In our empirical example, we use other well-known exploratory indices such as auto- and cross-correlations as an initial tool for assessing the presence of significant stationary time-based dynamics in the raw data. We then evaluate the assumption of stationarity implicated in such analyses by inspecting the codes in other layers of the HS algorithm.

1.2. C_2 and 0_2: Identifying Segments with Flags for Persistent Lack of Extreme Expressions

For the next level of the HS process, we use the inter-0-code-length, denoted below as l, to form a sequence, called the "subroutine" (*SR*) sequence. This sequence is represented in Figure 3 as the third line of each set of rows. A large value of l represents a long period of nonextreme expression ($l = 21$ is the largest value for the data depicted in Figure 3). The corresponding frequency distribution of l can then be used to choose two cutoff points and distinguish the long, medium, and short periods of nonextreme expressions. We represent these periods by 0_2, 1_2, and 2_2, respectively. In this way, the *SR* sequence is transformed into a code sequence C_2 in which the code 0_2 marks a period of persistent nonextreme affect. The cutoff value from the *SR*-sequence to yield 0_2 of the C_2 code was 11. Values of zero were recoded as 2_2 and all other values were recoded as 1_2. This new code C_2 is described in Figure 3 as the fourth line of each set of rows.

1.3. C_3 and 0_3: Identifying Segments with Flags for Frequent Recurrence of High Expressions

The code sequences C_1 and C_2 form the first two levels of the HS hierarchy. Similar to the steps taken in constructing C_2, we now use the inter-0_2-code-length to form a so-called "main program" (*MP*) sequence (fifth line of each set of rows in Figure 3). This top level is called the main program because it can be viewed as a computer program that contains information about other lower levels and such information can be used to simulate the code sequences C_2 and C_1, as well as regenerating the observed raw data. This is a bottom-up algorithmic composition (c.f., Geman, Potter, & Chi, 2002) that can be seen as a lossy data compression method (i.e., when decoding after the compression, some of the original data are lost; Hsieh et al., 2006). Here, a large value of inter-0_3-code-length indicates a segment with high frequency of code 0 in C_1, or extreme expression of the variable. To illustrate, the *MP*-sequence in Figure 3 is: 5, 1, 0, 15, 2. The first digit (5) was derived from the beginning of the segment C_2 (i.e., $1_2, 1_2, 2_2, 1_2, 2_2$). The second digit (1) was derived from the 1_2 located between the first two 0_2 codes. The third digit (0) was derived from the second and third 0_2 codes, and so on.

Finally, the collection of codes in the (*MP*) sequence is also used to construct a histogram on the top level of the hierarchy. From this frequency distribution, again, we choose two cutoff points and classify the (m) values into three cells representing large, medium, and short spacing, represented by 0_3, 1_3, and 2_3, respectively. The cutoff value on the *MP*-sequence to generate digital sequences of 0_3 was 15. Values of zero in the *MP*-sequence were recoded into 2_3 and all other values were recoded as 1_3. Thus, the *MP*-sequence is transformed into a code sequence C_3 on the top level of the hierarchy. The flags 0_3 identify segments of frequent recurrence of zero in C_1. This code is represented in the last line of each row of Figure 3.

1.4. Summary of Hierarchical Segmentation

Together, the three HS levels serve as a nonparametric tool for grouping together data segments with similar dynamics and, thus, mapping out the states underlying a given process. The first layer of coding, C_1, contains recoded values of the original time series, where the flag 0 is used to mark instances of unusual events (e.g., extremely high PA or NA). The second layer, C_2, contains the time intervals between successive 0 flags. The flag 0_2 is used to indicate unusually

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

long periods that lack any extreme event. The third layer, C_3, carries the time intervals between successive 0_2 flags. The flag 0_3 is used to mark periods that are characterized by frequent recurrence of extreme expression. The goal of these seemingly opposite flags is to cover the full range of affect, from extreme expressions to periods of complete absence. Whereas the C_1 sequence is simply a recoding of the raw data, the C_2 sequence may be regarded as a subroutine, *SR*, of the main program, *MP* (i.e., C_3). That is, the *SR* serves to execute a subsidiary function of the main program. Information extracted using the HS can be used alone as a data exploratory tool for pattern recognition, or the corresponding codes can be combined with other conventional analyses to answer specific empirical questions. In the next section, we describe how the HS codes can be used to conduct chi-square tests aimed at addressing questions pertaining to concurrent and lagged synchrony between dyadic members.

2. Coherence Measurement

In dyadic interactions we are interested in the relationship between two people (or any other two units in a dyad). One crucial question here concerns the identification of patterns manifested by one person when the other person shows a particular psychological state. In our data, some pertinent questions may include: what are the affective patterns in one person when the other is in, say, a long period of consistently low PA? Is there any synchrony in the affective structure of both people for this particular period? For this, we first apply the HS process to the two variables, for each person. We then align the codes (0_3 and 0_2) of both persons along a common time metric (e.g., codes for the two individuals within the same day) and evaluate the coherence between the resulting code sequences. Coherence here is defined as synchrony between the two individuals, as components of the same system. We quantify this synchrony via chi-square tests regarding both the entire process and specific segments identified by HS.

When two C_1 code sequences from both members of the dyad are aligned along the same time metric (e.g., days in our data), a day of coherence is defined as having the same digital codes on both sequences. That is, the count of an (i, j) code-pair, with $i = 0, 1, 2$ for one member and $j = 0, 1, 2$ for the other member, will form a 3×3 contingency table. Dependency between the two variables taking values 0, 1, or 2 can be examined using the chi-square test. Under the null hypothesis of no association among the cells, if the chi-square value is significant, then the positive determinant of the 3×3 matrix would indicate coherence, otherwise antagonism. For simplicity, this is termed concurrent coherence. For the two dyad members, concurrent coherence in a particular variable can be quantified via the chi-square statistic as

$$\chi^2 = \sum_{i=0}^{2}\sum_{j=0}^{2} \frac{\left[O_{ij} - N \cdot \frac{n_i}{N} \cdot \frac{m_j}{N}\right]^2}{N \cdot \frac{n_i}{N} \cdot \frac{m_j}{N}}, \tag{1}$$

where O_{ij} is the observed (i, j)th cell count, and n_i, $i = 0, 1, 2$ and m_j, $j = 0, 1, 2$, denote the marginal row and column sums, respectively; and N is the total number ($\sum_{i=0}^{2} n_i = \sum_{j=0}^{2} m_j = N$) in the contingency table. This statistic follows a chi-square distributed with degrees of freedom $(9 - 3 - 3 + 1) = 4$, under the null hypothesis of independence of the row and column variables.

By lagging the code sequence of one person, we can also evaluate lagged coherence. For example, in a lag-1 analysis, one member's C_1 code sequence is aligned with the other member's lagged code sequence offset by one day. Any lag analysis is performed in a similar fashion. Specifically, the chi-square statistic associated with a k-lag can be obtained as

$$\chi^{2(k)} = \sum_{i=0}^{2}\sum_{j=0}^{2} \frac{\left[O_{ij}^{(k)} - N^{(k)} \cdot \frac{n_i^{(k)}}{N^{(k)}} \cdot \frac{m_j^{(k)}}{N^{(k)}}\right]^2}{N^{(k)} \cdot \frac{n_i^{(k)}}{N^{(k)}} \cdot \frac{m_j^{(k)}}{N^{(k)}}} \tag{2}$$

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

with the total count being equal to $N^{(k)} = N - k$. These same formulas can be used to assess coherence within different segments in the data (i.e., specific periods of interest). In that case, the counts would be restricted to the particular segment under study. Whereas concurrent coherence is a symmetric measurement, lagged coherence is asymmetric and can be examined in both directions. For instance, the chi-square statistic indicating the association between the husband's PA code at time t and the wife's PA code at time $t - 1$ is generally not the same as the association between the wife's PA code at time t and the husband's PA code at time $t - 1$.

