

Time Alignment as a Necessary Step in the Analysis of Sleep Probabilistic Curves

Zuzana Roš'áková^{1,2}, Roman Rosipal¹

¹*Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovakia, zuzana.rostakova@savba.sk*

²*Faculty of Electrical Engineering and Information Technology, Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia*

Sleep can be characterised as a dynamic process that has a finite set of sleep stages during the night. The standard Rechtschaffen and Kales sleep model produces discrete representation of sleep and does not take into account its dynamic structure. In contrast, the continuous sleep representation provided by the probabilistic sleep model accounts for the dynamics of the sleep process. However, analysis of the sleep probabilistic curves is problematic when time misalignment is present. In this study, we highlight the necessity of curve synchronisation before further analysis. Original and in time aligned sleep probabilistic curves were transformed into a finite dimensional vector space, and their ability to predict subjects' age or daily measures is evaluated. We conclude that curve alignment significantly improves the prediction of the daily measures, especially in the case of the *S2*-related sleep states or slow wave sleep.

Keywords: Curve alignment, probabilistic sleep model, functional principal component analysis.

1. INTRODUCTION

Sleep is a continuous process that can be described by a finite number of sleep stages. The Rechtschaffen and Kales sleep model (R&K) [1] distinguishes five basic sleep stages. These include the *Wake* stage, or stage of full wakefulness; stages *S1* (light sleep), *S2*, *S3*, and *S4*, also called *nonREM* stages; and finally, the *REM* (rapid eye movement) stage, during which quick eye movements behind closed eyelids are typical. Stages *S3* and *S4* represent slow wave sleep *SWS* (or deep sleep), and in this study are considered together. The R&K model is mainly based on the analysis of EEG signals, which are divided into non-overlapping 30-second segments; each time segment is assigned to one of the above-mentioned sleep stages. This results in a discrete sleep representation known as a sleep hypnogram, where changes between a small number of sleep stages are not smooth.

The probabilistic sleep model (PSM) [2] is an alternative method of sleep process modelling. The model is EEG-based, but in contrast to the R&K model, only 3-second time segments are considered, and PSM distinguishes 20 sleep states called sleep microstates. Instead of strict assignment of a time segment to one of the 20 sleep states, a probability value is computed for each sleep microstate separately. Considering the probability values as a function of time, a sleep probabilistic curve is obtained (Fig. 1).

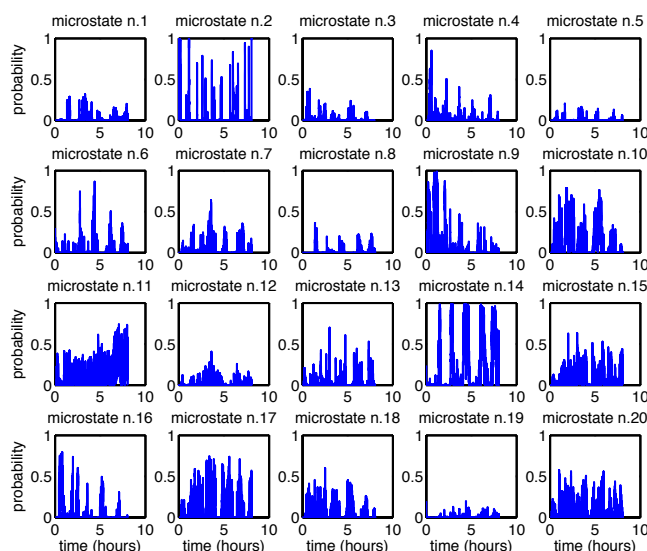


Fig. 1. An example of sleep probabilistic curves for 20 sleep microstates. The blue curves represent a whole night profile of a 42-year-old healthy man.

