Longitudinal estimation of stress-related states through bio-sensor data

Vadym Mozgovoy

Faculty of Law, Criminal Sciences and Public Administration,
Swiss Graduate School of Public Administration, University of Lausanne,
Lausanne, Switzerland

Abstract

Purpose – The authors aim to develop a conceptual framework for longitudinal estimation of stress-related states in the wild (IW), based on the machine learning (ML) algorithms that use physiological and non-physiological bio-sensor data.

Design/methodology/approach – The authors propose a conceptual framework for longitudinal estimation of stress-related states consisting of four blocks: (1) identification; (2) validation; (3) measurement and (4) visualization. The authors implement each step of the proposed conceptual framework, using the example of Gaussian mixture model (GMM) and K-means algorithm. These ML algorithms are trained on the data of 18 workers from the public administration sector who wore biometric devices for about two months.

Findings – The authors confirm the convergent validity of a proposed conceptual framework IW. Empirical data analysis suggests that two-cluster models achieve five-fold cross-validation accuracy exceeding 70% in identifying stress. Coefficient of accuracy decreases for three-cluster models achieving around 45%. The authors conclude that identification models may serve to derive longitudinal stress-related measures.

Research limitations/implications – Proposed conceptual framework may guide researchers in creating validated stress-related indicators. At the same time, physiological sensing of stress through identification models is limited because of subject-specific reactions to stressors.

Practical implications – Longitudinal indicators on stress allow estimation of long-term impact coming from external environment on stress-related states. Such stress-related indicators can become an integral part of mobile/web/computer applications supporting stress management programs.

Social implications – Timely identification of excessive stress may improve individual well-being and prevent development stress-related diseases.

Originality/value – The study develops a novel conceptual framework for longitudinal estimation of stress-related states using physiological and non-physiological bio-sensor data, given that scientific knowledge on validated longitudinal indicators of stress is in emergent state.

Keywords Bio-sensor, Conceptual framework, Stress identification, Longitudinal measure, Validation, Well-being

Paper type Research paper

1. Introduction

Digital technology, including bio-sensor devices and well-being applications, may facilitate coping with stress and improve individual well-being [1]. Ubiquitous computing literature examines remote stress pattern recognition, data processing and feedback to users. In the context of digital stress pattern recognition, it is particularly important to understand how algorithms for remote stress identification can be used effectively. Automatic remote identification of stress episodes requires training ML algorithms using related data [2, 3].

However, previous research has not clearly established the accuracy of measurements based on indirect measurements through unsupervised algorithms [3–6] and has not formalized the process of longitudinal evaluation of stress-related states.

In particular, validation is not always performed based on remote stress identification [4, 5]. However, longitudinal evaluation of stress-related states using bio-sensor data requires validated measurements. This study contributes to the stress pattern recognition literature by addressing the gap regarding stress validation and develops and implements a generalizable conceptual framework for the longitudinal evaluation of stress-related states and uses examples of clustering-based algorithms.

2. Background

Academic literature on pattern recognition, data mining and ML [7–9] provides a theoretical explanation of how stress identification algorithms operate. Algorithms can be trained by combining physiological and non-physiological (contextual) activations in the time and frequency domains to infer stress-related states. Algorithms for stress identification may follow supervised [10, 11] or unsupervised [2, 3] learning approaches. They involve individual [12, 13] and ensemble learning [14, 15] models. Algorithmic training can be conducted in a laboratory (LAB) setting [13, 16] or in an uncontrolled open environment, also known as "in the wild" (IW) [11, 15, 17]. Different stress identification models are designed for different uses, which correspond to specific types of environments: restricted, semi-restricted and non-restricted daily life environments [1].

The measurement of stress-related signals can be direct [18, 19] or indirect [3–6]. Direct measurement is usually conducted through isolated signals, while the stress level can be determined based on the level of variable(s) associated with stress. Longitudinal evaluations based on direct measurement usually follow (1) measurement and (2) visualization steps. In contrast, indirect measurement is concerned with stress pattern recognition through identification models (algorithms). Measurement involves physiological and/or non-physiological (contextual) variables that react to stress triggers. Physiological signals often involve the heart rate (HR), galvanic skin response (GSR) or blood pressure. Non-physiological variables include user physical activity, voice intensity and location.

