Metadata of the chapter that will be visualized online

Chapter Title	Automatic Sleep Scoring from a Single Electrode Using Delay Differential Equations	
Copyright Year	2012	
Copyright Holder	Springer-Verlag	
Author	Family Name	**Lainscsek**
	Particle	
	Given Name	**Claudia**
	Suffix	
	Email	claudia@salk.edu
Author	Family Name	**Letellier**
	Particle	
	Given Name	**Terrence J. Sejnowski,ValérieMessager,Christophe**
	Suffix	
	Email	christophe.letellier@coria.fr
Author	Family Name	**Muir**
	Particle	
	Given Name	**Adriana Portman,andJean-François**
	Suffix	
	Email	jean-francois.muir@chu-rouen.fr
Abstract	Sleep scoring is commonly performed from electroencephalogram (EEG), electrooculogram (EOG), and electromyogram (EMG) to produce a so-called hypnogram. A neurologist thus visually encodes each epoch of 30 s into one of the sleep stages (wake, REM sleep, S_1, S_2, S_3, S_4). To avoid such a long process (about 3–4 hours) a technique for automatic sleep scoring from the signal of a single EEG electrode located in the C_3/A_2 area using nonlinear delay differential equations (DDEs) is presented here. Our approach considers brain activity as resulting from a dynamical system whose parameters should vary according to the sleep stages. It is thus shown that there is at least one coefficient that depends on sleep stages and which can be used to construct a hypnogram. The correlation between manual hypnograms and the coefficient evolution is around 80%, that is, about the inter-rater variability. In order to rank sleep quality from the best to the worst, we introduced a global sleep quality index which is used to compare manual and automatic sleep scorings, thus using our ability to state about sleep quality that is the final goal for physicians.	

Automatic Sleep Scoring from a Single Electrode Using Delay Differential Equations

Claudia Lainscsek, Valérie Messager, Adriana Portman, Jean-François Muir, Terrence J. Sejnowski, and Christophe Letellier

Abstract Sleep scoring is commonly performed from electroencephalogram (EEG), electrooculogram (EOG), and electromyogram (EMG) to produce a so-called hypnogram. A neurologist thus visually encodes each epoch of 30 s into one of the sleep stages (wake, REM sleep, S_1, S_2, S_3, S_4). To avoid such a long process (about 3–4 hours) a technique for automatic sleep scoring from the signal of a single EEG electrode located in the C_3/A_2 area using nonlinear delay differential equations (DDEs) is presented here. Our approach considers brain activity as resulting from a dynamical system whose parameters should vary according to the sleep stages. It is thus shown that there is at least one coefficient that depends on sleep stages and which can be used to construct a hypnogram. The correlation between manual hypnograms and the coefficient evolution is around 80%, that is, about the inter-rater variability. In order to rank sleep quality from the best to the worst, we introduced a global sleep quality index which is used to compare manual and automatic sleep scorings, thus using our ability to state about sleep quality that is the final goal for physicians.

1 Introduction

Up to 2007, polysomnographic recordings were scored into sleep stages according to the rules introduced by Rechtschaffen and Kales [19] which are mainly based on a spectral analysis. The scoring, accomplished by well-trained neurologist, consists in scoring all 30 s epochs into one of the six stages of vigilance, namely awakeness,

C. Lainscsek (✉) • T.J. Sejnowski
Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA
e-mail: claudia@salk.edu; terry@salk.edu

V. Messager • C. Letellier
Université de Rouen, CORIA, Avenue de l'Université, BP 12, F-76801 Saint-Etienne du Rouvray cedex, France
e-mail: valerie.messager@coria.fr; christophe.letellier@coria.fr

A. Portman • J.-F. Muir
Hôpital de Bois-Guillaume, Bois-Guillaume, France
e-mail: adrianaportmann@yahoo.fr; jean-francois.muir@chu-rouen.fr

J. Awrejcewicz (ed.), *Applied Non-Linear Dynamical Systems*, Springer Proceedings in Mathematics & Statistics 93, DOI 10.1007/978-3-319-08266-0_27

rapid eyes movement sleep (REM), and sleep stages S1, S2, S3, and S4. RK rules were recently modified to overcome the inter-rater variability ([11]). The most important change was that stages 3 and 4 merged into a single stage, named slow-wave sleep or N3. In spite of that, recent studies only showed slight improvements with the new rules ([6]) with an inter-rater agreement slightly greater than 72% ([3]).

Automatic sleep scoring techniques are thus welcome. Most of the computer-assisted scoring techniques stages were based on RK rules ([10, 12, 18]). In fact, most of them try to reproduce what is done by neurologists and which can lead to an overall epoch-by-epoch agreement of 80%, and require a quite complex decisional tree (see Fig. 2 in [2]). With the emergence of "chaos theory," recurrence plots quantifiers, Lyapunov exponents, or correlation dimension were used to obtain hypnograms with an overall agreement which was rarely greater than 60 or 70% ([23]).

