

RESEARCH

Open Access



# Enhanced automatic sleep spindle detection: a sliding window-based wavelet analysis and comparison using a proposal assessment method

Xiaobin Zhuang<sup>1,2</sup>, Yuanqing Li<sup>1,2\*</sup> and Nengneng Peng<sup>1,2</sup>

\*Correspondence:  
auyqli@scut.edu.cn

<sup>1</sup> Center for Brain  
Computer Interfaces  
and Brain Information  
Processing, South China  
University of Technology,  
Guangzhou 510640, China  
Full list of author information  
is available at the end of the  
article

## Abstract

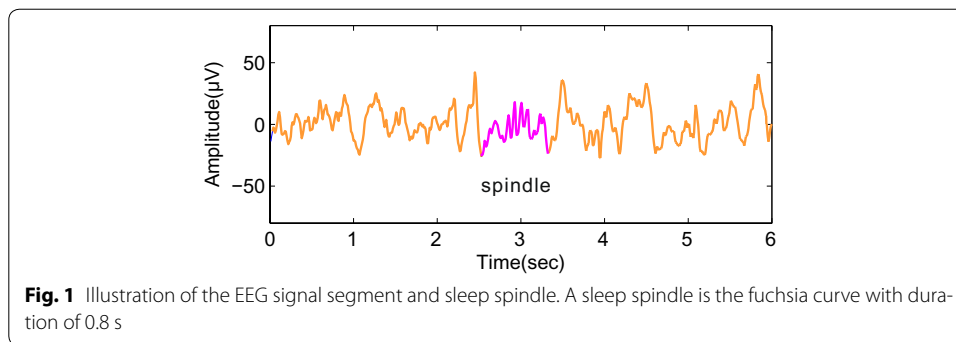
Sleep spindles are thought to be related to some sleep diseases and play an important role in memory consolidation. They were traditionally identified by physiology experts based on rules and recently detected by automatic algorithms. However, many automatic approaches were validated on the different electroencephalogram (EEG) using various assessment methods, making it difficult to appraise a method objectively and fairly. In this paper, we proposed a sliding window-based probability estimation (SWPE) method for sleep spindle detection. We performed a continuous wavelet transform with Mexican hat wavelet function, following by a sliding window to find out the candidate spindle points corresponding to the large wavelet coefficients at the frequencies of spindles and estimated their probabilities. To enhance the results, we used the envelope of the rectified signal to reject some false sleep spindle candidates. This was an enhanced method and we called it SWPE-E in this paper. Finally, we compared our approaches with four approaches on the same public available EEG database, and the result showed the significative improvement of our proposed approaches.

**Keywords:** Sleep spindle, Wavelet transform, Automatic detection, EEG

## Background

Sleep spindles are characteristic electroencephalogram (EEG) waves mainly observed during stage 2 non-rapid eye movement (NREM), but occasionally appearing in sleep stages 3 and 4 as well, as shown in Fig. 1. The sleep spindle is one of the few transient EEG events and is composed of a group of rhythmic waves with progressively increasing amplitude followed by a gradual decrease (De Gennaro and Ferrara 2003). Strictly speaking, a sleep spindle is a train of distinct waves with a frequency of 11–16 Hz (most commonly 12–14 Hz) (Berry et al. 2012). Interestingly, the characteristics of sleep spindles, such as density, amplitude, or duration, vary substantially between individuals but are reasonably stable for each individual (Warby et al. 2014; Tsanas and Clifford 2015).

Sleep spindles play an important role in clinical research. Sleep spindles are believed to be associated with synaptic plasticity, memory consolidation, and long-term storage of memory representations (Warby et al. 2014). In clinical practice, several disorders, such



as schizophrenia, autism, and sleep disorders, have been found to accompany a change in spindle density (Diekelmann and Born 2010). Thus, spindle detection is significant and may provide further guidance for clinical practice.