There is a relatively small body of stress recognition literature that is concerned with longitudinal estimation of stress. The existing literature focuses particularly on the longitudinal estimation steps, the ML algorithm and indicators. The studies invariably have reported two contrasting direct and indirect measurement types. The indirect measurement method differs from research on direct longitudinal measures of stress using measurement and visualization steps [18, 19] because it requires additional identification and validation steps to make the derived indices more reliable IW. Specifically, validation of the identification models may use precision, recall, F1-measure, or mean absolute error. Researchers used direct measures of the HR and blood pressure [7] and GSR [8]. Other researchers used indirect measurement with ML algorithms. Some indirect algorithm-based measures follow a fully unsupervised approach [4, 5] that does not include validation. Two recent studies include implied IW validation [3, 6]. Yet, the role of validation in the indirect measurement remains underexplored. The present study aims to address this gap in research by proposing and implementing a process model for the longitudinal evaluation of stress-related states.

3. Proposed process model

The stress pattern recognition literature suggests that longitudinal evaluation could be based on indirect measurement often involving (1) identification, (2) validation, (3) measurement and (4) visualization [3, 6]. This study develops and applies a process model for the

longitudinal evaluation of stress-related states consisting of four stages (Figure 1). They include identification through a constructed algorithm (model), model validation, measurement and visualization.

The first stage corresponds to the construction of a stress identification model from IW data. The Yerkes–Dodson law [20] and catastrophe theory of arousal [21, 22] may underpin clustering-based stress pattern recognition and subsequent analytic data labeling [23]. The Yerkes–Dodson law assumes the existence of three ranges of arousal: (1) low range of arousal, (2) middle range of arousal and (3) high range of arousal. The first range has a strong connection to relaxation, the second range to optimal performance and the third range to stress.

Validation is the second step performed by calculating convergent validity coefficients of accuracy and recall based on (1) transfer learning, and (2) LAB-collected data with labels. Transfer learning assumes training of GMM and K-means using data collected IW with GSR, HR and motion values as input. It identifies the parameters of the models and uses them to predict cluster labels using the LAB dataset based on the same input variables. The computational order of fitting the models (COM) at the identification step is as follows: two-cluster K-means and GMM models and three-cluster K-means and GMM models, where models are fit independently. The estimation and validation consist of four stages as follows:

(1) Standardize IW data;

(2) Standardize LAB data and create five splits;

(3) Fit models on IW data and save IW model parameters: two-cluster K-means and GMM; three-cluster K-means and GMM according to COM and

(4) Input LAB data to trained IW models, predict labels and compare to the LAB data labels based on five-fold splits according to COM.

Validation ends by comparing clustering labels and LAB-collected labels to establish accuracy and recall. The final steps are to (3) derive stress-related indicators from identification models and (4) visualize them. The dataset to which this model is applied and data analysis are described below.

3.1 Participants and data collection

To apply the process model (i.e. conceptual framework) for longitudinal evaluation of stress-related states, the current research draws upon two datasets. The first dataset (collected IW) serves to fit the models to the data. The second dataset (LAB data) serves to estimate the validation coefficients.

The IW dataset (Table 1) was collected from the Municipal Fiscal Administration (Switzerland). A total of 18 public service employees participated in the experiment, which included males and females of different ages. Participants wore wearable biometric devices

1 • Identification: Construction of stress identification model (algorithm)

2 • Validation: Validation of the model

3 • Measurement: Computation of stress-related indicators of incidence based on identification model
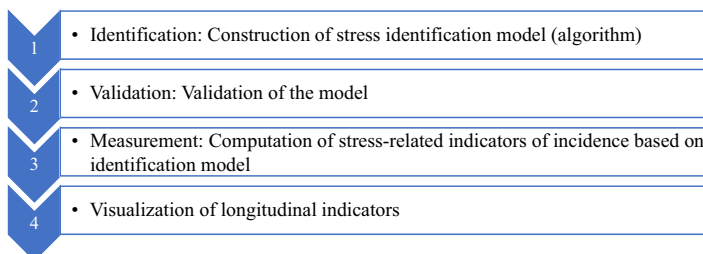
4 • Visualization of longitudinal indicators

Figure 1.
Process model for
longitudinal evaluation
of stress-related states

from 11/01/2017 to 12/18/2017. Overall, 16 of the 18 users wore wearable bio-sensor devices during working days and 2 users wore devices during weekends and working days. Registrations of wearable devices included GSR, HR, motion intensity (MI), and other variables. The sampling frequency was 1 Hz. Registrations that did not meet manufacturer quality values were automatically detected and deleted from the dataset.