We consider three different types of local coherence (i.e., within a certain phase). The first type is the associations between low, medium, and high expression of affects (i.e., codes 0, 1, and 2 from C_1) between the two people when one person is in a period of frequent extreme responses (e.g., frequent outbursts of high NA). The chi-square test of the corresponding 3×3 contingency table applied to this data segment helps to reveal concurrent coherence and lagged influences between both individuals during such periods of extreme affect.

The second type of local coherence is explored on data segments marked by the code 0_2. The chi-square of the corresponding 3×3 contingency table is informative of synchrony in the dyad when one person is in a period of persistent nonextreme affect (e.g., a long period of low or medium NA). Local coherence of the third kind is applied specifically to the non-0_3 portions of the data. Such non-0_3 portions include all the days on which a dyad member does *not* show frequent recurrence of extreme affects. The aim in this case is to discover possible patterns of association when one person shows only occasional spikes (if at all) of extreme affects, but otherwise shows either low or medium affect. Note that, whereas a 0_2 segment is also a non-0_3 segment, a dyad member does not have to be in a state of prolonged consecutive segments of low affect to be considered as being in a non-0_3 state. That is, an individual could simply be in a period with medium affect but without showing frequent segments of extreme affect.

In addition to the described inter-individual coherence, intraindividual coherence can also be examined in a similar way. For example, the C_1 code sequences of the two affects from one individual (i.e., PA and NA) can be aligned on a common time metric and represented in a 3×3 contingency table. Such alignment would describe daily correspondence between the two processes, and we term this the 1st-order (i.e., level-1 of HS) *intra*-individual coherence. In the case of two variables that complement each other, the chi-square statistic test of independence between them would be rejected. In the case of sequences of a given variable from two different individuals, the resulting matching would yield a 1st-order interindividual coherence. This would represent the concurrent associations between the two individuals on the given variable. By lagging the sequences, lag-1, lag-2, ..., lag-k, 1st-order inter-individual coherences can be examined.

3. Simulations

To illustrate the potential of HS in extracting systematic patterns from nonstationary data, we conducted a series of simulations. For this, we generated data under a number of conditions, defined a particular event of interest, and then used HS to extract patterns of the event. In the first condition, we simulated a time series with $t = 10{,}000$ data points using a sine curve model with initial periods of similar wavelength and periods of larger wavelength at the end of the series (see Figure 4a). We defined event 1 as the upper 10% of the time series. This event would be expressed by the peaks in the sine curve series. Such expressions correspond to the code 0 in the C_1 sequence and are summarized in the histogram of the series as the frequencies on the far right side (Figure 4b and thick line in panel a). Similarly, the low points of the series would correspond to the code 2 (frequencies on the far left side of panel b) and all the points in between would constitute the code 1 (central frequencies in panel b).

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

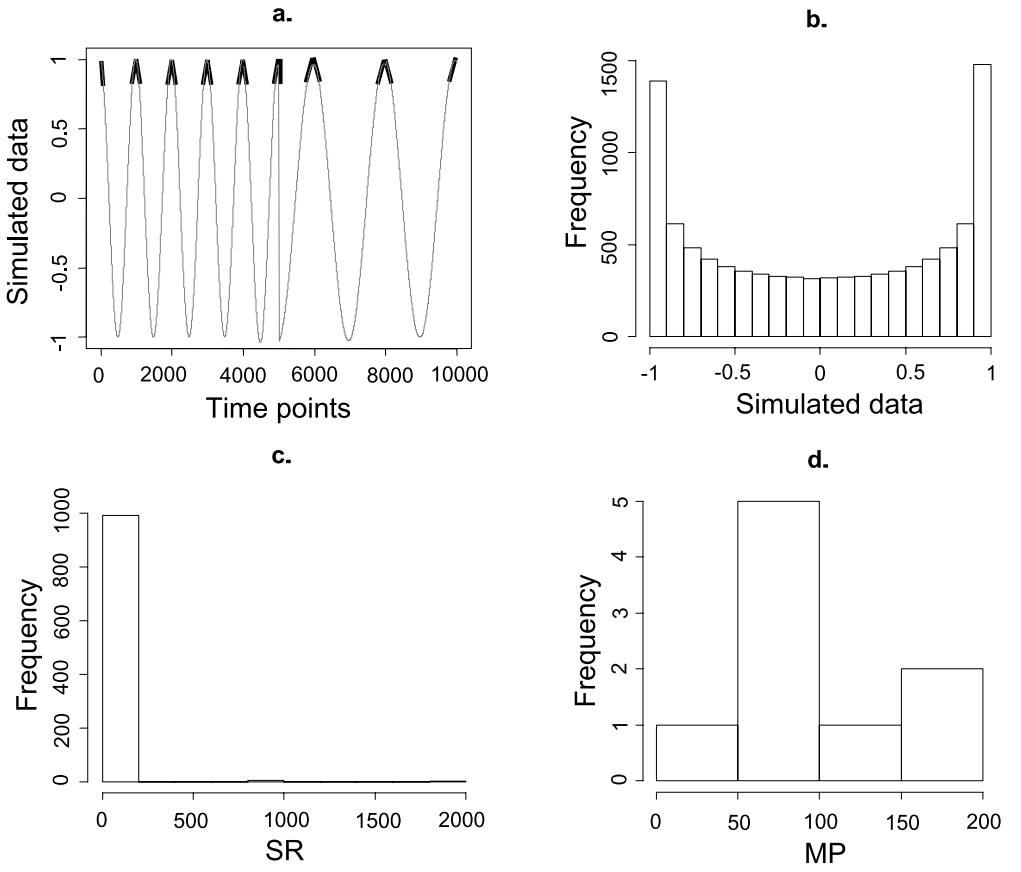


FIGURE 4.
Simulated time series from a sine curve model ($\text{Event}_1(X) = 10\%X(t)$). Panel (**a**) = simulated data. Panel (**b**) = histogram of raw scores. Panel (**c**) = histogram of HS subroutine. Panel (**d**) = histogram of HS main program.

Figure 4c presents the histogram of the second level of the HS (i.e., subroutine, *SR*). This figure illustrates the recurrence time for 0 in the C_1 code; that is, the time intervals between two data points with 0. The left bar of the histogram includes all those distances between adjacent data points with 0. These are the short recurrence times associated with neighboring points located near the peaks of the curves. The middle and right bars in panel c represent longer recurrence intervals. The middle bar includes six instances corresponding to the six periods between the first six peaks in the series, all of which have the same wavelength. Finally, the bar in the far right of panel c includes the two distances between nonadjacent data points with 0 in C_1 code. These two instances correspond to the time interval between the last three peaks in the series, which have a longer wavelength than the previous peaks.

