Accuracy of the Emfit and Empatica to detect activity pattern

Introduction

Wearable biosensors have become increasingly prevalent in our society, offering valuable insights into personal health metrics such as physical activity (Aroganam et al., 2019) and menstrual cycles (Alzueta et al., 2022; Lang et al., 2024; Maijala et al., 2019) as well as aiding in the detection of diseases and mental disorders (Dai et al., 2023). Of particular interest is their application in identifying sleep disorders (Kuo et al., 2017; Scott et al., 2020), which afflict approximately one third of the population (Pavlova & Latreille, 2019). To effectively meet this purpose, it is essential to create an adequate device that could be used daily (Seneviratne et al., 2017). Therefore, research in this field strives to strike a balance between affordability and accuracy, while ensuring that biosensors yield ecologically valid data without causing disruption.

Consensually, the gold standard for assessing sleep activity is polysomnography (PSG) (Rundo & Downey, 2019), which incorporates various sensors such as electroencephalography (EEG), electrocardiography (ECG), electromyography (EMG), and accelerometry. This array of sensors enables highly precise measurements, yet its complexity and cost make it primarily feasible within clinical settings, notably hospitals (Perez-Pozuelo et al., 2020). Moreover, the abundance of sensors and cables may disrupt sleep patterns, compromising the ecological validity of the data (Massar et al., 2021). Hence, there is a pressing need for alternative wearable sensors that offer affordable solutions while prioritizing comfort and qualitative measurements.

In the exploration for alternatives, wearable biosensors undergo rigorous accuracy testing against the gold standard, polysomnography (PSG). Some companies innovate with prototypes like the Oura ring. This wearable, despite its discreet and comfortable design as a ring, integrates various technologies including an accelerometer and autonomic nervous system-mediated peripheral signals, alongside circadian monitoring capabilities. Research on its accuracy revealed a 73% capability to categorize the four sleep stages compared to PSG (Altini et al., 2021). Conversely, a study on four different wearables (wristbands/watches) and three nearable devices yielded inconsistent results, indicating challenges in categorizing the four sleep stages but proficient detection of sleep and wakefulness states (Chinoy et al., 2022).

While these results show promise, there remains room for improvement. Additionally, studies indicate a significant decrease in performance of wearable and nearable devices when used by individuals with sleep disorders (Chinoy et al., 2022; Concheiro-Moscoso et al., 2023). This poses a significant challenge as these devices are intended to detect such disorders and provide actionable insights. Consequently, numerous studies are now directed towards enhancing performance for populations affected by specific disorders. In this paper, our focus will be on Periodic Limb Movement in Sleep (PLMS), a neurological disorder characterized by involuntary and recurrent movements of the limbs, typically the legs, during sleep. PLMS is often linked with sleep interruptions and micro-arousals, though affected individuals may not be consciously aware of them. Symptoms include excessive daytime fatigue, drowsiness, difficulty falling asleep, and frequent nighttime awakenings (Haba-Rubio et al., 2016; Hornyak et al., 2006). These symptoms significantly impact quality of life and daily functioning, underscoring the crucial need for accurate and reliable detection of this disorder, particularly through wearable sensors.

As an example, a sock crafted from smart textile was devised to gather data from the legs (Eguchi et al., 2019). The advantage lies in the wearer's convenience; they can wear it effortlessly like a regular sock, eliminating the need for prior knowledge about sensor placement. Promising results ensued, with accurate detection of muscular activation. However, it remains a case study, necessitating further research. Another study employed a motion sensor affixed to the foot (Kye et al., 2017). They then amalgamated the decomposed x, y, z axis data obtained from the accelerometer. This enabled the derivation of an activation pattern used to attempt PLMS detection in the wearer, employing classification algorithms such as a neural network, support vector machine, and k-nearest neighbors. Remarkably, the k-nearest neighbors algorithm achieved the highest classification accuracy at 97%. Overall, every measurement technique lacks validity or customizability, necessitating further research to address this gap.