The validation dataset (i.e. LAB) was collected under LAB conditions (Table 1), where 15 users (male and female) wore RespiBAN sensors under induced stress, amusement and baseline conditions (data labels). The devices produced GSR, HR, accelerometer and other data. The sampling frequency was 700 Hz. Further details on the LAB data collection can be found in the Wearable Stress and Affect Detection (WESAD) scientific report [24].

3.2 Pre-processing, feature extraction and measurements

The sampling process in the IW data entails choosing three days of device exploitation at random for each of the 18 users. In total, 1,018,497 observations appeared in the sample. This corresponds to approximately 283 h of device use for 14 working days in November and 10 working days in December. The sliding windows technique was then used, with a 60-s step size and no overlap, for calculating averages. After deleting incorrect values, the dataset was reduced to 14,133 instances.

Next, pre-processing using the sliding windows technique was performed (window length of one min and 55 s of overlap) on the validation LAB data. The window length corresponded to 42,000 observations, the overlap corresponded to 35,000 observations and the increment corresponded to 7,000 data points. The resulting dataset contained 6,475 instances.

Motion intensity and physiological signals (HR and GSR) were combined. Each instance was composed of GSR, HR and MI mean interval values. HR was measured in beats per minute (bpm) [25] and GSR was measured in kilo-ohms (kohms) [26, 27]. GSR reflects electrical conductivity, which changes when the skin glands produce ionic sweat [28]. It is linked to the emotional arousal state, where higher levels may increase the GSR [26, 27]. Motion is an additional measure reflecting the intensity of physical activity, and its score ranges on a continuous scale from 0 to 100 [29]. For example, GSR or HR may increase with increased physical activity. This, in turn, causes physiological arousal, which does not always imply stress. Therefore, combining physiological sensing with the additional contextual variable of motion may improve stress identification [29].

3.3 Data analysis

Data pre-processing and feature extraction for IW data were performed using Matrix Laboratory (MATLAB). Unsupervised models involved using the *sklearn* module and writing original code for labeling the output of K-means and GMM in Python.

3.4 Research design

3.4.1 Identification. The identification step involves fitting stress identification models, cluster mean characteristics for K-means algorithms and means and variance–covariance for

Dataset name	# of participants	Sex	Mean age	Data collection method	Device	Sampling frequency (Hz)
IW	18	6 males; 12 females	>18	Open use under variable work intensity conditions	Everion	1 Hz
LAB	15	12 males; 3 females	27.5	Induction of stress-related states	RespiBAN	700 Hz

Table 1.
Characteristics of datasets

GMM models using the IW dataset. In this study, the models were fit to standardized data. Therefore, scaling the data, fitting unsupervised models and assigning labels were required. Next, the K-means and GMM algorithms [7] were applied for two- and three-cluster problems. K-means performs hard assignments, optimizing the cost function defined in Eqn (1). GMM is a soft clustering algorithm that estimates the cluster means and covariance. It optimizes a cost function (Eqn (2)):

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \| x_n - \mu_k \|^2 \tag{1}$$

where, $r_{nk} \in \{0, 1\}$, $n = 1...N$ and $k = 1, ..., K$.

$$ln\, p(X|\pi, \mu, \Sigma) = \sum_{n=1}^{N} \ln \left\{ \sum_{k=1}^{K} \pi_k \mathcal{N}(x_n | \mu_k, \Sigma_k) \right\} \tag{2}$$

where, $n = 1...N$, $k = 1, ..., K$, π_k is a prior distribution and $\mathcal{N}(x_n | \mu_k, \Sigma_k)$ is a Gaussian distribution.