Figure 4d displays the histogram of the third segmentation level (i.e., main program, *MP*). The bars in this histogram represent the time intervals between two successive 0_2 flags, which identify portions of the data that lack extreme expressions (i.e., high scores). To ease presentation, we refer to a segment that is characterized by prolonged lack of extreme expression as a *tranquil* phase. A low value on the abscissa indicates that it does not take long for the system to enter a tranquil phase. In contrast, a large value on the abscissa indicates that it takes the system a long time to enter or return to a tranquil phase. The first bar includes one instance, which corresponds to the time series' transition to a tranquil phase shortly after the very first peak of the time

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

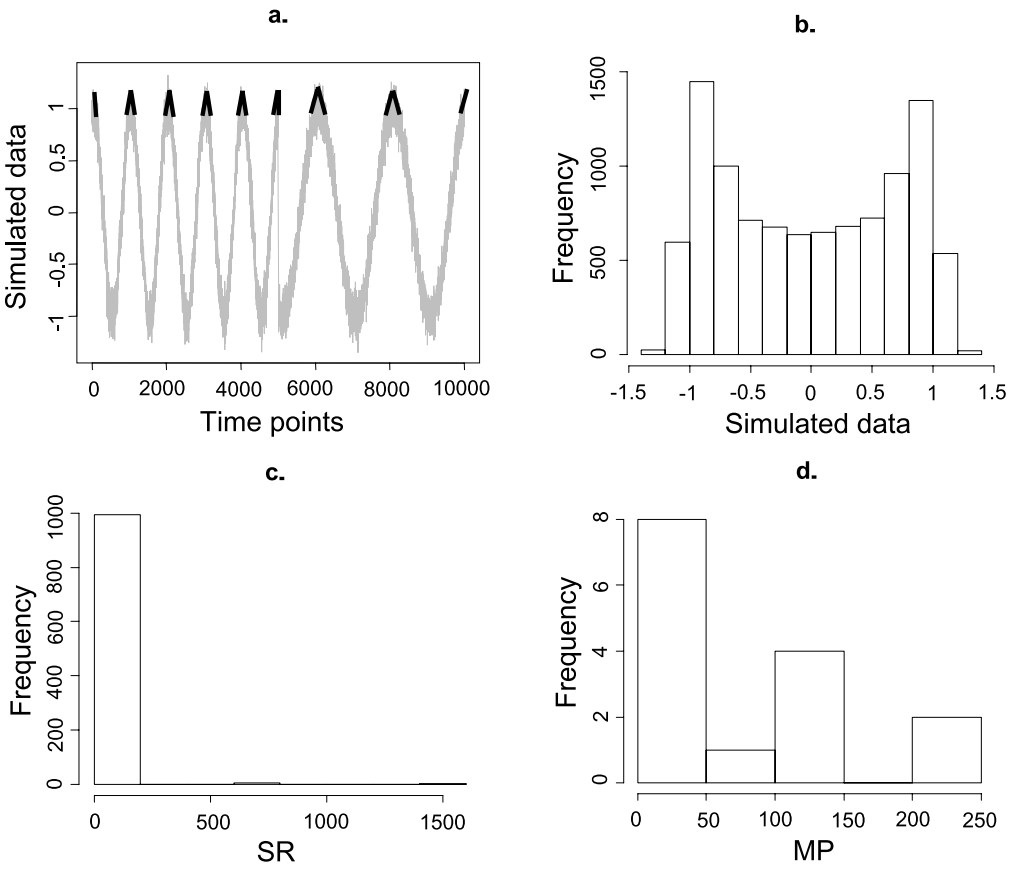


FIGURE 5.
Simulated time series from a sine curve model + noise ($\text{Event}_1(X) = 10\%X(t)$). Panel (**a**) = simulated data. Panel (**b**) = histogram of raw scores. Panel (**c**) = histogram of HS subroutine. Panel (**d**) = histogram of HS main program.

series (in panel a). The second bar includes five counts, which correspond to the time intervals between the tranquil phases locating near the five subsequent peaks with similar wavelength (see panel a). The third bar also contains one count and this corresponds to the time interval between the post-peak tranquil phase of the last high-frequency cycle, and the tranquil phase after the first peak of the low-frequency cycle (i.e., last peak in panel a). Finally, the fourth bar reflects the intervals corresponding to the tranquil phases near the last three peaks with slower wavelength (i.e., peaks 7 and 8 in panel a).

For the second condition, we simulated a time series of the same length $t = 10{,}000$ using a sine curve model similar to condition 1 but with added noise (Figure 5a). Similarly, we defined event 0 as the upper 10% of the time series. The resulting time series is similar to the one presented in Figure 4a but with added noise. The histogram of the series (panel b) is also similar to the one presented in Figure 4b but with intervals exceeding -1 and $+1$ due to the random noise. As before, the bar on the right represents the defined events, the upper 10% of the series, corresponding to 0 in the C_1 code sequence. The sequence of data points with such coding is also highlighted by the dark black lines marking the peaks in the series (panel a). Panels c and d display the histogram of recurrence time for 0 in the C_1 code and the periods containing high frequency of the defined event, respectively. The frequency distributions in these histograms are

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

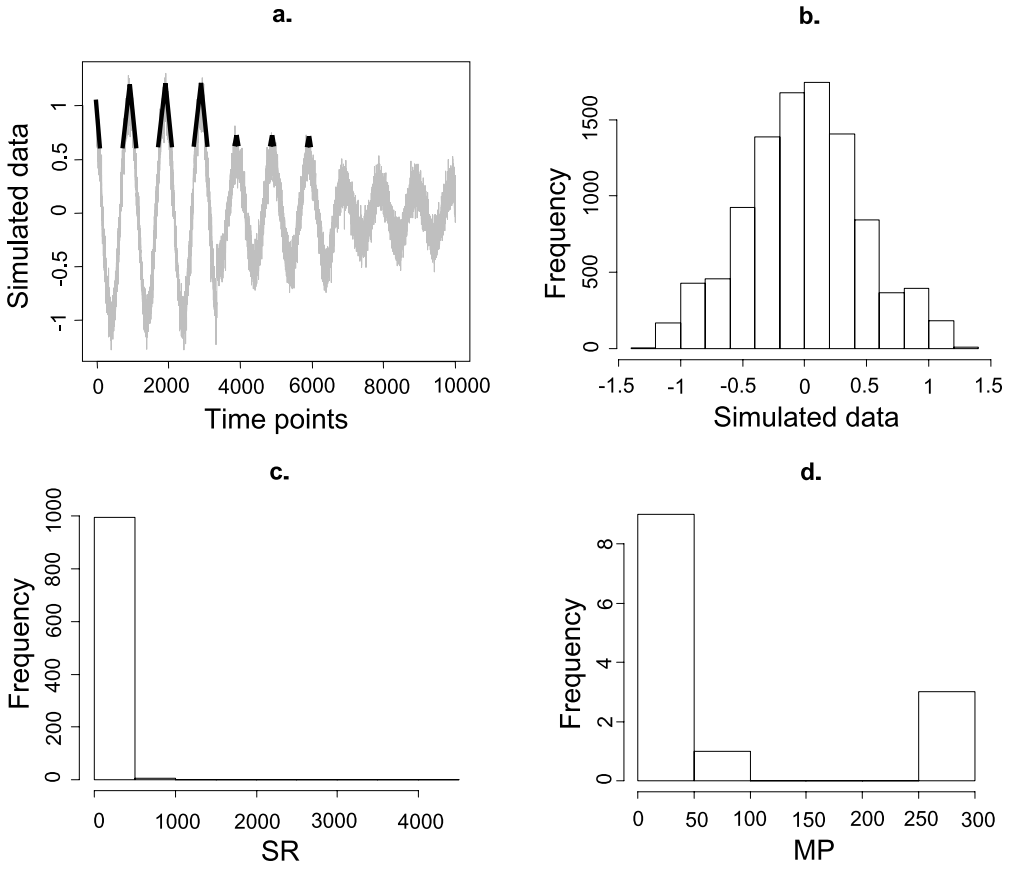


FIGURE 6.
Simulated time series from a sine curve model + noise + variable amplitude (Event$_1(X) = 10\%X(t)$). Panel (**a**) = simulated data. Panel (**b**) = histogram of raw scores. Panel (**c**) = histogram of HS subroutine. Panel (**d**) = histogram of HS main program.

similar to those displayed in Figures 4c and 4d. That is, noise in the data does not affect the HS coding, given that the same time series underlies the observed data.