In this paper, we broaden the exploration of qualitative wearable biosensors to include the Emfit and the Empatica. Emfit is a nearable device designed to be placed under the mattress, facilitating the measurement of movement in bed (Rauhala et al., 2009). Meanwhile, Empatica is a wristband equipped with motion (accelerometer) and electrodermal activity sensors (Regalia et al., 2019). The objective of this study is to compare these devices with the activity patterns derived from PSG and evaluate their accuracy through correlation analysis. To achieve this goal, we will utilize existing data.

Method

The data analyzed in this paper was collected in a previous study. In this original research, they gathered 115 participants from which 62 (54%) were females. They were aged from 18 to 85 years old with a mean of 47. It was assessed if the participant had any known sleep disorder. 55 individuals were diagnosed with Obstructive Sleep Apnea, 7 with Central Disorder of Hypersomnolence, 7 with Parasomnias, 6 with Insomnias, and 36 who had not yet been assigned to a specific disorder. From this original dataset we received the data from 71 participants. For the analysis of Emfit 05 and 19, we used the data of 36 of them that met the criteria for inclusion. The others had either missing measures or significant errors within their dataset. For Empatica, we included 57 participants, having to exclude the others for the same flaws in their dataset.

For the collection of data, the participants had to come for a few nights at laboratory where several measurements were taken during their sleep. As baseline, the polysomnography from SOMNOmedics was used with a sample rate of 4 Hz. The data saved as an .edf file entailed signals from EEG, ECG, EMG, Accelerometer, and measures of oxygen level. Additionally, the participants had to wear a Fitbit Inspire 2 which measures activity/calories, the heartrate, and the sleep stages with minute resolutions. They also wore the E4 Empatica wristband which is made of 4 sensors with a sample rate of 32 Hz: the photoplethysmography (PPG) for the blood volume pulse, the electrodermal activity (EDA) for the measure of the sympathetic nervous system arousal, the 3-axis accelerometer to capture motion-based activity, and lastly the infrared thermopile to monitor the skin temperature. Furthermore, the EMFIT QS consists of two quazi-piezoelectric sensors that need setup under the mattress in the bed, with a sample rate of 0,25 Hz. These sensors react to dynamic forces caused by heart contractions, breathing and body movement. The first mat (labeled emfit 05) is put behind the torso, while the second (labeled Emfit 19) is put below the legs of the participants. The signals from each device were sent to a local router that then saved the anonymized data on a cloud. The data was later accessed and processed using Python 3.11.5.

From the signals were derived many measures like the blood volume pulse, the accelerometer, the skin temperature and the electrodermal activity from which the activity patterns, the heartrate, and ultimately the sleep stages could be derived. In this paper, we will focus on using the activity patterns collected from the PSG, from the two Emfit sensors and from the Empatica wristband. We will first visualize the activity pattern, then normalize them to be able to superpose them on the PSG golden-standard, and finally we will analyze the cross-correlation of each device with the PSG to assess their accuracy.

Results

In this results section, we will analyze and process the raw data of the patients with the goal in mind to assess the accuracy of each device compared to the PSG golden-standard through cross-correlation. For illustration of the analysis, we will showcase the process on the data from the participant 029 and we will extrapolate to the whole dataset when required. We first visualize the overall activity pattern by plotting the raw data from one of the devices (the Empatica) to be able to select the time window of interest.