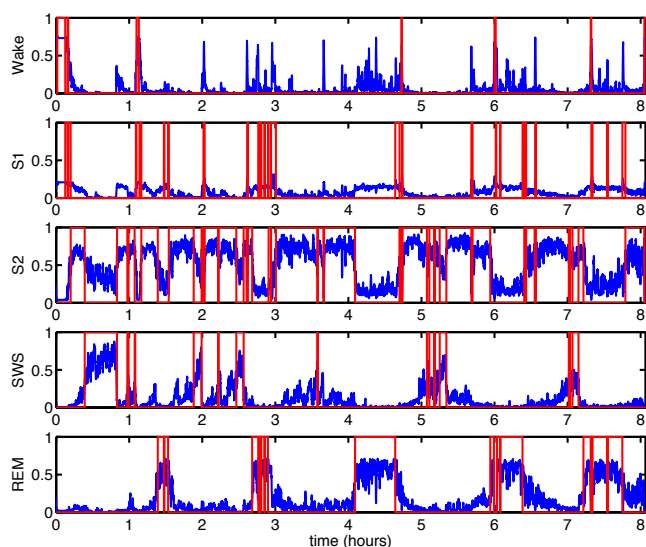


Fig. 2. An example of the sleep probabilistic curves for sleep stages *Wake*, *S1*, *S2*, *SWS*, *REM*. The blue curves represent a whole night profile of a 42-year-old healthy man. Corresponding Rechtschaffen and Kales scores [1] are depicted in red.

Physiological interpretation of sleep microstates is not straightforward. Therefore, PSM also estimates the probability (weights) of similarities between microstates and the standard sleep stages *Wake*, *S1*, *S2*, *SWS* or *REM*. For example, using the current analysis, sleep microstate 1 is similar to the *S2* stage, with the probability of 84.7%, and to the *REM* stage with the probability of 10.9%. Thus, the sum of probabilities to the other R&K sleep stages is 4.4%.

In addition to the interpretation of sleep microstates, estimated probabilities may be used as weights in a linear combination of the sleep microstates in order to reconstruct probabilistic curves for the standard R&K sleep stages (Fig. 2). In the analysis of the sleep probabilistic curves, we aimed to find typical overnight sleep profiles which significantly correlated with age or daily life performance. However, when the curves are misaligned in time (Fig. 3), the relationship with daily life measures is difficult to detect. Two curves X and Y observed on the approximately same time interval $T = [a, b]$ are misaligned if they are of similar shape, but important features like local maxima or minima are shifted in time.

To align a pair of curves X, Y means to find a strictly increasing time transformation $h: T \rightarrow T$ which minimises chosen similarity criterion C between the curves under the assumption of the common start and end point

$$h(a) = a \quad \text{and} \quad h(b) = b.$$

The area under the squared difference of two curves

$$C(X, Y \circ h) = \int_T (X(t) - Y \circ h(t))^2 dt \quad (1)$$

is an example of a similarity criterion which is used with small modifications in the majority of the curve alignment methods.

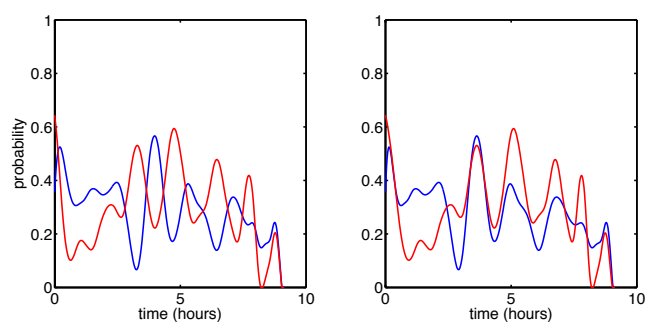


Fig. 3. An example of two smoothed sleep probabilistic curves with similar overnight profiles misaligned in time (left), and the aligned version (right).

Two of the many methods developed for curve alignment are self-modelling time warping [3] and pairwise curve synchronisation [4]. The first method aligns a set of curves to one common target curve, while the second approach aligns each pair of curves separately. The curve alignment by moments [5] defines a set of moments in a curve and aligns two curves in such a way that the difference between their moments is as small as possible. Following our previous practical experience with the presented sleep dataset, we prefer the elastic time warping method as described in [6]. The general idea of elastic warping was first mentioned in [7] or [8]. This method aligns transformations of each curve called the *square-root slope function* by using the criterion (1), rather than curves themselves. For more detail, see [6].

There may exist sleep features in which exact timing influences sleep quality and daily behaviour. Therefore, we aimed to detect sleep microstates or standard sleep stages where the time alignment of the sleep probabilistic curves significantly improves correlations between the sleep structure and daily measures, and conversely, where the curve alignment is counter-productive.

