# Wavelet packet modelling of infant sleep state using heart rate data

Guy P. Nason,[†] Theofanis Sapatinas [‡] and Andrew Sawczenko [§]

Submitted: 19th June 2000; Revised: 19th March 2001

**Abstract**

We show how a recently developed wavelet packet modelling methodology could be useful for infant sleep state classification using heart rate data. The suggested approach produces adequate classification rates when applied to recordings from an infant who was placed to bed at night at different ages. As well as classification, this approach gives us valuable information about the relationship between sleep state and heart rate. The statistical model tells us which sorts of wavelet packets of heart rate are most important for classifying sleep state.

*Keywords:* ANTEDEPENDENCE MODELS; INFANT SLEEP STATE CLASSIFICATION; LINEAR DISCRIMINANT ANALYSIS; VARIABLE SELECTION; WAVELETS; WAVELET PACKETS

## 1 INTRODUCTION

This article concentrates on the interesting medical problem of infant sleep state classification using heart rate data. It utilises the recent methodology of Nason & Sapatinas (2001) which advocates the non-decimated wavelet packet transform (NWPT) to model a response time series in terms of a (possibly) non-stationary explanatory time series; it is assumed that both time series have the same finite length. The suggested computational technique transforms the explanatory time series into a NWPT representation resulting in a situation having more "variables" than observations. Then, statistical variable selection techniques are sought to identify which wavelet packets (variables) are useful for modelling the response time series. The selected statistical model usually provides valuable information about which components in the explanatory time series drive the response time series.

The article is organised as follows: Section 2 provides a detailed history of the medical problem studied here and describes the actual data set used in our analysis. In Section 3, we give a brief description of (discrete) wavelets and wavelet packets and a brief explanation of the wavelet packet modelling methodology of Nason & Sapatinas (2001). In Section 4, we show how the NWPT representation of heart rate (the explanatory time series) could be useful for infant sleep state (the response time series) classification. In particular, a non-sophisticated variable selection technique is first adapted to tackle the problem of having more variables than observations. This step, although somewhat rough and ready, is computationally fast and provides an adequate initial dimension reduction prior to the exploration of standard statistical classification techniques. Furthermore, a linear discriminant analysis on the selected variables (non-decimated wavelet packets) produces adequate sleep state classification rates when applied to recordings from an infant who was placed to bed at night at different ages. As well as classification, this approach gives us valuable information about the relationship between sleep state and heart rate. The statistical model tells us which sorts of wavelet packets of heart rate are most important for classifying sleep state. Moreover, to evaluate the success

[†] Department of Mathematics, University of Bristol, BS8 1TW, England (Email: G.P.Nason@bristol.ac.uk).

[‡] *Address for correspondence:* Department of Mathematics and Statistics, University of Cyprus, P.O. Box 20537, CY 1678 Nicosia, Cyprus (Email: T.Sapatinas@ucy.ac.cy).

[§] Institute of Child Health, Royal Hospital for Sick Children, Bristol, England (Email: a1@abjs.demon.co.uk).

rates obtained with the above approach, the antedependence modelling method from Krzanowski *et al.* (1995) was adapted. This, more sophisticated and computational demanding, approach has been recently developed to perform discrimination when the number of variables is larger than the number of observations. It is shown that the first and second order antedependence models also produce adequate sleep state classification rates and improve on the ones obtained (on the same infant) by combining the rough variable selection technique with linear discriminant analysis. Some concluding remarks are made in Section 5.

# 2 BACKGROUND MATERIAL AND THE DATA SET

In this section, we give a detailed account on the history of the medical problem and describe the actual date set used in our analysis.

## 2.1 THE HISTORY OF THE PROBLEM

Sleeping and waking states are ubiquitous behavioural characteristics already present during fetal life and which continue to develop during post-natal life. During *adulcy*, sleep consists of two distinct types which alternated within a 90 minute cycle. If awoken during periods associated with *rapid eye movements* (REM sleep) volunteers reported vivid dreams but not during other periods of sleep, which are labelled as *non-rapid eye movements* (non-REM sleep) (see Dement & Kleitman, 1957). There are major physiological changes between awake and sleep but also within sleep between REM and non-REM.

During *infancy*, the terms ACTIVE SLEEP and QUIET SLEEP are used in an analagous manner to REM and non-REM sleep (Anders *et al.*, 1971). The different terminology emphasises that infancy is a dynamic period characterised by rapid growth, development and 'maturation'. Whilst the physiological patterns observed during ACTIVE SLEEP and QUIET SLEEP are similar to the adult equivalants, they are different and change with increasing age: thus physiological recordings made during ACTIVE SLEEP in the first month are not the same as those made at four months of age. The newborn spends the majority of its time sleeping but the awake periods lengthen and coalesce towards the day-time. Initially sleep is characterised by long periods of ACTIVE SLEEP interspaced with with shorter periods of QUIET SLEEP but the periods of QUIET SLEEP lengthen, whilst the duration of ACTIVE SLEEP periods either shorten or remain the same. The infant thus develops an approximate 50–60 minute sleep cycle consisting of alternate 20–30 minute periods of ACTIVE SLEEP and QUIET SLEEP. Other body systems also mature and change during this period. For example infants have faster heart rates and breathing rates than adults but the rates decrease with increasing age reflecting not only growth in the heart and lungs but also maturation in the overall controlling systems sited within the developing brain (see Harper *et al.*, 1976; Schechtman *et al.*, 1993).