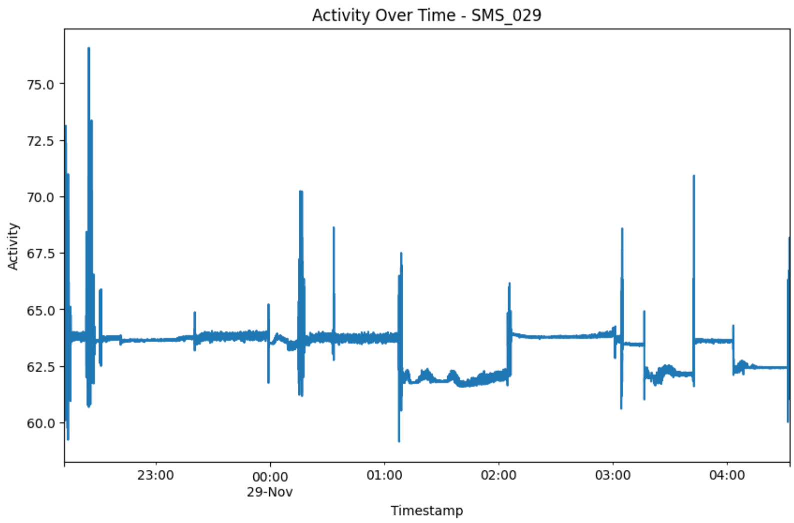

Figure 1. The activity pattern derived form the Empatica device against time

Next, to be able to compare the activity patterns of the devices, we normalize each one of them to be bounded between 0 and 1. In that way, we normalized the data for PSG (Figure 2a), Empatica(Figure 2b), Emfit 05 (Figure 2c) and Emfit 19 (Figure 2d).

