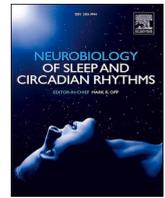




Contents lists available at ScienceDirect

# Neurobiology of Sleep and Circadian Rhythms

journal homepage: [www.elsevier.com/locate/](http://www.elsevier.com/locate/)

## Analysis and visualization of sleep stages based on deep neural networks

Patrick Krauss<sup>a,b,c,\*</sup>, Claus Metzner<sup>a,d</sup>, Nidhi Joshi<sup>a</sup>, Holger Schulze<sup>a</sup>, Maximilian Traxdorf<sup>e</sup>,  
Andreas Maier<sup>f</sup>, Achim Schilling<sup>a,b,g</sup>

<sup>a</sup> Neuroscience Lab, Experimental Otolaryngology, University Hospital Erlangen, Germany

<sup>b</sup> Cognitive Computational Neuroscience Group at the Chair of English Philology and Linguistics, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany

<sup>c</sup> Cognitive Neuroscience Center, University of Groningen, the Netherlands

<sup>d</sup> Biophysics, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany

<sup>e</sup> Department of Otolaryngology, Head and Neck Surgery, University Hospital Erlangen, Germany

<sup>f</sup> Machine Intelligence, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Germany

<sup>g</sup> Laboratory of Sensory and Cognitive Neuroscience, Aix-Marseille University, Marseille, France

### ARTICLE INFO

#### Keywords:

Sleep stage scoring  
Hypnodensity graphs  
Multidimensional scaling (MDS)  
Electroencephalography (EEG)  
Artificial neural networks  
Deep learning  
Polysomnography (PSG)  
Sleep cycle analysis

### ABSTRACT

Automatic sleep stage scoring based on deep neural networks has come into focus of sleep researchers and physicians, as a reliable method able to objectively classify sleep stages would save human resources and simplify clinical routines. Due to novel open-source software libraries for machine learning, in combination with enormous recent progress in hardware development, a paradigm shift in the field of sleep research towards automatic diagnostics might be imminent. We argue that modern machine learning techniques are not just a tool to perform automatic sleep stage classification, but are also a creative approach to find hidden properties of sleep physiology. We have already developed and established algorithms to visualize and cluster EEG data, facilitating first assessments on sleep health in terms of sleep-apnea and consequently reduced daytime vigilance. In the following study, we further analyze cortical activity during sleep by determining the probabilities of momentary sleep stages, represented as hypnodensity graphs and then computing vectorial cross-correlations of different EEG channels. We can show that this measure serves to estimate the period length of sleep cycles and thus can help to find disturbances due to pathological conditions.

### 1. Introduction

Sleep stage scoring is a standard procedure and part of every polysomnographic analysis (Bradley and Peterson, 2008; Burns et al., 2008). Up to now, sleep stage scoring based on physiological signals (EEG: electroencephalography, EMG: electromyography, EOG: electro-oculography) is performed by experienced clinicians, which do the classification by hand according to the AASM guidelines (Berry et al., 2012). However, this procedure is time consuming and highly prone to errors resulting in high inter-rater variability (Danker-Hopfe et al., 2004, 2009; Wendt et al., 2015). To overcome these limitations, machine learning algorithms were applied. These algorithms are still not fully trusted by clinicians as they act as a black-box, from which no one knows what the internal criteria for the different sleep stages really are. Thus, one core-problem of modern machine learning research has met clinical routines, the so called black-box problem (KrauseAdam and Ng, 2016) which in other scientific fields is also called the opacity debate

(De Laat, 2018). This problem can be tackled by simply using hand-crafted features such as time tags of K-complexes or sleep spindles as neural network input (Phan et al., 2018; Boostani et al., 2017), (for an example see (PhanQuan et al., 2013)). However, this procedure is very elaborate and the results are still not satisfying. In times of rising computing power and increasing storage capacities, a Big Data approach where neural networks are finding the best features in a self-organized way seems to be more promising. In order to analyze these features and the emerging internal representations, novel concepts for data visualization are needed. A sophisticated dimensionality reduction method for visualising high dimensional data is a demanding task, as certain popular methods can lead to pseudo clustering of data points caused by the initial conditions and the used hyper-parameters of the projection algorithm. A famous example is t-distributed stochastic neighbor embedding (t-SNE (van der Maaten and Hinton, 2008)), which is highly unstable and extremely dependent on the initialization and choice of parameters (Martin et al., 2016).

\* Corresponding author. Neuroscience Lab, Experimental Otolaryngology, University Hospital Erlangen, Germany.

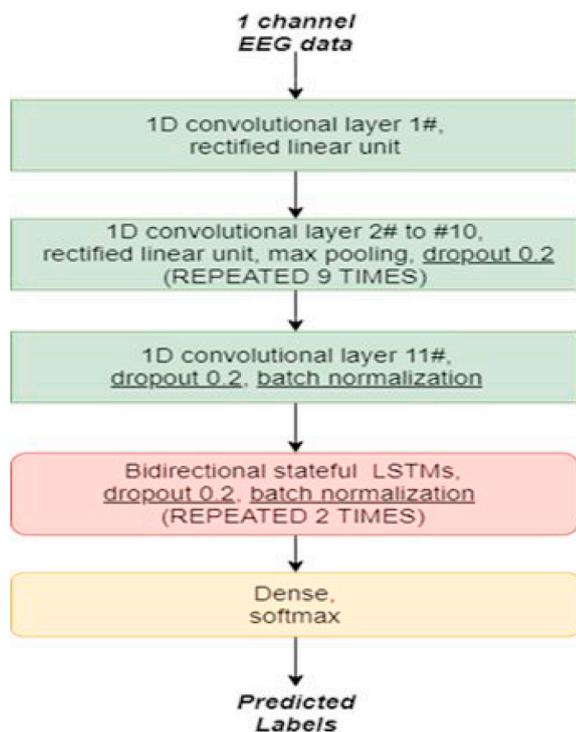
E-mail address: [patrick.krauss@uk-erlangen.de](mailto:patrick.krauss@uk-erlangen.de) (P. Krauss).