ACTIVE SLEEP is recognised by uneven respiration and sporadic body movements but with low muscle tone in between these movements. This reduced tone can lead to partial collapse of the upper airway and snoring, or even complete cessation of air flow (apnea). Rapid eye movements, smiles, frowns, grimaces, mouthing, sucking, sighs, and twitches are frequent and are associated with increased variability in heart rate. QUIET SLEEP is characterised by slow regular respiration, less variability in heart rate, an increased muscle tone and fewer movements. A major difficulty in any classification system is that individual infants do not show all these criteria all the time. Conventionally, sleep states are characterised primarily using electrophysiological measurements which involves the attachment of EEG (electro-encephalogram — "brain-waves") and EOG (electro-oculogram — eye movements) sensors. Sleep state is manually determined the next day by a trained observer visually interpreting predetermined time periods (eg each 30 second period) of the infant's EEG and EOG that had been concurrently recorded (see Anders *et al.*, 1971). Whilst this is an accurate and reproducible method of sleep state analysis (about 80% inter-observer agreement) the determination is time-consuming, laborious and expensive. The attachment of the recording sensors to the infants scalp (EEG) and face (EOG) may be distressing to both parents and infants, may lead to artifacts by interfering with the infants sleep, and is not practicable in the home environment. Thus such recordings must be performed in the hospital which further adds to cost and potential distress. By comparison, heart rate is automatically measured using standard commercial ECG (electro-cardiogram) monitors. The ECG recording is relatively unobtrusive, since the leads are attached to the infants chest and parents can be readily taught to do this. Moreover, heart rate is cheap to measure directly.

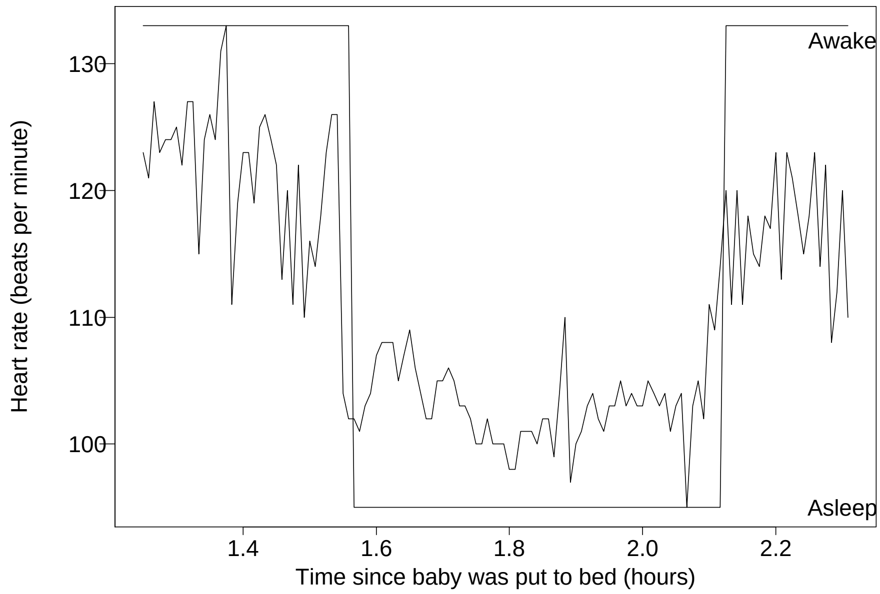


Figure 1: Time series of heart rate and sleep state for a four month old baby beginning at 23:09:42. (Heart rate is labelled by the left-hand axis. Sleep state is labelled by the right-hand axis and takes only two states: ASLEEP and AWAKE.)

## 2.2 THE DATA SET

Sawzenko *et al.* (1995) have recently completed a prospective study of nocturnal infant physiology in a sleep laboratory designed to be similar to a normal domestic bedroom. Five mothers and their healthy first-born infants slept in the thermally controlled room each month for the first 5 months. Conventional polsomnography including one channel of EEG and EOG, chest and abdominal movement and ECG (Oxcams, Oxford Ltd) as well as multiple temperature measusements (Squirrel 1200, Cambrige instruments Ltd) and infra-red video recordings were made. In the studies reported here all infants slept supine in a cot besides their mother, but mothers were free to care for their infants as they would at home (eg. feed, change nappy, etc) whilst recording continued to take place. Most studies comenced around 20:00–21:00 and finished around 08:00–09:00. Sleep staging was performed off line (see Stefanski *et al.*, 1984) and state was assigned to consecutive 30 second blocks of averaged heart rate. Three sleep states were recorded: AWAKE, ACTIVE SLEEP and QUIET SLEEP, and two consecutive minutes were needed for a transition to be recorded.

To simplify our analysis, we combined ACTIVE SLEEP and QUITE SLEEP into one category (called ASLEEP) and, thus, have concentrated on two sleeping states (AWAKE and ASLEEP). We have also considered recordings only from *one* infant who was placed to bed at night at different ages. Figure 1 shows a segment of two time series recorded from a four month old infant who was placed to bed at night. The time series shown are of heart rate and sleep state (ie whether the infant was AWAKE or ASLEEP) sampled every 30 seconds. After about 1.5 hours the infant eventually fell ASLEEP only to wake around half an hour later. Figure 1 shows that the heart rate is low when the baby is ASLEEP and high when it is AWAKE. It seems that the mean level of heart rate over certain time scales is likely to be important for determining sleep state. From the above a method which can reliably predict sleep state from heart rate would therefore be clinically valuable.

# 3 THE NON-DECIMATED WAVELET PACKET MODELLING METHODOLOGY

In this section, we give a brief description of (discrete) wavelets and wavelet packets and a brief explanation of the wavelet packet modelling methodology of Nason & Sapatinas (2001).

## 3.1 (DISCRETE) WAVELETS AND WAVELET PACKETS

Daubechies (1992) and Hess-Nielsen & Wickerhauser (1996) give excellent descriptions of wavelets and wavelet packets respectively, and explain how they reveal information about the variation of signals in time and frequency. A comprehensive description of wavelets and wavelet packets is beyond the scope (and length) of the present paper. To gain an overview of the methodology we use it is, however, enough to know that

- wavelets form a set of oscillatory basis functions that can be used to efficiently represent functions of interest (this set is usually constructed by dilating and translating a single *mother wavelet* function enjoying nice mathematical properties, such as compact support, high regularity and a number of vanishing moments);

- wavelet packets form a large "library" of oscillatory basis functions of which wavelets are a subset (these basis functions inherit nice mathematical properties from their generating wavelet basis functions). For any particular application, the "best basis" can be chosen from the library of oscillatory basis functions according to some user-defined criterion function (like the Shannon entropy measure which defines a "best basis" to be one which represents functions sparsely).

