Evaluation of a wireless home sleep monitoring system compared to Polysomnography

Qiang Pan, Damien Brulin, Eric Campo

Evaluation of a wireless home sleep monitoring system compared to Polysomnography

Qiang Pan[1], Damien Brulin[1], Eric Campo[1]
[1]LAAS-CNRS, University of Toulouse, CNRS, UT2J, Toulouse, France
qpan@laas.fr,dbrulin@laas.fr,ecampo@laas.fr

Abstract

Sleep is essential for human health. Bad sleep and sleep disorders have been increasingly prevalent and are gradually becoming a social problem that cannot be ignored. The current gold standard in sleep monitoring is polysomnography (PSG) allowing nearly complete approach. Unfortunately, this wealth of information is obtained at the cost of invasive system, only usable in hospital environment under the control of sleep experts. Therefore, we develop a wireless body networks for home sleep monitoring with effort on non-intrusiveness, portability and autonomy. In this paper, we present our global architecture from sensors to user display with a focus on main functions and hardware. Then, we introduce the chosen indicators for sleep monitoring and the algorithms developed for sleep stages classification. Finally we show the evaluation of our approach compared to PSG. We illustrate the sleep stage classification during one night in the sleep unit of Toulouse University Hospital and highlight correlation between body temperature on extremities and Periodic Limb Movement during Sleep. Results are promising but need to be reinforced with new tests in hospital with several volunteers.

Keywords: Sleep monitoring system, classification, PSG

I. INTRODUCTION

Sleep is crucial for human health and quality of life. However poor sleep and sleep disorders are increasingly prevalent among the world's older population [1]. Health professionals consider sleep as an important indicator of health status, poor sleep quality is indeed likely to be a sign of many diseases [2]. Being able to monitor sleep is then a crucial issue in order to detect and to prevent sleep disorders.

Polysomnography (PSG) is considered as the gold standard for quantifying sleep time, differentiating sleep stages, and assessing sleep fragmentation [3]. As a global solution (EEG, ECG, EMG, EOG, oral-nasal airflow, body position, thoracic and abdominal movements, pulse oximetry, and limb movements) PSG provides comprehensive physiological information (Fig. 1). However, the PSG device has several drawbacks if one wants to use it for long-term in-home monitoring. Indeed, the PSG device is highly invasive for the patient (a minimum of 22 wires is required) and uncomfortable which may disturb sleep. Besides, the PSG can only be used in the hospital on one night due to the cost of this medical analysis and the need for experienced technologists to perform and analyze the recordings.

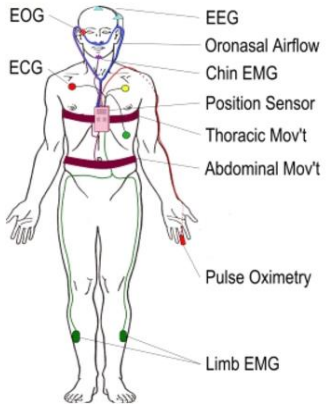

Figure 1. Polysomnography system [4]

Many sleep monitoring systems have been proposed to allow long-term monitoring with performance as close as possible to that of PSG. Readers are referred to our published review on the current status and future challenges of sleep monitoring systems [5] for more details. Given the state of the art in this research area, we proposed a new hardware and software architecture that enables long-term continuous sleep monitoring in a home environment. In order to evaluate the performance of our approach, we performed a one-night test in the sleep unit of the Toulouse University Hospital with a volunteer equipped with both our system and the PSG. Initially, we wanted to perform several tests with different users but due to the COVID pandemic, we only performed one night for one volunteer.

The paper is organized as follows. In section 2, we briefly detail the global architecture of our sleep monitoring system and the two classification algorithmic approaches we developed based on wrist movements. In section 3, we present the experimental conditions and discuss the first results comparing the decision of our classification algorithms to those of PSG. In section 4, we conclude the paper and propose improvements and future work.

