34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Consistency of Sleep Restoration Gain (SRG) as a Measure for Assessing Sleep Quality

Islam S. Badreldin, *Student Member, IEEE,* and Ahmed A. Morsy, *Senior Member, IEEE*

Abstract— **We propose a new sleep quality measure that assesses the sleep restorative gain of a polysomnography sleep record. In this preliminary investigation, we derive this new measure from manually scored sleep hypnograms. We compare the proposed measure to classical sleep indices such as TST, SE, and ArI, and demonstrate its self-consistency and degree of correlation with these measures . Using 47 sleep records from publicly available sleep databases, we graphically and quantitatively demonstrate the effectiveness of the proposed measure in summarizing the hypnogram of a sleep record.**

I. INTRODUCTION

Sleep is a physiological process by which the body restores its wellness and vitality. A good night of sleep can contribute positively to one's energy and efficiency the following day. Conversely, sleep fragmentation and shortening of sleep duration can lead to serious health implications, including sleepiness, hypertension, coronary artery disease, diabetes, obesity, and mood disturbances [1]. It is therefore desirable to have a sleep quality measure that can be used clinically to assess the restorative quality of a night of sleep. Restorative sleep in a normal subject has a well-defined structure and periodicity called sleep cycle, in which sleep alternates between non-rapid eye movement (NREM) sleep and REM sleep [2].

Sleep quality assessment can be done subjectively or objectively. Subjective assessment is done by help of the subjects themselves by answering standardized questionnaires about how restorative and undisturbed their sleep quality was. Objective assessment is performed by evaluating, either automatically or by a clinician, the neuro-physiological signals recorded from the subject during a sleep. Examples of subjective measures for sleep include the Epworth Sleepiness Scale (ESS) [3], which is intended to measure daytime sleepiness through a well-defined questionnaire. ESS is administered routinely by sleep clinics [4]. However, ESS is not intended to measure sleep at any particular night. Validation studies show that the ESS score is relatively maintained after 1 year [5], [6]. Another related popular subjective measure is the Pittsburgh Sleep Quality Index (PSQI) [7], which is also internally highly reproducible [6].

Many approaches to objective measurement of sleep quality rely on Polysomnography (PSG), which is a standard technique for evaluating normal and disturbed sleep [8]. Some of these approaches make use of sleep staging hypnograms scored according to the criteria of Rechtschaffen and Kales (R&K) [9], or, more recently, the criteria of the American Academy of Sleep Medicine (AASM) [10]. According to R&K, NREM sleep is further divided into stages 1, 2, 3, and 4. Sleep stage 1 represents drowsiness, while stage 2 represents light sleep, with stages 3 and 4 representing deep sleep. AASM grouped stages 3 and 4 together in one stage, called N3. Other approaches attempt to quantify sleep quality by calculating derived parameters from the neuro-physiological signals recorded in PSG, without depending on the scored sleep stages hypnograms [11].

Total Sleep Time (TST) is one of the commonly used measures of sleep quality, which is defined as the total time spent in actual sleep (REM and NREM) [8]. Another related measure is the Sleep Efficiency (SE), which is defined as the percentage ratio (TST/TTB)$\times$100%, where (TTB) is the Total Time in Bed, defined as the time spent on the bed from the 'lights off' when the recording starts in the night to the 'lights on' in the morning when the recoding ends [8]. A third measure is the Arousal Index (ArI), defined as the average number of arousal events per hour of sleep [8], where an arousal event is scored according to the criteria of AASM [10]. TST, SE, and ArI are three of the most clinically used measures to assess sleep quality.

In [12], the total number of transitions into a Wake state or Stage 1, divided by TST, was proposed as a Sleep Fragmentation Index (SFI). This was later modified in [13] to include all transitions from one stage to another. More recently, the authors in [8] extended the idea by proposing a 'weighted-transition' SFI. In [1] the authors criticized the former SFI sleep quality measures, as they fail to capture the temporal dynamics of the underlying sleep process. They proposed a new Entropy-based Measure to capture the sleep stage transition dynamics, and compared that measure to TST, SE, ArI, and SFI.