<https://doi.org/10.1016/j.nbscr.2021.100064>

Received 16 July 2020; Received in revised form 27 February 2021; Accepted 1 March 2021

Available online 12 March 2021

2451-9944/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



**Fig. 1. Network architecture.**

Building blocks of the deep neural network trained on sleep stage classification of EEG data. The network consists of eleven 1D convolutional layers, nine max pooling layers, 2 layers of bidirectional stateful LSTMs, and a fully connected classification layer with softmax output.

In previous studies, we developed several approaches to statistically analyze and visualize high-dimensional neural data (Krauss et al., 2018a; Schilling et al., 2018). We developed a statistical method for analyzing and comparing high-dimensional spatio-temporal cortical activation patterns for different auditory and somatosensory stimulus conditions in rodents and humans (Krauss et al., 2018a). The cortical activity patterns were represented by amplitude vectors calculated via a sliding window method (for the exact procedure see (Krauss et al., 2018a)). We could already demonstrate that this method can discriminate different sleep stages in human EEG recordings (Krauss et al., 2018b). Furthermore, we could analyze the microstructure of cortical activity during sleep and found that it reflects respiratory events and the state of daytime vigilance (Traxdorf et al., 2019). Recently, our method has been generalized, and can now be used to analyze and compare representations of artificial neural networks (Schilling et al., 2018). In the following study, we first illustrate how to visualize the representations of EEG data gained from different layers of artificial neural networks (sleep stage embeddings). Furthermore, we demonstrate that these complex representations cluster better in higher layers of the artificial neural networks, quantified by the generalized discrimination value (GDV, see also (Schilling et al., 2018)). Subsequently, we visualize the output of the last layer of the neural network, and thus the momentary probabilities of the predicted sleep stages, as so called hypnodensity graphs (as introduced by (Jens et al., 2018)). Finally, we use these probability vectors to calculate vectorial cross-correlations, in order to analyze the period length of sleep cycles.

## 2. Methods

### 2.1. Data base

The study was conducted in the Department of Otorhinolaryngology, Head Neck Surgery, of the Friedrich-Alexander University Erlangen-

Nürnberg (FAU), following approval by the local Ethics Committee (323–16 Bc). All 68 participating subjects, 46 male and 22 female, mean age  $32.5 \pm 11.5$  years, were recruited by the Department of Otorhinolaryngology, Head and Neck Surgery. Written informed consent was obtained from the participants before the cardiorespiratory polysomnography (PSG). Inclusion criterion for this study was age between 21 and 80 years. Exclusion criteria were a positive history of misuse of sedatives, alcohol or addictive drugs and untreated sleep disorders. Data analysis was carried out during time in bed of the subjects, accumulating to a total recording time of approximately 510 h.

Each of the 68 data sets comprising a full night of sleep consisted of 3 channels of EEG recordings (F4-M1, C4-M1, O2-M1) and the corresponding sleep stages. Sleep stages were analyzed and scored visually in 30 s epochs according to the AASM criteria (Version 2.1, 2014) by a sleep specialist accredited by the German Sleep Society (DGSM) (Conradt, 2007; Berry et al., 2012). Thereby, typical artifacts (Tatum et al., 2011) have been marked and removed subsequently for further analysis and processing steps.

### 2.2. Sleep stage classification with neural networks

A deep neural network was trained on single channel sleep EEG data. Therefore, we used the total number of 68 data sets and split them randomly in 54 training data sets (80%) and 14 test data sets (20%), as this is the standard approach in machine learning and pattern recognition (Bishop, 2006; Goodfellow et al., 2016). For each data set, the three different EEG channels were concatenated, yielding a corresponding single channel data set with three times the duration of the original data set. As labels for supervised training and as ground truth for the test data set we used the sleep stages from visual analysis and scoring.

The network consisting of several convolutional layers (LeCun Bengio et al., 1995) and two bidirectional stateful LSTM layers (Hochreiter and Schmidhuber, 1997) was trained with error back propagation to classify the sleep stages (for exact network architecture see Fig. 1). The convolutional layers allow to extract relevant features for classification in an unsupervised way (LeCun et al., 2015), i.e. without the need for manually extracting features like frequency power spectra or amplitudes. Recurrent network layers like the here used LSTM layers allow to extract and capture temporal information [] from the sequences of sleep stages.

### 2.3. Software resources

The software is written using the programming language python 3.6, together with the KERAS AI library (Chollet, 2018) and tensorflow backend. The data is pre-processed using Numpy (Stéfan van der Walt and Varoquaux, 2011). For multi-dimension-scaling (MDS) the scikit-learn library (Pedregosa et al., 2011) is used and the visualization of the results is done using the matplotlib library (Hunter, 2007).