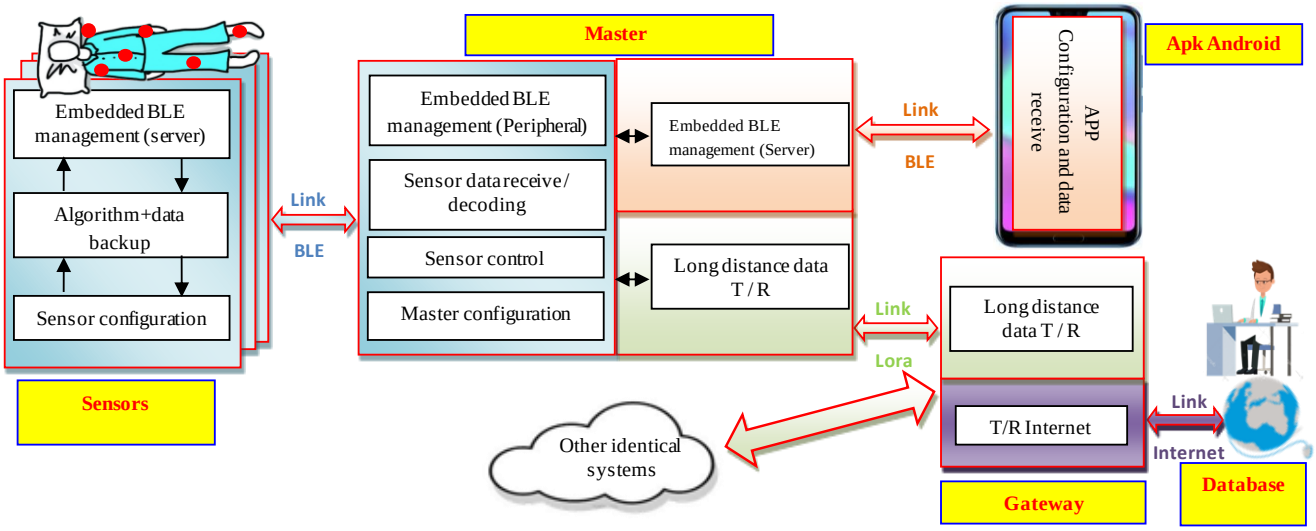

Figure 2. Hardware architecture

II. SLEEP MONITORING ARCHITECTURE

A. *Overview of the architecture*

The global architecture we proposed and developed to carry out sleep monitoring in a home environment is illustrated in Fig. 3.

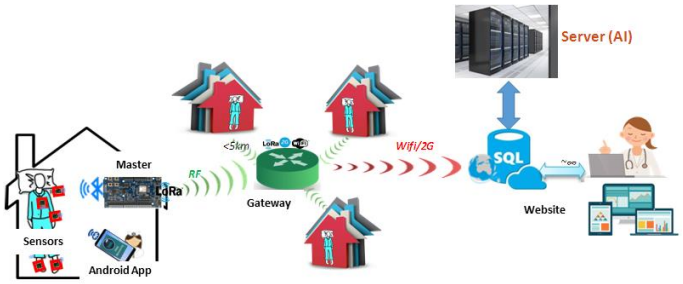

Figure 3. Overall architecture of the proposed sleep monitoring system

After discussions with researchers, technicians and doctors, we specified the requirements that the system should meet. We proposed and built a communicating wearable system that communicates in a network architecture, which can include several people at home or in the hospital. This system is fully configurable locally by the person concerned and also remotely by the doctor. Once configured, the system sends the monitoring data of each patient (at home or at the hospital) to a server-based database to be visualized and analyzed by the doctors on an adapted interface.

1) *Main functions and services*

The system is divided into several sub-systems: "Sensors", a "Master" board, a "Gateway", an Android application, a database and a website (Fig. 2).

Sensors

The system's data acquisition is performed by different sleep monitoring modules, each module using specific sensors to collect important physiological data related to sleep. The sensors we use are integrated in a miniaturized electronic board (Fig. 4) designed at LAAS-CNRS [6].

The board is a system-on-chip, connected, and powered by a button cell (3V). The main components are: (1) an NRF51822 microcontroller containing a 32-bit ARM Cortex M0 processor and a 256kB flash memory, equipped with a BLE V4 LE module; (2) a 16kB non-volatile FRAM memory for data backup during standby; (3) a low-power ADXL 362 triaxial accelerometer.