We give a few examples of the types of functions that we are referring to. For example, the Haar mother wavelet in *continuous time* is given by the function

$$\psi(t) = \begin{cases} 1/\sqrt{2} & \text{if } t \in (0, \frac{1}{2}) \\ -1/\sqrt{2} & \text{if } t \in (\frac{1}{2}, 1) \\ 0 & \text{otherwise.} \end{cases}$$

The wavelets are all scaled and shifted versions of the mother wavelet. For example, the wavelet at scale $j$ and location $2^{-j}k$ is given by $\psi_{jk}(t) = 2^{j/2}\psi(2^j t - k)$. Wavelet coefficients at scale $j$ and location $2^{-j}k$ are found by forming the inner product of $f$ with $\psi_{jk}$. Wavelets in *discrete time* can be formed from the continuous ones. For example, the Haar mother wavelet in discrete time is given by the vector $(1/\sqrt{2}, -1/\sqrt{2})$, at the next coarser scale by $(1/2, 1/2, -1/2, -1/2)$, and so on. For more information about discrete wavelets see Nason *et al.* (2000), in particular the formula definition for discrete wavelets, $\psi_{jk}$ for $j > 0$ and $k = 0, \ldots, N_j - 1$, is given by

$$\psi_{1n} = \sum_k g_{n-2k}\delta_{0k} = g_n$$

and

$$\psi_{j+1,n} = \sum_k h_{n-2k}\psi_{jk}, \tag{1}$$

where $N_j = (2^j - 1)(N_h - 1) + 1$ and $g_k, h_k$ are the quadrature mirror filters of Daubechies' (1992) compactly supported wavelets of length $N_h$.

Wavelets are all obtained from simple scalings (by factors of 2) and translations $2^{-j}k$ of one mother function. The frequency response of the mother wavelet is a local bandpass filter and the frequency response of wavelets at other scales $j$ bandpasses a signal at different octaves. The time-locality of the wavelet filter is controlled by $k$. So for example, the Littlewood-Paley or Shannon wavelet is defined in the Fourier domain to be $\hat{\psi}(\omega) = (2\pi)^{-1/2}$ for $\omega \in (\pi, \pi/2)$ and zero elsewhere (see Daubechies 1992, page 115). The finest scale Littlewood-Paley wavelet coefficients correspond to an exact bandpass filtering in the frequency range $(\pi/2, \pi)$ (highest frequency band), the next finest scale to an exact bandpass filter at $(\pi/4, \pi/2)$ and so on. As one can see wavelets do not necessarily provide good frequency resolution at some scales: wavelet packets were introducted partly to correct this deficiency. For example, a resolution of, say, $(3\pi/4, \pi)$ could be achieved with a wavelet packet. Alternatively, algorithmically discrete wavelet packets could be formed by replacing $h_{n-2k}$ by $g_{n-2k}$ in (1) and, for example, producing a discrete wavelet packet such as $(1/2, -1/2, 1/2, -1/2)$ which is clearly not derived from the Haar mother wavelet by a simple scaling or translation as it contains two complete oscillations not one. In fact, wavelet packets can be indexed by scale, location and additionally *number of oscillations* and in a sense provide a well-spaced cover of functions spanning the time-frequency plane. *Non-decimated* wavelets or wavelet packets just means that the functions can be placed at *any* location. In the standard *decimated* transform wavelets/wavelet packets are restricted to lie at dyadic locations depending on their scale (we mentioned above that $\psi_{jk}$ was located at $2^{-j}k$). For non-decimated continuous wavelets, $\psi_{jk}$ is located at $2^{-J}k$, i.e. independent of $j$ (but dependent on the finest scale of the observed data points $J$). For discrete wavelets, non-decimation means just that $\psi_{jk}$ can lie at any time point $t$, i.e. $\psi_{jk}(t) = \psi_{j,k-t}$.

For the modelling described in Section 3.2 we compute the wavelet packet coefficients by filtering the time series in question with the appropriate wavelet packet function. In other words we are forming combinations of the explanatory time series such as:

$$W_t^{(1)} = (X_t - X_{t-1})/\sqrt{2}$$

for the finest scale Haar wavelet,

$$W_t^{(2)} = (X_t + X_{t-1} - X_{t-3} - X_{t-4})/2$$

for the next coarsest scale Haar wavelet, and

$$W_t^{(3)} = (X_t - X_{t-1} + X_{t-3} - X_{t-4})/2$$

for the wavelet packet mentioned above. This filtering is done recursively with an extension of the discrete wavelet transform algorithm (see Mallat, 1989), compared with a direct calculation of all the filtered series, very fast.

For further details on the computational algorithms associated with wavelets and wavelet packets we refer, for example, to Nason & Sapatinas (2001). For recent surveys on the use of wavelets (mainly) and wavelet packets in statistics, time series and related subjects we refer, for example, to Antoniadis (1997), Nason & von Sachs (1999), Vidakovic (1999), Abramovich *et al.* (2000) and Percival & Walden (2000).

**A note on classification with wavelets and wavelet packets:** Wavelet and wavelet packet methods have been recently used in classification problems following the standard "training-predicting" paradigm (see, for example, Coifman & Saito, 1994; Learned & Willsky, 1995). The next section, however, briefly describes the modelling methodology of Nason & Sapatinas (2001) that is of a somewhat different type as models are built *in situ* rather than have a large set of training samples.