Neural networks were also used to distinguish different features exhibited in the spectral domain but were not able to distinguish more than the REM sleep from non-REM sleep ([9]). Another technique was correctly scoring sleep stages but required two EEG channels, one horizontal electrooculogram channel and one chin electromyogram channel ([20]). An automatic sleep classification was able to distinguish wake, slow-wave sleep and rapid eye movements sleep stages ([22]), but a specific sensor, a head accelerometer, was required and must be added to conventional sensors.

Our aim is to develop a reliable automatic technique using a single EEG signal for scoring hypnograms. The subsequent part of this paper is organized as follows. In Sect. 2 the pool of patients which were recorded is described. Section 3 is devoted to our automatic sleep scoring technique and to a new global sleep quality index used to rank a set of hypnograms. In Sect. 4 the results are presented and Sect. 5 gives a conclusion.

2 Patients

This retrospective observational study was conducted at the sleep laboratory at the medical university hospital Intensive Care Unit in Rouen. We selected 38 recordings, but only 35 were associated with a reliable sleep scoring. These patients were long-term ventilated for chronic respiratory failure and grouped into two types. The first type corresponds to an obesity hypoventilation syndrome (OHS) commonly seen in severely overweight people who fail to breathe normally resulting in low blood oxygen levels and high blood carbon dioxide (CO_2) levels. Many of these patients have increased upper airway resistances during sleep (obstructive sleep apnea). This induces a significant amount of wake after sleep onset (WASO) leading to abnormal daytime sleepiness. This disease puts strain on the heart, possibly resulting in heart failure, leg swelling, and various other related symptoms. The second group of respiratory failure, considered here, is associated with chronic obstructive pulmonary disease (COPD). This refers to small airway obstructions

Table 1 Main clinical characteristics of the patients ($n = 34$)

Demographics and respiratory parameters	Mean	(SD)
Age (year)	64.5	(11.7)
Male:female	24:11	
Body mass index ($kg.m^{-2}$)	42.0	(10.5)
PaO_2 (cmH_2O)	9.5	1.1
$PaCO_2$ (cmH_2O)	5.8	(0.9)

Normal values: $(10.7 < PaO_2 < 12.0)\ cmH_2O$, $PaCO_2 \approx 5.3\ cmH_2O$, $(18.5 < BMI < 25)\ kg.m^{-2}$ and obesity is defined by $BMI > 30\ kg.m^{-2}$

and emphysema, two commonly coexisting pulmonary diseases in which the airways progressively narrow inducing shortness of breath. In these patients, the airflow limitation is usually nonreversible when treated with bronchodilators and progressively becomes more and more severe. One efficient treatment is to put these patients under noninvasive mechanical ventilatory assistance. In the present case, all patients were ventilated with the bilevel ventilator ResMed VPAP III. All patients included in this study were in stable condition, as assessed by clinical examination and arterial blood gases.

Main characteristics of the thirty-five patients for which the sleep was scored during one night under mechanical ventilation are reported in Table 1. Twenty patients (57%) had OHS and 15 patients (43%) had COPD. Thirteen patients (38%) were diagnosed with obstructive sleep apnea syndrome (defined as more than 10 apneas per hour). Upon study inclusion, the patients were ventilated for a few months. Nineteen patients (56%) were hypercapnic ($PaCo_2 > 5.6\ cmH_2O$).

3 Method

3.1 Automatic Sleep Scoring

A nonlinear delay differential equation has the general form

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_3} + \cdots + a_{i-1} x_{\tau_n} + a_i x_{\tau_1}^2 + a_{i+1} x_{\tau_1} x_{\tau_2} + a_{i+2} x_{\tau_1} x_{\tau_3} + \cdots + a_{j-1} x_{\tau_n}^2 + a_j x_{\tau_1}^3 + a_{j+1} x_{\tau_1}^2 x_{\tau_2} + \ldots\cdots + a_l x_{\tau_n}^m \quad (1)$$

where $x = x(t)$ and $x_{\tau_j} = x(t - \tau_j)$. In this general form, the DDE has n delays, l monomials with their corresponding coefficients a_i, and a degree of nonlinearity equal to m. In the subsequent part of this paper, we will define a k-term DDE as an equation with only $k < l$ monomials selected from the right-hand side of the general form (1). As for any global modeling technique, there is a significant improvement of capturing main characteristics of the underlying dynamics from observed data by carefully selecting the structure of the DDE model ([1, 14–16]). The minimal mean

squared error is used for this process. By structure selection, we mean retaining only monomials in the DDE that have the most significant contribution to classify the data. An equally important task is to select the right Time-delays, since linear terms are directly related to the fundamental timescales and nonlinear terms to the nonlinear couplings between them ([16]). This can be performed by using a genetic algorithm ([15]) or by an exhaustive search for the best model among the general form with $n = 2$ and $m = 3$ resulting in $l = 9$ monomials as performed in [16].