Cluster labeling is based on the mean GSR values for each cluster. It attributes a label corresponding to a greater level of arousal to a cluster having a greater average according to the GSR (acting as the decider variable). For a two-class problem, it attributes the label "Highest Arousal" to observations having the highest average GSR. Similarly, the cluster index having the lowest average GSR value is assigned the "Lowest Arousal" label.

For a three-class problem, it assigns the label "Emotional Overarousal" to observations having the highest average GSR value and "Arousal" to observations with the second-highest average GSR. Similarly, it assigns the "Relaxation" label to the cluster with the lowest average GSR.

3.4.2 Validation. A five-fold cross-validation step compares the predictions of the clustering output to the LAB data labels. Once cluster labels are assigned, this step follows by computing the proportion of cases that match the "laboratory ground truth labels." For two-class learning, convergent validity coefficients are computed by matching clustering-based outputs with relevant and similar classes from the LAB dataset: "High Arousal" = = "Stress" and "Low Arousal" = = "Non–Stress," where "Non-Stress" comprises "Amusement" and "Baseline Condition." For the three-class learning problem, the convergent validity is based on the comparison rule: "Emotional Overarousal" = = "Stress," "Arousal" = = "Amusement" and "Relaxation" = = "Baseline." A match between clustering labels and LAB labels is defined as a True Positive.

The IW training dataset consisted of 14,132 observations (90.84%). The LAB test dataset consisted of 1,295 observations in each of the five splits (9.16%).

Next, validation accuracy and recall are computed to establish convergent validity and to evaluate to what extent the identification models reflect stress. The accuracy and recall are defined using (4) and (5), respectively, as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$\text{recall} = \frac{TP}{TP + FN} \tag{5}$$

where, TP is true positive, FP is false positive, TN is true negative and FN is false negative.

Finally, the motion variable for the LAB dataset corresponds to the MI as follows:

$$\text{Motion} = \sqrt{\text{acc}_x^2 + \text{acc}_y^2 + \text{acc}_z^2} \tag{6}$$

3.4.3 Measurement and visualization. Finally, the incidence indicators are computed and visualized. Incidence indicators count the average number of specific stress-related incidents per time window. Stress-related realization corresponds to the identification of the clustering algorithm output. For example, it is possible to count the average number of arousal episodes in a given time window. These counts distributed over time serve to reflect longitudinal stress-related incidence. The identification models serve to derive the high-arousal index and low-arousal index for the two-cluster problem, whereas the emotional over-arousal index, arousal index and relaxation index apply to the three-cluster problem. This approach allows us to obtain a longitudinal index of stress-related incidents associated with accuracy measures. They allow evaluation of the impact of the external environment on stress-related states over time. Furthermore, such indices should account for the number of hours that a bio-sensor device is active. For example, the number of stress episodes divided by the amount of time when the device is active during a time frame should ensure comparability between data points. This is important when device use is not uniform over time. In the final step, the visualization output is evaluated to account for the longitudinal realizations of stress-related episodes.

4. Results
This section presents the process model (Figure 1) application results.

4.1 Identification
To describe the data characteristics, means and standard deviations were computed for HR ($M = 76.17$; SD $= 11.57$), GSR ($M = 3.75$; SD $= 2.96$) and motion ($M = 1.75$; SD $= 1.17$) and average cluster characteristics for two-cluster K-means and GMM models in standardized form and variance–covariance matrices for GMM (Table 2). Model characteristics were then used to make predictions using the LAB-collected dataset based on similar inputs: scaled HR, GSR and MI. In total, 50 iterations (i.e. repeated experiments) were performed to fit the data and determine the most common solution. Clusters with lowest mean GSR values received "Low Arousal" labels and clusters with highest GSR values received "High Arousal" labels. Results indicate that a greater GSR is associated with a higher HR.

This step computes the average cluster characteristics for three-cluster K-means and GMM models in standardized form and variance–covariance matrices for GMM (Table 3).