### 3.2 STATISTICAL MODEL BUILDING USING NON-DECIMATED WAVELET PACKETS

The basic statistical modelling idea is very simple. Rather than build a statistical model directly between a response time series $Y_t = (Y_1, \ldots, Y_T)'$ and an explanatory time series $X_t = (X_1, \ldots, X_T)'$, for some fixed integer $T > 0$, Nason & Sapatinas (2001) proposed to build a model between $Y_t$ and a NWPT version of $X_t$. The NWPT representation of $X_t$ generates $K = 2T - 2$ derived time series (wavelet packets), each one having $T$ observations. We can subsequently model $Y_t$ in terms of the matrix $\mathbf{W} = (\mathbf{X}_1, \ldots, \mathbf{X}_T)'$, where $\mathbf{X}_i$ is $K$-dimensional and each dimension corresponds to a particular wavelet basis function. Each variable of $\mathbf{X}_t$ quantifies how similar $X_t$ is to a particular wavelet packet at time $t$. In other words each component of $\mathbf{X}_t$ tells us "how much" of each wavelet packet there is in $X_t$ at any particular time $t$. The decomposition of $X_t$ into $K$ different wavelet packet components is extremely useful since we can subsequently model $Y_t$ in terms of the components using standard statistical methodology. To summarize: $X_t$ is the "explanatory" time series and $\mathbf{X}_t$ is the "collection of NWPT coefficients" of $X_t$. We refer to Nason & Sapatinas (2001) for the modelling advantages of using *non-decimated* wavelet packets against (*decimated*) wavelets and/or wavelet packets, and for an S-Plus function from the freeware `WaveThresh` package that implements the NWPT.

The number of variables ($K = 2T - 2$) generated by the NWPT is always larger than the number of observations ($T$) and, hence, the problem of having more variables than observations arises (a problem as many standard statistical techniques require $K < T$). In our infant sleep state classification problem (see Section 4), however, we have tackled the problem of having more variables than observations by considering the following two strategies.

1. First, we consider what we call the "naive" method. This approach selects an arbitrary number of variables, say $K_1 < T$, which correlate best with $Y_t$. Although this step is somewhat rough and ready, it is computationally fast and provides an adequate initial dimension reduction. The variables that exhibit the largest $K_1$ correlations then form the *working set*. Then standard statistical techniques can be used to build a model between $Y_t$ and the working set variables.

2. Secondly, we consider the antedependence modelling method. The antedependence models were introduced by Gabriel (1962) as a nested series of models suitable for handling data that are serially correlated (and exhibit the general features of a non-stationary time series) and used by Kenward (1987) in the analysis of repeated measurements. A set of $p$ ordered variables is said to have an antedependence structure of order $r$ if the $i$th variable ($i > r$), given the preceding $r$, is independent of all further preceding variables. Complete independence ($r = 0$) and general dependence ($r = p - 1$) are special cases of this structure. Under the antedependence structure of order $r$, the inverse of the variance-covariance matrix has non-zero elements only on the leading diagonal and on the $r$ diagonals immediately above and immediately below it. The antedependence models were recently developed by Krzanowski *et al.* (1995) in the discriminant context to circumvent the problem of singular variance-covariance matrices (when the number of variables is larger than the number of observations) and successfully applied to spectroscopic data.

# 4 INFANT SLEEP-STATE CLASSIFICATION

This section applies the statistical modelling methodology explained in Section 3.2 to sleep state and heart rate recordings from an infant who was placed to bed at night at different stages of development. For computational reasons, we mainly used segments of length $T = 128$ but longer segments of $T = 512$ were subsequently used.

We start our analysis by considering the medical example described in Section 2.2 and shown in Figure 1. Recall that this example concerns sleep states ($Y_t$) and heart rates ($X_t$) for a four month old infant. We transform $X_t$ with the NWPT using Daubechies' (1992) extremal-phase mother wavelet with 10 vanishing moments and form the matrix $\mathbf{W} = (\mathbf{X}_1, \ldots, \mathbf{X}_T)'$. (There are no hard rules about the choice of the mother wavelet — however a choice has to be made.) The matrix consists of $T = 128$ observations on $K = 254$ variables. Since we have more variables than observations we first reduce the dimensionality of the variables enough so that we can subsequently use standard statistical classification techniques. We apply the "naive" approach mentioned in Section 3.2 to select a subset $K_1 = 13$ (the "best" 5%) of the $K = 254$ variables. The resulting "top five" variables were labelled by $S1, \ldots, S5$ and identified in Table 1 along with their correlations.

Table 1: Resolution levels and frequency indices of the "top five" (non-decimated) wavelet packets that were identified as being important for relating $Y_t$ to $X_t$. The "correlation" column shows the correlation between $Y_t$ and the particular wavelet packet coefficients.

| | *Wavelet packet* | | |
|---|---|---|---|
| *Packet ID* | *Resolution level $j$* | *Frequency index* | *Correlation with $Y_t$* |
| $S1$ | 4 | 0 | 0.92 |
| $S2$ | 5 | 0 | 0.89 |
| $S3$ | 3 | 0 | 0.89 |
| $S4$ | 6 | 0 | 0.89 |
| $S5$ | 1 | 0 | 0.79 |

Note that the best variables discovered by the "naive" variable selection strategy all have frequency index 0 (indeed, the next best, not shown here, also has frequency index 0 at scale 2). The wavelet packets at frequency index 0 are *father wavelets* which resemble statistical kernel functions. This can be easily seen in Figure 2 which shows the three father wavelets corresponding to resolution levels of 3 ($S3$), 4 ($S1$) and 5 ($S2$). The appearance of the father wavelets suggests that averaging over resolution levels 1 ($2^{(7-1)} = 64$ minutes, coarsest) to 6 ($2^{(7-6)} = 2$ minutes, finest) in the immediate past is important for determining sleep state. This corresponds with the observation earlier than the level of the heart rate over these scales is an important determining factor. The prospects for real-time prediction (*on-line*) are probably not as good as the father wavelets average a short time into the future as well — e.g. $S1$ requires about 6 minutes, $S2$ about 2.5 minutes and $S3$ about 11 minutes.