Here only models with up to three terms were considered (see Table 2 in [16]). The variable x corresponds to the signal provided by the electrode located in the C_3/A_2 area of the scalp. We ran a genetic algorithm to minimize the least square error of 30 s data windows to select the best models and delays for each 30 s window ([8, 15]). For 95% of the data windows (corresponding to the 35 patients), the four models

$$\dot{x} = a_1\, x_{\tau_1} + a_2\, x_{\tau_2} + a_3\, x_{\tau_1}^2\,; \tag{2}$$

$$\dot{x} = a_1\, x_{\tau_1} + a_2\, x_{\tau_2} + a_4\, x_{\tau_1}\, x_{\tau_2}\,; \tag{3}$$

$$\dot{x} = a_1\, x_{\tau_1} + a_2\, x_{\tau_2} + a_6\, x_{\tau_1}^3\,; \tag{4}$$

$$\dot{x} = a_1\, x_{\tau_1} + a_2\, x_{\tau_2} + a_7\, x_{\tau_1}^2\, x_{\tau_2}\,; \tag{5}$$

were selected as well as delays between $1\,\delta t$ and $4\,\delta t$ with $\delta t = \frac{1}{64}$ s. Among these four models, model (5) is the best to distinguish wake, REM, and S1 from the sleep stages S2, S3, and S4 (see left panel from Fig. 1). Delay $\tau_1 = 1$ is useful to distinguish wake, S2, S3, and S4 from REM and S1 (right panel from Fig. 1). Delay $\tau_2 = 3$ allows to distinguish wake from sleep stages. Thus, combining model (5) with delays $\tau_1 = 1$ and $\tau_2 = 3$ provides the model with the most discriminative ability. Among the three coefficients of model (5), parameter a_2 was found to be the most correlated ($r = 0.95$) to the manually scored hypnogram, as exemplified in Fig. 2 in the case of patient 15. We then used this model and this coefficient to score the sleep for our 35 patients.

It was then necessary to convert the a_2-time series which corresponds to the time evolution of a real number sampled at 0.1 Hz (one point per 10 s) into a sequence of integers from 1 (stage S1) to 6 (awake). This is the tricky part of our technique. In the case of patient 15, we got an automatically scored hypnogram which was quite close to the manually scored one (Fig. 2b).

3.2 Assessing the Sleep Quality

Since patients with chronic respiratory failures are ventilated during their sleep, it is important to assess whether the ventilation improves the sleep quality or,

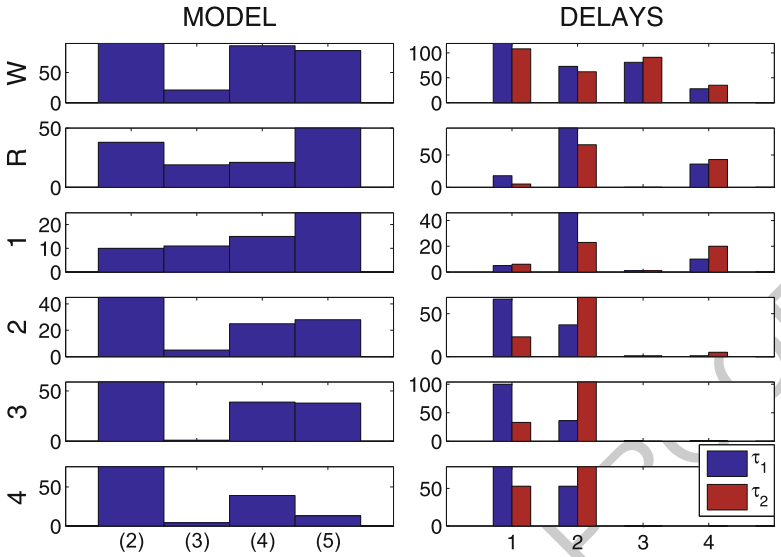


Fig. 1 Histograms of the number of time each of the four selective DDEs (*left*) and each delays (*right*) were selected with minimum error for each sleep stage