2. SUBJECT AND METHODS

In this study, the polysomnographic (PSG) recordings of 146 healthy subjects spending two consecutive nights in the sleep laboratory were used. These recordings represent a subset of PSG data collected in the European sleep project SIESTA [9].

After awakening, the subjects participated in a battery of neuropsychological tests. They were also asked to subjectively score their sleep quality or level of drive and drowsiness [10]. In addition, in the morning and evening, their pulse rate and blood pressure were recorded. The whole set of daily measures collected is listed in Table 1.

In addition to the original daily measures, we considered three artificial factors – factor of subjectively scored sleep quality (FA1), physiological factor (FA2), and neuropsychological factor (FA3). These factors were obtained by applying the factor analysis method to the set of all available daily measures [10].

Table 1. The list of daily measures and their abbreviations used in the article. For more details about the cognitive tests or questionnaires see [10].

Abbreviation	Measure
s_qua	subjectively scored sleep quality
a_qua	subjectively scored awakening quality
s_com	subjectively scored somatic complaints
wb_m, wb_e	well-being morning/evening
drive	level of drive in the morning
drows	level of drowsiness in the morning
aff	level of affectivity in the morning
mood	level of mood in the morning
pul_m, pul_e	pulse rate in the morning/evening
dia_m, dia_e	diastolic blood pressure in the morning/evening
sys_m, sys_e	systolic blood pressure in the morning/evening
num_m	numerical memory test
ad_ts	alphabetical cross-out test, total score
ad_sv	alphabetical cross-out test, attention variability
errp	alphabetical cross-out test, percentage of errors
fma_r, fma_l	fine motor activity test (right and left hand)

First, PSM was applied to the PSG data, and sleep probabilistic curves of 20 sleep microstates were extracted. Because of a high time variation between the beginning of sleep (lights off) and falling asleep, the beginning of all probabilistic sleep curves was set to the sleep latency, which is defined as three consecutive 30-second periods of the *S1* stage, or the first period of the *S2* stage, whichever comes first.

In the second step the probabilistic curves for the standard sleep stages *Wake*, *S1*, *S2*, *SWS* or *REM* were reconstructed by using the sleep probabilistic curves of 20 sleep microstates, and similarity weights estimated by the PSM.

The sleep probabilistic curves were smoothed by applying the functional principal component analysis (FPCA) method with smoothing covariance surface [11]. The smoothing step and the procedure described below were performed for each sleep microstate or standard sleep stage separately.

The FPCA method is also able to predict the sleep probabilistic curves profile at the end of the night according to the behaviour of the whole database. Therefore, we can define curves of all subjects over the same time interval.

A by-product of the FPCA method is the transformation of a functional half-space of all non-negative curves into a finite dimensional vector space of principal component scores. Each smoothed sleep probabilistic curve X_i , $i = 1, \dots, N$ where $N = 2 \times 146 = 292$ can be expressed as a sum of an overall mean curve μ and a linear combination of K functional principal components ϕ_1, \dots, ϕ_K

$$X_i(t) = \mu(t) + \sum_{j=1}^K a_{ij} \phi_j(t). \quad (2)$$

Functions $\phi_j : T \rightarrow \mathbb{R}$, $j = 1, \dots, K$ are normalised and mutu-

ally orthogonal

$$\int_T \phi_i(t) \phi_j(t) dt = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

The vector of principal component scores

$$a_i = (a_{i1}, \dots, a_{iK})^T, \quad i = 1, \dots, N$$

was used for further analysis, as the representative of the curve X_i .

To relate principal component scores of the sleep probabilistic curves with daily measures, a linear regression model was applied. For a chosen sleep state, the dataset was divided into a training and a testing part. A daily measure m was modelled as a linear combination of principal component scores belonging to the training dataset D_{train}

$$m_i = \beta_0 + \sum_{j=1}^K \beta_j a_{ij} + \varepsilon_i, \quad i \in D_{train},$$

where the estimators $\hat{\beta}_0, \dots, \hat{\beta}_K$ for the unknown parameters were obtained by the standard method of least squares. Then we tested whether the estimated model is significantly better than a constant model. After that, values of the daily measure for the testing dataset D_{test} were predicted by the estimated linear model

$$\hat{m}_l = \hat{\beta}_0 + \sum_{j=1}^K \hat{\beta}_j a_{lj}, \quad l \in D_{test}.$$

Finally, Spearman's correlation coefficient was computed between real values m_l , $l \in D_{test}$ and predicted values \hat{m}_l , $l \in D_{test}$.