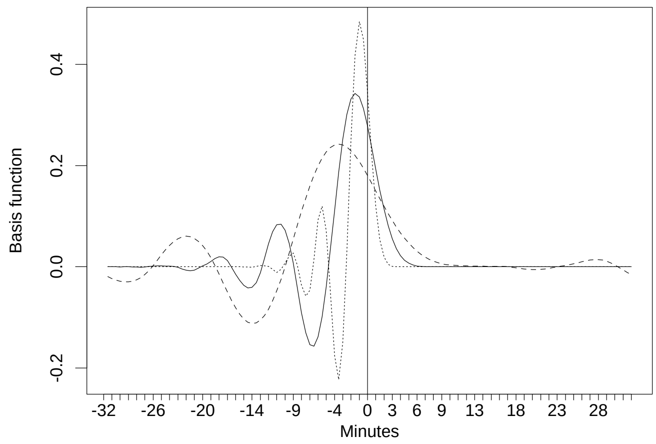


Figure 2: The top three wavelet packets for classifying the baby sleep state from heart rate. The figure shows Daubechies' (1992) extremal phase father wavelets with 10 vanishing at resolution levels 4 ($S1$, solid line), 5 ($S2$, dotted line), and 3 ($S3$, dashed line). The vertical line shows current time $t$. (Only three are shown for clarity.)

The top five variables ($S1, \ldots, S5$) could be then used as an input in various standard statistical techniques to identify which wavelet packets are useful for modelling the sleep state. In all we experimented with three statistical modelling methods: linear discriminant analysis (LDA), logistic regression, and classification and regression trees. Of these three methods LDA working on the log-transformed absolute values of coefficients was most successful and is described here. The log transform is of use when local oscillatory *power* is thought to be important in driving the response time series. Taking the log of the absolute values is like squaring and then taking logs which is analogous to forming the log-periodogram in classical stationary time series analysis. For instance, in our infant sleep state classification problem, it is the power of oscillation itself that is related to changes of sleep state. Using power-based statistics is a standard signal processing manoeuvre (see, for example, Learned & Willsky, 1995; Nason *et al.*, 2000) also advocate the use of power-based wavelet coefficients in local time-scale modelling.

The LDA analysis determines which linear combinations are best for discrimination. The best linear combination turned out to be

$$12\ \mathrm{S1} - 0.78\ \mathrm{S2} - 0.37\ \mathrm{S3} + 4.5\ \mathrm{S4} - 2.5\ \mathrm{S5}. \tag{2}$$

Thus, for this data set, $S1$ is very influential and corresponds to averaging over periods of about 10 minutes (looking at the solid curve in Figure 2). Interestingly enough, this period of oscillation was found in analyses carried out earlier by Stoffer (1991) to be present in infants unexposed to maternal alcohol. Although here we are saying that the 10 minute cycle is important for linking heart rate and sleep state. The analysis made by Stoffer (1991) identifies a 9 minute cycle in spectral analysis of just sleep state.

## 4.1 Prediction and evaluation

To exercise our model we took the next 128 heart rate values, performed the NWPT analysis, extracted the same top five variables and used the linear combinations determined by the LDA in (2) to predict the sleep state for the next 128 time periods (additionally, the mean of the next 128 heart rate values was adjusted to be the same as the previous 128 values to prevent this affecting the analysis as it adds no discriminatory value). Figure 3 shows the new NWPT values projected onto the first two discriminant axes, the location of the discriminant rule and 13 misclassified observations. With this classification we achieved a 90% overall success rate (13 observations misclassified the infant to be ASLEEP when it was really AWAKE). Figure 4 shows the new heart rate series with the true and predicted heart rate. Our method is fooled into thinking that the baby has gone to sleep just after 2.4 hours, probably by the sharp drop in heart rate. Likewise, just around 3.1 hours our method is a bit slow in noticing that the baby woke up, but the delay in noticing

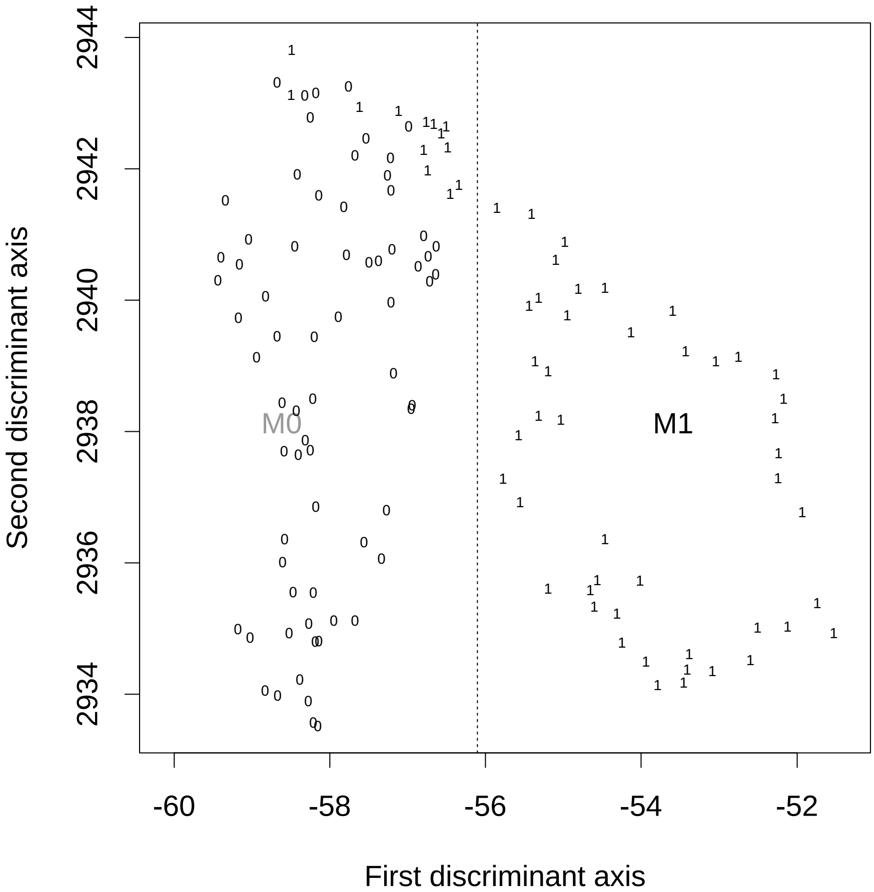


Figure 3: NWPT values from the new heart rate time series projected onto first and second linear discriminant axes. The label of each point shows its true group membership. The vertical dashed line shows the discriminant rule: observations to the left are assigned to the ASLEEP group (0), those to the right are assigned to the AWAKE group (1). The 13 misclassified observations appear in the top-left of the plot. The M0 and M1 labels refer to the means of the ASLEEP and AWAKE groups used to build the discriminant model.

