# Multivariate Time Series Visualization for Sleep Electroencephalograms

Rosa Paccotacya Yanque[1], Olenka Vargas Lazarte[2], Roberto Rodríguez Urquiaga[3]

[1,2,3]Universidad Nacional de San Agustín de Arequipa, Perú

e-mail(s): [1]rpaccotacya@unsa.edu.pe, [2]ovargasl@unsa.edu.pe, [3]rrodriguezu@unsa.edu.pe

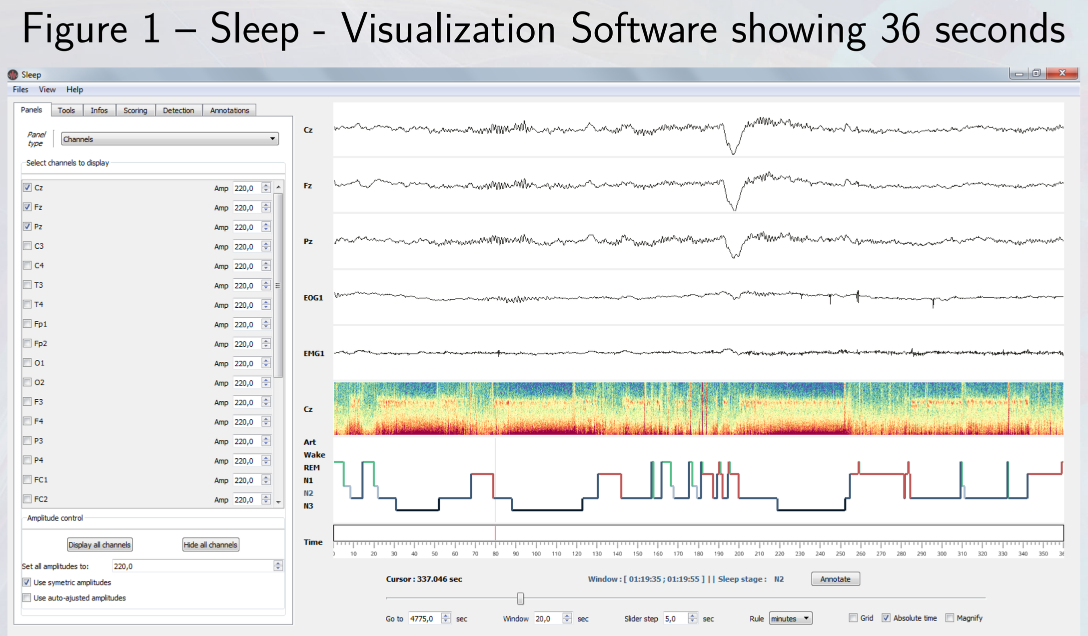
## ABSTRACT

In this work, we present in-progress research where time series of sleep electroencephalograms(EEG) are represented using Statistical Methods in Machine Learning. The performance of the Dynamic time warping (DTW) and Eros(Extended Frobenius norm) similarity measures is compared, using mT-SNE (a variant of typical t-SNE T-distributed Stochastic Neighbor Embedding) as projection for dimensionality reduction. We got better results with EROS and mTSNE, however, the challenge is based on an adequate visualization, in order to be able to interpret the data without difficulty.

## INTRODUCTION

Clinical data is mainly composed of Multivariate Time Series, leading to complexity in its analysis (REDDY; AGGARWAL, 2015) . The traditional method to analyze the electroencephalogram (EEG) during sleep has been based on visual inspection, sheet by sheet, the records of several hours of sleep (FERNÁNDEZ-MAS et al., 1998). Sleeping is very important for our health, and therefore its monitoring. Many studies are being conducted using different approaches such as ML, data mining, statistics, health informatics, etc. (COMBRISSON et al., 2017) (CHE et al., 2017). Figure (1) shows an example of current visualization for sleep data.

Figure 1 – Sleep - Visualization Software showing 36 seconds



Source: **Sleep: An Open Source Python Software for Visualization**; **Analysis**; **and Staging of Sleep Data**

EEG is composed of Multivariate Time Series (MTS) and its Visualization is difficult due to the complexity and run-time for distance measures. Our goal is to develop an interactive for comprehensive, reliable analysis of MTS data of sleep electroencephalograms. The used approach is based on MT-sne (NGUYEN et al., 2017) (MAATEN; HINTON, 2008) using Eros and comparing it to DTW a common measure for Multivariate Time Series.