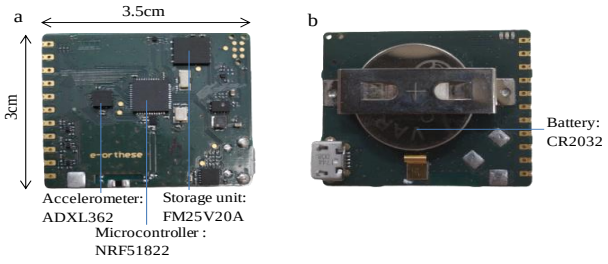

Figure 4. The sensor's basic electronic board. (a) Front side; (b) Back side

Master board

The master board is the control and data collection center of the proposed SMS which carries out five tasks:

- Reception of operating commands: the search of sensors (discovery phase), connection and disconnection of sensors (connecting phase) via BLE from a custom smartphone application.
- Reception of control commands to set sensor operating modes. Sensor operating modes include turn on, turn off and data transmission (data exchange).
- Reception and gathering data from sensors.
- Gathering ambient luminosity and temperature data from the sensors on the board.
- Sending all collected data to the gateway via the LoRa network.

Gateway

The gateway is the data transfer station of the system. It is connected to (1) the master board via LoRa in order to receive sleep monitoring data and (2) the WiFi network to transmit all collected data to a remote database. The gateway is useful when there is no internet connection possible at the patient's home.

2) Sleep monitoring modules

The sleep monitoring modules are the essential elements of the SMS we propose. They collect the raw data needed for sleep monitoring. We briefly present them below.

Wrist module

The wrist module consists of the basic electronic board, presented previously, and a temperature sensor (negative temperature coefficient, NTC). The module is worn on the non-dominant wrist like a watch. The temperature sensor is attached to the index finger with medical tape.

Chest module

The chest module consists of the basic electronic board and a temperature sensor (NTC). It is wrapped in soft paper and attached to the front of the chest with medical tape. This module is designed to measure chest temperature and detect the sleep positions.

Foot module

The foot module consists of two sub-modules: (1) one with the basic electronic board and a temperature sensor; (2) one with only the basic electronic board. The two sub-modules are wrapped in soft paper and attached to two insteps ((1) on the left instep, (2) on the right instep). The temperature sensor is attached to the big toe with medical tape. These modules are designed to measure feet movement and extremity temperature.

Sound module

The sound module collects sound data every second, using a MAX9814 microphone powered by two 03-2032 batteries to ensure continuous sound data collection throughout the night. It is placed next to the head within one meter during monitoring. This module is used to detect snoring.

Ambient module

The ambient module measures the temperature and luminosity of the sleeping environment. Both sensors are integrated in the master board. The temperature sensor is also a NTC sensor and the luminosity sensor is a TSL2591. The temperature and luminosity data are collected every minute.

B. Principle of algorithms

1) Choice of sleep indicators

Using the hardware architecture presented previously, we can compute relevant indicators for sleep monitoring, determined by referring to the Pittsburgh Sleep Quality Index (PSQI) [7] and the recommendations of the sleep experts of the sleep unit at the Center Hospital of Toulouse. These are sleep stages: sleep position, snoring, periodic leg movement index (PLMI), body temperature (finger, toe and chest) and ambient conditions (luminosity and temperature). Among these indicators, sleep stages, PLMI and temperatures are the most interesting considering sleep monitoring. Firstly, obtaining the time spent in the different sleep stages can provide better information to guide behavioral changes and recommendations to improve sleep quality [8]. Then PLMI can be used to predict Restless Legs Syndrome (RLS) which is a sensorimotor disorder that often has a profound impact on sleep [9]. The typical symptom of RLS is Periodic Limb Movement during Sleep (PLMS), so by detecting leg movements PLMI can be determined. Finally, sleep experts suggest that there may be a link between PLMS and extremity (finger and toe) temperature, but no one has yet investigated this hypothesis.