### 2.4. Hypnodensity correlations

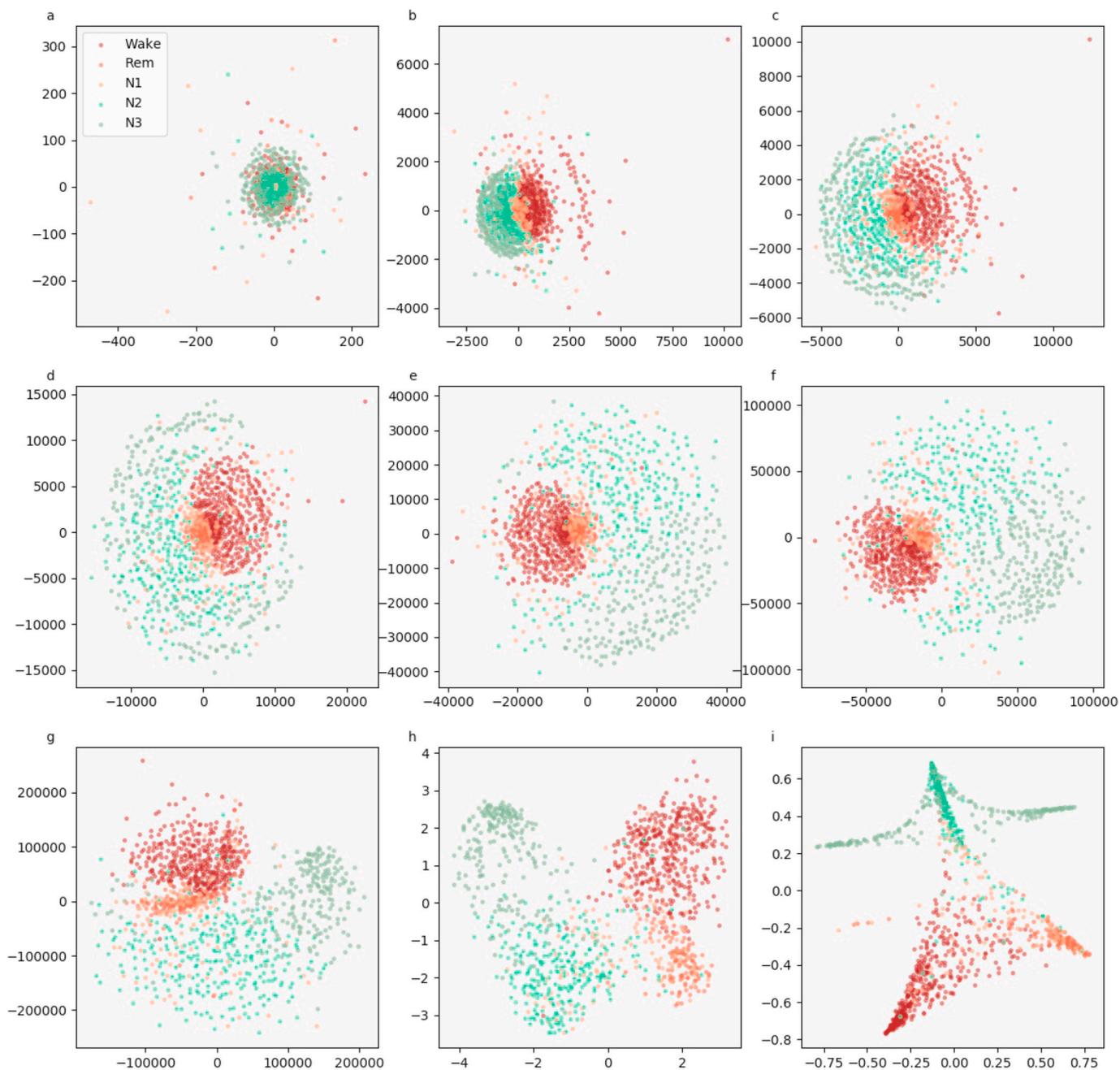
The hypnodensity cross-correlation for lag-time  $\tau$  is defined as

$$r_{x,y}(\tau) = \frac{\frac{1}{T-\tau} \sum_{t=1}^{T-\tau} (\mathbf{x}_t - \bar{\mathbf{x}}) \circ (\mathbf{y}_{t+\tau} - \bar{\mathbf{y}})}{\sqrt{\frac{1}{T} \sum_{t=1}^T \|\mathbf{x}_t - \bar{\mathbf{x}}\|^2} \cdot \sqrt{\frac{1}{T} \sum_{t=1}^T \|\mathbf{y}_t - \bar{\mathbf{y}}\|^2}} \quad (1)$$

and the hypnodensity auto-correlation for lag-time  $\tau$  is defined as

$$r_{x,x}(\tau) = \frac{\frac{1}{T-\tau} \sum_{t=1}^{T-\tau} (\mathbf{x}_t - \bar{\mathbf{x}}) \circ (\mathbf{x}_{t+\tau} - \bar{\mathbf{x}})}{\sqrt{\frac{1}{T} \sum_{t=1}^T \|\mathbf{x}_t - \bar{\mathbf{x}}\|^2}} \quad (2)$$

with  $\mathbf{x}_t = (x_1, x_2, x_3, x_4, x_5)_t^T$  and  $\mathbf{y}_t = (y_1, y_2, y_3, y_4, y_5)_t^T$  being 5-dimensional vectors that contain the momentary probabilities of all sleep



**Fig. 2. Sleep stage embeddings.**

MDS visualization of EEG data representations (embeddings) within the hidden layers of the deep neural network: input layer (a), max pooling layers (b–g), last LSTM layer (h), and softmax layer (i). The hidden layers (b–g) lead to increasingly better separability of sleep stages compared to z-scored raw EEG data (a). The softmax classification layer (i) normalizes the data to a value range of [0,1]. Note that absolute coordinates of points in MDS projections have no particular meaning other than scaling relative distances between any pair of points.

stages (indices 1 to 5) at time  $t$  for EEG channels  $x$  and  $y$ , respectively, and the mean probability vectors  $\bar{x} = \frac{1}{T} \sum_{t=1}^T x_t$  and  $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$ . Here,  $\cdot$  is the dot product, and  $\| \cdot \|$  the vector norm.