Traditionally, sleep spindles are identified by physiology experts, who rely on their experience to visually identify them. This approach is a very time-consuming and difficult task; furthermore, different experts agree on only approximately 70% of the identifications of sleep spindles (Wendt et al. 2012). In recent years, many researchers have proposed different approaches for detecting spindles. Generally, previous researchers either considered the raw signal as isolated sample points or epochs. Schönwald et al. (2003) applied a Gabor transforms to analyze EEG signals. Duman et al. (2009) introduced a technology combining Short-Time Fourier Transform, Multiple Signal Classification, and the Teager Energy Operator to locate spindles. In addition, Wendt et al. (2012) processed the signals from two channels with two equiripple low-pass filters and obtained individual envelopes. Though these approaches achieved a high level of performance, they only work on NREM data or require signals from multi-channels, which is not flexible enough for practical application. Pattern recognition methods, such as neural network (Shimada et al. 2000) and Bayesian classification (Babadi et al. 2012), were also applied to develop automatic sleep spindle detection methods (Kabir et al. 2015). However, the downside to these approaches is the required use of training data to generate a classifier for the analysis of subjects. Tsanas and Clifford (2015) developed two similar approaches based on the Morlet wavelet transforms and combined spindles with appropriate duration. However, to some extent, dividing the signal into fixed-length epochs or focusing on a sample point in isolation is not suitable for spindle detection.

Based on this point of view, we employed a sliding window-based probability estimation (SWPE) method for sleep spindle detection, aiming to improve the performance of the spindle detection method. In our work, sample points were estimated in relation to their neighboring points. Moreover, we presented a novel assessment method to evaluate and compare the performance of the different approaches for detecting sleep spindles.

## Materials and methods

In this section, we summarize the database we used in this work and then introduce our approach in two steps. Step 3 was the enhancement of our approach, using the envelope of the rectified signal. A suggested assessment method is also discussed at the end of this section.

## Materials

The database in this work was obtained from the DREAM project (Devuyst et al. 2011). The full database consisted of whole-night polysomnogram (PSG) recordings collected from eight subjects with different pathologies. PSG recordings usually included several bioelectricity signals but only the EEG signals were analyzed in our study. These recordings were sampled at different frequencies, namely 50, 100, or 200 Hz. As described, a segment of 30 min from the central EEG channel of each subject was extracted for spindles scoring, rather than analyzing the recording from the whole night. Two experts independently annotated the onset of sleep spindles in an effort order to make the annotations more reliable. However, one expert scored eight segments and provided the exact duration of spindles, while the other expert scored only six segments and set all of the durations to 1 second. The database can be accessed from <http://www.tcts.fpms.ac.be/devuyst/Databases/DatabaseSpindles>.

## Methods

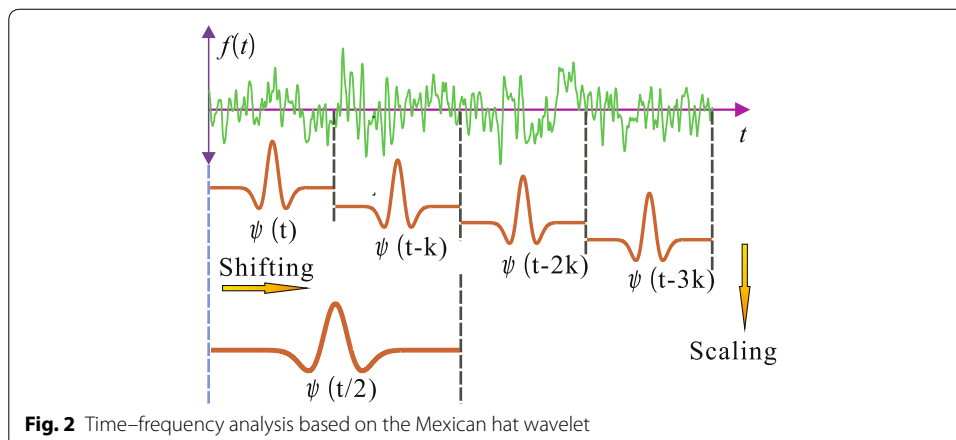
### Step 1: Continuous wavelet transform with Mexican hat wavelet function

A signal is defined as non-stationary if its statistics (including mean, variance, and higher-order statistics) changed with time, so the EEG signals used in this study are non-stationary according to this definition. Many classical frequency analysis methods, such as the Fourier transform, will fail to capture the dynamics of the underlying events and are not appropriate for sleep spindle detection. Short-Time Fourier Transform (STFT) can be applied for non-stationary signal processing, but it uses an equal sized window for all ranges of frequency. This is a weakness when analyzing signals in multi-resolution. In this paper, the Mexican hat wavelet transform was used to perform a time–frequency analysis, as shown in Fig. 2.