2) Sleep indicators computation

Sleep stages

We propose two approaches to automatically determine falling asleep/waking up times and sleep stages ("Awake", "Light sleep", "Deep sleep" and "REM" (Rapid Eye Movement also known as paradoxical sleep) [10]. The first one (called T1) is based on the thresholds of M_i movement levels (obtained by the 3-axis accelerometer of the wrist module, see Eq. 1) over periods PM of 19 epochs (30s for one epoch) considering 9 epochs before and 9 epochs after the current epoch. Indeed, as sleep is a constantly evolving process, it is necessary to associate the previous and following epochs when analyzing the sleep state at a given time. The second approach (called 5km2) is a k-mean based approach to perform clustering. The Thresholds approach is based on 3 thresholds

TS/W, TDLR, TLR to respectively discriminate epochs classified as sleep or awake, deep sleep or light sleep/REM and light sleep or REM. These thresholds are applied on PM values and standard deviation of PM values of 6-epoch group. The 5km2 method use a 2 dimension features composed of PM value and the standard derivation of PM values of 6-epoch group. It consists of 5 iterations ok k-means clustering with k=2. For each iteration, two clusters are determined, one corresponding to a sleep stages, the other one gathering others. More details on these two approaches are available in [11]

PLMI

According to standard criteria [12], PLMS are only considered if they are part of a series of four or more consecutive movements of 0.5-10 seconds duration with an inter-movement interval of 5-90 seconds and an amplitude greater than 8 mV above the basic an electromyograph (EMG) signal.

Based on the standard PLMS criteria, the PLMS detection rule using the foot module is defined as follows:

- The movement level M_i > 21 (see Eq. 1) is considered as the movement emergence.
- When the number of consecutive samples with M_i > 21 is between 1 and 10, it should be considered as a *movement group*.
- Adjacent movement groups with an interval between 5 to 90 seconds are considered as *significant movement groups*. The interval is from the end of the movement group to the beginning of the next movement group.
- A series of four or more consecutive significant movement groups will be considered *a PLMS group*, the number of significant movement groups being the number of PLMS in that PLMS group.

$$M_i = |Ax_{i+1} - Ax_i| + |Ay_{i+1} - Ay_i| + |Az_{i+1} - Az_i| \quad (1)$$

Body temperature

We define chest temperature as Tc_i, finger temperature as Tf_i, toe temperature as Tt_i. Then we calculate the sum of their respective differences, denoted $SD3T_i$ (see Eq. 2), where i is the sample index with temperature sampled every second.

$$SD3T_i = |Tc_i - Tf_i| + |Tc_i - Tt_i| + |Tf_i - Tt_i| \quad (2)$$

We also compute the first-order difference for finger and toe temperatures overnight, denoted DTf and DTt, respectively (see Eq. 3 and Eq. 4) as there appears to be some correlation between the onset of PLMS and temperature changes at the extremities.

$$DTf_i = |Tf_i - Tf_{i-1}| \quad (3)$$

$$DTt_i = |Tt_i - Tt_{i-1}| \quad (4)$$

III. EVALUATION OF THE PROPOSED SLEEP MONITORING SYSTEM WITH A PSG

A. *Description of the experimental protocol*

1) *Environmental conditions*

The test was performed in a sleep unit of a standard ward of the Center Hospital of Toulouse. The volunteer is equipped with the PSG system and our five sleep monitoring devices (Fig. 5a). The corresponding position of the PSG sensors and our modules (chest, wrist and foot module) on the body is illustrated in Fig. 5b. Generally speaking, the time required to wear the modules we proposed is roughly within ten minutes, and it can be done alone. In contrast, it takes about half an hour to install the PSG to the body, and it must be installed by at least one professionally trained medical staff member. The volunteer recruited was a 28-year-old man with a BMI (body mass index) of 18.3. We were only able to achieve one-night test for a primary evaluation of the performance as it has been carried out just before the first containment in France in March 2020.

2) *Synchronization protocol*

Each of the five sleep monitoring modules and the PSG has its own independent control unit and data acquisition. In practice, it is difficult to ensure that they would all start collecting data at the same time. Therefore, it is essential to synchronize the data collected by all the sleep monitoring devices and the PSG after all data have been collected in order to obtain meaningful results.