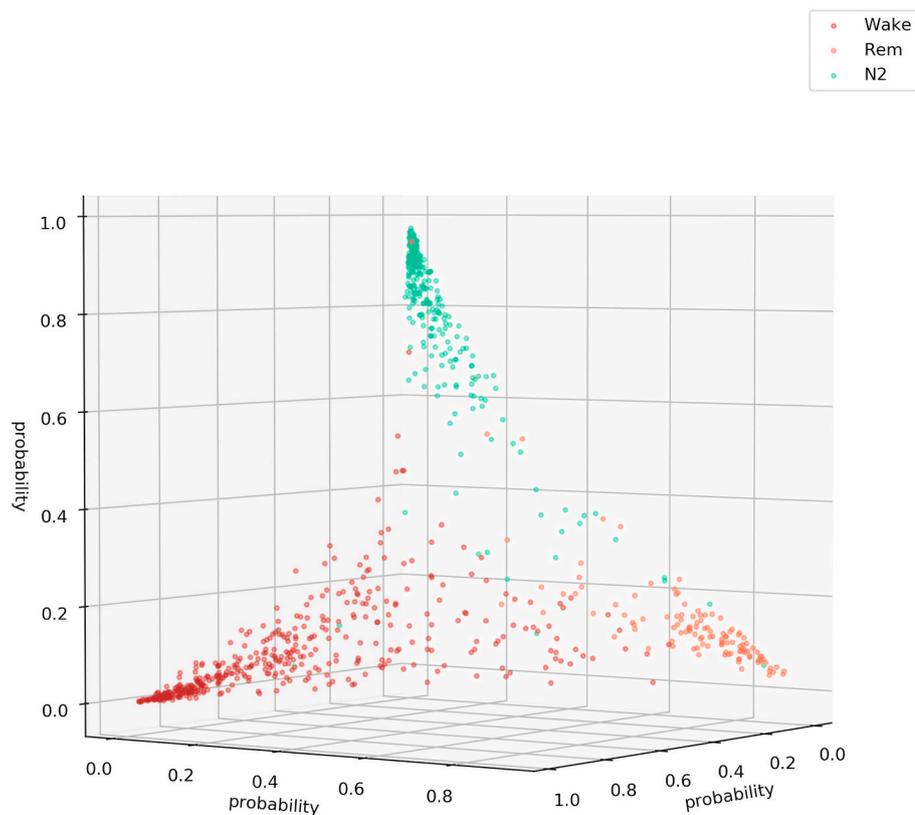
### 3. Results

#### 3.1. Sleep stage embeddings

We could show that even a single EEG channel contains enough information to classify different sleep stages. We used several convolutional blocks, which transform the input EEG data into high-dimensional complex features. A procedure referred to as *feature space embedding*.

Thus, in contrast to handcrafted features we make the neural network to find its own features. This higher order features lead to a good separability of the transformed EEG vectors (neural network layer output or representation) belonging to different sleep stages. In the following, we refer to these self-organized representations as *sleep stage embeddings*.

We visualize the sleep stage embeddings by a dimensionality reduction into 2D, using multidimensional scaling (MDS) (Krauss et al., 2018a) (Fig. 2). Moreover, we evaluate the generalized discrimination value (GDV), which quantifies separability of data classes in high-dimensional state spaces (Schilling et al., 2018). The MDS plots show that the convolutional layers lead to better separability (Fig. 2b–g) compared to z-scored raw EEG data (network input, Fig. 2a).



**Fig. 3. Softmax output for three sleep stages (Wake, REM, N2).**

This figure illustrates that the softmax output spans a hyper-plane (in 3D a 2D hyper-plane with 3 corners). Note that the possible outputs lie in the volume below the hyper-plane. Not all points lie on the plane as confusions with sleep stage N3 or N1 cannot be shown in a 3D plot. Thus, if the softmax output would for example point to sleep stage N3 the point would lie near the origin (0,0,0). Note that absolute coordinates of points in MDS projections have no particular meaning other than scaling relative distances between any pair of points.

Furthermore, the LSTM layer transforms the data so that the different classes are linearly separable (Fig. 2h). The softmax classification layer normalizes the data (value range [0,1]). The shape of the clusters in Fig. 2i can be explained by the fact that the softmax layers places the different classes at the 5 edges of a 4D-hyper-plane (example for 3D projections shown in Figs. 3 and 4). As the probabilities of all 5 sleep stages have to sum up to a value of 1, the range of possible softmax outputs is located on this hyper-plane. The MDS projection of the hyper-plane in 2D leads to the shown patterns (Fig. 2i). Note that, Figs. 2–4) were generated using one sample training data set. Labels are neither used during embedding in the neural network layers, nor for projection, but they serve only for color coding the data points after embedding and projection.

The quantification of the class separability via the GDV (Fig. 2, red dots: input, 6 max-pooling layers, the second LSTM layer, the softmax output layer), shows that the convolutional layers perform data preprocessing. In a previous study, we could demonstrate that the GDV has to overcome an “energy barrier”, when the data structure is complex. Thus, the number of layers where the GDV only slightly decreases depends on data complexity. As complex features are needed to classify the data, there are relatively many initial layers with small GDV decrease (cf. Fig. 5). The LSTM layer leads to a clear increase of separability, i.e. drop in the GDV value (8th red point in Fig. 5, compare also Fig. 2 i). The softmax layer causes a further decrease of the GDV value.

### 3.2. Hypnodensity graphs

An efficient illustration of the softmax layer output are so called hypnodensity graphs (Jens et al., 2018). Here, the probabilities for all sleep stages at each time point are shown instead of only the most probable sleep stage, as it is done in traditional hypnograms. This is an elegant approach, since the probabilities (output of softmax layer) can be compared with the uncertainties, i.e. different assessments of different somnologists, which cause the relatively high

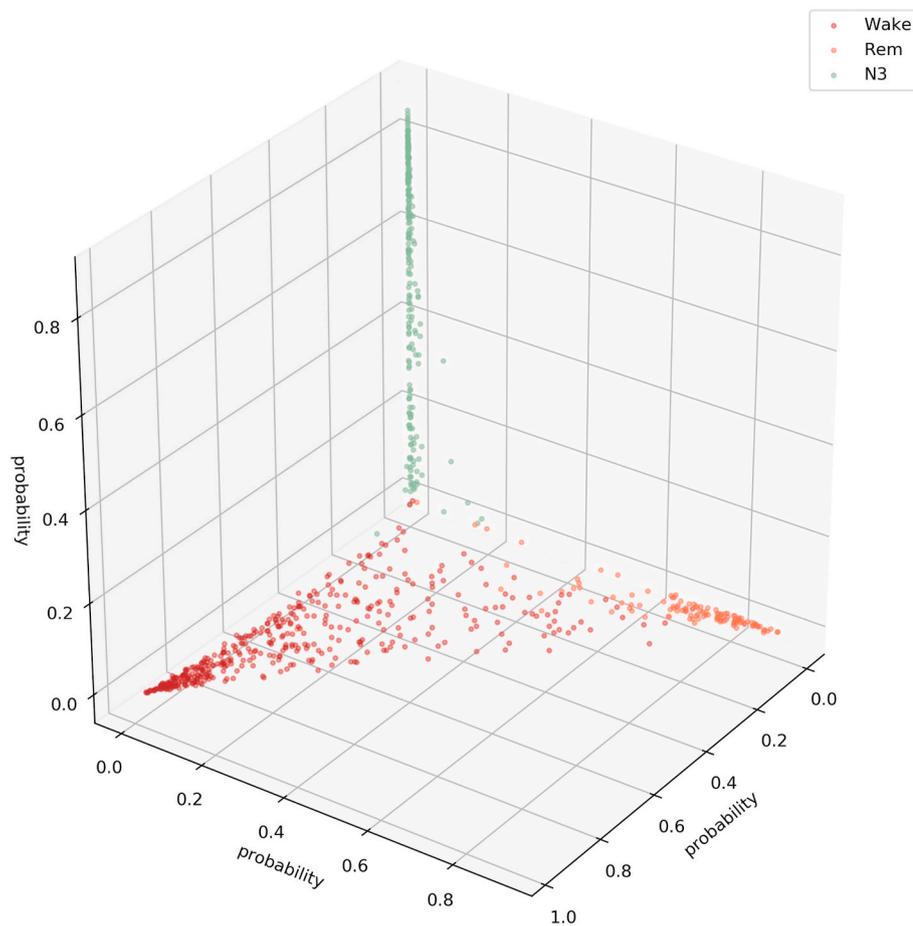
inter-rater-variability (Danker-Hopfe et al., 2004, 2009; Wendt et al., 2015).