In the third simulation, we generated a time series with a pure sine curve with equal wavelength but varying amplitude and also random noise. The simulated series is displayed in Figure 6a and the resulting histogram is in Figure 6b. This histogram shows more normality than those from the previous simulations. Specifying the top 10% of the data as extreme events resulted in the first 7 peaks being identified as instances of extreme expression (see highlight area in panel a). As before, panel c, displays the recurrence time for events coded as 0 in C_1. Unlike in the other conditions, this histogram has only two bars. The bar on the left represents those distances between adjacent extreme events, encompassing all events in the highlighted area. The small bar to the right represents the recurrent time between two nonadjacent defined events. These correspond to the distances between the highlighted areas. Because the wavelength is constant across the series, there are not more different recurrent times at this level. Finally, panel d displays the histogram of third segmentation level. Here, we have three different types of periods for entry or transition into a tranquil phase. Because of the differences in amplitude throughout the series, this histogram is different compared with the ones displayed in the previous simulations.

In the last simulation, we generated bivariate time series with three time segments varying in the degree of coherence between both series. In the first period ($t = 0$ to 300; first segment in

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

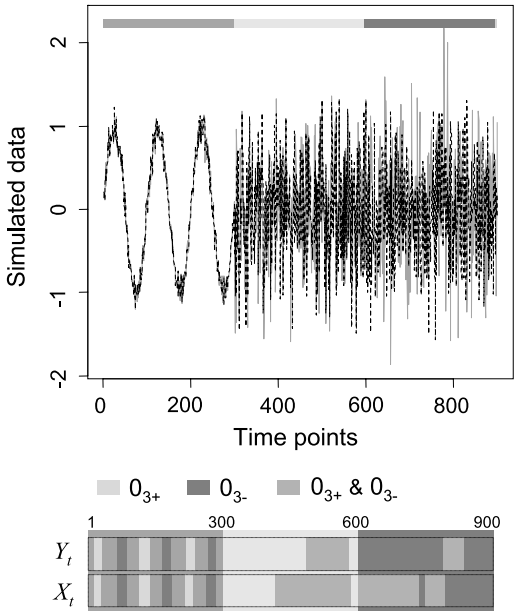


FIGURE 7.
Simulated bivariate series with three time segments and varying degree of dynamics.

Figure 7), both series follow a sine curve model $X_t = \sin(2\pi_t) + ke_{1t}$ and $Y_t = \sin(2\pi_t) + ke_{2t}$, with similar wavelength and amplitude, for $0 < t < 300$, where k is a constant and (e_1, e_2) is bivariate normal with mean $\mu = 0$, correlation $\rho = 0.9$, and common variance $\sigma^2 = 0.1$. In the second period (segment denoted by light colored line near top of figure) both series follow a model $X_t = ke_{1t}$ and $Y_t = ke_{2t}$, for $300 < t < 600$, and with (e_1, e_2) as bivariate normal and similar properties as in the first segment. In the last period (segment denoted by darker colored line), both series are also $X_t = ke_{1t}$ and $Y_t = ke_{2t}$, for $600 < t < 900$, and with (e_1, e_2) also bivariate normal but with $\rho = 0$. The panel below the data represents the HS code C_3 for each series and the overlap between the codes for both series.

In this simulation, we modified our specification for the HS to highlight areas with unusually frequent and unusually rare recurrence of extreme expressions. We used the flags 0_{3+} and 0_{3-} to mark occurrences of positive and negative value, respectively. Identifying the valence of such occurrences could be useful, for example, if the researcher is interested in distinguishing between positive and negative values. To find these values, we first considered the number of extreme events, irrespective of the valence. At the third level of HS, when a segment was coded as 0_3, we counted the number of positive and negative events for that segment. If the number of positive events was greater than the number of negative, the segment was coded as 0_{3+} and vice versa. In the first segment, the codes 0_{3+} and 0_{3-}, representing the high and low extremes, are perfectly aligned, reflecting the potential existence of a high degree of concurrence between the two series. In the second segment, such concurrence or overlap does not exist in the data and the HS codes are only aligned in part (for the 0_{3+} & 0_{3-}), representing the correlation between e_1 and e_2. In the last segment, which comprises random noise only, the HS coding does not detect any alignment or concurrent coherence between the two series. Follow-up chi-square tests can then be performed to formalize whether statistically significant associations are present in different portions of these simulated data. The corresponding results are omitted here because the simulated changes in association patterns are clearly evident in the HS plots.

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

These simulations show that HS is helpful for detecting patterns from nonstationary data. This appears to hold for data conditions including noise, variation in frequency and amplitude, and correlation between series. Some restrictions of the HS, however, should be noted. For example, if a spectral analysis method had been chosen in the third simulation, the peaks in the last waves would have been detected. This was not the case with the HS because the event of interest was the upper 10% of the series, which was below the peak of such waves. Although such peaks are the highest expression of the area, they are below the selected cutoff for the overall series. Alternative ways of extracting nonstationarity associated with changes in amplitude will be highlighted in the Discussion section.

Note that the HS can theoretically be applied to time series of any length because there is no formal hypothesis testing in the execution of the HS algorithm. In our simulations, we specified a very large number of time points to better highlight the strengths and limitations of the HS in extracting different kinds of changes in dynamics. In our empirical application (to be elaborated in detail next), we applied the HS to time series of a moderate length, with data measured daily over 182 days. In practice, a very short time series may risk bypassing important trends in a system and would likely show no sign of non-stationarity. In such cases, other standard techniques that assume stationarity can be used instead. Of course, if other standard techniques (e.g., the chi-square tests in our coherence analyses) are used in conjunction with the HS codes, the corresponding assumptions underlying these techniques still have to be factored into consideration.

4. Empirical Results

In the first set of analyses, we ran exploratory auto-, partial auto- and cross-correlation analyses using the four time series (i.e., PA and NA, for each person). The husband's PA and NA showed statistically significant partial autocorrelations up to lag 2, indicating that an autoregressive model of order 2, AR(2) model, may be an appropriate model for these data, assuming stationarity. In contrast, the wife's NA showed a significant lag-1 autocorrelation only and her PA showed significant lag-1 partial autocorrelation. This suggested that a moving average model of order 1 (i.e., MA(1) model) and an AR(1) model may be possible candidates for representing these time series if, again, stationarity is assumed.

An examination of the intraindividual cross-correlations revealed a significant lag-1 cross-correlation between the husband's PA at time t and NA at time $t - 1$. The wife, in contrast, did not show significant lag-1 cross-correlations between her PA and NA. Instead, significant cross-correlations were observed sporadically at lags 6 and 14. With regard to inter-individual cross-correlations, no significant estimates were found between the PA of both individuals. However, there was a significant lag-1 cross-correlation between the husband's NA at time t and the wife's NA at time $t - 1$. The lag-1 cross-correlation in the reverse direction (i.e., between the husband's NA at time $t - 1$ and the wife's NA at time t) was not significant. This suggested that, overall, how negative the wife felt on a given day tended to be associated with how negative the husband felt on the following day.

In the next step, we applied the HS algorithm to the four time series. Figure 8 depicts the corresponding histograms (i.e., positive and negative affect, for husband and wife, respectively). Whereas PA scores appear to be normal, the NA data are positively skewed, with a high frequency of small values and rare occurrences of high values, which is typical in emotion research. Because of these responses, the middle category of the HS coding is of small substantive importance for the NA series, so only a small portion of each series was classified into this specific category. Specifically, to recode the raw data in level 1 of the HS, we chose cutoff points separately for each of the four time series. For each dyad member's NA, we used the individual's 75th and 85th percentiles as the cutoff points for defining low, medium, and high expression. The PA

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

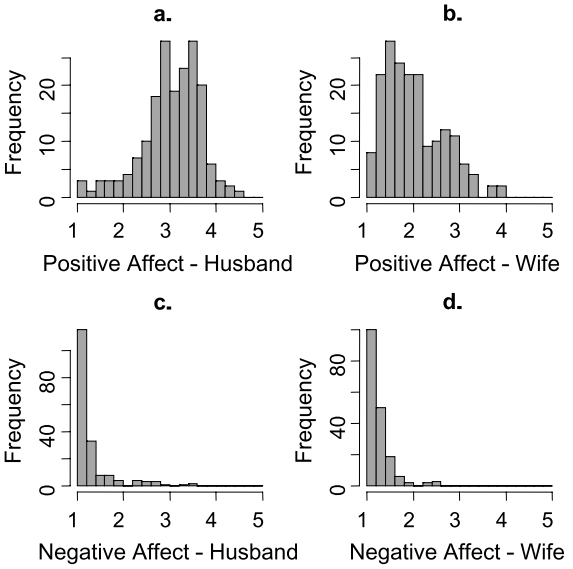


FIGURE 8.
Raw data histograms of positive and negative affect for each person.

data showed a more symmetrical distribution so we used each dyad member's 50th and 85th percentiles as the cutoff points. The resulting *SR* and *MP* histograms from the HS algorithm for each of the four variables are displayed in Figures 9 and 10.