The synchronization method we use consists of two steps. First, the five sleep monitoring devices are synchronized with each other. Since each sleep monitoring device has an accelerometer that collects movement data, we intend to synchronize them by shaking them simultaneously and looking for the same marker in the collected movement data. The second step is the synchronization between the five sleep monitoring devices and the PSG. Since all five sleep monitoring devices were synchronized in the first step, it was only necessary to synchronize one of the sleep monitoring devices with the PSG in this step. Since both our foot device and the EMG of the PSG collect leg movement data, we intend to compare the waveforms of the leg movement data collected by the two devices to perform the synchronization between them (Fig. 6).

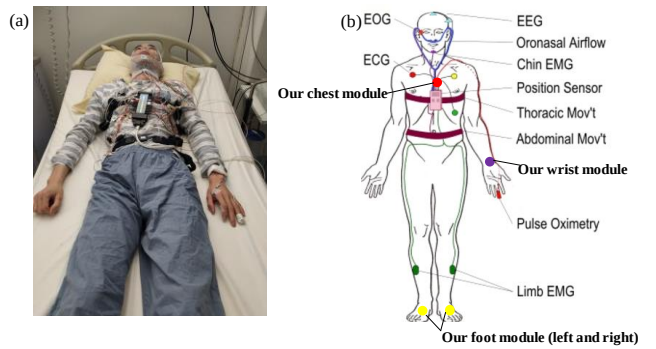

Figure 5. (a) Volunteer with the PSG and our wristband; (b) Schematic diagram of the corresponding location of the PSG sensors and our modules

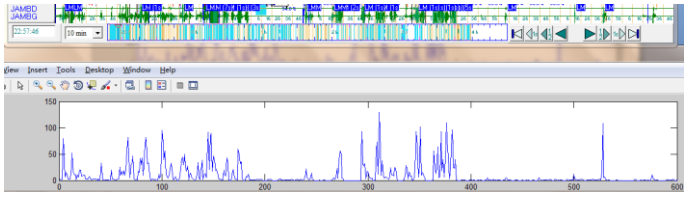

Figure 6. Synchronization between the EMG and the left foot module

B. *Performance evaluation of sleep indicators*

1) *Sleep stage classification*

We compare the hypnograms obtained by the T1 and 5km2 methods with the hypnogram obtained by the PSG as shown in Fig. 7. The hypnogram data of the methods we propose are obtained directly by operating the algorithms on Matlab by programming. The hypnogram of the PSG is read by the software "DeltaFree EEG reader". It is important to note that the software can automatically generate an initial hypnogram and various events from the data collected by the PSG. However, they are not completely accurate and must be manually checked and corrected by a physician.

We have compared the sleep stage classification results of the two proposed methods with the PSG, epoch by epoch. Table I and Table II illustrate the confusion matrices between T1 and PSG, and between 5km2 and PSG, respectively. From a physiological significance point of view, deep sleep is very different from awake and light sleep. Therefore, confusion between deep sleep and awake, and confusion between deep sleep and light sleep can be considered serious.