Additionally, the automatic sleep stage classification can be used to evaluate the data with higher temporal resolution. Even though, the network has been trained on labeled data with a temporal resolution of 30 s which is in accordance with AASM scoring rules (Berry et al., 2012), we evaluate sleep stage probabilities predicted by the neural network with a temporal resolution of 5 s, using a sliding window approach (Fig. 6 and Fig. S1 – S10). Thus, we still analyzed 30 s windows, yet with 5 s steps. The predicted sleep stage probabilities were assigned to the respective center of each window of 30 s width.

The hypnodensity plots show that each EEG channel is sufficient to perform sleep stage classification, as the hypnodensity graphs are very similar (Fig. 6) for all the channels. Interestingly, the sleep stage *wake* seems to be slightly over-represented in channel C4 (6a), whereas in channel O2, N3 seems to slightly dominate (6c). The similarity between channels can be quantified, for example using pairwise cross-correlations.

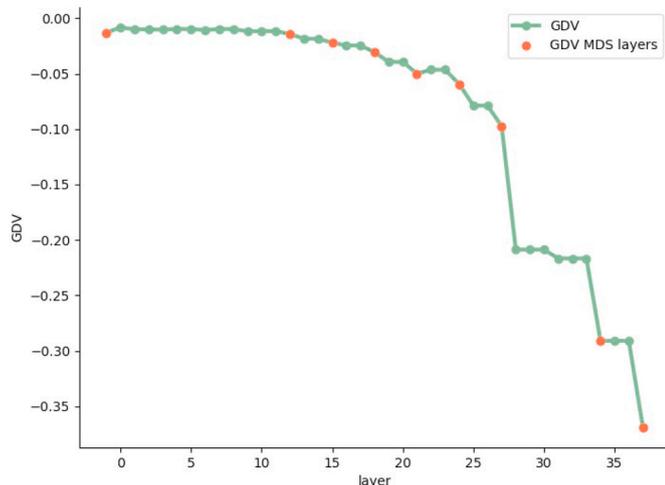
### 3.3. Hypnodensity based sleep cycle period length analysis

On the basis of these EEG channel specific hypnodensity graphs, sleep stage probabilities across and within channels can be compared (see Methods). For this purpose we calculate generalized vector auto- and cross-correlations of the 5-dimensional probability vectors for the 3 EEG channels of each data set. This correlations provide valuable information on neuronal activity patterns during sleep. In particular, all auto- as well as cross-correlations have a global maximum, which is near to the value 1, at zero lag time. This means that there is no time shift of the neuronal patterns between the different channels. Furthermore, the high correlation coefficients at lag time zero indicate that all channels are sufficient to perform sleep scoring, which could already been predicted by the hypnodensity graphs. Local side-maxima indicate repetitions of neuronal patterns as they occur in repeating sleep-cycles.



**Fig. 4. Softmax output of for three sleep stages (Wake, REM, N3).**

This figure illustrates that the softmax output spans a hyper-plane (in 3D a 2D hyper-plane with 3 corners). It can be seen that the sleep stage N3 is not confused with the Wake or the REM state. Note that absolute coordinates of points in MDS projections have no particular meaning other than scaling relative distances between any pair of points.



**Fig. 5. Separability of sleep stages.**

Separability increases with increasing layer depth. Note that, a GDV of 0 corresponds to non-separable data classes, whereas a GDV of  $-1$  corresponds to a perfect data class separation. Red dots refer to the layer outputs shown in 2 (first red dot: input, 2nd-7th: max pooling layers, 8th: 2nd LSTM layer, 9th: softmax output layer). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

For subject 55, the auto and cross-correlations show a local maximum at a lag-time of about 100 min (Fig. 7). This means that after this period of time the sleep stages repeat. In contrast, data of subject 59

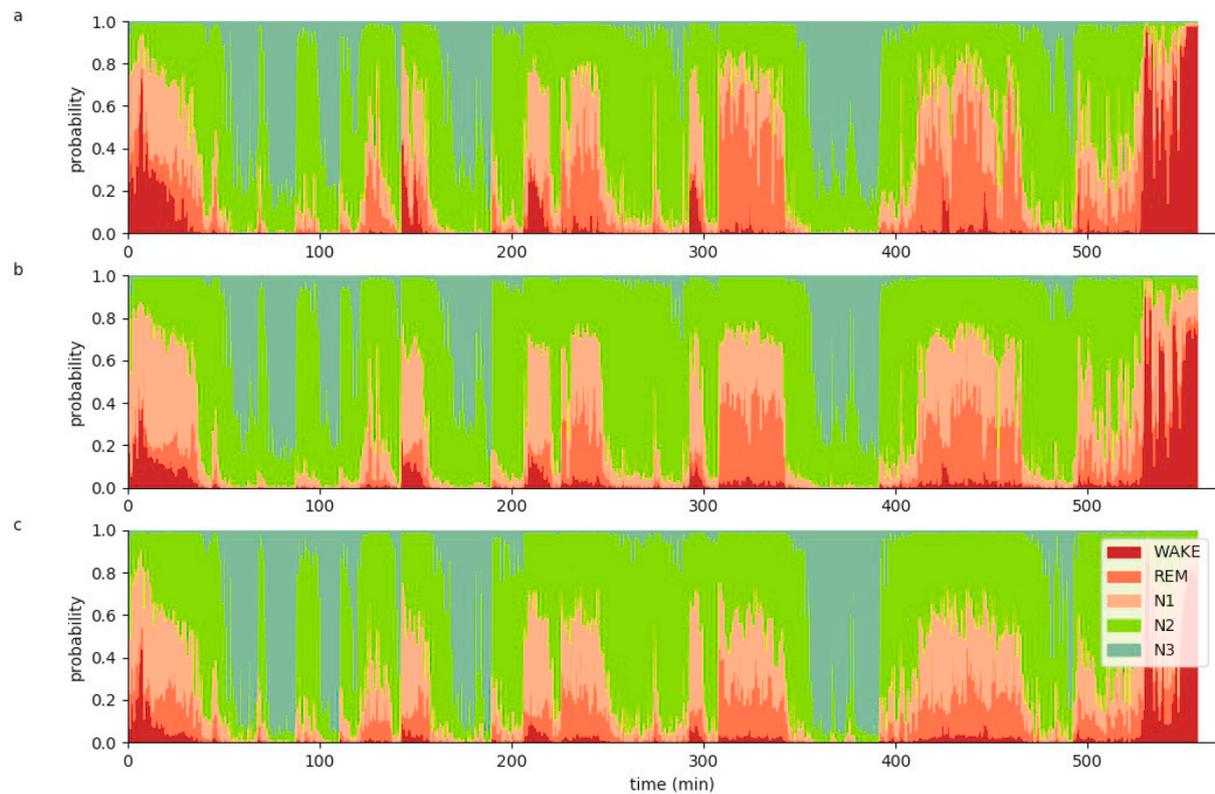
show two local maxima at lag-times of 75 and 150 min, respectively (Fig. 8). Hence, this subject’s sleep stage period length is about 75 min. These estimated period lengths have been confirmed by the sleep stage scoring based on visual analysis performed by sleep specialists. This means that our novel method helps to reliably estimate the individual period length of sleep cycles. Further data are shown in Figs. S11 – S19.