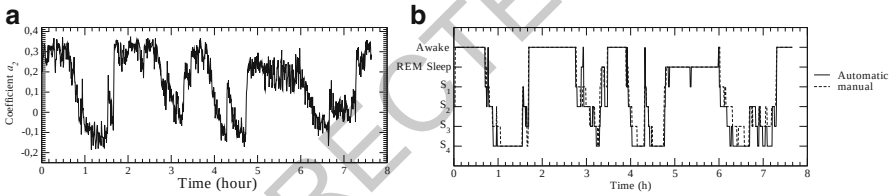


Fig. 2 Time series of coefficient a_2 of the delay differential equation (5) and the corresponding hypnogram. Case of patient 15 (male, 76 years, BMI=50 kg.m^{-2}). The manually scored hypnogram (*green*) is also reported for comparison. (**a**) Raw a_2 time series (**b**) Sequence of integers

at least, that it does not degrade it. In order to do that, it is necessary to be able to rank hypnograms according to sleep quality. From a subjective point of view, sleep quality refers to patient feelings about the refreshing effect of sleep which can be assessed using some sleep diary or the Pittsburgh Quality Index ([4]). The characteristics commonly taken into account in such evaluation are sleep latency, sleep duration, regular sleep efficiency, sleep disturbances (including sleep disruptive events such as snoring, apnea, or pains), use of sleeping medication, and daytime dysfunction ([4]).

Up-to-now, the objective evaluation of sleep quality was based on the same characteristics but directly measured from hypnograms ([11]). Also considered are the arousal index (number of arousals per hour) and the number of various respiratory events. To assess the evolution of sleep quality, all these quantities are then subjectively combined and compared since none of them can alone allow

to rank hypnograms according to sleep quality (see [17] for details). In order to avoid this last subjective step, we introduced a new index which combines the most important sleep characteristics. Thus, our global sleep quality index takes into account the number of sleep cycles (each cycle, between 90 and 120 min, contains some slow-wave sleep restoring physical functions and some rapid eye movements restoring cognitive functions), the fraction of WASO, the fraction of stable sleep, the number of micro-arousals, and the number of stage transitions. The global sleep quality index η_{GSQ} is defined as

$$\eta_{\mathrm{GSQ}} = \eta_{\mathrm{cycle}} \cdot \eta_{\mathrm{restoring}} \cdot \eta_{\mathrm{stability}} \cdot (1 - \eta_{\mathrm{M-frag}}) \cdot (1 - \eta_{\mu-\mathrm{frag}}) \tag{6}$$

where $\eta_{\mathrm{cy}} = \mathrm{Max}\left(\frac{N_{cy}}{6}, 1\right)$ and N_{cy} is the number of sleep cycles that saturates to one when it exceeds 6 cycles; the restoring capacity of sleep is evaluated according to

$$\eta_{\mathrm{restoring}} = \mathrm{Min}\left(\frac{5}{2}\,\frac{\tau_{S3} + \tau_{S4} + \tau_R}{\tau_{S1} + \tau_{S2} + \tau_{S3} + \tau_{S4} + \tau_R}, 1\right) \tag{7}$$

with τ_i being the time duration spent in the ith sleep stage (i = S1, S2, S3, S4, and R) and saturates to 1 when the restorative sleep (S3, S4, and R) exceeds $\frac{2}{5}$ of the effective sleep; the sleep stability is evaluated according to

$$\eta_{\mathrm{stability}} = \frac{\tau'_{S1} + \tau'_{S2} + \tau'_{S3} + \tau'_{S4} + \tau'_R}{\tau_{\mathrm{effective\ sleep}}} \tag{8}$$

with τ'_i being the time spent in the ith sleep stage without any micro-arousal and not corresponding to an epoch connexe to a stage transition, and $\tau_{\mathrm{effective\ sleep}}$ being the time duration of sleep stages ($\tau_{S1} + \tau_{S2} + \tau_{S3} + \tau_{S4} + \tau_R$); the sleep macro-fragmentation is evaluated according to

$$\eta_{\mathrm{M-frag}} = \frac{\tau_{\mathrm{waso}}}{\tau_{\mathrm{waso}} + \tau_{\mathrm{effective\ sleep}}}; \tag{9}$$

the sleep micro-fragmentation is evaluated according to

$$\eta_{\mu-\mathrm{frag}} = \frac{(\tau_{S1} - \tau'_{S1}) + (\tau_{S2} - \tau'_{S2}) + (\tau_{S3} - \tau'_{S3}) + (\tau_{S4} - \tau'_{S4}) + (\tau_R - \tau'_R)}{\tau_{\mathrm{effective\ sleep}}} \tag{10}$$

with $\tau_i - \tau'_i$ being the time spent in an epoch of the ith sleep stage with a micro-arousal or connection to a stage transition.

4 Results

The time series of coefficient a_2 were found quite well correlated to the corresponding hypnograms ($\overline{r} = 0.86 \pm 0.1$). To assess the quality of our sleep scorings using the coefficient a_2 we computed the confusion matrix ([13]) which is a specific table layout used to assess performance of classifier. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. The confusion matrix for all epochs of all patients is reported in Fig. 3. To get a graphical representation the numbers were also converted to a percentage. A dark diagonal from the upper-left corner to the lower-right corner with all other squares in white would indicate perfect scoring of each data window into the correct sleep stage.