is 2 minutes. However, the true record does note that during this period the human judge was uncertain about the true sleep state.

Furthermore, we evaluate the classification performance by building models at various time-intervals during a particular night and also on recordings at different stages of the infant's life. As infants mature their EEG and EOG become easier to classify and conventionally determined sleep state becomes more accurate with less disagreement between observers. This was reflected by our LDA models which became better at predicting. Table 2 shows success (interval) rates for the infant at different stages of development and suggests that better classification may be possible with the older infant.

The LDA and Figure 3 would appear to be "inappropriate" because of the highly correlated nature of the data. As an alternative, we now use the first ($r = 1$) and second ($r = 2$) order antedependence models from Krzanowski *et al.* (1995) mentioned in Section 3.2 to build discriminatory models for the infant sleep state. These models, obviously, result in an *off-line* classification since some of the wavelet packets average a short time into the future as well. Although more sophisticated, these models are computationally more demanding. Furthermore, they cannot be used directly to identify which wavelet packets are useful for modelling the sleep state and, therefore, we cannot

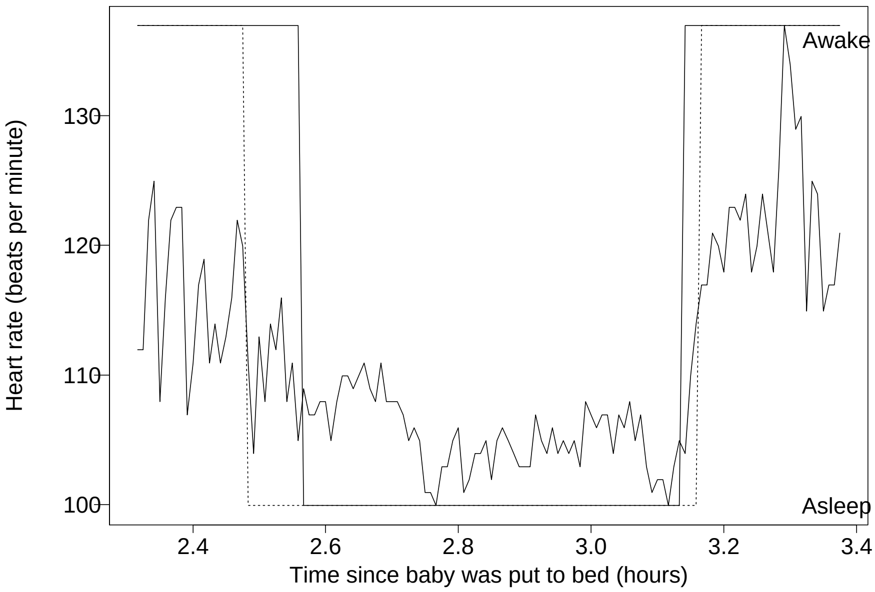


Figure 4: New heart rate series with true sleep state (solid) and predicted sleep state (dashed).

Table 2: Success (interval) rates of infant sleep state with increasing age with our method

| *Infant Age (Months)* | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| *Success Rate* | 75-78% | 75-80% | 80-90% | 82-90% |

easily attach physical and scientific interpretations to the selected models. However, these models, by construction, are suitable for data that are serially correlated and exhibit the general features of a non-stationary time series. Moreover, they serve as a basis for evaluating the success rates obtained when the "naive" variable selection was considered and LDA was subsequently used for infant sleep state classification.

To also evaluate the effect of the length $T$, we now consider longer segments of length $T = 512$. In this case, the NWPT representation results in $K = 1022$ variables. Table 3 shows the leave-one-out cross-validated success rates for the infant at different stages of development and suggests that this approach improves on the classification rates obtained earlier with the "naive" method. It also suggests that better classification may be possible with the older infant; a view that concurs with the one drawn using the "naive" method.

It could be argued that the leave-one-out cross-validation might not be very appropriate for evaluating the classification performance, since future values are also been used to predict the past. However, when the first and second order antedependence models were built on the segment shown in Figure 1 and their classification performance were subsequently evaluated on the time series given in Figure 3, similar conclusions were drawn. This is also true, although not reproduced here, when we evaluated the classification performance of the first and second antedependence models by building classifiers at various time-intervals and throughout different stages of the infant's life. Moreover, a limited comparison of one step ahead predicted with one step ahead observed values (which is more appropriate for time series data) produces similar results, although a more detailed analysis should be made to draw definite conclusions.