Mathematically, the Mexican hat wavelet is defined as the negative normalized second derivative of a Gaussian function,

$$\psi(t) = -c(e^{-t^2/2})'' = c(1 - t^2)e^{-t^2/2} \quad (1)$$

where  $c$  is a constant for normalization and  $c = 2\pi^{1/4}/\sqrt{3}$ .



Before analyzing the time–frequency characteristics, we calculated the power spectral density of the raw sleep EEG signals. The results showed that the amplitude of the signals generally decreased as the frequency increased, roughly fitting a negative power equation. Moreover, the most concentrated frequency band of the sleep EEG signal was below 10 Hz. However, it has been introduced that the frequency band of a sleep spindle is predominantly between frequencies of 11–16 Hz. Thus, the spindle would be submerged by the low-frequency components, making spindle detection difficult. Based on this point of view, we focused on frequencies from 8–25 Hz with a resolution of 0.2 Hz in this paper. As a result, the signals were analyzed at 86 adjacent scales and a wavelet coefficient matrix with 86 rows was ultimately obtained.

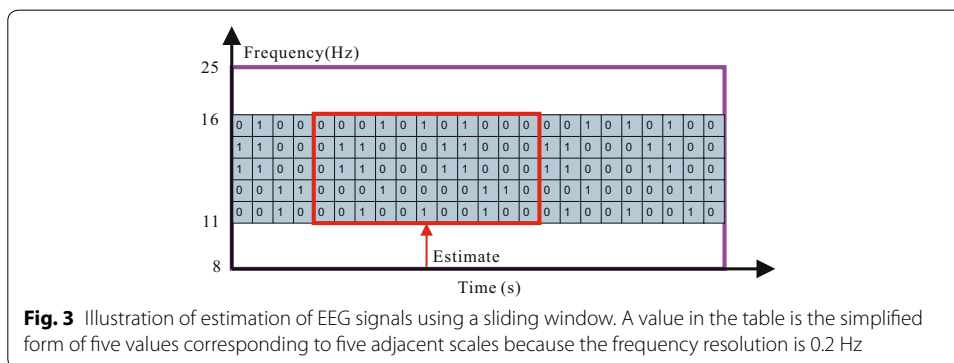
### Step 2: Sleep spindles identification with a sliding window

After obtaining the wavelet coefficient matrix, we employed a sliding window-based method to estimate the sample points. To find out the candidate spindle points corresponding to large wavelet coefficients at the frequencies, we sorted the wavelet coefficients of each sample point in descending order and set the top 10% of the coefficients to 1 and the rest to 0. Here, 10% is an empirical value in our study.

A sliding window was then placed on the matrix, as shown by the red rectangle in Fig. 3. We denoted the width and height of the sliding window by *row* and *col*, separately. For a signal sampled at a frequency of *f*, the *row* will be  $0.1 \times f$ , while the *col* will be a constant value of 25 (representing a frequency band from 11 to 16 Hz with a resolution of 0.2 Hz). For instance, for signals sampled at 100 Hz, the width of the sliding window will be 10 ( $10 = 0.1 \times 100$ ) and the height of the sliding window will be a constant value of 25. Additionally, a sample point was set to 1 only when the sum of its binary coefficients in the sliding window was in the top 10%. As a result, 90% of sample points will be set to 0 according to this rule. The raw EEG signal was represented by a binary signal.

In our study, the binary signal was used to generate sleep spindle candidates. For a binary signal *b* with a length of *N*, the probability that the sample point belongs to a spindle is estimated by

$$p(b_i) = \sum_{k=i-nf}^{i+nf} b_k \quad i \in [1 + nf, N - nf] \quad (2)$$



where  $b_i$  is a sample point of a binary signal  $b$ ,  $f$  is the frequency of the raw signal,  $N$  is the length of the binary signal, and  $n$  is an empirical value equated to 0.1 in this paper.