In this paper, we propose a measure for sleep restorative ability. We call this new measure the Sleep Restoration Gain (SRG) measure for a night of sleep, based on the concept of 'Sleep Debt' explained in the next section. Similar to [1], we capture the temporal dynamics of sleep evident in the hypnogram. Unlike [1], we propose a simplified dynamical model to derive this new measure quantitatively from sleep hypnograms. We demonstrate its self-consistency as well as its consistency with some of the previously reported sleep quality measures and discuss its advantages over them.

II. SLEEP PROCESSES

Attempts to explain the underlying processes of sleep date back to several decades. In [14], Borbély proposed a

---

*This work was not supported by any organization

I. S. Badreldin and A. A. Morsy are with the Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University, Egypt `{ibadreldin,amorsy} at ieee.org`

two-process sleep model, in which the S-process determines sleep propensity, and builds up during wakefulness and declines during sleep. Sleep or wakefulness happens when S is above or below a threshold, respectively. This threshold itself is assumed to follow a circadian rhythm, called the C-process. Later, Johns [3] proposed a four-process sleep model where two processes contribute to the total sleep drive, and two other processes contribute to the total wake drive. The wake drive is modeled as the summation of a primary wake and a secondary wake drive. The primary wake drive follows a circadian rhythm, while the secondary wake drive is influenced by body posture, behavior, physical activity, feelings, mental activity, etc. Similarly, the sleep drive is modeled as the summation of primary and secondary sleep drives. Sleep propensity at a particular time instance is dependent on the total wake drive and the total sleep drive at that time instance.

More recently, the author of [15] elaborated more on the four-process model. He associated the two sleep drives with the NREM and REM states. He proposed, through qualitative examples, that sleep occurs when the total wake drive falls below the NREM sleep drive, after-which the NREM sleep drive starts to discharge (decline). The author also introduced the concept of ‘Sleep Debt’, which is defined [15] as the vertical distance of the NREM drive above the primary wake drive, during a state of wakefulness. Sleep debt can later cause a period of sleep propensity whenever the second wake drive fails to raise the total wake drive higher than the NREM sleep drive. This way, the concept of sleep debt is related to alertness [15].

III. METHODS

A. Restorative Sleep Gain Measure

Since the concept of sleep debt is related to alertness, we can argue that a positive value for sleep debt means bad restoration of the night of sleep, while negative values of sleep debt indicate good restoration. This way, we define a Sleep Restoration Gain (SRG) measure to be the negation of sleep debt discussed in [15]. Sleep debt was defined in [15] as the difference between NREM sleep drive and the primary wake drive. NREM sleep drive builds up in all states of consciousness other than NREM and discharges in NREM state. The rate of discharge of NREM is higher during deep sleep (Stages 3 and 4) than in lighter sleep (Stages 1 and 2). In the same manner, we postulate that SRG builds up more quickly during deep sleep than in lighter sleep, and we propose the dynamics illustrated in Algorithm (1) to be able to derive SRG. We measure SRG in hours and increment SRG by a unit of 30 seconds, since this is the duration of one sleep epoch according to both R&K and AASM. Deeper sleep stages increment the SRG value more than lighter sleep stages. Wakefulness lowers the value of SRG, and indicates an accumulation of sleep debt. Moreover, since an arousal event can severely disturb sleep quality, a transition from any sleep stage into wakefulness is highly penalized. For the purpose of this study, SRG incrementation and decrementation factors are chosen in an arbitrary but rationalized way. The results presented in this paper must then be seen as a semi-qantitative approach for calculating SRG.

Algorithm 1: Deriving SRG measure from a hypnogram

```
input : sleep hypnogram
output: SRG value
prevstage ← $hypnogram[0]$;
increment ← $0$;
SRG ← $0$;
unit ← $30/3600$;
while not at end of hypnogram do
    read currentstage;
    switch currentstage do
        case WAKE
            if prevstage $\neq$ WAKE then
                increment ← $-1 \times$ unit $\times 15$;
            else
                increment ← $-1 \times$ unit;
            end
        case REM  increment ← $0$;
        case S1  increment ← unit$/1.5$;
        case S2  increment ← unit;
        case S3  increment ← unit $\times 1.5$;
        case S4  increment ← unit $\times 2$;
        otherwise  increment ← $0$;
    endsw
    SRG ← SRG $+$ increment;
    prevstage ← currentstage;
end
```

B. Data Sets

To conduct a preliminary validation of the SRG measure, we applied Algorithm (1) on publicly available sleep databases from PhysioBank [16]. We validated the approach on the 25 records of the University College Dublin Database (`ucddb`), the 8 records of the Sleep-EDF Database (`sleep-edf`) [17], and all the 18 records of the MIT-BIH Polysomnographic Database (`slpdb`) [18] except for records `slp01a`, `slp01b`, `slp02a`, and `slp02b`, since they constitute fragmented partial sleep records.