## METHODOLOGY AND EXPERIMENTS

Figure (2) shows the pipeline of our research. Initially we extract the data in matrices that represent multivariate time series. Then, the similarity between the series is calculated, finally moving to the low-dimensional projection, preserving the original characteristics of the data. Implementation was done in Python.

- DTW Measure

$$C_d(i,j) = dist(i,j) + \min\{C_d(i-1,j-1), C_d(i-1,j), C_d(i,j-1)\} \quad (1)$$

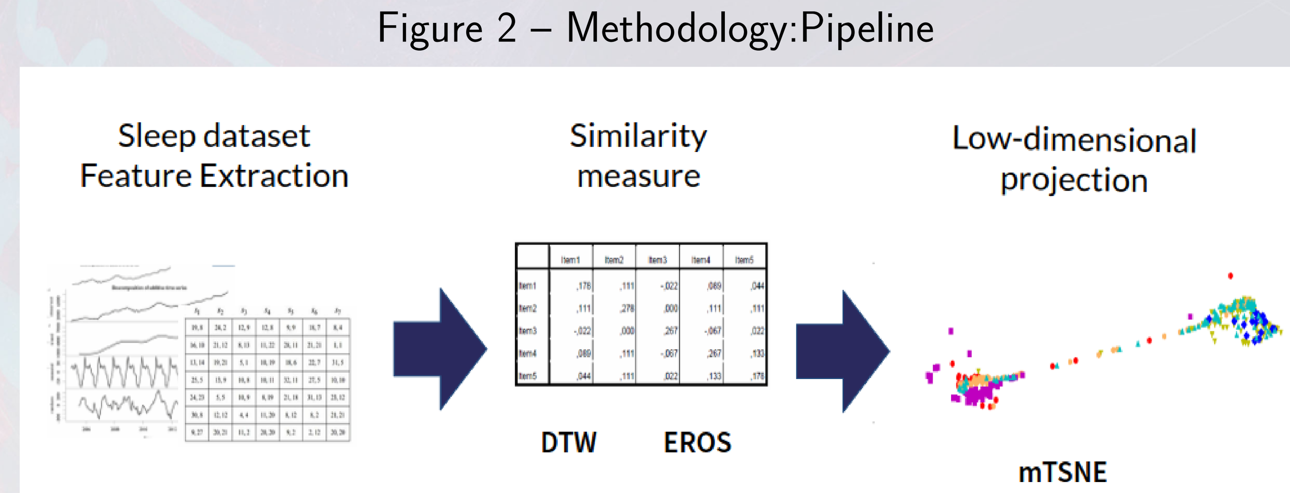
- EROS Measure

$$EROS(x_i, x_j, w) = \sum_{l=1}^{n} w_l | < v_{il}, v_{jl} > | \quad (2)$$

- TSNE Projection

$$p_{j|i} = \frac{exp(\frac{-||x_i - x_j||^2}{2\sigma_i^2})}{\sum_{k \neq i} exp(\frac{-||x_i - x_k||^2}{2\sigma_i^2})}. \quad (3)$$

Figure 2 – Methodology:Pipeline




## EXPERIMENTS

Dataset: Physionet "Sleep Recordings and Hypnograms in European Data Format (EDF)" (data from 70 healthy patients and 23 patients with sleep problems, age range between 20 to 66 years old). From the 7 EEG signals provided, 4 to 100 Hz were used(100 records per second). It was extracted a sample from the complete dataset of around five hours divided into segments of 30 seconds.Each of these segments has a label: W (awake), Phase 1, 2, 3, 4, and R (REM).

## RESULTS

Figures 3 and 4 show results obtained. First, comparing visualization between EROS and DTW, then the found pattern in patients with sleep problems.The Bar at the right represents time in hours, it is from 0 to 7 hrs approx.