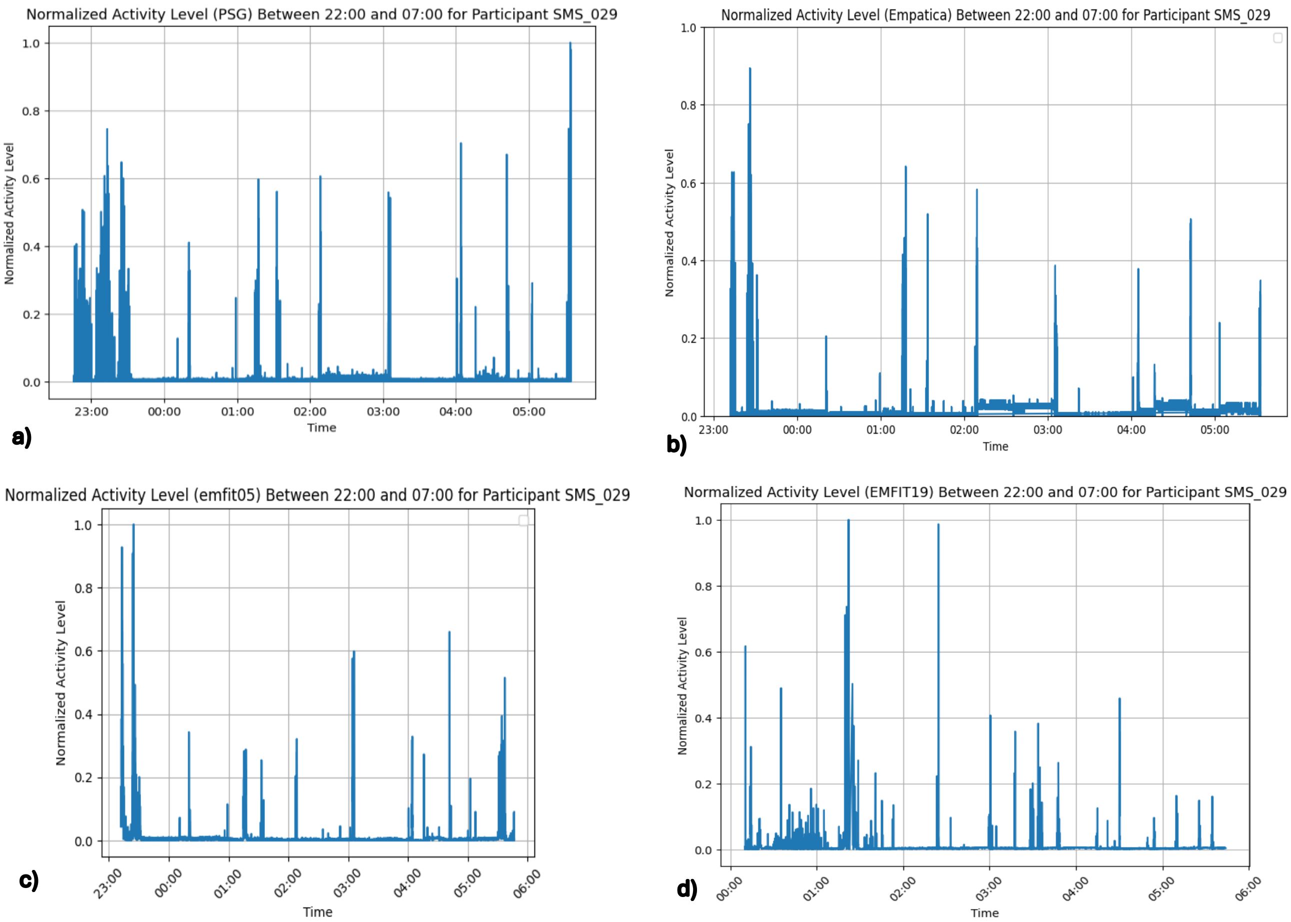

Figure 2. The normalized activity pattern for a) PSG, b) Empatica, c) Emfit 05 and d) Emfit 19 against time

As the next step, we superpose the activity pattern of each device with the activity pattern from the PSG. This allows us to compare visually the accuracy of the different devices and to see if the cross-correlation analysis is feasible (Figure 3a, 3b and 3c).