The `ucddb` contains 25 full overnight polysomnograms from adult subjects with suspected sleep-disordered breathing [16]. The `sleep-edf` recordings were obtained from Caucasian males and females (21 – 35 years old) without any medication [17]. The 4 `sc*` recordings were obtained in 1989 from ambulatory healthy volunteers during 24 hours in their normal daily life, using a modified cassette tape recorder, while the other 4 `st*` recordings were obtained in 1994 from subjects who had mild difficulty falling asleep but were otherwise healthy, during a night in the hospital, using a miniature telemetry system with very good signal quality [17]. In the `slpdb` database, all 16 subjects were male, aged 32 to 56 (mean age 43), with weights ranging from 89 to 152 kg (mean weight 119 kg) [18].

C. Implementation and Benchmarking

Access to the PhysioBank databases was done in C++ by the help of the PhysioToolkit [16] using the GNU Compiler

TABLE I

SPEARMAN CORRELATION COEFFICIENTS BETWEEN SRG AND CLASSICAL SLEEP INDICES

Database	TST	SE	ArI
`ucddb`	0.6674	0.5569	-0.9605
`sleep-edf`	0.4940§	0.8415	-0.4029†
`slpdb`	0.6436	0.8399	-0.6587
ALL	0.6777	0.6307	-0.8359

For all entries, $p < 0.01$. $^{\S}p = 0.21$ $^{\dagger}p = 0.32$

Collection version 4.4.5. The proposed SRG measure, using Algorithm (1), as well as the TST, SE, and ArI measures were implemented in GNU Octave version 3.2.4. The Spearman correlation coefficients between the SRG values and each of the TST, SE, and ArI values were calculated for all the records.

For `ucddb` and `slpdb`, the 'lights off' time and the 'lights on' time were assumed to be the start of the record and the end of the record, respectively. For `sleep-edf`, the 'lights off' and the 'lights on' times of the 24-hour `sc*` records were assumed heuristically by looking at the hypnogram for sleep onset and sleep offset. Also, for the `st*` records that have many 'unscored' epochs before sleep onset and after sleep offset, these 'unscored' epochs were assumed as WAKE.

IV. RESULTS

Figure 1 shows the calculated SRG curves based on the hypnograms. Table I shows the Spearman correlation coefficients between SRG and other classical indices for the different databases.

V. DISCUSSION & CONCLUSIONS

In Figure 1, record `ucddb002` shows how frequent arousal events are penalized in the SRG measure, ultimately resulting in a large negative value for SRG by the end of the record. This effectively expresses a case of sleep debt. In this way, SRG values are correlated with ArI values. Records `ucddb006` and `ucddb015` show how deep sleep (Stage 4) can quickly build up a restoration gain, yet this restoration gain is later lost by frequent arousal events in the second half of the sleep records. Records `ucddb010` and `ucddb027` show that it is possible to build a moderate restoration gain using a sleep record of mostly light sleep (Stage 2), despite the presence of a moderate number of arousal events. This way SRG values are correlated with TST and SE values. Record `slp45` shows a typical restorative night of sleep that can build up a high SRG value by the end of the record. In Table I, it is evident that the proposed SRG measure is correlated with classical sleep indices, with the added advantage of capturing the temporal dynamics of hypnograms, and summarizing them in a numerical value that can be an indicator of the restorative quality of sleep.

In this preliminary investigation, the consistency of the proposed SRG measure was demonstrated. SRG can effectively summarize a sleep hypnogram, and can be used as a good indicator of sleep restoration (positive values) or sleep debt (negative values). SRG values were derived using a simple dynamical model. Future work includes devising a more mathematically rigorous algorithm for deriving SRG values and a wider clinical validation of its effectiveness.