The hierarchy of flags obtained from applying HS on the four time series (i.e., 182 days of measurements) is plotted in Figure 11. Based on the HS codes, the wife showed frequent recurrence of high PA (i.e., 0_3 segment) in the first 20 days. This diverged from the husband's extreme-PA profile, which is characterized by frequent recurrence of high PA between days 72 and 115, but not on the first 20 days. Frequent recurrence of high NA was observed in the husband's NA on the first 59 days of the study. The wife's NA was observed to show frequent high recurrence between days 11 and 69. This figure also suggests that during the first 69 days of the study, the days with frequent recurrence of high NA also coincided with the days with prolonged lack of high PA in both the husband's as well as the wife's PA (see the 0_2 regions).

Based on visual inspection of the codes in Figure 11, we deduced three periods of distinct affective features in the dyad's data. The three periods may be regarded as representing three different regimes—or latent phases—that are characterized by specific dynamics. The first period (between days 1 and 69) was characterized by a large number of days on which both individuals' NA ratings were identified as being in a 0_3 state and their PA ratings were identified as being in an 0_2 state. The second period fell between the 70th and the 114th days, and corresponded to a segment of observations with different dynamics from the first 69 days. During this period, the husband showed frequent recurrence of high PA (0_3), the wife showed a prolonged period of low PA (0_2), and yet, both individuals manifested a prolonged period of low expressions in their NA (0_2 in Figure 11). The remaining data (between 115 and 182 days) were then classified into the third regime. Collectively, we identified three distinct periods or regimes with different affective dynamics using HS.

4.1. Interindividual Coherence Analyses

Next, we examined coherence between the affect of the two individuals during specific phases of the data. Table 1 contains the chi-square results from these coherence analyses. The

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

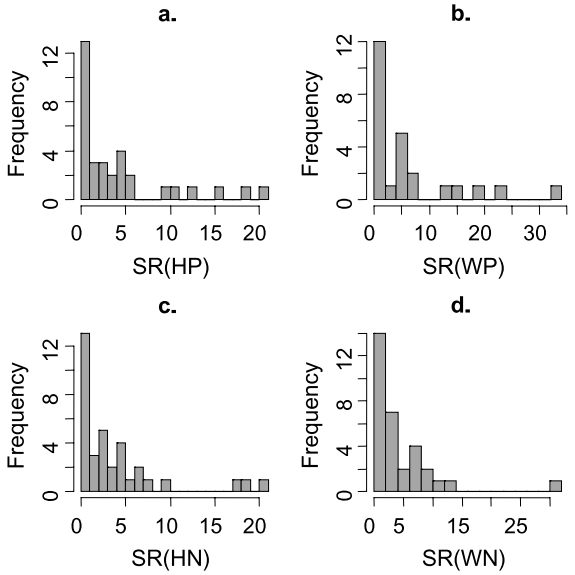


FIGURE 9.
Subroutine histograms of positive and negative affect for each person. *HP* = husband's positive affect, *WP* = wife's positive affect. *HN* = husband's negative affect, *WN* = wife's negative affect.

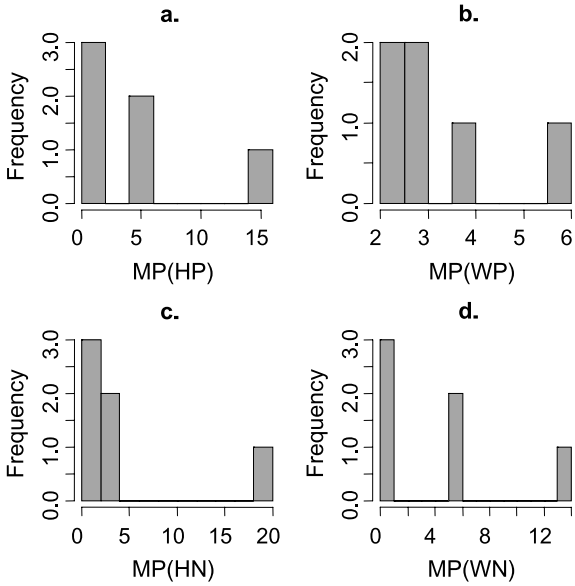


FIGURE 10.
Main Program histograms of positive and negative affect for each person.

four columns represent whether the reference for the coherence was either the husband ($W \rightarrow H$) or the wife ($H \rightarrow W$), for positive and negative affect, respectively. The first set of results, marked as "overall", contains results for the entire time series ($t = 0$ to 182). All the lag-0 tests were significant. In contrast, none of the lagged analyses (lag-1 to lag-3) was significant, except for a small lag-1 coherence ($p = 0.048$) between the wife's NA from time $t - 1$ and the hus-