Parameters	Method	Clusters	Variables	HR	Normalized GSR	Motion
Means	K-means	Low arousal	–	−0.36	−0.18	−0.36
		High arousal		1.16	0.59	1.16
	GMM	Low arousal		−0.21	−0.28	−0.37
		High arousal		0.55	0.74	0.98
Covariance	GMM	Low arousal	HR	0.72	0.05	0.11
			GSR	0.05	0.27	0.02
			Motion	0.11	0.02	0.14
		High arousal	HR	1.33	−0.20	0.80
			GSR	−0.20	2.15	−0.67
			Motion	0.80	−0.67	1.91

Table 2.
Two-cluster models

Parameters	Method	Clusters	Variables	HR	Normalized GSR	Motion
Means	K-means	Relaxation	–	−0.34	−0.28	−0.36
		Emotional over-arousal		0.22	2.26	0.12
		Arousal		1.33	−0.19	1.48
	GMM	Relaxation		0.04	1.45	0.23
		Emotional over-arousal		0.83	−0.31	1.33
		Arousal		−0.24	−0.28	−0.43
Covariance	GMM	Relaxation	HR	1.00	0.49	0.27
			GSR	0.49	2.07	0.02
			Motion	0.27	0.02	0.64
		Emotional over-arousal	HR	1.24	0.03	0.71
			GSR	0.03	0.35	0.17
			Motion	0.71	0.17	2.06
		Arousal	HR	0.68	0.07	0.10
			GSR	0.07	0.27	0.02
			Motion	0.10	0.02	0.10

Table 3.
Three-cluster models

Similar to the two-cluster models, predictions on the LAB dataset were performed by running 50 iterations and determining the most common solution. Clusters with the lowest mean GSR values received "Relaxation" labels; clusters with second-highest mean GSR values received "Arousal" labels and clusters with highest GSR values received "Emotional Overarousal" labels. The results indicate that the highest average motion values were associated with either the arousal or emotional over-arousal classes.

4.2 Validation
Two-cluster and three-cluster K-means and GMM models were trained using the IW dataset (Table 4). Data labels were then predicted using the LAB dataset, where the labels obtained from models trained on the IW data were compared to the LAB clustering-based labels (true states). Results indicate that the two-cluster model is more relevant for reflecting stress, whereas the three-cluster model may be more relevant for reflecting arousal. However, two-cluster models distinguish between stress and any other state with accuracy greater than 70%, whereas the three-cluster models only achieve accuracies slightly above 45%.

4.3 Measurement and visualization
Measurement is defined as an average of stress-related incidence during a time interval using K-means clustering. It corresponds to deriving "High Arousal" and "Low Arousal" indices based on the two-cluster model and "Relaxation," "Arousa," and "Emotional Overarousal" indices for the three-cluster model. Low-arousal episodes dominated stress-related dynamics up to 450 min (Figure 2); however, there was a mixture of high and low arousal states after 450 min. Figure 2 also shows some emotional over-arousal episodes in later stages.

| Indicator | Models | | | |
	Two-cluster K-means	Two-cluster GMM	Three-cluster K-means	Three-cluster GMM
Mean accuracy, %	73.39	60.79	45.74	47.38
Mean recall score	0.63	0.52	0.40	0.42

Table 4.
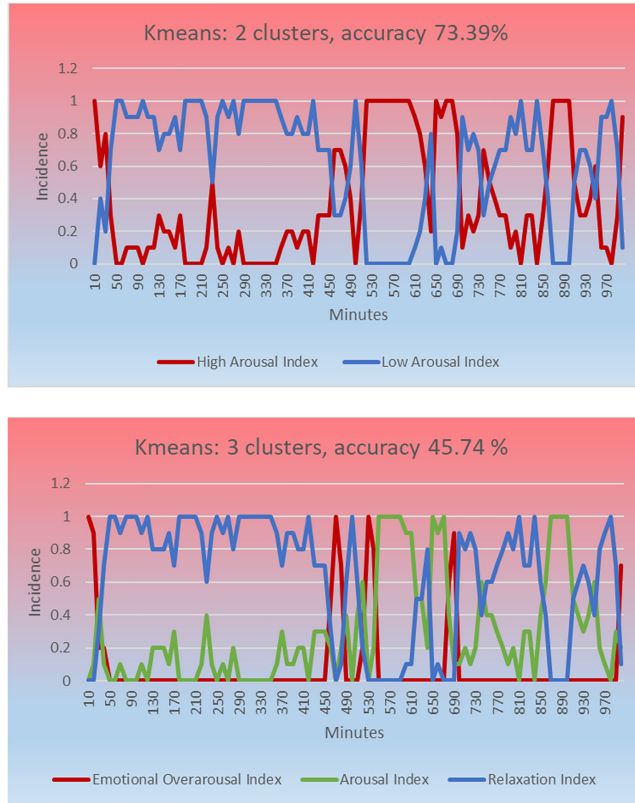Validation coefficients
for K-means and GMM
models