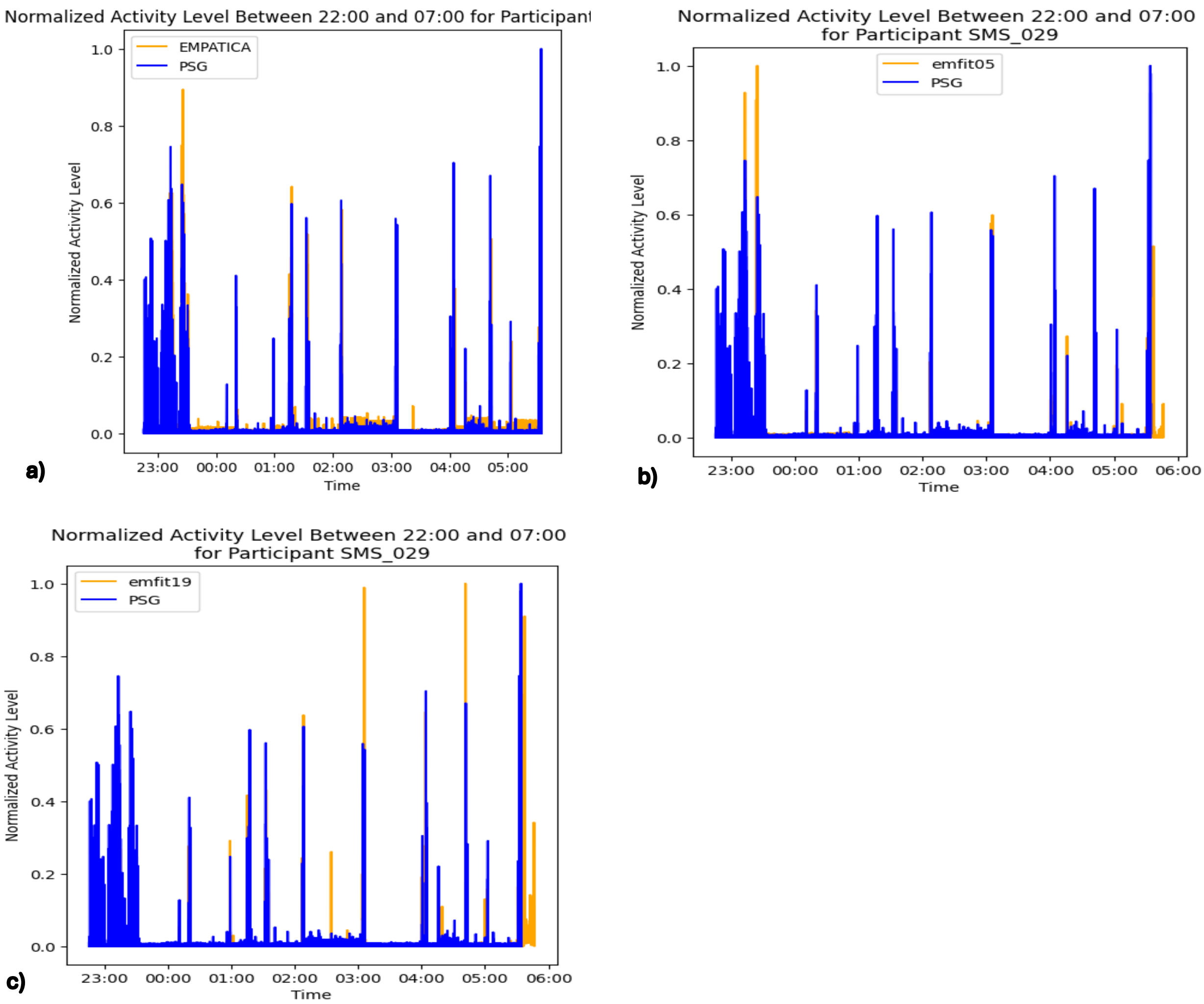

Figure 3. The superposed normalized activity patterns of a) Empatica, b) Emfit 05 and c) Emfit 19 over PSG against time

Finally we proceed with the cross-correlation analysis for each device. The process is the following. First we apply a standard cross-correlation algorithm from the library scipy on the normalized data of each device against the PSG. This shifts the signal of the device, while extending it with padding to match the accurate length of the PSG signal. Then we align the data with regard to the x-axis by looking at the highest point of the cross correlation analysis. This gives us a correlation without alignment and a correlation after alignment that we averaged for all patients to a mean and standard-deviation. For the Emfit 05 (Figure 4a), we obtained a correlation without alignment of $\mu = 0{,}26$, $\sigma = 0{,}13$ and a cross-correlation after alignment of $\mu = 0{,}99$, $\sigma = 0{,}05$. For the Emfit 19 (Figure 4b), we obtained a correlation without alignment of $\mu = 0{,}32$, $\sigma = 0{,}17$

and a cross-correlation after alignment of μ = 0,97, σ = 0,09. Finally, for the Empatica (Figure 4c), we had some issue with the lags of the correlation without alignment which means we have only the plot to visualise. However we have the cross-correlation after alignment of μ = 0,58, σ = 0,13.