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

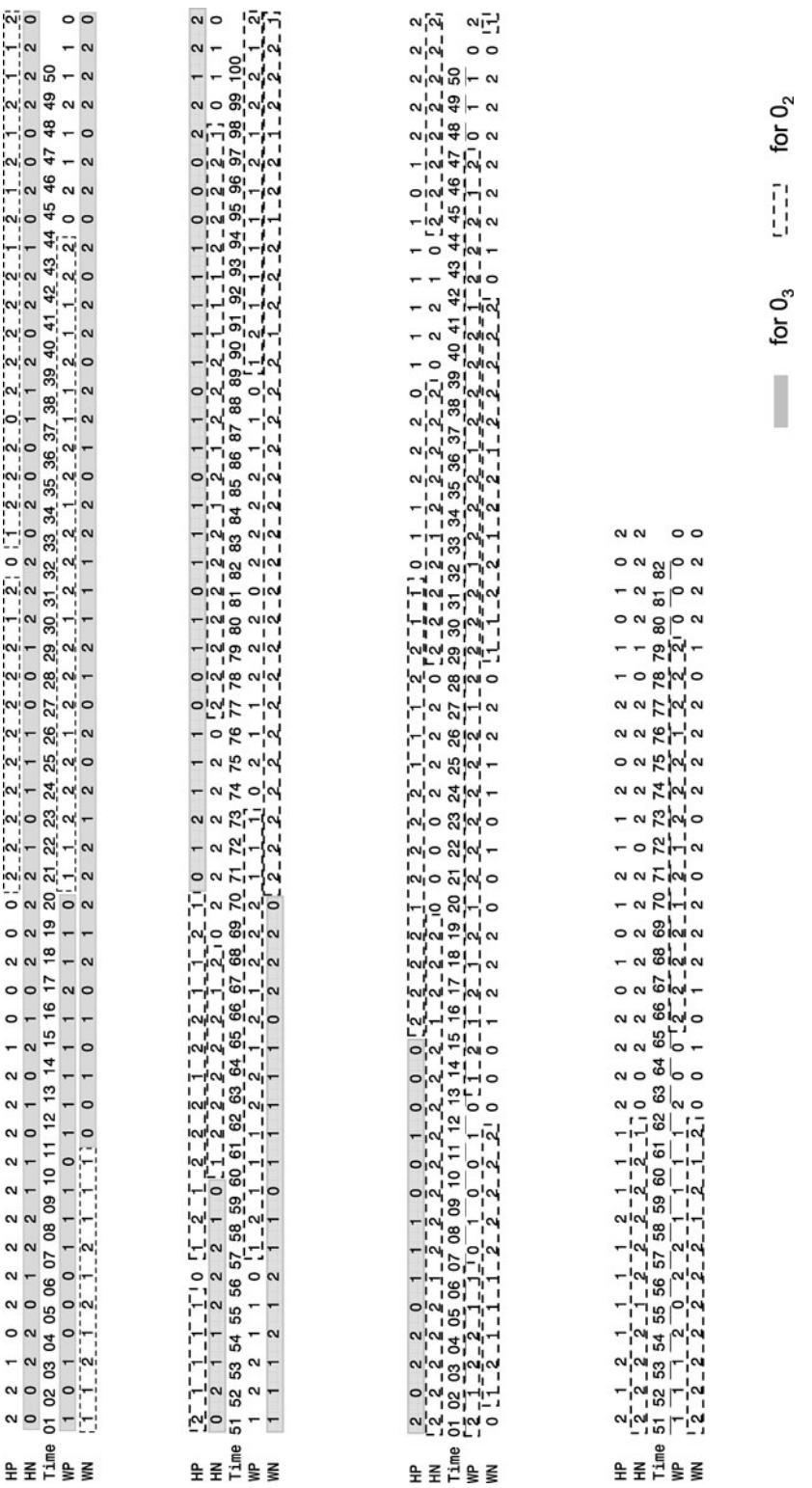


FIGURE 11.
Flags (0_2 and 0_3) for PA and NA of both individuals in the top two levels of the HS hierarchy.

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

TABLE 1.
Chi-square test results from coherence analysis.

		$WP \rightarrow HP$ $(t-k \rightarrow t)$		$HP \rightarrow WP$ $(t-k \rightarrow t)$		$WN \rightarrow HN$ $(t-k \rightarrow t)$		$HN \rightarrow WN$ $(t-k \rightarrow t)$	
	lag (k)	χ^2	p	χ^2	p	χ^2	p	χ^2	p
overall	0	12.48	0.014	12.48	0.014	26.24	<0.001	26.24	<0.001
	1	4.013	0.404	4.019	0.403	9.574	0.048	7.314	0.120
	2	3.673	0.452	4.151	0.386	6.878	0.142	4.814	0.307
	3	0.670	0.955	4.98	0.289	2.766	0.598	7.523	0.110
0_3	0	6.576	0.160	4.433	0.351	7.779	0.100	6.228	0.183
	1	3.422	0.490	5.523	0.238	4.220	0.377	4.790	0.309
	2	1.827	0.767	9.376	0.052	4.400	0.355	1.102	0.894
	3	4.302	0.367	0.8	0.938	3.531	0.473	2.082	0.721
non-0_3	0	8.689	0.069	11.835	0.019	16.85	0.002	20.48	0.001
	1	6.062	0.195	7.620	0.107	6.283	0.179	10.14	0.038
	2	6.759	0.149	8.025	0.091	4.643	0.326	3.069	0.546
	3	1.315	0.859	5.889	0.208	2.434	0.657	4.560	0.335
0_2	0	1.882	0.390	9.096	0.011	5.720	0.057	4.501	0.105

Note. $W \rightarrow H$ indicates coherence of both individuals' affect using the husband's data as reference. $H \rightarrow W$ indicates coherence of both individuals' affect using the wife's data as reference. P = positive affect. N = negative affect.

band's NA at time t. This result is consistent with the cross-correlation analyses between these individuals' NA across all 182 days. The significant lag-1 coherence from wife's NA toward the husband's NA is also in line with the significant lag-1 cross-correlation between the husband's NA at time t and the wife's NA at time $t-1$. In addition, there were several lagged coherences between NAs that approached p-values around 0.1. Although not significant, this pattern of coherences for NA was tied to χ^2 values that were notably higher than those associated with the values for PA. These results indicate that there was a strong overall concurrent coherence in PA and NA between both individuals, as well as some degree of lagged coherence in their negative—but not positive—affect.

The second set of analyses includes coherences during periods when one of the affect variables was in a state of frequent extreme recurrence (0_3), as determined from the HS code. For example, the lag-0 χ^2 value of 6.576, for $WP_{t-k} \rightarrow HP_t$, represents the association between the husband and wife's PA when the *husband's PA* was in a state of frequent extreme recurrence. This value would generally be different from the lag-0 χ^2 value for $HP_{t-k} \rightarrow WP_t$ (4.433); this latter value would represent the same association when the *wife's PA* was in a state of frequent extreme recurrence. Neither the wife's NA or PA showed signs of concurrent synchrony with the husband's NA ($p > 0.05$). That is, when one person was in a state of frequent recurrence in either high PA or high NA, the other person tended *not* to "imitate" this affective state, either concurrently or through lags. Thus, the driving force of this high affective volatility may lie outside of the couple's dynamic system; this may also reflect the couple's own way of stabilizing interpersonal dynamics.

The next block of analyses, labeled non-0_3 (i.e., all the data minus the 0_3 period), represents results from the nonextreme periods associated with each of the four affect variables. In these nonextreme periods, the synchrony of PA and NA of both individuals is evident. The husband showed significant lag-0 coherence toward the wife's NA as well as PA, but directional coherence from husband to wife was only observed in NA. During the periods when the wife did not show

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

frequent recurrence of high affect (either PA or NA), the husband tended to show synchrony in his affect. There was also evidence of the husband's NA from time $t-1$ relating to the wife's lack of extreme high NA at time t. In contrast, the wife tended to show only concurrent coherence toward the husband's lack of frequent extreme NA, but not PA.

The last section in Table 1, labeled "0_2", represents analyses from the period of persistent non-extreme affect (segments with the flag of 0_2). When the wife's PA is in this period, there was a concurrent association between both individuals' prolonged lack of extreme PA ($p = 0.011$). When the husband's NA is in this period, there was a marginally significant association between both individuals' prolonged lack of extreme NA ($p = 0.057$). These results point to the specificity of the synchronization patterns between the dyad members during periods of nonextreme affect. That is, there was a tendency for the husband's PA to synchronize with the wife's lack of extreme PA concurrently, whereas the wife tended to synchronize with the husband's lack of extreme NA, also on the same day.

In sum, the HS analyses showed that the concurrent and lagged coherences (and lack thereof) between these two individuals varied depending on the specific affective regimes of the dyad. Analysis of the entire series and those from segments characterized by prolonged lack of extreme NA indicated a lag-1 coherence in the wife's NA synchronizing toward the husband's NA. In contrast, lag-1 coherence in the reverse direction (i.e., from husband to wife) was found during the relatively tranquil phases (i.e., the non-0_3 data segments, which included segments with medium expression). This distinction would be bypassed if standard cross-correlation analysis is performed without first testing the tenability of the stationarity assumption. With HS we were able to capture changes in directionality of the lead-lag relationships and other affective features over time.

5. Discussion

In this article, we presented a technique for exploring the dynamics underlying multivariate processes in dyadic interactions. This technique is potentially useful for pattern recognition in interactions that involve nonstationary and/or nonlinear dynamics. As applied to our empirical data, we first partitioned the time series into segments with a presumed underlying state of each person in the dyad. We then examined the synchrony between the states of both individuals by aligning the corresponding sequences and evaluating their coherence via chi-square tests. Identifying systematic patterns from data with potential nonstationarity and/or nonlinearity has been a central goal in many scientific pursuits. Examples range from detecting gene expression in DNA sequences (Ewens & Grant, 2005) to uncovering social patterns using small-world networks (Amaral, Diaz–Guilera, Moreira, Goldberger, & Lipsitz, 2004; Watts & Strogatz, 1998). We believe our proposed approach could add to the existing applications in social and behavioral research.