Figure 2.
Stress-related indices
based on two-cluster
and three-cluster
K-means and GMM
models

5. Discussion and conclusion

5.1 Findings

This study proposed a conceptual framework for the longitudinal estimation of stress-related states through bio-sensor data consisting of four stages: (1) identification, (2) validation, (3) measurement and (4) visualization. The proposed process model facilitated the development of stress-related indices and could be generalized to supervised and unsupervised learning models. The results indicate that it is possible to use unlabeled datasets collected IW for two-class stress pattern recognition, achieving an accuracy of approximately 70%. The findings reveal that it is possible to visualize longitudinal stress-related incidence using line plots. Thus, the evaluation framework provides additional information regarding the extent to which a longitudinal index reflects stress.

A comparison between the current study and the background literature reveals both similarities and divergence points (Table 5). The approach used in this study is comparable in longitudinal estimation steps to those used by other researchers. The present research reflects Study 3 [4] and Study 4 [5] involving three estimation steps: identification, measurement and visualization. The evaluation framework is similar to that found in Study 5 [3] and Study 6 [6] involving four estimation steps.

The ML algorithms used are also aligned with extant stress recognition literature. Clustering algorithms were used in Study 4 [5] and Study 5 [3], similar to the proposed evaluation framework. The HR measure was previously used in Study 1 [18], Study 5 [3] and

Characteristics		Study (s) Study 1 [18]	Study 2 [19]	Study 3 [4]	Study 4 [5]	Study 5 [3]	Study 6 [6]	Present Study
Formalization of the longitudinal estimation process		No	No	No	No	No	No	Yes
Identification via model	Learning approach	No	No	Unsupervised	Unsupervised	Unsupervised	Supervised	Unsupervised
	Data	No	No	Physiological	Physiological; contextual	Physiological; contextual	Physiological; contextual	Physiological; contextual
Measurement	Direct/indirect	Direct	Direct	Indirect	Indirect	Indirect	Indirect	Indirect
	Indicators	HR; arterial blood pressure	GSR	Daily stress incidence	Index of arousal	Stress level Abnormal degree of HR	Long-term stress score	Low arousal High arousal Emotional over-arousal Arousal Relaxation
Validation procedure	Validation data for computing validation metrics	No	No	No	No	Social media posts	PANAS questionnaires	Stress-related labels under laboratory induction
	Transfer learning	No	No	No	No	No	No	Yes
Visualization		Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 5.
Comparison table

Study 6 [6] GSR was used in Study 2 [19] and motion was used in Study 3 [4]. Most importantly, the application of clustering algorithms is consistent with the methodological literature on pattern recognition, data mining and ML [7–9].

Simultaneously, the present study diverges from extant research and broader methodological studies on stress recognition. Compared with the present study, which proposes a formalized process model for longitudinal estimation involving four explicit steps, Study 5 [3] and Study 6 [6], also using four estimation steps, do so implicitly and without deliberately formalizing them. Additionally, Study 3 [4] and Study 4 [5], focusing on indirect measures, do not include a validation step.

Furthermore, the ML algorithm differs from Study 5 [3] and Study 6 [6]. The present study is concerned with the validation of the IW data analysis results based on a LAB setting, while Study 5 [3] and Study 6 [6] focus on IW validation. More specifically, the present study applies transfer learning when a model trained using IW data is tested using labeled data collected in a LAB setting. Moreover, the proposed evaluation framework uses emotional arousal and relaxation indicators, diverging from the extant literature except for Study 4 [5], drawing upon an index of arousal. However, our results on stress indicators are more consistent with a recent study using a theoretically justified model for identifying three stress-related states [23].