REFERENCES

[1] M. R. Kirsch, K. Monahan, J. Weng, S. Redline, and K. A. Loparo, "Entropy-based measures for quantifying sleep-stage transition dynamics: Relationship to sleep fragmentation and daytime sleepiness," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 3, pp. 787–796, Mar. 2012.

[2] H. Kryger, T. Roth, and W. Dement, *Principles and practice of sleep medicine*, 3rd ed. W. B. Saunders, Philadelphia, 2000.

[3] M. Johns, "Sleepiness in different situations measured by the Epworth Sleepiness Scale," *Sleep*, vol. 17, no. 8, pp. 703–710, Dec. 1994.

[4] D. Martinez, T. C. Breitenbach, M. S. Lumertz, D. L. Alcântara, N. S. da Rocha, C. M. Cassol, and M. D. C. S. Lenz, "Repeating administration of Epworth Sleepiness Scale is clinically useful," *Sleep & breathing = Schlaf & Atmung*, vol. 15, no. 4, pp. 763–73, Dec. 2011.

[5] H. Kumru, J. Santamaria, and R. Belcher, "Variability in the Epworth sleepiness scale score between the patient and the partner," *Sleep Medicine*, vol. 5, no. 4, pp. 369–371, Jul. 2004.

[6] K. L. Knutson, P. J. Rathouz, L. L. Yan, K. Liu, and D. S. Lauderdale, "Stability of the Pittsburgh sleep quality index and the Epworth sleepiness questionnaires over 1 year in early middle-aged adults: the CARDIA study," *Sleep (Rochester)*, vol. 29, no. 11, pp. 1503–1506, 2006.

[7] D. J. Buysse, C. F. R. III, T. H. Monk, S. R. Berman, and D. J. Kupfer, "The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research," *Psychiatry Research*, vol. 28, no. 2, pp. 193–213, 1989.

[8] V. Swarnkar, U. R. Abeyratne, C. Hukins, and B. Duce, "A state transition-based method for quantifying EEG sleep fragmentation," *Medical & Biological Engineering & Computing*, vol. 47, no. 10, pp. 1053–1061, 2009.

[9] A. Rechtschaffen and A. Kales, *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. US Government Printing Office, US Public Health Service, 1968.

[10] C. Iber, S. Ancoli-Israel, and A. Chesson, "The AASM manual for the scoring of and associated events: Rules, terminology and technical specifications," Tech. Rep., 2007.

[11] T. Penzel, N. Wessel, M. Riedl, J. W. Kantelhardt, M. Glos, and I. Fietze, "Cardiovascular and respiratory dynamics in patients with sleep apnea," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, Aug 31-Sep 4 2010, pp. 276–279.

[12] A. Kishi, Z. R. Struzik, B. H. Natelson, F. Togo, and Y. Yamamoto, "Dynamics of sleep stage transitions in healthy humans and patients with chronic fatigue syndrome," *American journal of physiology Regulatory integrative and comparative physiology*, vol. 294, no. 6, pp. R1980–R1987, 2008.

[13] M. J. Morrell, L. Finn, H. Kim, P. E. Peppard, M. S. Badr, and T. Young, "Sleep fragmentation, awake blood pressure, and sleep-disordered breathing in a population-based study," *American Journal of Respiratory and Critical Care Medicine*, vol. 162, no. 6, pp. 2091–2096, 2000.

[14] A. Borbély, "A two process model of sleep regulation," *Human Neurobiology*, vol. 1, no. 3, pp. 195–204, 1982.

[15] P. T. George, "Sleepiness, troika of consciousness cycle, and the Epworth Sleepiness Scale," *Sleep & breathing = Schlaf & Atmung*, vol. 5, no. 4, pp. 181–91, Dec. 2001.

[16] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and Physionet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–E220, 2000.

[17] B. Kemp, A. Zwinderman, B. Tuk, H. Kamphuisen, and J. Oberye, "Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG," *Biomedical Engineering, IEEE Transactions on*, vol. 47, no. 9, pp. 1185–1194, Sep. 2000.

[18] Y. Ichimaru and G. B. Moody, "Development of the polysomnographic database on CD-ROM," *Psychiatry and clinical neurosciences*, vol. 53, no. 2, pp. 175–7, Apr. 1999.