5.1. Application to Dyadic Interactions

Our application of HS to the data from one dyad illustrates one of the various possibilities of this approach for modeling dyadic interactions. This application revealed patterns of interactions with insights about the couple's affective dynamics. Such interactions were not obvious in the observed data and can add to information detectable with other existing methods. For instance, we showed that computing cross-correlation using the entire multivariate time series assuming stationarity masked information about the lead-lag relationship between the dyad members. In this case, alternative methods such as windowed cross-correlation techniques (Boker et al., 2002) could be a better tool for exploring the time-varying coherence between the two individuals. The

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

HS supplements these techniques by providing a nonparametric approach to extracting time-varying coherence.

One method that shares ideas with HS is recurrence—and cross-recurrence—analysis (Eckmann et al., 1987; Marwan & Kurths, 2002; Shockley, Butwill, Zbilut, & Webber, 2002), a technique for identifying repeating patterns in nonlinear dynamical systems. Recurrence analysis was first developed as a visual technique to detect hidden rhythms, patterns, and nonstationarities in experimental data (Eckmann et al., 1987). It involves the quantification of features of recurrence plots. Our method, however, is an exploratory approach that deals with multivariate interactions in a different way. HS generates a schema of sequence codes at different levels of the data that helps to detect information in complex interacting systems. The revealed information is, at this point, exploratory but it could serve as the basis for creating confirmatory models with which to evaluate the dynamic system as a whole.

Some of these confirmatory models could be of the family of change point models and regime switching models. In the former (Carlin et al., 1992; Hinkley, 1970, 1971; Stephens, 1994), the time points at which changes occur can be specified a priori or are estimated from the data. However, the change points still have to be defined within the context of a confirmatory dynamic model and explicit predictions have to be made regarding the dynamics of the system before and after the change point(s). In regime switching models (Kim & Nelson, 1999; Dolan et al., 2005), regime is essentially defined as a discrete latent variable indicating the current state of the system. That is, the system is hypothesized to exhibit different dynamics contingent upon this latent indicator. The probability that the system is in a particular regime or state at time t is then estimated from the data. As in the case of change point models, the researcher needs to have a preconceived notion of the differences in dynamics among different regimes, and thus a confirmatory model for each regime. In most instances, even though external criteria such as the Bayesian Information Criterion (BIC) can be used to help deduce the number of regimes in a data set, the researcher still needs to specify a range for the possible number of regimes in a data set. In practice, estimating the parameters for models with more than three regimes is a particularly challenging task given the existence of multiple local minima in the associated likelihood function. Such estimation difficulties may in turn, limit the types of dynamic models researchers use to describe their data.

The information revealed by HS could be helpful for constructing such models of change. For example, in our empirical data, we found three periods or "regimes" (i.e., pre-$HP0_3$, $HP0_3$, and post-$HP0_3$). Different affective dynamics for PA and NA were observed within each regime. In particular, the couple's NA showed emerging coherence and synchronization toward a mutually low NA state beginning from the second period. During the third period, the couple retained this synchrony, while the frequent recurrence in the husband's PA observed during the second period was no longer observed. This information could then be used to specify a bivariate dynamic model with regime-specific coupling parameters. This exploratory-confirmatory modeling approach can bring new light to the analysis of dyadic interactions.

The nonparametric nature of coupling with the coding of this methodology appears to be particularly suitable for psychological measurements, especially for ordinal measures with blurred information about the true construct. Models applied to such kind of data can easily fail. This is because validating the underlying assumption with time series of discrete data is particularly difficult. For example, time series models such as autoregressive (AR) or moving average (MA), and their hybrid (ARMA) models, are known to be difficult to adapt to discrete data (Stoffer, 1991). Whereas Markov models for categorical data exist (Frühwirth-Schnatter, 2006), these models rely on various types of stationary assumptions that might not hold with real data, especially of the kind used here (Manuca & Savit, 1996).

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

5.2. *Methodological Issues and Extensions*

As any incipient technique, the full range of applications of HS has yet to be studied. Similarly, there is a number of methodological features that need detailed examination. One such aspect concerns the conditions needed for reliably capturing the dynamics of a system. These include, for example, the minimum number of measurement occasions required to discriminate dynamic signal from random noise, or the maximum number of dimensions that can be explored simultaneously. Similarly, more work is needed to study measures of complexity in this type of time-series data and their use to map out the underlying mechanism of the dynamic system (Chaitin, 1987; Lempel & Ziv, 1976; Ziv & Lempel, 1977; Hsieh et al., 2006).

In one of our simulation examples, we noted that the current specifications used in the HS did not detect the second reduction in amplitude in a cycle that shows invariant frequency over time. That is, after the second decrease in amplitude, the peaks of the cycle were no longer considered extreme event (i.e., "0") and information from this region was not given much emphasis in other layers of the HS hierarchy. Cyclic processes that show continual or abrupt shifts in amplitude are commonly observed in social and behavioral sciences (Bisconti, Bergeman, & Boker 2004; Chow, Ram, Boker, Fujita, & Clore, 2005; Chow, Hamaker, Fujita, & Boker, 2009; Molenaar & Newell, 2003). While the current formulation of the HS algorithm is meant to extract certain kinds of systematic patterns, this formulation can certainly be adapted and new layers can be added to tailor to the nature of empirical applications. For instance, if an extreme event is defined to be all the points at which the first derivative of the time series is zero, all the peaks would have been identified by the HS. Extending the current formulation of the HS to allow for localized extreme event detection by means of sliding windows also provides an alternative way of detecting such changes.

Another aspect of research that deserves attention is the possible extensions of HS for examining complex nonlinear dynamics. For example, it would be possible to include in the HS procedure information from the data that could be used to detect states leading to, say, attractors or discontinuities in the dynamics. Such information could come from both inside and outside the dyadic system. Examples of information from inside the system would be a fight between both individuals. Examples of information from outside the system would be something as important as losing one's job or as seemingly trivial as the weather. These events can be considered as "perturbations" in the system and can be used to detect the dynamics of the dyad during a given period. Such information can be further used to examine possible changes in the dyadic interactions, changes in which the dyad engages to recover from the perturbations. The ability to detect these discontinuities would be one of the advantages of HS.

In sum, in this article, we introduced a nonparametric technique based on computation modeling and illustrated its utility for examining affective processes in couples. We hope that this flexible approach opens new possibilities to the study of dynamics in dyadic interactions.

Acknowledgements

This study was supported in part by grants from the National Science Foundation (BCS-05-27766, BCS-08-27021 and BCS-08-26844) and NIH-NINDS (R01 NS057146-01).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

References

Amaral, L.A.N., Diaz–Guilera, A., Moreira, A.A., Goldberger, A.L., & Lipsitz, L.A. (2004). Emergence of complex dynamics in a simple model of signaling network. *Proceedings of the National Academy of Sciences*, *101*, 155551.

Bisconti, T.L., Bergeman, C.S., & Boker, S.M. (2004). Emotional well–being in recently bereaved widows: A dynamical systems approach. *Journal of Gerontology: Psychological Sciences*, *59*(B), 158–167.

Bobitt, R., Gourevitch, V., Miller, L., & Jensen, G. (1969). Dynamics of social interactive behavior: A computerized procedure for analyzing trends, patterns, and sequences. *Psychological Bulletin*, *71*, 110–121.

Boker, S.M., & Laurenceau, J.-P. (2006). Dynamical systems modeling: An application to the regulation of intimacy and disclosure in marriage. In Walls, T., & Schafer, J. (Eds.), *Models for intensive longitudinal data* (pp. 195–218). New York: Oxford University Press.