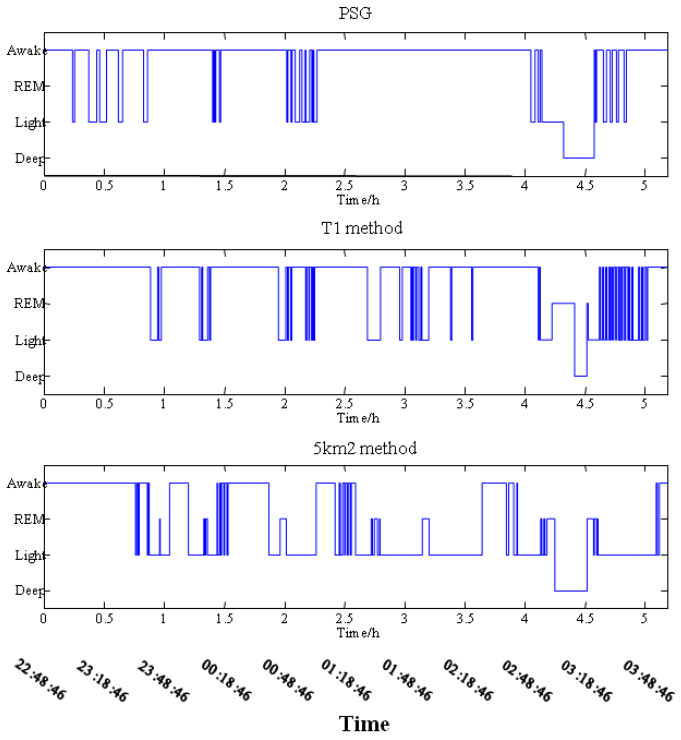

Figure 7. Hypnograms obtained with the PSG, T1 and 5km2 methods

TABLE I. Confusion matrix between T1 and PSG

		Predicted (T1 method)			
		Awake	***Light Sleep***	***Deep Sleep***	***REM***
True (PSG)	***Awake***	423	85	0	0
	Light sleep	51	21	0	11
	Deep sleep	0	6	12	13
	REM	0	0	0	0

TABLE II. Confusion matrix between 5km2 and PSG

		Predicted (5km2 method)			
		Awake	***Light Sleep***	***Deep Sleep***	***REM***
True (PSG)	***Awake***	211	272	0	25
	Light sleep	27	38	8	10
	Deep sleep	0	1	24	6
	REM	0	0	0	0

To evaluate more precisely the agreement between the two proposed methods and the sleep stages classification using the PSG method, Cohen's Kappa coefficient (κ) is calculated. Landis & Koch [13] characterized $\kappa < 0$ as indicating no agreement and 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 as almost

perfect agreement [13]. The T1 method showed fair agreement with the PSG (κ = 0.24), and the 5km2 method showed slight agreement with the PSG (κ = 0.09).

TABLE III. Evaluation indexes for the T1 and 5km2 methods

Index	Method	Awake	REM	Light Sleep	Deep Sleep
Sensitivity	T1	0.83	0.00	0.25	0.39
	5km2	0.42	0.00	0.46	0.77
Specificity	T1	0.55	0.96	0.83	1.00
	5km2	0.76	0.93	0.49	0.99
Accuracy	T1	0.78	0.96	0.75	0.97
	5km2	0.48	0.93	0.49	0.98
Precision	T1	0.89	0.00	0.19	1.00
	5km2	0.89	0.00	0.12	0.75
Balanced accuracy	T1	0.66	0.50	0.53	0.98
	5km2	0.56	0.50	0.49	0.87
F1-score	T1	0.86	0.00	0.22	0.56
	5km2	0.57	0.00	0.19	0.76

Six performance assessment indexes based on the confusion matrix are calculated and presented in Table III. In our experiment, the number of samples included in the different classes is uneven and usually varies greatly. At the same time, we consider the correct detection of positive and negative samples should be of equal importance. Therefore, among all the performance assessment indexes, we believe that the balanced accuracy is the best one to evaluate the overall performance of the proposed methods. Method T1 has the highest or equal balanced accuracy in each class.

T1 seems to be the best method but it will have to be confirmed by other tests involving several people. Indeed, we carried out only one night test due to the COVID pandemic and the volunteer had many difficulties to fall asleep which explain that the recording only lasts 5 hours with no REM phases. However, as a preliminary test, we can see that our approaches, in particular T1, have promising results, closed to those of PSG, with fewer sensors and greater comfort for the user. The K-mean approach needs more data to provide better results.

2) PLMS detection performance

The number of PLMS per hour during sleep detected by this rule is defined as the PLMS index (PLMI), which is the diagnostic indicator for PLMS based on the foot module. The number of PLMS distributed in each sleep stage detected by the PSG and our left foot module is shown in Table IV.