The current study contributes to the stress pattern recognition literature in two ways. First, it is the first study that formalizes four steps of longitudinal evaluation of stress-related states based on indirect measurement in a distinguishable and consecutive manner. Drawing upon the convergent validity criteria, this model applies methodological rigor to the process of longitudinal evaluation of stress-related states. Demonstrating the possibility of using transfer learning is the second contribution of the present study to the stress pattern recognition literature.

5.2 Limitations

Despite its contributions to the stress recognition literature, the present study has limitations. The visualization includes data windows with uniform hours of use, but variable user activity over disjoint time frames should be analyzed for requiring adjustments. Visualization could also involve benchmark curves for high and low stress, helping to situate excessive stress, optimum and relaxation states, although the gradient color background used for visualization of regions with high stress intensity and relaxation already enhances the graphical interface, helping to situate those states. Although there is a large sliding window overlap in the LAB-collected dataset, it allows the creation of large numbers of instances for cross-validation. Finally, the motion indicator for the IW dataset is automatically generated based on manufacturer data and may be computationally different from the MI derived from LAB data.

5.3 Implications for practice

The findings involving indirect measurement of stress-related states through unsupervised learning models imply that indirect measurement may support longitudinal stress monitoring. Digital monitoring may become an integral part of digital applications embedded in mobile, portable, wearable devices and/or context-based recommendation systems. Finally, the present study confirms that digital stress management is relevant in both work and private environments.

5.4 Future research

Future studies should consider the identification of stress-related episodes through deep learning. Notably, ensemble learning involving several models and indicators can improve the theoretical representation of stress constructs and enhance the reliability of stress pattern

recognition. Future research should examine this approach. Furthermore, predictive time series models of stress could be developed to identify pre-stress conditions. More research should be dedicated to validating remote stress pattern recognition algorithms in real-life contexts. Future research could address individual reactions to stressors, as well as consider complementarity, the pros and cons of training the algorithms in several environments.

5.5 Conclusion
This study concludes that indirect stress identification models may serve to derive valid stress-related indices. This study proposed a conceptual framework for the longitudinal estimation of stress-related states through bio-sensor data for any type of stress identification model and confirmed its convergent validity. The findings show that indirect measurement may support longitudinal stress monitoring in real time. Consequently, the present study confirms the feasibility of digital stress management in professional and private settings.

References

1. Can YS, Arnrich B, Ersoy C. Stress detection in daily life scenarios using smart phones and wearable sensors: a survey. J Biomed Inform. 2019, 103139.

2. Kalas MS, Momin BF. Stress detection and reduction using EEG signals. 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016, 471–475.

3. Feng Z, Li N, Feng L, Chen D, Zhu C. Leveraging ECG signals and social media for stress detection. Behav Inf Technology. 2021; 40(2): 116–133.

4. Stepanovic S, Mozgovoy V, Mettler T. Designing visualizations for Workplace stress management: results of a pilot study at a Swiss municipality. International Conference on Electronic Government, 2019: 94–104.

5. Vadym M. Exploration of scenario-based simulations for stress benchmarking in Swiss public service. 2020 7th Swiss Conference on Data Science (SDS), 2020: 23–28.

6. Muaremi A, Arnrich B and Tröster G. Towards measuring stress with smartphones and wearable devices during workday and sleep. BioNanoScience. 2013; 3(2): 172–183.

7. Bishop CM, Pattern recognition and machine learning. New York: Springer, 2006.

8. Han J, Pei J, Kamber M, Data mining: concepts and techniques. Waltham: Elsevier, 2011.

9. Bonaccorso G, Machine Learning Algorithms: popular algorithms for data science and machine learning. Birmingham: Packt Publishing, 2018.

10. Jin N, Zhang X, Hou Z, Sanz-Prieto I, Mohammed BS. IoT based psychological and physical stress evaluation in sportsmen using heart rate variability. Aggress Violent Behav. 2021: 101587.

11. Li R and Liu Z. Stress detection using deep neural networks. BMC Med Inform Decis Making. 2020; 20(11): 1–10.

12. Greco A *et al.* Acute stress state classification based on electrodermal activity modeling. IEEE Transactions on Affective Computing. Early Access. 2021; 1-1.

13. Zubair M and Yoon C. Multilevel mental stress detection using ultra-short pulse rate variability series. Biomed Signal Process Control. 2020; 57: 101736.