To avoid misinterpretation of results caused by random splitting into training and testing datasets, 10-fold cross-validation was considered.

In the next step, the sleep probabilistic curves were aligned in time by the elastic time warping method [6] immediately after smoothing. The alignment was carried out for each sleep state separately. Transformation of aligned curves into principal component scores and modelling of daily measures were done in the same way as described above. Finally, the Wilcoxon test was performed to detect whether the difference between correlation coefficients obtained from the original and aligned curves was significant.

Different division of the data into 10 folds may lead to slightly different results. Therefore, the whole procedure was repeated 100 times and differences in correlations based on misaligned or aligned curves were considered as significant if the Wilcoxon test rejected the null hypothesis (no difference between correlation coefficients) in more than 40 trials.

3. RESULTS

3.1. Sleep microstates

Time synchronisation of the sleep probabilistic curves of 20 sleep microstates resulted in a few changes of correlations between real and predicted values of daily measures (Table 2).

Table 2. Average correlation coefficients for daily measures and sleep microstates. Only results where a significant difference between original and aligned curves was detected by the Wilcoxon test are presented. The percentage of linear models which were significantly better than a constant model is depicted in brackets.

Daily meas.	Microstate	p-value	$\bar{\rho}$, misaligned curves	$\bar{\rho}$, aligned curves
wb_m	5 (65% SWS)	0.011	-0.14 (0%)	0.10 (91%)
	8 (73% REM)	0.017	0.02 (0%)	0.13 (37%)
	9 (76% S2)	0.017	-0.04 (7%)	0.16 (94%)
	19 (88% W)	0.014	-0.13 (0%)	0.09 (0%)
FA2	14 (72% REM)	0.021	0.25 (54%)	0.10 (0%)
	20 (63% S2)	0.014	0.05 (9%)	0.25 (100%)
pul_m	8 (73% REM)	0.038	0.02 (2%)	0.17 (99%)
sys_m	14 (72% REM)	0.054	0.21 (100%)	0.03 (0%)
dia_m	14 (72% REM)	0.021	0.15 (69%)	-0.02 (0%)
dia_e	1 (85% S2)	0.045	-0.08 (0%)	0.13 (87%)
FA3	6 (85% Wake)	0.017	0.25 (99%)	0.05 (11%)
	8 (73% REM)	0.054	0.02 (0%)	0.17 (80%)
ad_sv	5 (65% SWS)	0.001	-0.13 (0%)	0.14 (63%)

In the case of subjectively scored sleep and awakening quality, the only significant difference was observed in the case of morning well-being (*wb_m*) and affectivity (*aff*) tests.

The sleep microstate 5 represents deeper sleep (65% SWS, 35% S2) and the principal component scores of the original, in time misaligned sleep probabilistic curves, were not able to predict the values of *wb_m*. An increment in average correlation values was observed after curve alignment, and the percentage of linear models being significantly better than a constant model was higher (0% for misaligned curves and 91% for aligned curves).

The average correlations between the real level of affectivity in the morning and its values predicted by using principal component scores of misaligned curves corresponding either to microstate 8 (73% REM) or 19 (88% Wake) were close to zero, and none of the fitted models was better than a constant model. Thus, it was impossible to predict the level of affectivity by using the information from these microstates. After curve alignment, the average correlations increased in both cases, but were still low, see Table 2. Improvement in percentage of linear models being significantly better than a constant model was observed only in the case of microstate 8.

No relationship between the structure of the sleep microstate 9 (76% S2) and the level of affectivity in the morning was detected when misaligned sleep probabilistic curves were used. The average correlation was $\bar{\rho} = -0.04$ and the ratio of trained linear models which outperformed a constant model was less than 10% (Table 2). In contrast, after curve alignment, the average correlation increased to 0.16, and the percentage of linear models being significantly better than a constant model was also higher.