#### 4. Discussion

In this study, we present novel approaches for the evaluation and visualization of neural data recorded during human sleep. In contrast to numerous previous studies and network architectures used to perform automatic sleep stage scoring (e.g. (Ebrahimi et al., 2008; Koley and Dey, 2012; Zhang and Wu, 2017; Amiya et al., 2018; Alickovic and Subasi, 2018; Zhang et al., 2016)), we provide approaches helping to interpret and understand the network output.

Thus, we illustrate how deep neuronal networks automatically find features that help to separate different sleep stages based on single channel EEG data. That this representations actually lead to a good separability of the different sleep stages is visualized using MDS plots and quantified using the GDV value.

The softmax output, i.e. the predicted probabilities for each sleep stage at every time step, is visualized with hypnodensity graphs (Jens et al., 2018) a sophisticated way to show the uncertainties in sleep stage classification. The calculation of the vectorial auto- and cross-correlations within and between channels, as introduced in this study, is a tool to gain more insight into the architecture of sleep, e.g. sleep cycles. We propose the auto-correlation plots as novel tool to estimate the average period length of sleep stage cycles, and to find disturbances in the period length of sleep cycles, e.g. as an indicator for



**Fig. 6. Hypnodensity graph.**

Hypnodensity graph of subject 55 with a temporal resolution of 5 s separately evaluated for the three different EEG channels C4 (a), F4 (b) and O2 (c). The probabilities for all sleep stages at each time point are shown. Each EEG channel is sufficient to perform sleep stage classification, as the hypnodensity graphs are similar.

pathological conditions, for example in the case of patients with obstructive sleep apnea (OSAS). Here, sleep is dominated by events like arousals and short periods of wakefulness. These effects “destroy” the ordered sequence of sleep stages of normal sleep cycles. The cross-correlation plots between different recording channels may serve to shed light into phenomena like local or fragmented sleep. The global maximum of cross-correlation plots, especially if the corresponding lag-time is different from zero, indicates travelling waves across the cerebral cortex, or the order at which the different regions change sleep stages.

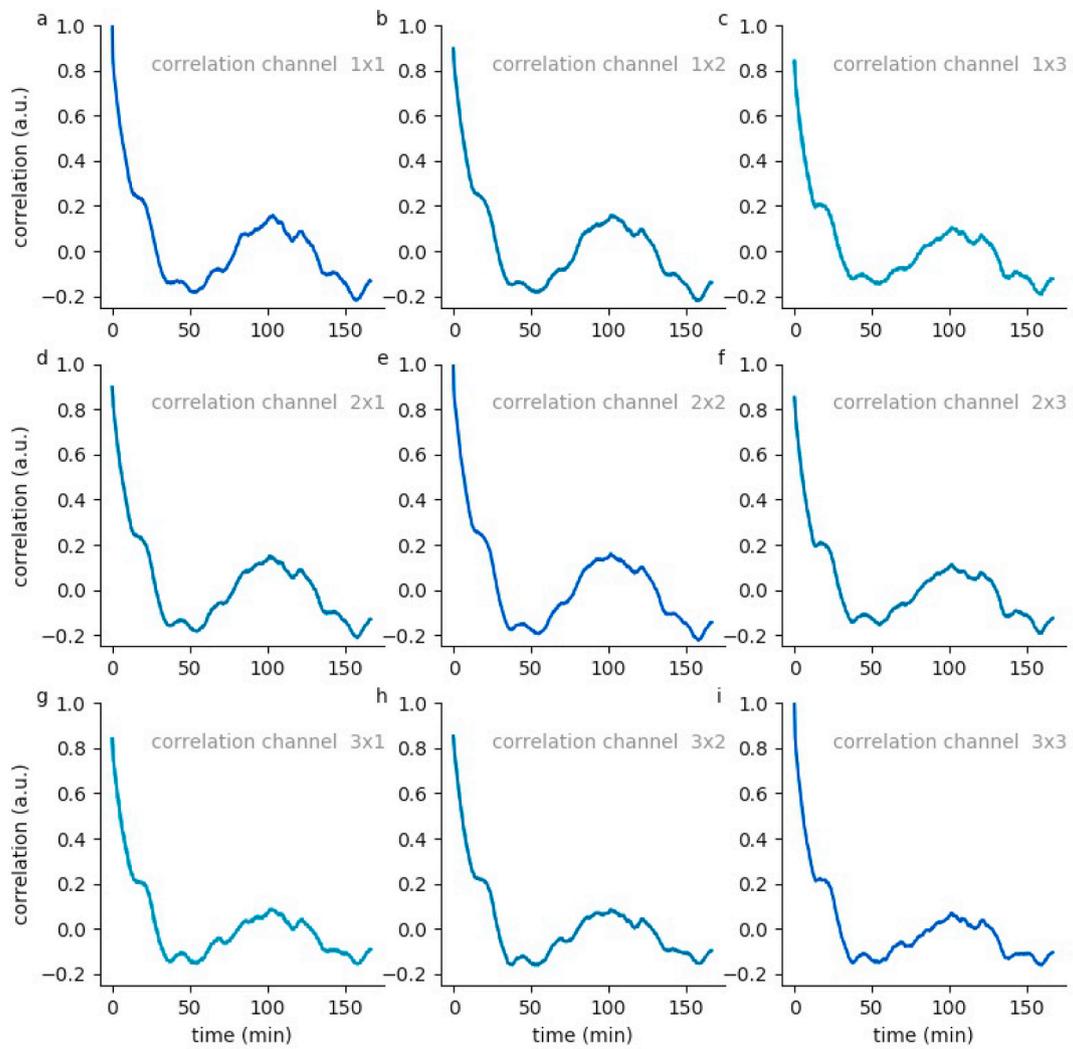
Another interesting extension of the correlation method for future studies would be the application to non-stationary cases, i.e. changing period lengths over night. For instance, calculating the correlation only for subsequent parts of the entire recording time would yield different lag-times at maximum correlation, i.e. period length estimates, for each different part. If correlation analysis is initially performed over the complete night, as shown above, this yields the average cycle time, and hence gives a hint on how to choose the duration of the different parts to be analyzed. Alternatively, a sliding window approach could be applied, yielding a time-dependent period length, i.e. lag-time at maximum correlation. Even more sophisticated would be the application of super-statistical methods (Metzner et al., 2015), which are designed to detect the temporal change of statistical parameters in complex systems.