Then, we used a threshold of 0.5 to determine sleep spindle points. A sleep spindle point can be regarded as a sample point from the whole sleep spindle in this paper.

$$b_i = \begin{cases} \text{spindle point} & p(b_i) > 0.5 \\ \text{non-spindle point} & p(b_i) \leq 0.5 \end{cases} \quad (3)$$

Finally, the successive sleep spindle points were combined to generate the spindle candidates, with the constraint condition that the length of a spindle is between 0.4 and 1.6 s.

### **Step 3: Enhancement with the envelope of the rectified signal**

After generating the spindle candidates, we used the envelope of the rectified signal to remove the false spindle candidates. First, we used a filter with a pass band between 11 and 16 Hz (the frequency band of sleep spindles) to obtain a rectified signal  $y(n)$ , which can be expressed as follows:

$$y(n) = x(n) * h_{11-16}(n) \quad n = 1, \dots, N \quad (4)$$

where  $x(n)$  is the raw EEG signal and  $h_{11-16}(n)$  is the impulse response of the filter. Then, we used a low-pass filter with a cutoff frequency of 2 Hz to obtain the envelope of the rectified signal, which is expressed as follows:

$$e(n) = |y(n)| * h_{\text{lowpass}}(n) \quad n = 1, \dots, N \quad (5)$$

where  $e(n)$  is the envelope of the rectified signal and  $h_{\text{lowpass}}$  is the impulse response of the low-pass filter mentioned above.

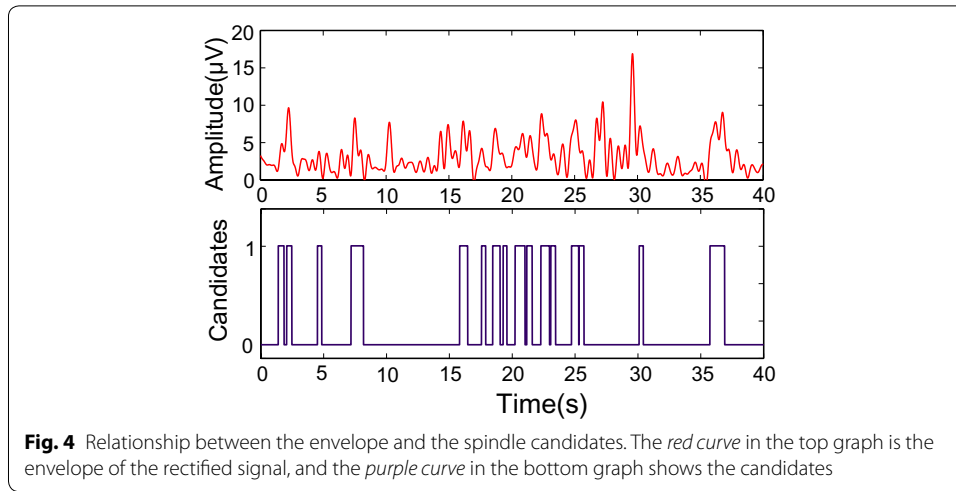
The relationship between the envelope and the spindle candidates is illustrated in Fig. 4. It shows that most of the candidates are located at positions corresponding to the local maximums of the envelope. Therefore, we applied (6) to estimate the reliability of each sleep spindle candidate  $c_i$  and removed 10% of candidates with the lowest reliabilities.

$$r(c_i) = \frac{1}{d_i} \sum_{k=o_i}^{o_i+d_i} e_k \quad i = 1, \dots, C \quad (6)$$

where  $C$  is the number of candidates,  $c_i$  represents a single candidate,  $d_i$  and  $o_i$  are the duration and the onset of  $c_i$ , respectively.