As additional measure of performance we used Cohen's kappa κ [5, 7, 21] which can be computed directly from the confusion matrix as [13]. $\kappa = \frac{p_a - p_e}{1 - p_e}$, where $p_a = \sum_{k=1}^{q} p_{kk}$, and $p_e = \sum_{k=1}^{q} p_{k+} p_{+k}$ where $q = 6$ for the 6 classes, p_a is the observed percentage of agreement, p_e is the expected percentage of agreement, p_{k+} is the percentage of actual classification, and p_{+k} is the percentage of predicted classification. We got $\overline{\kappa} = 0.51 \pm 0.1$ when comparing automatically scored hypnograms with the manually scored ones. Detailed results are reported in Table 2.

The global sleep quality index η_{GSQ} was first computed from the hypnograms scored by the neurologist. Patients were then ranked according to a decreasing η_{GSQ} (Fig. 4). The hypnogram of the patient with the largest η_{GSQ} (35.4 %) is shown in Fig. 5a: it presents 3 sleep cycles quite well structured. Contrary to this, the hypnogram of patient 22 with the smallest η_{GSQ} (0.1 %) is shown in Fig. 5b: it does not present a single well-structured sleep cycle and the effective sleep time duration is small ($\tau_{\text{effective sleep}} = 146$ min).

The rates of each sleep stage was computed for each hypnograms which were ranked according to decreasing η_{GSQ} (Fig. 6). The best hypnogram (patient 34, $\eta_{\mathrm{GSQ}} = 35.4$) presents a good proportion of restorative sleep. Contrary to this, the worst hypnogram (patient 22, $\eta_{\mathrm{GSQ}} = 0.1$) associated with a very small fraction

actual \ predicted	W	R	1	2	3	4
W	3026	1642	147	29	1	4
R	225	1288	762	62	5	11
1	61	733	3841	2388	607	164
2	7	13	191	703	524	528
3	1	34	394	790	1669	97
4	27	54	166	557	1050	8615

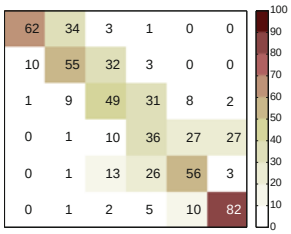


Fig. 3 Confusion matrix for all subjects: the table on the *left side* shows the numbers of predicted and actual sleep stage windows. The plot on the *right side* visualizes the percentage of predicted and actual sleep stage windows

Table 2 Correlation coefficient r and Cohen's κ between the manually scored hypnograms and the time series of coefficient a_2 of model (5) for each subject

#	r	κ	#	r	κ	#	r	κ	#	r	κ	#	r	κ	
1	0.82	0.36	9	0.91	0.53	17	0.70	0.28	24	0.95	0.65	32	0.91	0.55	t5.2
2	0.95	0.61	11	0.80	0.36	18	0.81	0.44	25	0.78	0.41	33	0.84	0.41	t5.3
3	0.89	0.59	12	0.90	0.64	19	0.87	0.53	26	0.82	0.50	34	0.93	0.66	t5.4
5	0.91	0.63	13	0.91	0.57	20	0.78	0.59	27	0.92	0.64	35	0.82	0.37	t5.5
6	0.92	0.51	14	0.76	0.36	21	0.79	0.39	29	0.94	0.61	36	0.89	0.55	t5.6
7	0.92	0.68	15	0.95	0.67	22	0.80	0.40	30	0.87	0.51	37	0.91	0.59	t5.7
8	0.79	0.43	16	0.90	0.54	23	0.80	0.41	31	0.79	0.43	38	0.91	0.51	t5.8

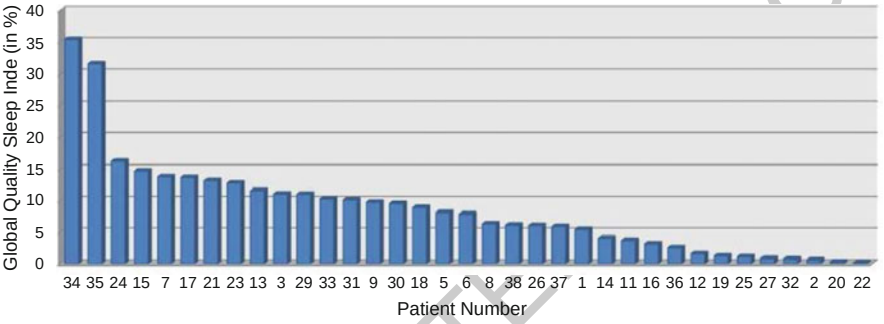
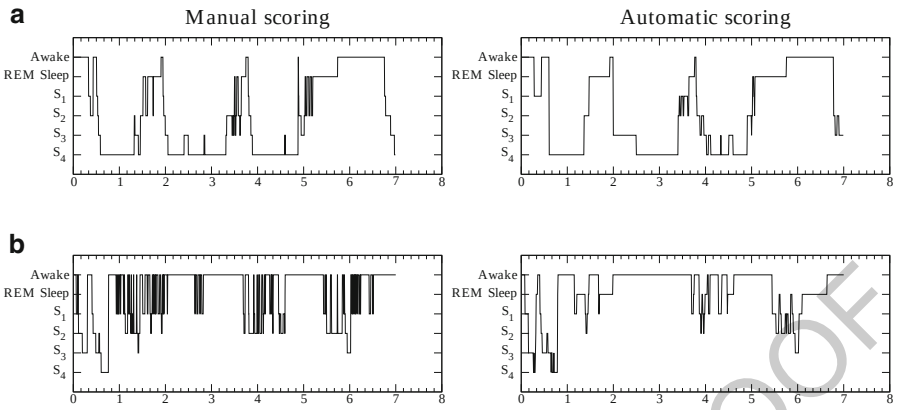