Boker, S.M., & Rotondo, J. (2003). Symmetry building and symmetry breaking in synchronized movement. In M. Stamenov, & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language* (pp. 163–171). Amsterdam: John Benjamins.

Boker, S.M., Xu, M., Rotondo, J.L., & King, K. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods*, *7*(1), 338–355.

Campbell, L., & Kashy, D. (2002). Estimating actor, partner, and interaction effects for dyadic data using proc mixed and hlm: A guided tour. *Personal Relationships*, *9*, 327–342.

Carlin, B., Gelfand, A., & Smith, A. (1992). Hierarchical Bayesian analysis of changepoints problems. *Applied Statistics*, *41*, 389–405.

Casdagli, M. (1997). Recurrence plots revisited. *Physica D*, *108*, 12–44.

Castellan, N. (1979). The analysis of behavior sequences. In R. Cairns (Ed.), *The analysis of social interactions: Methods, issues, and illustrations* (pp. 81–116). Hillsdale: Lawrence Erlbaum Associates.

Chaitin, G. (1987). *Algorithmic information theory*. Cambridge: Cambridge University Press.

Chow, S.-M., Ferrer, E., & Nesselroade, J. (2007). An unscented Kalman filter approach to the estimation of nonlinear dynamical systems models. *Multivariate Behavioral Research*, *42*, 283–321.

Chow, S.-M., Hamaker, E.J., Fujita, F., & Boker, S.M. (2009). Representing time-varying cyclic dynamics using multiple-subject state-space models. *British Journal of Mathematical and Statistical Psychology*, *62*, 683–716.

Chow, S.-M., Ram, N., Boker, S.M., Fujita, F., & Clore, G. (2005). Emotion as thermostat: Representing emotion regulation using a damped oscillator model. *Emotion*, *5*(2), 208–225.

de Rooij, M., & Kroonenberg, P. (2003). Multivariate multinomial logit models for dyadic sequential interaction data. *Multivariate Behavioral Research*, *38*, 463–504.

Dolan, C., Schmittmann, V., Lubke, G., & Neale, M. (2005). Regime switching in the latent growth curve mixture model. *Structural Equation Modeling*, *12*(1), 94–119.

Eckmann, J.-P., Kamphorst, S., & Ruell, D. (1987). Recurrence plots of dynamical systems. *Europhysics Letters*, *5*, 973–977.

Ewens, J.W., & Grant, G.R. (2005). *Statistical methods in bioinformatics*. New York: Springer.

Felmlee, D.H., & Greenberg, D.F. (1999). A dynamic systems model of dyadic interaction. *Journal of Mathematical Sociology*, *23*(3), 155–180.

Ferrer, E., Chen, S.-C., Chow, S.-M., & Hsieh, F. (2010). Exploring intra-individual, inter-individual and inter-variable dynamics in dyadic interactions. In S.M. Chow, E. Ferrer, & F. Hsieh (Eds.), *Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 381–411). New York: Taylor and Francis.

Ferrer, E., & Nesselroade, J.R. (2003). Modeling affective processes in dyadic relations via dynamic factor analysis. *Emotion*, *3*(4), 344–360.

Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models*. New York: Springer.

Gardner, W. (1993). Hierarchical continuous-time sequential analysis: A strategy for clinical research. *Journal of Consulting and Clinical Psychology*, *61*, 975–983.

Gardner, W., & Griffin, W. (1989). Methods for the analysis of parallel streams of continuously recorded social behaviors. *Psychological Bulletin*, *105*, 446–455.

Geman, S., Potter, D., & Chi, Z. (2002). Composition systems. *Quarterly of Applied Mathematics*, *60*, 707–736.

Goodman, L. (1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. *Journal of the American Statistical Associations*, *65*, 226–256.

Gottman, J. (1979). Detecting cyclicity in social interaction. *Psychological Bulletin*, *86*, 338–348.

Gottman, J., Murray, J., Swanson, C., Tyson, R., & Swanson, K. (2002). *The mathematics of marriage: Dynamic nonlinear models*. Cambridge: MIT Press.

Granic, I., & Hollenstein, T. (2003). Dynamic systems methods for models of developmental psychopathology. *Development and Psychopathology*, *15*, 641–669.

Hamilton, J. (1988). Rational-expectations econometric analysis of change in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control*, *12*, 385–423.

Hinkley, D.V. (1970). Inference about the changepoint in a sequence of random variables. *Biometrika*, *57*, 1–17.

Hinkley, D.V. (1971). Inference about the changepoint from cumulative sum tests. *Biometrika*, *58*, 509–523.

Hsieh, F., Hwang, C.-R., Lee, H.-C., Lan, Y.-C., & Horng, S.-B. (2006). Testing and mapping non-stationarity in animal behavioral processes: A case study on an individual female bean weevil. *Journal of Theoretical Biology*, *238*(4), 805–816.

Kashy, D., & Kenny, D. (2000). The analysis of data from dyads and groups. In H. Reiss, & C. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 451–477). New York: Cambridge University Press.

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press

Kim, C.-J., & Nelson, C.R. (1999). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge: MIT Press.
Kitagawa, G. (1981). A non-stationary time series model and its fitting by a recursive filter. *Journal of Time Series Analysis*, *2*, 103–116.
Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on Information Theory*, *22*, 21–27.
Levenson, R., & Gottman, J. (1983). Marital interactio: Physiological linkage and affective exchange. *Journal of Personality and Social Psychology*, *45*(3), 587–597.
Manuca, R., & Savit, D. (1996). Stationarity and nonstationarity in time series analysis. *Physica D*, *99*, 134–161.
Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A*, *302*, 299–307.
Molenaar, P.C.M., & Newell, K.M. (2003). Direct fit of a theoretical model of phase transition in oscillatory finger motions. *British Journal of Mathematical and Statistical Psychology*, *56*, 199–214.
Newsom, J. (2002). A multilevel structural equation model for dyadic data. *Structural Equation Modeling*, *9*, 431–447.
Newtson, D. (1993). The dynamics of action and interaction. In L. Smith, & E. Thelen (Eds.), *A dynamic systems approach to development: Applications* (pp. 241–264). Cambridge: MIT Press.
Priestley, M. (1988). *Nonlinear and non-stationary time series*. New York: Academic Press.
Raudenbush, S., Brennan, R., & Barnett, R. (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology*, *9*, 161–174.
Sackett, G. (1979). The lag sequential analysis of contingency and cyclicity in behavioral interaction research. In J. Osofsky (Ed.), *Handbook of infant development*. New York: Wiley.
Shockley, K., Butwill, M., Zbilut, C., & Webber, C. (2002). Cross recurrence quantification of coupled oscillators. *Physics Letters A*, *305*(1–2), 59–69.
Shumway, R., & Stoffer, D. (2006). *Time series analysis and its applications: With R examples*. New York: Springer.
Stephens, D. (1994). Bayesian retrospective multiple changepoint identification. *Applied Statistics*, *43*, 159–178.
Stoffer, D. (1991). Walsh–Fourier analysis and its statistical application. *Journal of the American Statistical Association*, *86*, 461–479.
Watson, D., Lee, A.C., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scale. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.
Watts, D., & Strogatz, S. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*, 440–442.
Weber, E., Molenaar, P., & Van der Molen, M. (1992). A nonstationarity test for the spectral analysis of physiological time series with an application to respiratory sinus arrhythmia. *Psychophysiology*, *29*(1), 55–65.
West, B. (1985). *An essay on the importance of being nonlinear*. Berlin: Springer.
Ziv, J., & Lempel, A. (1977). A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory*, *23*, 337–343.

Manuscript Received: 12 DEC 2008
Final Version Received: 6 AUG 2009
Published Online Date: 24 MAR 2010

https://doi.org/10.1007/s11336-009-9146-8 Published online by Cambridge University Press