### **A suggested assessment method for evaluating the detection method**

As discussed above, the first expert defined both the onsets and exact durations of the eight recordings, while the second expert only gave the onsets of the six recordings. Additionally, these annotations are not consistent, which caused researchers to have divergent views. For example, some researchers only considered the annotations from the first expert. Moreover, the criteria to evaluate the detection capabilities were quite different. Only considering the absolute difference between the detected spindle onset and the spindle onset determined by the experts (Tsanas and Clifford 2015) or only considering the overlap of the spindles (Causa et al. 2010) are both inappropriate in fact. In



this study, we suggested comparing the center offset between the detected spindle and the spindle determined by the experts. We regarded the detected spindle as correct if the absolute difference between these two centers is  $<0.5$  s.

To evaluate the performance of an algorithm, the precision (also known as positive predictive value) and the recall (also known as sensitivity) are usually used. However, the sensitivity–specificity metric was commonly used in the previous researches, so we cautiously include them in this paper. The metrics involved in our study are defined as follows:

1. Sensitivity: This metric is also known as recall, and it measures the proportion of sleep spindles correctly identified by an automatic method. We want the sensitivity (or recall) to be high or even equal to 100%:

$$\text{sensitivity} = \frac{TP}{TP+FN} . \quad (7)$$

2. Specificity: This metric measures the proportion of non-spindles correctly identified by an automatic method. We want the specificity to be high or even equal to 100%:

$$\text{specificity} = \frac{TN}{TN+FP} . \quad (8)$$

3. False Discovery Rate (FDR): This metric measures the proportion of wrong detection results. Ideally, we want FDR to be 0:

$$\text{FDR} = \frac{FP}{TN+FP} \quad (9)$$

$$\text{precision} = \frac{TP}{TN+FP} = 1 - \text{FDR} . \quad (10)$$

4. F1-score: The F1-score can be used as a single measure of the performance of an algorithm. It is the harmonic mean of precision and recall that introduced to measure the overall performance of an approach:

$$\text{score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} . \quad (11)$$

## Results and discussion

In our study, we compared our approach with four different approaches denoted by A1, A2, A3, and A4. A1 was proposed by Ferrarelli et al. (2007), A2 was proposed by Wendt et al. (2012), and A3 and A4 were successively proposed by Tsanas and Clifford (2015). The results are shown in Tables 1, 2, 3, 4, and 5, with the best performing approach for each case highlighted in italic.

When comparing these six methods, we can see that none of these methods has absolute superiority in sensitivity, specificity, and FDR. Although the forth approach (A4) performed best in relation to sensitivity, its specificity was much lower than our proposed approach SWPE. The first approach (A1) has a high value on specificity but its performance in sensitivity is very poor. Among these methods, our approach SWPE has distinct advantages in FDR.

The F1-score is a harmonic mean of precision and recall, which is usually applied as a single measure of a detector, especially in retrieval work. Therefore, the F1-score was

**Table 1 Sensitivity (%) of the six methods**

Methods	S1	S2	S3	S4	S5	S6	S7	S8	Mean
A1	6.59	2.68	6.12	2.90	15.49	20.75	22.22	6.25	10.38
A2	68.26	56.25	71.43	52.17	60.56	67.92	77.78	75.00	66.17
A3	69.46	58.04	67.35	65.22	51.41	50.31	88.89	83.33	66.75
A4	76.65	64.29	75.51	75.36	71.83	67.92	94.44	95.83	77.73
SWPE	45.51	46.43	67.35	46.38	46.48	48.43	61.11	56.25	52.24
SWPE-E	44.15	45.54	65.30	45.25	45.85	47.01	60.23	54.48	50.98
Mean	51.77	45.54	58.84	47.88	48.60	50.39	67.45	61.86	54.04

**Table 2 Specificity (%) of the six methods**

Methods	S1	S2	S3	S4	S5	S6	S7	S8	Mean
A1	99.82	99.88	99.89	99.42	99.16	99.39	98.88	99.32	99.47
A2	87.32	81.16	88.35	87.00	92.46	92.93	80.64	82.76	86.58
A3	95.04	93.84	94.75	87.98	95.66	97.62	94.16	79.57	92.33
A4	80.07	77.61	78.81	70.36	81.48	87.14	76.60	58.68	76.34
SWPE	98.01	98.52	99.09	97.75	98.49	98.72	98.21	96.58	98.17
SWPE-E	99.50	99.52	99.54	98.52	99.21	98.94	99.24	97.57	99.01
Mean	93.29	91.76	93.41	90.17	94.41	95.79	91.29	85.75	91.98