Fig. 4 Global sleep quality index computed from the manually scored hypnograms for the 35 patients of our protocol

of restorative sleep and a large one of WASO. Hypnograms are rather well ranked since the rate of WASO and sleep micro-fragmentation are anticorrelated to η_{GSQ} ($r = -0.65$, $p < 0.0001$ and $r = -0.75$, $p < 0.0001$, respectively). The rate of slow-wave sleep (S3 and S4) and the rate of REM sleep are correlated to η_{GSQ} ($r = 0.83$, $=< 0.0001$ and $r = 0.59$, $p < 0.0001$, respectively). These features and others that are outside the scope of this paper correspond to an increase of the sleep quality with η_{GSQ}.

We now computed the global sleep quality index from the automatically scored hypnograms with our technique (Fig. 7). They were ordered in a slightly different order than the manual hypnograms. In order to quantify this disagreement between these two orders, let us designate by n the rank (n') the rank obtained by computing η_{GSQ} from the manual (automatic) hypnograms. Thus $\Delta n = |n - n'|$ corresponds to the rank shift observed between these two orders. We thus have $\overline{\Delta n} = 4.6 \pm 5.4$, meaning that, in average, the good (bad) hypnograms remain the good (bad) ones. There are four notable exceptions with the hypnograms for patients 11, 15, 24, and 35 for which Δn equals to -23, $+15$, $+20$, and $+11$, respectively.



AQ2

Fig. 5 Hypnograms for two of the 35 patients corresponding to the largest and the smallest global sleep quality index. The gender, age, body mass index, and the rate of synchronous breathing cycles are also reported. (**a**) Patient 34 : male, 82 years, BMI=44.1, 2.1% of asynchronous cycles, and $\eta_{GSQ} = 35.4\%$. (**b**) Patient 22 : male, 83 years, BMI=36.3, 8.0% of asynchronous cycles, and $\eta_{GSQ} = 0.1\%$

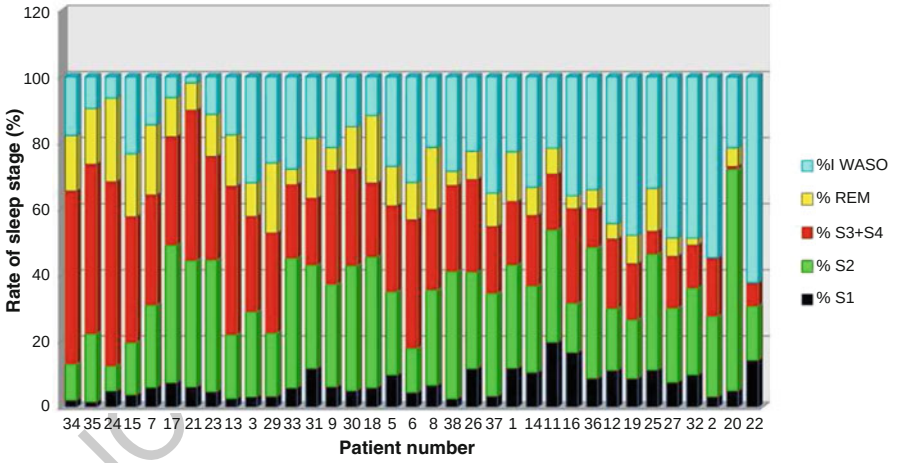


Fig. 6 Fraction of time duration of each sleep stage. Patients are ranked according to the global sleep quality index η_{GSQ}

The manually scored hypnogram of patient 11 (Fig. 8a) presents many fluctuations between wake and stage S1 and a very few epochs in stages S3 or S4 and REM sleep, thus associated with a small global quality sleep index ($\eta_{GSQ} = 3.7\%$). The evolution of the coefficient of the DDE fluctuates a lot between the values corresponding to wake and S1 stages. Consequently, since REM sleep is between these two stages from EEG, our technique returns too often REM sleep (and not WASO). This is significantly increasing the global sleep quality index to 24.9.