TABLE IV. Number of PLMS distributed in each sleep stage

	Total	Awake	REM	Light Sleep	Deep Sleep
PSG ref	56	47	0	6	3
LAAS solution	57	48	0	9	0

As we can see, the total number of PLMS given by the PSG report is very close to the result of our foot module. Furthermore, the number of PLMS distributed at each sleep stage obtained with our foot module is also very close to the PSG result. The main difference between the two is the PLMS distribution in light and deep sleep. The PSG detects 3 PLMS in deep sleep but our foot module detects no PLMS in deep sleep. In light sleep, our foot module detects 3 more PLMS than the PSG. The reason why our foot module does not detect any PLMS during deep sleep may be that the limb movement is very slight during deep sleep so the movement level of some foot movements does not reach the threshold of foot movement. Therefore, PLMS is not detected during deep sleep. In addition, during this night, the duration of deep sleep is also very short, which also increases the difficulty of detecting PLMS.

3) Link between temperature, sleep stages and PLMS

Body temperature and sleep stages

The synchronous comparison between SD3T and the PSG hypnogram is shown in Fig. 8. In Fig. 8a, the dashed box corresponds to the period when sleep is continuous and most of these periods are deep sleep. Similarly, in Fig. 8b, the red dashed box shows the period when the SD3T remains low and stable. This suggests that a stable and low SD3T may correspond to continuous restful sleep.

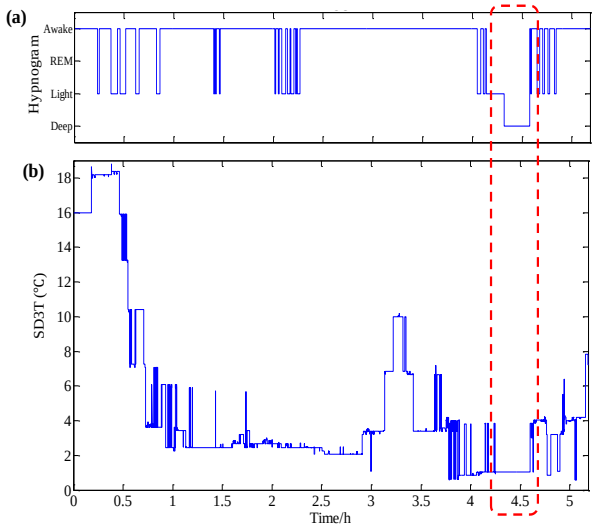

Figure 8. (a) Hypnogram obtained by the PSG. (b) SD3T

Body temperature and PLMS

The synchronous comparison between DTt, DTf and PLMS overnight is shown in Fig. 9.

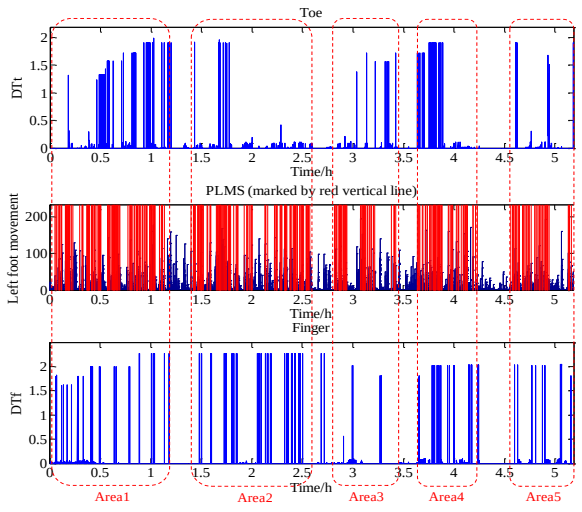

Figure 9. Synchronous comparison between DTt, DTf and PLM overnight

By observing Fig. 9, the distribution of PLMS is lowly correlated with DTt but highly correlated with DTf. In Fig. 9, five relatively independent groups can be found by observation, so we divide the night time into 5 areas, each area containing a relatively concentrated set of PLMS events. Each area contains a relatively concentrated set of high DTf values, i.e., a relatively large change in finger temperature. In addition, the density of the PLMS distribution and the density of DTf are similar in each area. In areas 1, 2 and 4, the PLMS emergence is very dense, and the emergence of the high DTf value is also relatively dense. In areas 3 and 5, the PLMS emergence is relatively low, and the emergence of high DTf value is also relatively low. In part of areas 1 and 2, part of areas 3 and 4, part of areas 4 and 5, there is no PLMS emergence, and no high DTf value.