14. Can YS, Chalabianloo N, Ekiz D, Ersoy C. Continuous stress detection using wearable sensors in real life: algorithmic programming contest case study. Sensors. 2019; 19(8): 1849.

15. Wang K, Guo P. An ensemble classification model with unsupervised representation learning for driving stress recognition using physiological signals. IEEE Trans Intell Transportation Syst. 2020; 22(6): 3303-3315. June 2021.

16. Can YS, Chalabianloo N, Ekiz D, Fernandez-Alvarez J, Riva G, Ersoy C. Personal stress-level clustering and decision-level smoothing to enhance the performance of ambulatory stress detection with smartwatches. IEEE Access. 2020; 8: 38146–38163.

17. Mou L *et al.* Driver stress detection via multimodal fusion using attention-based CNN-LSTM. Expert Syst Appl. 2021; 173: 114693.

18. Finke JB, Zhang X, Plein D, Schilling TM, Schächinger H, Larra MF. Combining mental and physical stress: synergy or interference?. Physiol Behav. 2021; 233: 113365.

19. Chae J, Hwang S, Seo W, Kang Y. Relationship between rework of engineering drawing tasks and stress level measured from physiological signals. Autom ConStruct. 2021; 124: 103560.

20. Corbett M. From law to folklore: work stress and the Yerkes-Dodson Law. J Managerial Psychol, 2015; 30(6): 741-752.

21. Hardy L. A catastrophe model of performance in sport. In: Jones JG, Hardy L. (Eds). Stress and performance in sport. John Wiley & Sons,; 1990. 81-106.

22. Hardy L, Parfitt G. A catastrophe model of anxiety and performance. Br J Psychol. 1991; 82(2): 163–178.

23. Mozgovoy V. Stress pattern recognition through wearable biosensors in the workplace: experimental longitudinal study on the role of motion intensity. 2019 6th Swiss Conference on Data Science (SDS), 2019: 37–45.

24. Schmidt P, Reiss A, Duerichen R, Marberger C, van Laerhoven K. Introducing wesad, a multimodal dataset for wearable stress and affect detection. Proceedings of the 20th ACM International Conference on Multimodal Interaction, 2018: 400–408.

25. de Santos A, Sánchez-Avila C, Guerra-Casanova J, Bailador-Del Pozo G. Real-time stress detection by means of physiological signals. Recent application in biometrics: InTech, 2011.

26. Angus F, Zhai J, Barreto A. Front-end analog pre-processing for real-time psychophysiological stress measurements. Proceedings of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 05), 2005: 218–221.

27. Zhai J, Barreto AB, Chin C, Li C. Realization of stress detection using psychophysiological signals for improvement of human-computer interactions. SoutheastCon, 2005. Proceedings. IEEE, 2005: 415–420.

28. de Santos Sierra A, Avila CS, Casanova JG,del Pozo GB, Vera VJ. Two stress detection schemes based on physiological signals for real-time applications. Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on, 2010: 364–367.

29. Garcia-Ceja E, Osmani V, Mayora O. Automatic stress detection in working environments from smartphones' accelerometer data: a first step. IEEE J Biomed Health Info. 2015; 20(4): 1053–1060.

About the author

Dr. Vadym Mozgovoy obtained his PhD in Public Administration from the University of Lausanne, Switzerland, in 2021. He developed his PhD thesis as Swiss National Science Foundation (SNSF) research assistant at Swiss Graduate School of Public Administration (IDHEAP). The thesis by Dr. Vadym Mozgovoy is called "Physiolytics at the Workplace: Longitudinal Evaluation of Organizational Effects at Swiss Public Administrations". He received the Best Paper Award at the Swiss Conference on Data Science the 2019. He also received the Best Doctoral Student Conference Paper Award at the Public and Nonprofit (PNP) Division of the Academy of Management (AOM) in 2021. He holds a BSc in Business Administration from the University of Geneva. He finished a MSc program in Statistics with a methodological Master thesis at the University of Geneva in 2017. Previously, he worked as a Junior Statistician in the context of Biostatistics and Epidemiology in an advisory firm providing consultancy to multinationals in Geneva from 2013 to 2016. Dr. Vadym Mozgovoy can be contacted at: mozgovoy.vadym@gmail.com