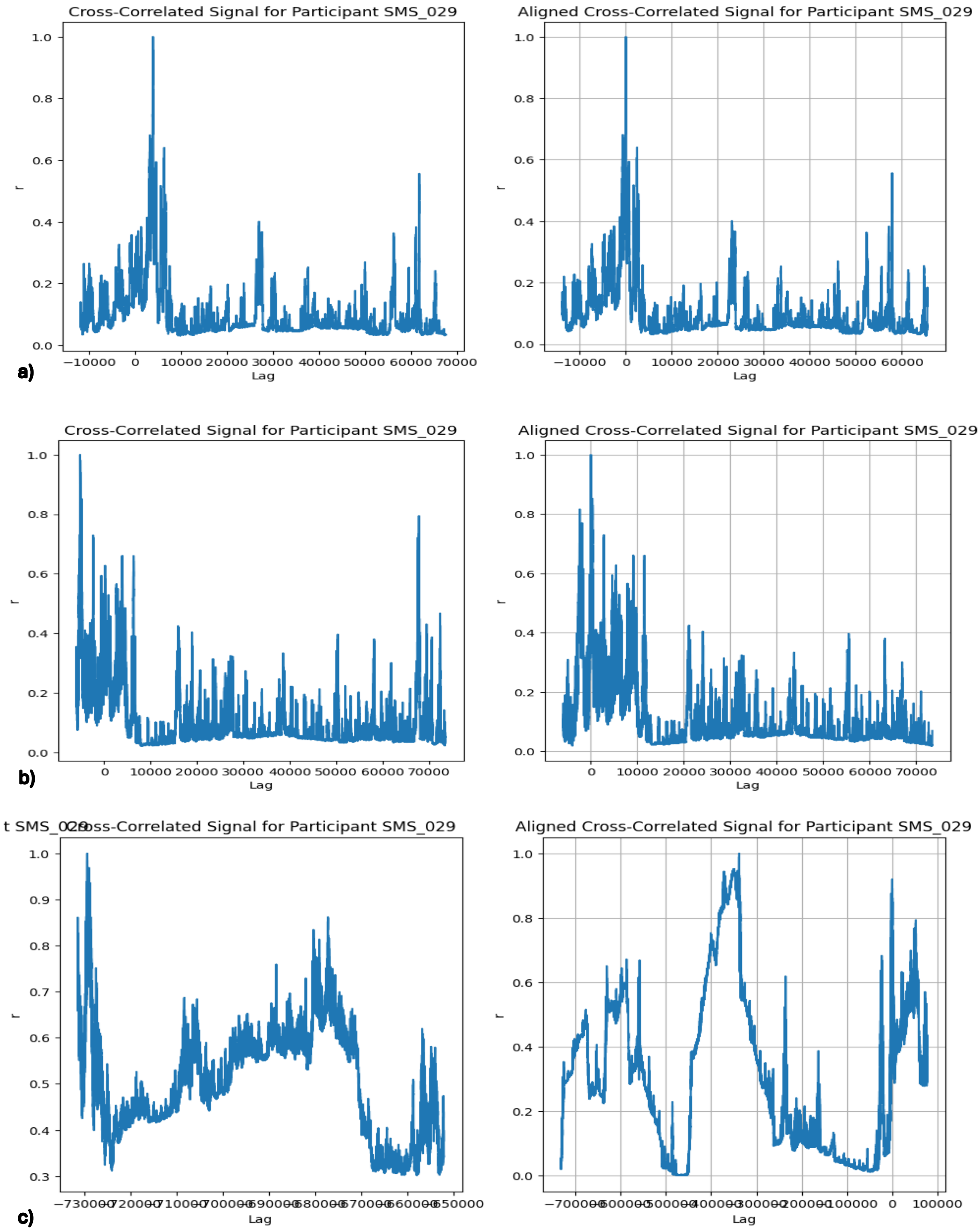

Figure 4. Cross-correlation with PSG: without and after alignment for a) Emfit 05, b) Emfit 19, and c) Empatica

Discussion

The prospect of this paper is to add to the research of cheap, discreet, but accurate measuring devices to detect sleep disorders like the restless leg syndrome. We examined two devices: the EMFIT QS, which has two quazi-piezoelectric sensors placed behind the torso (05) and below the legs (19), and the Empatica E4 wristband. We compared their activity patterns with the gold standard PSG (SOMNOmedics) using cross-correlation.

The results show a low mean correlation without alignment for the Emfit 05 and for the Emfit 19. But as we align the signals to maximize the correlation, we obtain a very high ($r < 0{,}95$) mean correlation for both Emfit 05 and Emfit 19. For the Empatica, we have no data for the correlation without alignment, but we obtain a moderate ($r \approx 0{,}50$) mean correlation after the alignment. From our analysis, we can conclude that the Empatica E4 can detect the activity patterns to some extent, but that the accuracy is unsatisfactory. On the other hand, the Emfit quazi-piezoelectric sensors, with their high correlation, seem very accurate for detection of activity patterns during sleep. This makes them ideal for data collection of PLMS for detection of the restless leg syndrome through machine learning algorithm (Athavale et al., 2019).

Limitations of our analysis include primarily the unusually high participants exclusion due to missing measures or flawed datasets, possibly from device malfunctions, sensor setup errors, or data processing mistakes. Without more metainformation about the data, it is hard to pinpoint the source of the problem. Secondly, the correlations we obtain for the Emfit sensors before alignment seem very low and the correlations after alignment seem especially high. It is possible that accuracy of the sensors is indeed very good as we can visually see in the superposition plots (Figure 3), but it could also be due to a mistake made during the cross-correlation analysis (only thing I didn't do myself in this script). Literature supports Emfit's high accuracy (Rauhala et al., 2009), but the abnormal results of this paper warrant replication. Similar caution applies with the rather low results of the Empatica wristband.