**Table 3 FDR (%) of the six methods**

Methods	S1	S2	S3	S4	S5	S6	S7	S8	Mean
A1	21.43	40.00	40.00	83.33	38.89	23.26	83.33	80.00	51.28
A2	64.49	83.46	85.36	86.21	59.24	51.79	96.10	89.35	77.00
A3	41.12	61.54	73.60	82.21	49.66	32.77	86.67	89.95	64.69
A4	71.74	84.00	90.93	90.80	75.06	66.14	96.08	94.03	83.60
SWPE	19.15	32.47	32.65	54.93	27.47	21.43	74.42	68.97	41.44
SWPE-E	12.23	20.56	20.45	41.52	18.44	16.25	60.25	54.25	30.49
Mean	38.36	53.67	57.17	73.17	44.79	35.27	82.81	79.43	58.08

**Table 4 Precision (%) of the six methods**

Methods	S1	S2	S3	S4	S5	S6	S7	S8	Mean
A1	78.57	60.00	60.00	16.67	61.11	76.74	16.67	20.00	48.72
A2	35.51	16.54	14.64	13.79	40.76	48.21	3.90	10.65	23.00
A3	58.88	38.46	26.40	17.79	50.34	67.23	13.33	10.05	35.31
A4	28.26	16.00	9.07	9.20	24.94	33.86	3.92	5.97	16.40
SWPE	80.85	67.53	67.35	45.07	72.53	78.57	25.58	31.03	58.56
SWPE-E	87.77	79.44	79.55	58.48	81.56	83.75	39.75	45.75	69.51
Mean	61.64	46.33	42.84	26.83	55.21	64.73	17.19	20.58	41.92

**Table 5 F1-score (%) of the six methods**

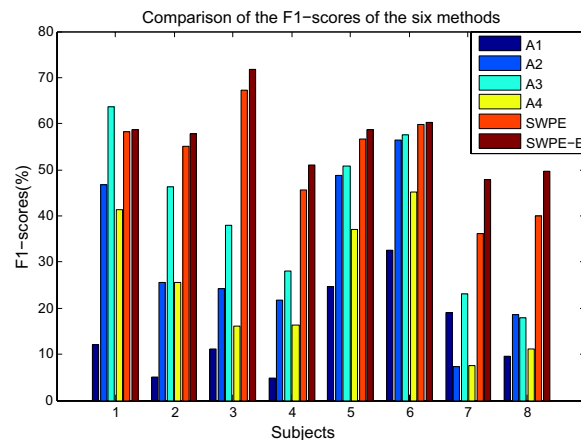
A1	12.16	5.13	11.11	4.94	24.72	32.67	19.05	9.52	17.11
A2	46.72	25.56	24.30	21.81	48.73	56.39	7.43	18.65	34.14
A3	63.73	46.26	37.93	27.95	50.87	57.55	23.18	17.94	46.19
A4	41.29	25.62	16.19	16.40	37.02	45.19	7.53	11.24	27.09
SWPE	58.24	55.03	67.35	45.72	56.65	59.92	36.06	40.00	55.22
SWPE-E	58.75	57.89	71.72	51.02	58.70	60.22	47.89	49.73	58.82
Mean	56.28	45.93	49.58	34.39	51.69	56.67	27.40	30.88	47.21

employed to measure the overall performance of an algorithm in this paper. We can see that the F1-score of our approach, SWPE-E, was significantly higher than the other approaches, as shown in Fig. 5.

When comparing the approach SWPE and the enhanced method SWPE-E, we can see that SWPE-E has a lower FDR because it rejected some false candidates.

## Conclusions

This paper proposes a novel approach for automatic detection of sleep spindles and introduced a method for comparing the performance of different detection methods. Our approach relied on only one single EEG channel and does not need the specified NREM stage signals for detection. We performed a continuous wavelet transform

**Fig. 5** Comparison of the F1-scores of the six methods



following by a sliding window to find out the candidate spindle points corresponding to the large wavelet coefficients at the frequencies of spindles. We estimated their probabilities with neighboring sample points. Finally, we compared our detection method with four other approaches using a novel assessment method based on the center offset, and the results showed that our detection method performed significantly better than the other methods.