This study is in line with numerous previous studies trying to use

artificial neural networks as a model (KriegeskortePamela, 2018; Schilling et al., 2020a; Gerum et al., 2020; Krauss et al., 2019a, 2019b; Krauss and Maier, 2020), and as a tool (Schilling et al., 2020b) to analyze and understand brain activity. This combined approach at the intersection of neuroscience and artificial intelligence is crucial to make further progress in the field, as traditional popular analysis methods have been proven to be not sufficient to understand brain dynamics (Jonas and Paul Kording, 2017; Brown, 2014; Lazebnik, 2002).

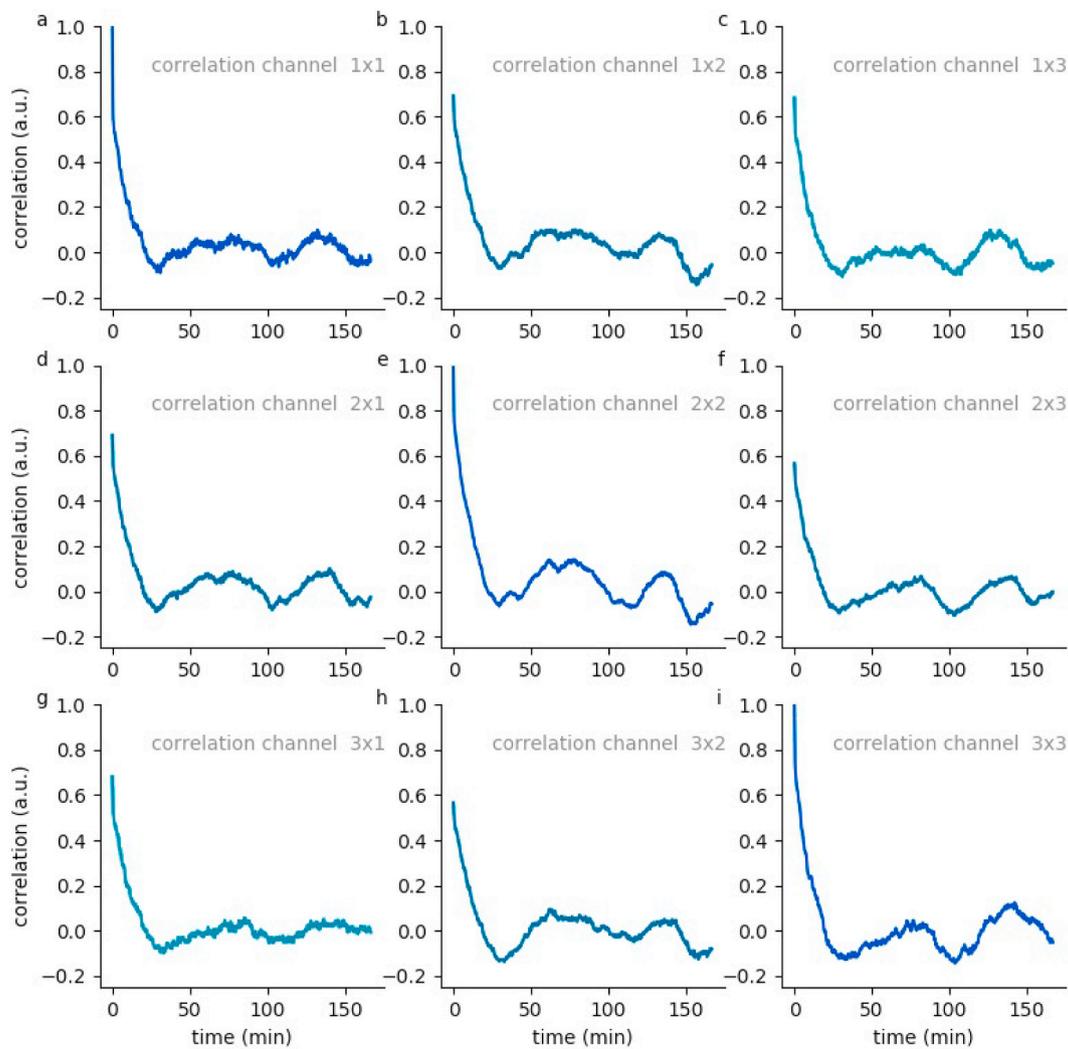
#### CRediT authorship contribution statement

**Patrick Krauss:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Claus Metzner:** Formal analysis, Investigation, Methodology, Software, Validation, Writing – review & editing. **Nidhi Joshi:** Formal analysis, Software, Visualization. **Holger Schulze:** Resources. **Maximilian Traxdorf:** Data curation, Resources. **Andreas Maier:** Investigation, Resources, Supervision, Validation. **Achim Schilling:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft.