Table 3: Cross-validated classification success rates using antedependence models of order 1 (AD1) and 2 (AD2) for an infant at different ages. Rates show how accurately the categories of ASLEEP and AWAKE were classified as well as the overall classification success rate. Cross-validation was performed with the leave-one-out method of Lachenbruch & Mickey (1968).

| *Infant Age (Months)* | *Model* | *Success rate* | | |
|---|---|---|---|---|
| | | ASLEEP | AWAKE | *Overall* |
| 2 | AD1 | 0.89 | 0.90 | 0.90 |
| | AD2 | 0.89 | 0.89 | 0.89 |
| 3 | AD1 | 0.95 | 0.88 | 0.94 |
| | AD2 | 0.96 | 0.88 | 0.95 |
| 4 | AD1 | 0.94 | 0.86 | 0.89 |
| | AD2 | 0.95 | 0.89 | 0.91 |
| 5 | AD1 | 0.97 | 0.95 | 0.96 |
| | AD2 | 0.97 | 0.95 | 0.96 |

# 5 DISCUSSION

This article demonstrates how the recently developed modelling methodology of Nason & Sapatinas (2001) could be useful for infant sleep state classification using the non-decimated wavelet packet transform of the heart rate. Although we have had some success in classifying sleep state by building a model in various parts of a night and predicting what the sleep state is in later periods it may not be possible to transfer the exact model to the same infant at different ages. However, the same father wavelets nearly always recur in the best model suggesting that averaging over certain time-scales is important. Only the coefficients in the sleep state/heart rate model differ across nights. We were a little surprised that more complex wavelets did not seem important to the sleep state classification.

In our analysis we have, of course, concentrated on two sleeping states (ASLEEP, AWAKE). As discussed in Section 2, in the literature attention has focussed on further subdividing ASLEEP into ACTIVE SLEEP and QUITE SLEEP during which dreaming and many upper airway breathing disorders occur (see DeHann *et al.*, 1977; Schechtman *et al.*, 1988). For example, Harper *et al.* (1987) modified the technique of Welch & Richardson (1973) and developed an *off-line* system based on cardiac (4 variables from heart rate) and respiratory (3 variables from respiration) measures. They quote success rates of 85% using all 7 variables, 82% for the 4 cardiac variables, and 80% for the 3 respiratory variables.

We stress, however, that a more detailed sleep state categorization could be undertaken, if necessary, when the non-decimated wavelet packet transform modelling methodology is adopted. This is only, of course, a matter of choice of the statistical classification technique we select and *not* because we are using wavelet methods. Preliminary analysis, however, has shown that with more than two sleep states both the “naive” method and the antedependence modelling methodology do not seem to produce satisfactory results. Therefore, more appropriate statistical analyses are needed once the non-decimated wavelet packet transform representation of the heart rate has been built. It might be the case that these approaches could produce satisfactory classification rates when a more detailed infant sleep state categorization is considered, and also transfer the exact model to different infants and at different stages of their lives. Further, a referee has pointed out that measuring success in terms of *mean time to detect a state change* might be a more appropriate measure of success. We agree that this quantity might be more important in this, and other situations.

We finally mention that overnight infant *rectal temperature* falls rapidly around bed-time, reaches a trough and then gradually rise in the early morning. This fall in rectal temperature becomes greater with increasing age (see Lodemore *et al.*, 1991); it has recently become apparent that there are clear variations in infant rectal temperature with sleep state (see Tappin *et al.*, 1996). There is also evidence that disordered thermoregulation (see Sawczenko & Fleming, 1996) may be responsible for some cases of the *Sudden Infant Death Syndrome* (SIDS), which has a peak incidence around 3 months. Two thirds of SIDS now appear to occur during the early morning i.e. during the rise in rectal temperature (see, Fleming *et al.*, 1996). Whilst hypothermia or hyperthermia can directly kill it is more likely thermal stress causes instability in the control of the respiratory system (see Fleming *et al.*, 1993). A unifying 'triple risk' model suggests that SIDS occurs during a critical developmental period, in vulnerable infants, exposed to exogenous stressors. In a longitudinal series of home studies of rectal temperature infants differed in the age at which rectal temperature fell; infants at the highest epidemiological risk of SIDS dropped their rectal temperature at an older age (see Lodemore *et al.*, 1992). Other studies have suggested differences in cardiovascular recordings in infants who later die from SIDS compared to controls (see Kludge *et al.*, 1988; Schechtman *et al.*, 1992). This work supports the suggestion that environmental factors, probably including parental actions, may adversely affect baseline infant physiology. There is thus a need to longitudinally record, in the home, the parallel and interrelated developmental patterns of several physiological systems during early infancy.

Obviously, analysis of the above scientific factors could be proved useful to the sleep state classification. This problem is an interesting topic for further research and we intend to address it in the future.

## ACKNOWLEDGEMENTS

Guy Nason and Theofanis Sapatinas were supported in part by EPSRC grant GR/K70236 and Nason was supported by EPSRC grants GR/M10229, AF/001664 and DERA grant CU016 0000007644. Andrew Sawczenko was supported by grant 186 from the Foundation for the Study of Infant Deaths, Shield Nationwide Ltd, The United Bristol Healthcare NHS Trust Medical Research Fund and Cot Death Research.

The authors would like to thank Peter Fleming, Jeanine Young and K Pollard of the Institute of Child Health, The Royal Hospital for Sick Children, Bristol for supplying data. Many thanks are due to Wojtek Krzanowski, School of Mathematical Sciences, Exeter University, who kindly provided software for computing and evaluating antedependence models. We would like to thank Bernard Silverman and Peter Green for reading an earlier version of this manuscript and for providing helpful comments. Helpful comments of three anonymous referees are gratefully acknowledged.

# References

[1] Abramovich, F., Bailey, T.C. & Sapatinas, T. (2000). Wavelet analysis and its statistical applications. *The Statistician*, **49**, 1–29.

[2] Anders, T.F., Emde, R.N. & Parmelee, A.H. (eds). (1971). *A Manual of Standardized Terminology, Techniques and Criteria for Scoring of States of Sleep and Wakefullness in Newborn Infants*. Los Angeles: UCLA Brain Information Service.

[3] Antoniadis, A. (1997). Wavelets in statistics: a review (with discussion). *J. Ital. Statist. Soc.*, **6**, 97–144.

[4] Coifman, R.R. & Saito, N. (1994). Constructions of local orthonormal bases for classification and regression. *Compt. Rend. Acad. Sci. Paris*, Series A, **319**, 191–196.