this figure will be printed in b/w

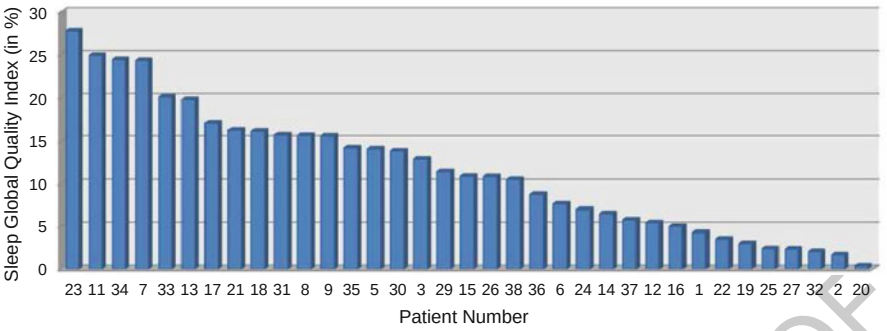


Fig. 7 Global sleep quality index computed from the automatically scored hypnograms for the 35 patients of our protocol

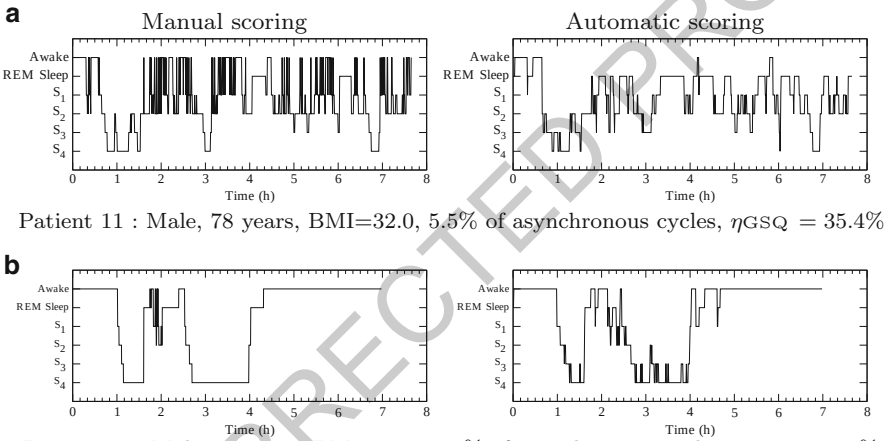


Fig. 8 Hypnograms for two badly scored using our automatic technique

It is important to note that a neurologist uses a lot the electrooculogram and the electromyogram to distinguish REM sleep from awake and S1, two signals which are not considered by our technique.

Contrary to this, the automatically scored hypnograms for patient 24 is characterized by a global sleep quality index $\eta_{\mathrm{GSQ}} = 7.0\%$ is significantly smaller than the value (16.2%) obtained from the manual hypnograms (Fig. 8b). There are few reasons explaining such a large departure between these two η_{GSQ}-values. The global sleep duration (between the first and the last sleep epoch) is larger than the one obtained from the automatic scoring (221.5 min and 198 min, respectively), but the number of sleep cycles is 2 in both cases. The rate of WASO in the automatic hypnogram is about three times the rate obtained from the manually scored hypnogram (19.9 and 6.6, respectively). The rate of micro-fragmentation obtained with our technique is about three times the rate returned by the neurologist

(31.8 and 11.1, respectively). The stability is smaller in the hypnogram provided by our technique than in the one scored by the neurologist (38.2 and 58.3, respectively). All these modifications tend to increase the global sleep quality index.

5 Conclusions

In 88% of subjects the overall sleep quality index computed from the DDE hypnograms are in agreement with the sleep quality index computed from the visually scored hypnograms. The difference in 12% of all patients results from converting the real number outputs of the DDE to the integers used for indexing sleep stages (S1, S2, S3, S4, R, and wake). This is the weakest part of the present version of our technique. In spite of this, our hypnograms are already sufficiently close to the manual hypnograms that are used to assess the sleep quality. Importantly, this first study has led to the identification of possible improvements that are currently being developed.

Our automatic scoring technique using DDEs is well correlated to the corresponding visually scored hypnograms ($\overline{r} = 0.86 \pm 0.1$). This excellent agreement becomes even more impressive when considering the use of only one scalp electrode for the DDE method. Indeed, the most promising aspect of our technique is that only one scalp electrode is sufficient to accurately score sleep stages.