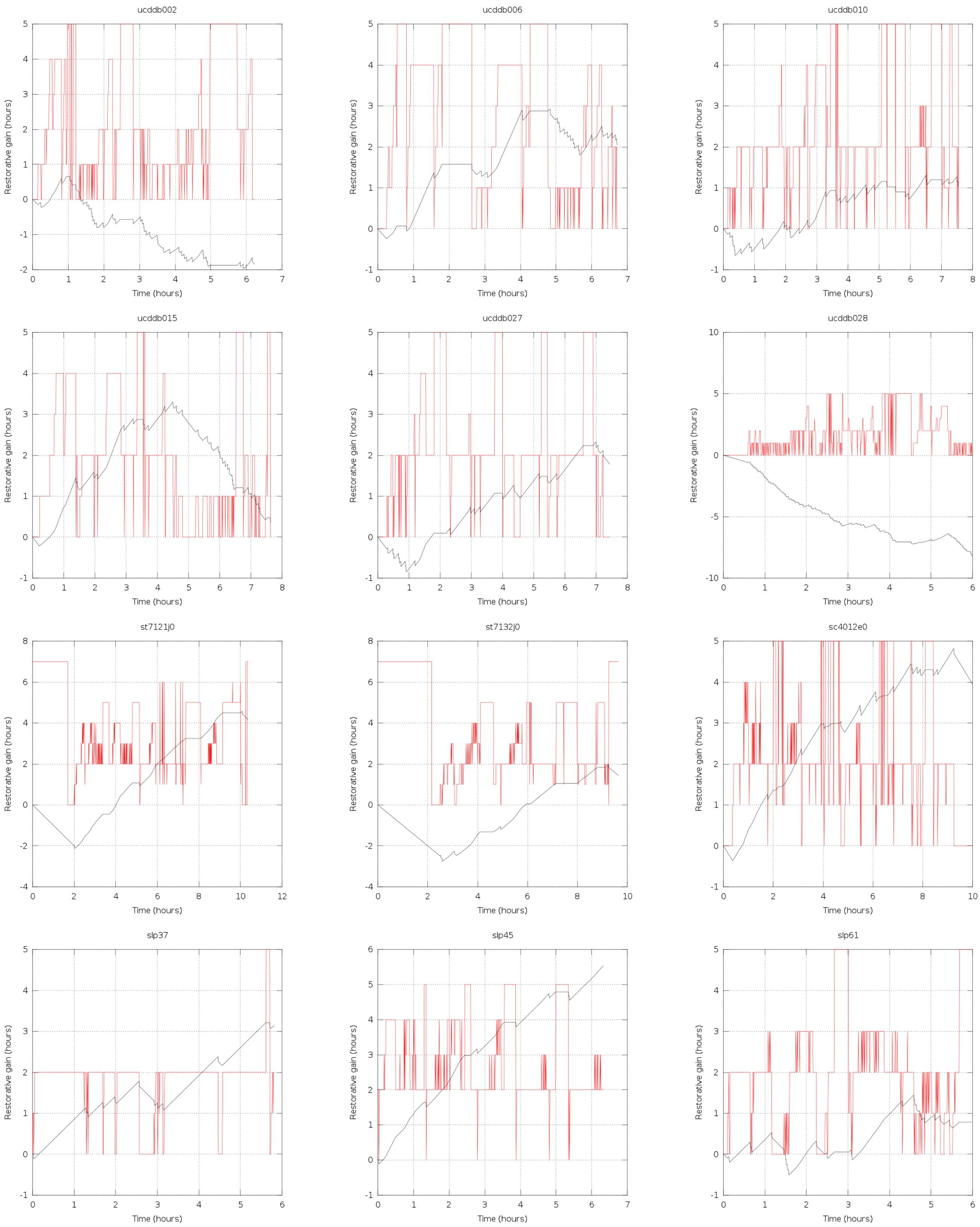

Fig. 1. Hypnograms (red) and the corresponding SRG curves (black). Hypnogram values: WAKE (0), S1 (1), S2 (2), S3 (3), S4 (4), REM (5), ARTIFACT (6), UNSCORED (7)

]