Considering the physiological measures, the curve alignment produced significant changes in correlations, especially in REM-related sleep microstate 14. The average correlations between the real and predicted values of FA2, systolic and diastolic pressure in the morning decreased after curve alignment (Table 2). The ratio of linear models outperforming a

constant model was equal to 0. However, for the misaligned cases, the percentage was above 50% in all three cases. This indicates that curve alignment is counterproductive for microstate 14.

Microstate 8 is also similar to the REM stage, but its characteristics differ from those of microstate 14. As depicted in Table 2, the average correlations with pulse rate in the morning changed from insignificant for in time misaligned curves to significant (≈ 0.17) after the curves were aligned. The curve alignment also increased the percentage of linear models outperforming a constant model (99%), indicating that there may exist a relationship between the structure of microstate 8 and the pulse rate in the morning.

The last difference was observed in the case of the S2-related sleep microstates 1 and 20 (Table 2) and FA2 or diastolic blood pressure in the evening. In both cases, curve alignment produced higher correlations in contrast to the case of misaligned curves. The percentage of linear models being better than a constant model was above 87% after curve alignment.

Significant average correlations (≈ 0.25) between the FA3 factor score and the structure of the microstate 6 diminished after curve alignment. This microstate characterises full awakening during the night or in the morning. The amount of time spent awake during the night influences our cognitive performance in the morning. However, after curve alignment the information about the exact amount of time spent awake is missing, and therefore, a decrement in average correlation values and in the percentage of linear models better than a constant model was observed.

For FA3 or *ad_sv* the curve alignment improved the prediction ability of linear models fitted to the principal component scores of either microstate 8 (73% REM) or 5 (65% SWS). In addition, the percentage of linear models outperforming a null model increased after curve alignment from 0% for misaligned curves to 80% for microstate 8 or 63% for microstate 5. However, the difference between average correlations produced by either misaligned or aligned curves of microstate 8 was at the edge of significance ($p\text{-value} = 0.054$).

3.2. Standard R&K sleep stages

The alignment of the sleep probabilistic curves caused changes in correlations between real and predicted results of daily measures, especially for the REM stage (Table 3).

Improvement was observed in the case of age, physiological factors, or the results of the numerical memory test. In the last case, the insignificant correlation between daily measures and misaligned curves became statistically significant after curve alignment. Moreover, the percentage of linear models which were significantly better than a constant model increased from 0% for misaligned curves to 85% for aligned curves.

In the case of the level of mood or affectivity, the average correlations between real and predicted values were significantly lower for aligned curves for the REM stage. However, no estimated linear model was better than a constant model

Table 3. Average correlation coefficients for daily measures and sleep stages. The same notation as in Table 2 is used.

Daily meas.	Sleep stage	p-value	$\bar{\rho}$ misaligned curves	$\bar{\rho}$ aligned curves
age	REM	0.026	0.40 (100%)	0.58 (100%)
drive	Wake	0.021	0.14 (60%)	−0.05 (0%)
mood	REM	0.006	0.02 (0%)	−0.23 (0%)
aff	REM	0.003	0.01 (0%)	−0.22 (0%)
	S2	0.021	0.03 (8%)	0.24 (100%)
FA2	REM	0.021	0.23 (83.2%)	0.42 (100%)
num_m	REM	0.014	−0.03 (0%)	0.13 (85%)

for either aligned or misaligned curves. We hypothesize that the relationship between the *REM* stage and the level of mood or affectivity simply does not exist, or we are not able to detect it either with original or in time aligned curves.

For the *Wake* stage, the only significant change observed was a decreased correlation with the level of drive in the aligned curves. More than half of the linear models were significantly better than a constant model when using principal component scores of the misaligned curves. This indicates the existence of a relationship between the *Wake* stage profile and the level of drive. However, after alignment, this relationship disappeared (the percentage was exactly 0).

Regarding the sleep stages *S1*, *S2* or *SWS*, the correlations between real and predicted values of daily measures were higher for aligned curves in several cases, but not significantly.

4. DISCUSSION AND CONCLUSIONS

In this study, we demonstrated the benefit of time alignment of sleep probabilistic curves when detecting the relationship between sleep structure and daily measures. The sleep probabilistic curves of either 20 sleep microstates or five standard sleep stages were aligned using the elastic warping method.