References

1. Aguirre, L.A., Billings, S.A.: Improved structure selection for nonlinear models based on term clustering. Int. J. Contr. **62**(3), 569–587 (1995)
2. Anderer, P., Gruber, G., Parapatics, S., Woertz, M., Miazhynskaia, T., Klösch, G., Saletu, B.,Zeitlhofer, J., Barbanoj, M.J., Danker-Hopfe, H., Himanen, S.-L., Kemp, B., Penzel, T., Grözinger, M., Kunz, D., Rappelsberger, P., Schlögl, A., Dorffner, G.: An E-health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the somnolyzer 24x7 utilizing the SIESTA database. Neuropsychobiology **51**(3), 115–133 (2005)
3. Basner, M., Griefahn, B., Penzel, T.: Inter-rater agreement in sleep stage classification between centers with different backgrounds. Somnologie **12**,75–84 (2008)
4. Buysse, D.J., Reynolds, C.F., Monk, T.H., Berman, S.R., Kupfer, D.J.: The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatric Res. **28**, 193–213 (1989)
5. Cohen, J.: A coefficient of agreement for nominal scales. Educ. Psychol. Meas. **20**(1), 37–46 (1960)
6. Danker-Hopfe, H., Anderer, P., Zeitlhofer, J., Boeck, M., Dorn, H., gruber, G., Heller, E., Loretz, E., Moser, D., Parapatics, S., Saletu, B., Schmidt, A. Dorfner, G.: Interrater reliability for sleep scoring according to the Rechtschaffen and Kales and the new AASM standard. J.Sleep Res. **18**(1), 74–84 (2009)
7. Fleiss, J.L., Cohen, J.: The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ. Psychol. Meas. **33**, 613–619 (1973)

8. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, Wokingham (1998)
9. Grözinger,M., Wolf, C., Uhl, T., Schäffner, C., Röschke, J.: Online detection of REM sleep based on the comprehensive evaluation of short adjacent EEG segments by artificial neural networks. Progr. Neuro-Psychopharmacol. Biol. Psychiatr. **21**(6), 951–963 (1997)
10. Harper, R.M., Schechtman, V.L., Kluge, K.A.: Machine classification of infant sleep state using cardiorespiratory measures. Electroencephalogr. Clin. Neurophysiol. **67**(4), 379–387 (1987)
11. Iber, C., Ancoli-Israel, S., Chesson, A., Quan, S.F. (eds.): The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology, and Technical Specification. American Academy of Sleep Medicine, Westchester (2007)
12. Jansen, B.H., Dawant, B.M.: Knowledge-based approach to sleep EEG analysis-a feasibility study. IEEE Trans. Biomed. Eng. **36**(5), 510–518, (1989)
13. Kohavi, R., Provost, F.: Glossary of terms. Mach. Learn. **30**(2/3), 271–274 (1998)
14. Lainscsek, C., Letellier, C.,Gorodnitsky, I.: Global modeling of the Rössler system from the *z*-variable. Phys. Lett. A, **314** 409–127 (2003)
15. Lainscsek, C., Rowat, P., Schettino, L., Lee, D., Song, D., Letellier, C., Poizner, H.: Finger tapping movements of Parkinson's disease patients automatically rated using nonlinear delay differential equations. Chaos, **22**, 013119 (2012)
16. Lainscsek, C., Sejnowski, T.J.: Electrocardiogram classification using delay differential equations. Chaos **23**(2), 023132 (2013)
17. Messager, V., Portmann, A., Muir, J.-F., Letellier, C.: A global sleep quality index for ranking hypnograms. in preparation (2013)
18. Principe, J.C., Gala, S.K., Chang, T.G.: Sleep staging automaton based on the theory of evidence. IEEE Trans. Biomed. Eng. **36**(5), 503–509 (1989)
19. Rechtschaffen, A., Kales, A. (eds.): A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subject. US Government Printing Office, National Institute of Health Publication, Washington (1968)
20. Schaltenbrand, N., Lengelle, R., Toussaint, M., Luthringer, R., Carelli, G. Jacqmin, A., Lainey, E., Muzet, A., Macher, J.P.: Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients. Sleep **1**, 26–35 (1996)
21. Scott, W.: Reliability of content analysis: The case of nominal scale coding. Publ. Opin. Q. **19**(3), 321–325 (1955)
22. Sunderam, S., Chernyy, N., Peixoto, N., Mason, J.P., Weinstein, S.L., Schiff, S.J., Gluckman, B.J.: Improved sleep-wake and behavior discrimination using MEMS accelerometers. J.Neurosci. Meth. **163**(2), 373–383 (2007)
23. Šušmáková, K., Krakovská, A.: Discrimination ability of individual measures used in sleep stages classification. Artif. Intell. Med. **44**(3), 261–277 (2008)

AQ3

AUTHOR QUERIES

AQ1. The term “contents” has been changed to “contains” in the sentence “..some rapid eye movements restoring ..” Please check if okay.
AQ2. Please check the figure caption of Fig. 5.
AQ3. Please update the reference [17].