In conclusion, it seems that the EMFIT QS provides very accurate measures of the activity patterns during sleep and is thus very promising towards detecting PLMS cases. On the other hand, the Empatica E4 showed poorer results, questioning its suitability to achieve the same purpose. Due to high limitations of this paper, it is advisable to carefully interpret these results and reproduction of the experimental paradigm is recommended.

Bibliography

Altini, M., Altini, M., Altini, M., Altini, M., Kinnunen, H., Kinnunen, H., & Kinnunen, H. (2021). The Promise of Sleep : A Multi-Sensor Approach for Accurate Sleep Stage Detection Using the Oura Ring. *Sensors*. https://doi.org/10.3390/s21134302

Alzueta, E., de Zambotti, M., Javitz, H., Dulai, T., Albinni, B., Simon, K. C., Sattari, N., Zhang, J., Shuster, A., Mednick, S. C., & Baker, F. C. (2022). Tracking Sleep, Temperature, Heart Rate, and Daily Symptoms Across the Menstrual Cycle with the Oura Ring in Healthy Women. *International Journal of Women's Health*, *14*, 491-503. https://doi.org/10.2147/IJWH.S341917

Aroganam, G., Manivannan, N., & Harrison, D. (2019). Review on Wearable Technology Sensors Used in Consumer Sport Applications. *Sensors*, *19*(9), Article 9. https://doi.org/10.3390/s19091983

Athavale, Y., Krishnan, S., Raissi, A., Kirolos, N., Jairam, T., Murray, B. J., & Boulos, M. I. (2019). Actigraphic detection of periodic limb movements : Development and validation of a potential device-independent algorithm. A proof of concept study. *Sleep*, *42*(9), zsz117. https://doi.org/10.1093/sleep/zsz117

Chinoy, E. D., Chinoy, E. D., Cuellar, J. A., Cuellar, J. A., Jameson, J., Jameson, J. T., Markwald, R. R., & Markwald, R. R. (2022). Performance of Four Commercial Wearable Sleep-Tracking Devices Tested Under Unrestricted Conditions at Home in Healthy Young Adults. *Nature and Science of Sleep*. https://doi.org/10.2147/nss.s348795

Concheiro-Moscoso, P., Groba, B., Álvarez-Estévez, D., Miranda-Duro, M. del C., Pousada, T., Nieto-Riveiro, L., Mejuto-Muiño, F. J., & Loureiro, J. P. (2023). Quality of Sleep Data Validation From the Xiaomi Mi Band 5 Against Polysomnography : Comparison Study. *null*. https://doi.org/10.2196/42073

Dai, R., Kannampallil, T., Kim, S., Thornton, V., Bierut, L., & Lu, C. (2023). Detecting Mental Disorders with Wearables : A Large Cohort Study. *Proceedings of the 8th ACM/IEEE*

Conference on Internet of Things Design and Implementation, 39-51. https://doi.org/10.1145/3576842.3582389

Eguchi, K., Nambu, M., Kamikawa, T., Ueshima, K., & Kuroda, T. (2019). Smart Textile Device with Embedded Fabric Electrodes Targeting Periodic Limb Movements Monitoring at Home : A Case Report. *Journal of Fiber Science and Technology*, *75*(11), 164-180. https://doi.org/10.2115/fiberst.2019-0020

Haba-Rubio, J., Marti-Soler, H., Marques-Vidal, P., Tobback, N., Andries, D., Preisig, M., Waeber, G., Vollenweider, P., Kutalik, Z., Tafti, M., & Heinzer, R. (2016). Prevalence and determinants of periodic limb movements in the general population. *Annals of Neurology*, *79*(3), 464-474. https://doi.org/10.1002/ana.24593