[5] Daubechies, I. (1992). *Ten Lectures on Wavelets.* Philadelphia: SIAM.

[6] DeHann, R.J., Patrick, Chess, G. & Jaco, N. (1977). Definition of sleep state in the newborn infant by heart rate analysis. *Am. J. Obstet. Gynecol.*, **127**, 753–758.

[7] Dement, W. & Kleitman, N. (1957). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalogr. Clin. Neurophysiol.*, **9**, 673–690.

[8] Fleming, P., Blair, P., Bacon, C., Bensley, D., Smith, I., Taylor, E., Berry, J., Golding, J. & Tripp, J. (1996). Environment of infants during sleep and the risk of the Sudden Infant Death Syndrome: results of 1993-1995 case-control study for the confidential inquiry into stillbirths and deaths in infancy. *Br. Med. J.*, **313**, 191–195.

[9] Fleming, P.J., Levine, M.R., Azaz, Y., Wigfield, R. & Stewart, A.J. (1993). Interactions between thermoregulation and the control of respiration in infants: possible relationship to sudden infant death. *Acta Paediatr.*, **82**, Suppl(389), 57–59.

[10] Gabriel, K.R. (1962). Antedependence analysis of a set of ordered variables. *Ann. Math. Statist.*, **33**, 201–212.

[11] Harper, R.M., Schechtman, V.L. & Kluge, K.A. (1987). Machine classification of infant sleep state using cardiorespiratory measures. *Electroencephalogr. Clin. Neurophysiol.*, **67**, 379–387.

[12] Harper, R.M., Hoppenbrouwers, T., Sterman, M.B., McGinty, D.J. & Hodgman, J. (1976). Polygraphic studies of normal infants during the first six months of life. 1. Heart rate and variability as a function of state. *Pediatr. Res.*, 10, 945–51.

[13] Hess-Nielsen, N. & Wickerhauser, M.V. (1996). Wavelets and time-frequency analysis. *Proc. IEEE*, **84**, 523–540.

[14] Kenward, M.G. (1987). A method for comparing profiles of repeated measurements. *Appl. Statist.*, **36**, 296–308.

[15] Kludge, K.A., Harper, R.M., Schechtman, V.L., Wilson, A.J., Hoffman, H.J. & Southall, D.P. (1988). Spectral analysis assessment of respiratory sinus arrhythmia in normal infants and infants who subsequently died of Sudden Infant death Syndrome. *Pediatr. Res.*, **24**, 677–682.

[16] Krzanowski, W.J., Jonathan, P., McCarthy, W.V. & Thomas, M.R. (1995). Discriminant analysis with singular covariance matrices: methods and application to spectroscopic data. *Appl. Statist.*, **44**, 101–115.

[17] Lachenbruch, P.A. & Mickey, M.R. (1968). Estimation of error rates in discriminant analysis. *Technometrics*, **10**, 1–11.

[18] Learned, R.E. & Willsky, A.S. (1995). A wavelet packet approach to transient signal classification. *Appl. Comput. Harm. Anal.*, **2**, 265–278.

[19] Lodemore, M., Petersen, S.A. & Wailoo, M.P. (1991). Development of night time temperature rhythms over the first six months of life. *Arch. Dis. Child.*, **66**, 521–524.

[20] Lodemore, M.R., Petersen, S.A. & Wailoo, M.P. (1992). Factors affecting the development of night time temperature rhythms. *Arch. Dis. Child.*, *67*, 1259–1261.

[21] Mallat, S.G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattn Anal. Mach. Intell.*, **11**, 674–693.

[22] Nason, G.P. & Sapatinas, T. (2001). Wavelet packet transfer function modelling of nonstationary time series. *Statist. Comput.*, **11** (to appear).

[23] Nason, G.P. & von Sachs, R. (1999). Wavelets in time series analysis. *Phil. Trans. R. Soc. Lond.*, Series A, **357**, 2511–2526.

[24] Nason, G.P., von Sachs, R. & Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. Roy. Statist. Soc.*, Series B, **62**, 271–292.

[25] Percival, D.B. & Walden, A.T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press.

[26] Schechtman, V.L., Harper, R.K. & Harper, R.M. (1993). Analysis of beat-to-beat heart rate changes during sleep-waking states in normal infants. *J. Dev. Physiol.*, **19**, 263–271.

[27] Sawczenko, A. & Fleming, P. (1996). Thermal stress, sleeping position and the Sudden Infant Death Syndrome. *Sleep*, **19**, S267–S270.

[28] Sawczenko, A., Galland, B., Young, J. & Fleming, P. (1995). Night time mother-infant interactive behaviour and physiology: a longitudinal comparison of room sharing versus bed sharing (‘co-sleeping’). *Pediatr. Pulmonol.*, **20**, 341.

[29] Schechtman, V.L., Kluge, K.L. & Harper, R.M. (1988). Time-domain system for assessing variation in heart-rate. *Med. Biol. Eng. Comput.*, **26**, 367–373.

[30] Schechtman, V.L., Raetz, S.L., Harper, R.K., Garfinkel, A., Wilson, A.J., Southall, D.P. & Harper, R.M. (1992). Dynamic Analysis of Cardiac R-R Intervals in Normal Infants and in Infants who subsequently succumbed to the Sudden Infant Death Syndrome. *Pediatr. Res.*, **31**, 606–612.

[31] Stoffer, D.S. (1991). Walsh-Fourier analysis and its statistical applications (with comments). *J. Am. Statist. Ass.*, **86**, 461–485.

[32] Tappin, D., Ford, R., Nelson, K., Price, B., Macey, P., Dove, R., Larkin, J. & Slade, B. (1996). Breathing, sleep state, and rectal temperature oscillations. *Arch. Dis. Child.*, **74**, 427–431.

[33] Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. New York: John Wiley & Sons.

[34] Welch, A. & Richardson, P. (1973). Computer sleep stage classification using heart rate data. *Electroencephalogr. Clin. Neurophysiol.*, **34**, 145-152.