Figure 3 – Visualization with DTW and ERO

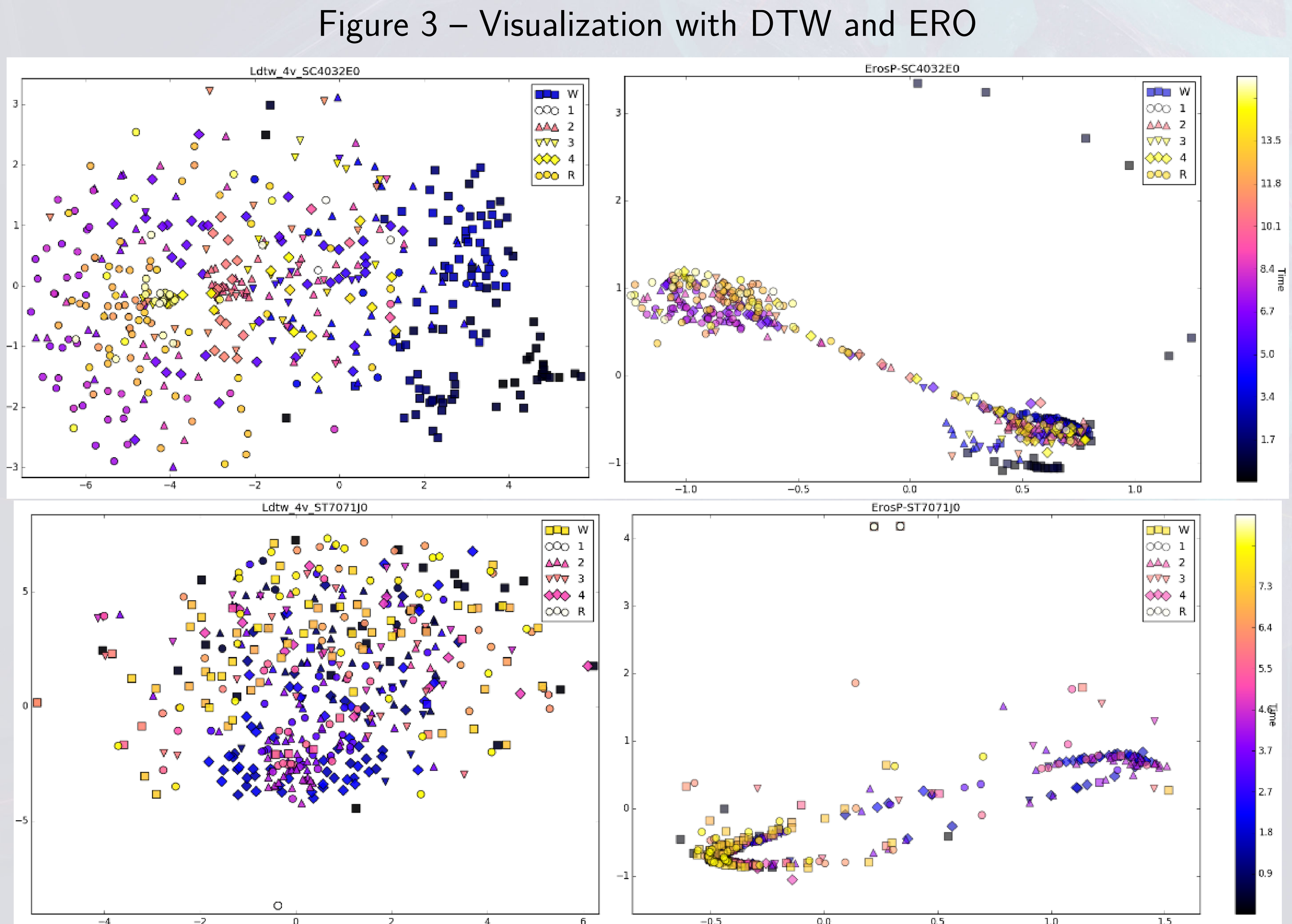
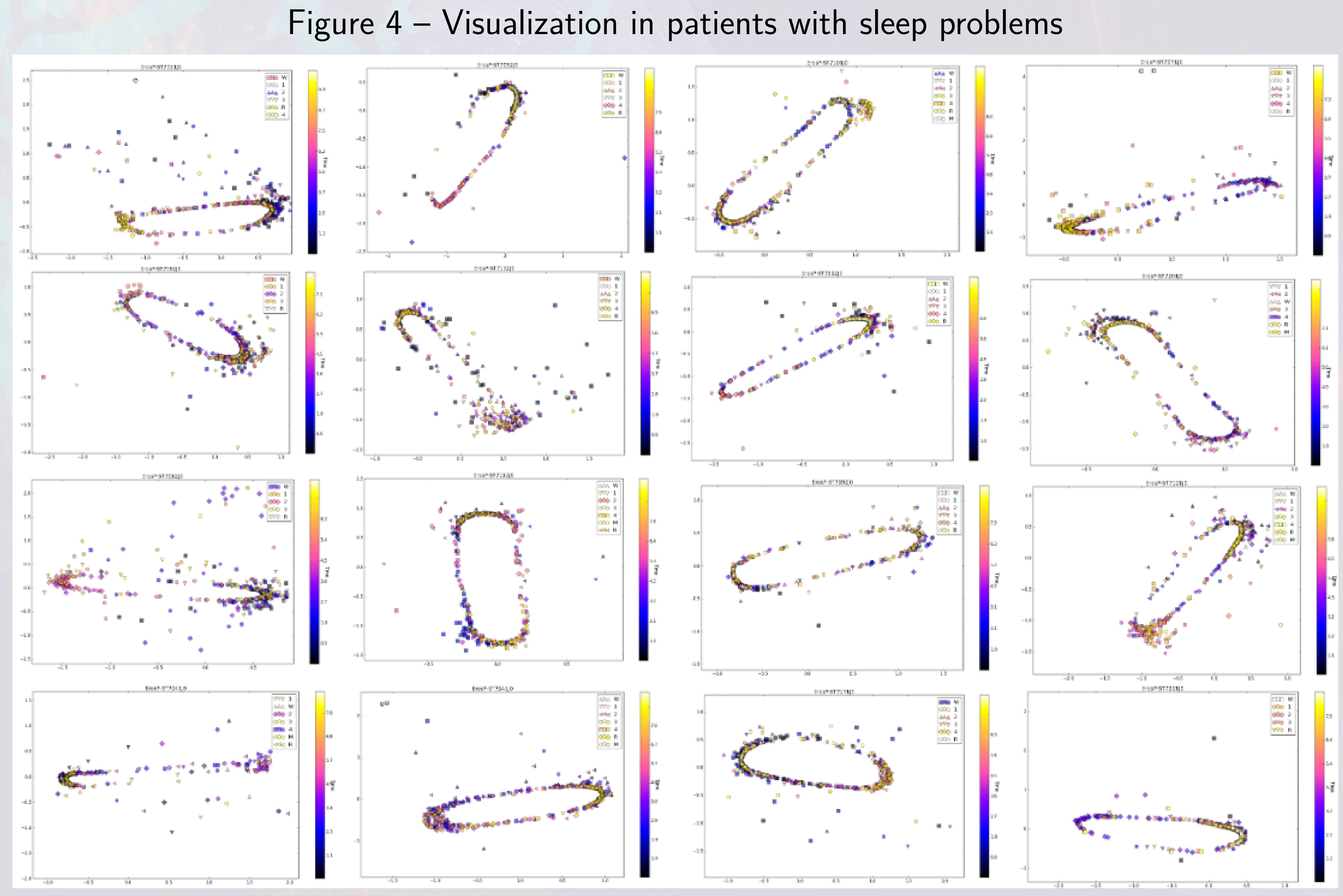