#### Authors' contributions

XZ, NP and YL participated in the design of the study, performed in the data analysis, and drafted the manuscript. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup> Center for Brain Computer Interfaces and Brain Information Processing, South China University of Technology, Guangzhou 510640, China. <sup>2</sup> Guangzhou Key Laboratory of Brain Computer Interaction and Applications, South China University of Technology, Guangzhou 510640, China.

#### Acknowledgements

This work was supported by the National Key Basic Research Program of China (973 Program) under Grant 2015CB351703, the National Natural Science Foundation of China under Grants 61633010, 91420302, and 61573150, and Guangdong Natural Science Foundation under Grant 2014A030312005.

#### Competing interests

The authors declare that they have no competing interests.

Received: 27 September 2016 Accepted: 22 November 2016

Published online: 08 December 2016

#### References

- Babadi B, McKinney S, Tarokh V, Ellenbogen J (2012) A data-driven Bayesian algorithm for sleep spindle detection. In: SLEEP, vol. 35, pp. 134–134 (2012). AMER ACAD SLEEP MEDICINE ONE WESTBROOK CORPORATE CTR, STE 920, WESTCHESTER, IL 60154 USA
- Berry RB, Brooks R, Gamaldo CE, Harding SM, Marcus C, Vaughn B (2012) The AASM manual for the scoring of sleep and associated events: rules terminology and technical specifications. American Academy of Sleep Medicine, Darien
- Causa L, Held CM, Causa J, Estévez PA, Perez CA, Chamorro R, Garrido M, Algarín C, Peirano P (2010) Automated sleep-spindle detection in healthy children polysomnograms. *IEEE Trans Biomed Eng* 57(9):2135–2146
- De Gennaro L, Ferrara M (2003) Sleep spindles: an overview. *Sleep Med Rev* 7(5):423–440
- Devuyst S, Dutoit T, Stenuit P, Kerkhofs M (2011) Automatic sleep spindles detection-overview and development of a standard proposal assessment method. In: 2011 Annual international conference of the IEEE engineering in medicine and biology society, pp. 1713–1716. IEEE
- Diekelmann S, Born J (2010) The memory function of sleep. *Nat Rev Neurosci* 11(2):114–126
- Duman F, Erdamar A, Eroglu O, Telatar Z, Yetkin S (2009) Efficient sleep spindle detection algorithm with decision tree. *Expert Syst Appl* 36(6):9980–9985
- Ferrarelli F, Huber R, Peterson MJ, Massimini M, Murphy M, Riedner BA, Watson A, Bria P, Tononi G (2007) Reduced sleep spindle activity in schizophrenia patients. *Am J Psychiatry* 164(3):483–492
- Kabir MM, Tafreshi R, Boivin DB, Haddad N (2015) Enhanced automated sleep spindle detection algorithm based on synchrosqueezing. *Med Biol Eng Comput* 53(7):635–644
- Schönwald SV, Gerhardt GJ, Emerson L, Chaves ML (2003) Characteristics of human EEG sleep spindles assessed by gabor transform. *Phys A Stat Mech Appl* 327(1):180–184
- Shimada T, Shiina T, Saito Y (2000) Detection of characteristic waves of sleep eeg by neural network analysis. *IEEE Trans Biomed Eng* 47(3):369–379
- Tsanas A, Clifford GD (2015) Stage-independent, single lead eeg sleep spindle detection using the continuous wavelet transform and local weighted smoothing. *Front Hum Neurosci* 9:181
- Warby SC, Wendt SL, Welinder P, Munk EG, Carrillo O, Sorensen HB, Jennum P, Peppard PE, Perona P, Mignot E (2014) Sleep-spindle detection: crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nat Methods* 11(4):385–392
- Wendt SL, Christensen JA, Kempfner J, Leonthin HL, Jennum P, Sorensen HB (2012) Validation of a novel automatic sleep spindle detector with high performance during sleep in middle aged subjects. In: 2012 Annual international conference of the IEEE engineering in medicine and biology society, pp. 4250–4253. IEEE