Hornyak, M., Feige, B., Riemann, D., & Voderholzer, U. (2006). Periodic leg movements in sleep and periodic limb movement disorder : Prevalence, clinical significance and treatment. *Sleep Medicine Reviews*, *10*(3), 169-177. https://doi.org/10.1016/j.smrv.2005.12.003

Kuo, C.-E., Liu, Y.-C., Chang, D.-W., Young, C.-P., Shaw, F.-Z., & Liang, S.-F. (2017). Development and Evaluation of a Wearable Device for Sleep Quality Assessment. *IEEE Transactions on Biomedical Engineering*, *64*(7), 1547-1557. https://doi.org/10.1109/TBME.2016.2612938

Kye, S., Moon, J., Lee, T., Lee, S., Lee, K., Shin, S., & Lee, Y. S. (2017). Detecting periodic limb movements in sleep using motion sensor embedded wearable band. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1087-1092. https://doi.org/10.1109/SMC.2017.8122756

Lang, A.-L., Bruhn, R.-L., Fehling, M., Heidenreich, A., Reisdorf, J., Khanyaree, I., Henningsen, M., & Remschmidt, C. (2024). Feasibility Study on Menstrual Cycles With Fitbit Device (FEMFIT) : Prospective Observational Cohort Study. *JMIR mHealth and uHealth*, *12*(1), e50135. https://doi.org/10.2196/50135

Maijala, A., Kinnunen, H., Koskimäki, H., Jämsä, T., & Kangas, M. (2019). Nocturnal finger skin temperature in menstrual cycle tracking : Ambulatory pilot study using a wearable Oura ring. *BMC Women's Health*, *19*(1), 150. https://doi.org/10.1186/s12905-019-0844-9

Massar, S. A. A., Chua, X. Y., Soon, C. S., Ng, A. S. C., Ong, J. L., Chee, N. I. Y. N., Lee, T. S., Ghosh, A., & Chee, M. W. L. (2021). Trait-like nocturnal sleep behavior identified by combining wearable, phone-use, and self-report data. *Npj Digital Medicine*, *4*(1), 1-10. https://doi.org/10.1038/s41746-021-00466-9

Perez-Pozuelo, I., Zhai, B., Palotti, J., Mall, R., Aupetit, M., Garcia-Gomez, J. M., Taheri, S., Guan, Y., & Fernandez-Luque, L. (2020). The future of sleep health : A data-driven revolution in sleep science and medicine. *Npj Digital Medicine*, *3*(1), 1-15. https://doi.org/10.1038/s41746-020-0244-4

Rauhala, E., Virkkala, J., & Himanen, S.-L. (2009). Periodic limb movement screening as an additional feature of Emfit sensor in sleep-disordered breathing studies. *Journal of Neuroscience Methods*, *178*(1), 157-161. https://doi.org/10.1016/j.jneumeth.2008.11.019

Regalia, G., Onorati, F., Lai, M., Caborni, C., & Picard, R. W. (2019). Multimodal wrist-worn devices for seizure detection and advancing research : Focus on the Empatica wristbands. *Epilepsy Research*, *153*, 79-82. https://doi.org/10.1016/j.eplepsyres.2019.02.007

Rundo, J. V., & Downey, R. (2019). Chapter 25—Polysomnography. In K. H. Levin & P. Chauvel (Éds.), *Handbook of Clinical Neurology* (Vol. 160, p. 381-392). Elsevier. https://doi.org/10.1016/B978-0-444-64032-1.00025-4

Scott, H., Lack, L., & Lovato, N. (2020). A systematic review of the accuracy of sleep wearable devices for estimating sleep onset. *Sleep Medicine Reviews*, *49*, 101227. https://doi.org/10.1016/j.smrv.2019.101227

Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., Hassan, M., & Seneviratne, A. (2017). A Survey of Wearable Devices and Challenges. *IEEE Communications Surveys & Tutorials*, *19*(4), 2573-2620. https://doi.org/10.1109/COMST.2017.2731979