The improvement in average correlations between real and predicted values of daily measures was observed after curve alignment in the *S2*- and *SWS*-related sleep microstates, which is consistent with the results observed in the case of the standard sleep stages. We can conclude that the whole structure of the sleep states related to the *S2* stage or *SWS* is more important than the exact timing of their periods.

The PSM distinguishes two sleep microstates similar to the *REM* stage. In the case of microstate 8 (73% *REM*), curve alignment helps to detect existing relationships between the structure of the microstate and morning pulse rate or level of affectivity. However, alignment of the sleep probabilistic curves of microstate 14 (72% *REM*) led to a decrement in average correlations. When considering the standard *REM* stage, new or improved correlations with daily measures were observed after curve alignment. The benefit of curve alignment for the *REM* stage is therefore questionable. We recommend the use of both misaligned and aligned versions of sleep probabilistic curves of microstates similar to the *REM* stage, and careful interpretation of the results.

These results indicate different structures of the two *REM*-related sleep microstates and confirm the necessity of a larger set of sleep states considered by PSM, in contrast to the standard R&K model.

Considering either aligned sleep probabilistic curves of the *Wake* stage or sleep microstates related to wakefulness, the average correlations were significantly lower in comparison to the misaligned case. Moreover, the percentage of linear models outperforming a constant model decreased. We hypothesize that, for the *Wake* stage, the exact timing of sleep features is important for the sleep quality and cognitive performance in the morning.

Finally, we can conclude that curve alignment is useful as a pre-processing step, especially when analysing the structure of *S2*, *SWS*, or related microstates.

ACKNOWLEDGEMENT

This research was supported by the Slovak Research and Development Agency (grants number APVV-0668-12 and APVV-16-0202) and by the VEGA 2/0011/16 grant.

REFERENCES

- [1] Rechtschaffen, A., Kales, A. (1968). *A Manual of Standardized Terminology Techniques and Scoring System for Sleep Stages of Human Subjects*. Public Health Service, U.S. Government Printing Office, Washington, D.C.
- [2] Lewandowski, A., Rosipal, R., Dorffner, G. (2012). Extracting more information from EEG recordings for a better description of sleep. *Computer Method and Programs in Biomedicine*, 108 (3), 961–972.
- [3] Gervini, D., Gasser, T. (2004). Self-modeling warping functions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66 (4), 959–971.
- [4] Müller, H. G., Tang, R. (2008). Pairwise curve synchronisation for functional data. *Biometrika*, 95 (4), 875–889.
- [5] James, G.M. (2007). Curve alignment by moments. *The Annals of Applied Statistics*, 1 (2), 480–501.
- [6] Tucker, J.D., Wu, W., Srivastava, A. (2013). Generative models for functional data using phase and amplitude separation. *Computational Statistics and Data Analysis*, 61, 50–60.
- [7] Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I.H. (2011). Shape analysis of elastic curves in euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33 (7), 1415–1428.
- [8] Srivastava, A., Wu, W., Kurtek, S., Klassen, E., Marron, J. S. (2011). Registration of functional data using Fisher-Rao metric. *arXiv:1103.3817 [math.ST]*.
- [9] Klösch, G., Kemp, B., Penzel, T., Schlögl, A., Rappelsberger, P., Trenker, E., Gruber, G., Zeitlhofer, J., Saletu, B., Herrmann, W., Himanen, S., Kunz, D., Barbanoj, M., Röschke, J., Varri, A., Dorffner, G. (2001). The SIESTA project polygraphic and clinical database. *Medicine and Biology Magazine*, 20 (3), 51–57.

- [10] Rosipal, R., Lewandowski, A., Dorffner, G. (2013). In search of objective components for sleep quality indexing in normal sleep. *Biological Psychology*, 94 (1), 210–220.
- [11] Yao, F., Müller, H. G., Clifford, A. J., Dueker, S. R., Follet, J., Lin, Y., Buchholz, B. A., Vogel, J. S. (2003). Shrinkage estimation for functional principal component scores, with application to population kinetics of plasma folate. *Biometrics*, 59 (3), 676–685.

Received September 9, 2017.
Accepted February 12, 2018.