Figure 4 – Visualization in patients with sleep problems



## CONCLUSIONS

- For the EEG visualization, we based on m-TSNE, using as metrics the EROS distance. The evaluation of the electroencephalogram data sample showed that the used approach provides interpretable information through visualization.
- Using Eros as similarity measure has a run-time in 75x speedup over DTW. We discovered an interesting pattern in visualization, an elliptical-shape for people with sleep problems. These ideas could help health professionals evaluate the performance of their patients.

More details can be found at:

## REFERENCES

CHE, Z. et al. DECADE: a deep metric learning model for multivariate time series. In: **KDD workshop on mining and learning from time series**. [S.l.: s.n.], 2017.

COMBRISSON, E. et al. Sleep: An Open-Source Python Software for Visualization, Analysis, and Staging of Sleep Data. **Frontiers in Neuroinformatics**, Frontiers, v. 11, p. 60, 2017.

FERNÁNDEZ-MAS, R. et al. Visualización gráfica de las transiciones de las fases del sueño en el hombre: métodos de representación tridimensional. **Salud mental**, v. 21, n. 6, p. 27–32, 1998.

MAATEN, L. v. d.; HINTON, G. Visualizing data using t-SNE. **Journal of machine learning research**, v. 9, Nov, p. 2579–2605, 2008.

NGUYEN, M. et al. m-tsne: A framework for visualizing high-dimensional multivariate time series. **arXiv preprint arXiv:1708.07942**, 2017.

REDDY, C. K.; AGGARWAL, C. C. **Healthcare data analytics**. [S.l.]: Chapman and Hall/CRC, 2015.