Based on these phenomena we can assume that PLMS emergence is positively correlated with DTf value. PLMS is maybe more correlated with finger temperature than with toe temperature.

IV. CONCLUSION AND PERSPECTIVES

In this paper, we presented a global home sleep monitoring system (from sensors to user display) as an alternative solution to PSG for long-term monitoring. We have performed a preliminary validation of two methods proposed for the sleep stages classification with reference to the PSG gold standard. Based on the confusion matrix analysis, the results show that the proposed threshold approach T1 method has a fair agreement with the PSG while k-mean based approach has a slight agreement with the PSG. The T1 method is efficient for the detection of awake and deep sleep in particular. However, all the proposed methods are relatively less efficient for the detection of REM and light sleep. In general, the T1 method is the most efficient among the two proposed methods. For PLMS detection, we define the detection rules based on the foot movement data acquired by our proposed foot module. The results show that the total number of PLMS and the number of PLMS distributed in each sleep stage detected by our foot module are both very close to the PSG. Furthermore, we explore the links between body temperature and hypnogram and between body temperature on extremities and PLMS. We have found that the lower and flat continuous SD3T corresponds to continuous sleep and even deep sleep, that the emergence of PLMS is positively correlated to the DTf value and that PLMS is more correlated to finger temperature than to toe temperature. This experiment has shown that it would be possible to predict PLMS based on the change in finger temperature. Nevertheless, further investigative work over several nights and several subjects should be carried out to confirm these first observations. Indeed, due to COVID-19, we only performed one night with one volunteer who did not bear well all PSG devices and so had difficulties to fall asleep.

ACKNOWLEDGMENT

We would like to thank Dr. Debs and all the medical staff of the sleep unit of Purpan Toulouse Hospital who allowed us to carry out the evaluation test of our approach.

REFERENCES

[1] Van de Straat V, Buffel V, Bracke P. Medicalization of Sleep Problems in an Aging Population: A Longitudinal Cross-National Study of Medication Use for Sleep Problems in Older European Adults. Journal of Aging and Health; 2018

[2] Dregan A, Armstrong D. Cross-country variation in sleep disturbance among working and older age groups: an analysis based on the European Social Survey. International psychogeriatrics, 2011, 23(9): 1413.

[3] Jafari, B., & Mohsenin, V. (2010). Polysomnography. Clinics in chest medicine, 31(2), 287-297.

[4] Montplaisir J, Boucher S, Gaétan Poirier, et al. Clinical, polysomnographic, and genetic characteristics of restless legs syndrome: A study of 133 patients diagnosed with new standard criteria. Movement Disorders, 1997, 12(1):61-65

[5] Pan Q, Brulin D, Campo E. Current Status and Future Challenges of Sleep Monitoring Systems: Systematic Review. JMIR Biomedical Engineering, 2020, 5(1): e2092.

[6] Charlon Y., Conception de dispositifs électroniques portés pour le suivi de l'état de santé des personnes âgées, Toulouse, LAAS-CNRS, 2014.

[7] Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry research, 28(2), 193-213.

[8] Daskalova N, Lee B, Huang J, et al. Investigating the effectiveness of cohort-based sleep recommendations. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(3): 1-19.

[9] Walters AS International Restless Legs Study Group. (1995). Toward a better definition of the restless legs syndrome. Mov Disord, 10, 634-42.

[10] Altevogt, B. M., & Colten, H. R. (Eds.). (2006). Sleep disorders and sleep deprivation: an unmet public health problem.

[11] Pan Q., Brulin D., Campo E. Home sleep monitoring based on wrist movement data processing//10th International Conference of Information and Communication Technology (ICICT-2020). November 13-15th, 2020.

[12] Michaud M, Paquet J, Lavigne G, et al. Sleep laboratory diagnosis of restless legs syndrome. Eur Neurol 2002;48:108-113.

[13] Landis, J. R., Koch, G. G. The measurement of observer agreement for categorical data. biometrics, 1977, 159-174.