**Fig. 7. Hypnodensity cycles.**

Temporal auto- and cross correlations of 5-dimensional hypnodensity probability vectors of sleep stages. In subject 55, a local maximum at a lag-time of about 100 min can be observed, indicating an individual period length of sleep cycles of 100 min.



**Fig. 8. Hypnodensity cycles.**

Temporal auto- and cross correlations of 5-dimensional hypnodensity probability vectors of sleep stages. In subject 59, two local maxima at lag-times of about 75 and 150 min can be observed, indicating an individual period length of sleep cycles of 75 min.

#### Declaration of competing interest

All authors declare no competing financial interests.

#### Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): grant KR5148/2-1 to PK – project number 436456810, and the Emerging Talents Initiative (ETI) of the University Erlangen-Nuremberg (grant 2019/2-Phil-01 to PK). We thank NVidia for the donation of GPU devices for Machine Learning purposes.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nbscr.2021.100064>.

#### References

- Alickovic, Emina, Subasi, Abdulhamit, 2018. Ensemble svm method for automatic sleep stage classification. *IEEE Trans. Instrument. Meas.* 67 (6), 1258–1265.
- Amiya, Patanaik, Ong, Ju Lynn, J Gooley, Joshua, Ancoli-Israel, Sonia, Chee, Michael WL., 2018. An end-to-end framework for real-time automatic sleep stage classification. *Sleep* 41 (5), zsy041.
- Berry, Richard B., Brooks, Rita, Charlene, E Gamaldo, Harding, Susan M., Marcus, C., Bradley, V Vaughn, et al., 2012. The aasm manual for the scoring of sleep and associated events. *Rules Terminol. Techn. Specifi. Darien, Illinois, Am. Acad. Sleep Med.* 176, 2012.
- Bishop, Christopher M., 2006. *Pattern Recognition and Machine Learning*. Springer.
- Boostani, Reza, Karimzadeh, Foroozan, Nami, Mohammad, 2017. A comparative review on sleep stage classification methods in patients and healthy individuals. *Comput. Methods Progr. Biomed.* 140, 77–91.
- Bradley, V Vaughn, Peterson, Giallanza, 2008. Technical review of polysomnography. *Chest* 134 (6), 1310–1319.
- Brown, Joshua W., 2014. The tale of the neuroscientists and the computer: why mechanistic theory matters. *Front. Neurosci.* 8, 349.
- Burns, Joseph W., Crofford, Leslie J., Chervin, Ronald D., 2008. Sleep stage dynamics in fibromyalgia patients and controls. *Sleep Med.* 9 (6), 689–696.
- Chollet, François, et al., 2018. Keras: the python deep learning library. *ascl. S. ascl: 1806.022*.
- Conrad Iber, 2007. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification*.
- Danker-Hopf, Heidi, Kunz, Dieter, Gruber, Georg, Klösch, Gerhard, José, L Lorenzo, Sari-Leena, Himanen, Kemp, Bob, Penzel, Thomas, Rösche, Joachim, Dorn, Hans, et al., 2004. Interrater reliability between scorers from eight european sleep laboratories in subjects with different sleep disorders. *J. Sleep Res.* 13 (1), 63–69.
- Danker-Hopf, Heidi, Anderer, Peter, Zeitlhofer, Josef, Boeck, Marion, Dorn, Hans, Gruber, Georg, Heller, Esther, Loretz, Erna, Moser, Doris, Parapatics, Silvia, et al., 2009. Interrater reliability for sleep scoring according to the rechtschaffen & kales and the new aasm standard. *J. Sleep Res.* 18 (1), 74–84.
- De Laat, Paul B., 2018. Algorithmic decision-making based on machine learning from big data: can transparency restore accountability? *Philos. Technol.* 31 (4), 525–541.
- Ebrahimi, Farideh, Mikaeili, Mohammad, Estrada, Edson, Nazeran, Homer, 2008. Automatic sleep stage classification based on eeg signals by using neural networks

- and wavelet packet coefficients. In: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, pp. 1151–1154.
- Gerum, Richard C., Erpenbeck, André, Krauss, Patrick, Schilling, Achim, 2020. Sparsity through evolutionary pruning prevents neuronal networks from overfitting. *Neural Network*.
- Goodfellow, Ian, Bengio, Yoshua, Courville, Aaron, Bengio, Yoshua, 2016. *Deep Learning*, vol. 1. MIT press Cambridge.
- Hochreiter, Sepp, Schmidhuber, Jürgen, 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780.
- Hunter, John D., 2007. Matplotlib: a 2d graphics environment. *Comput. Sci. Eng.* 9 (3), 90–95.
- Jens, B Stephansen, Olesen, Alexander N., Olsen, Mads, Ambati, Aditya, Eileen, B Leary, Moore, Hyatt E., Carrillo, Oscar, Lin, Ling, Han, Fang, Han, Yan, et al., 2018. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. *Nat. Commun.* 9 (1), 1–15.
- Jonas, Eric, Paul Kording, Konrad, 2017. Could a neuroscientist understand a microprocessor? *PLoS Comput. Biol.* 13 (1), e1005268.
- Koley, B., Dey, Debangshu, 2012. An ensemble system for automatic sleep stage classification using single channel eeg signal. *Comput. Biol. Med.* 42 (12), 1186–1195.
- Krause, Josua, Adam, Perer, Ng, Kenney, 2016. Interacting with predictions: visual inspection of black-box machine learning models. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 5686–5697.
- Krauss, Patrick, Maier, Andreas, 2020. Will we ever have conscious machines? *Front. Comput. Neurosci.* 14.
- Krauss, Patrick, Metzner, Claus, Schilling, Achim, Tziridis, Konstantin, Traxdorf, Maximilian, Wollbrink, Andreas, Rampf, Stefan, Pantev, Christo, Schulze, Holger, 2018a. A statistical method for analyzing and comparing spatiotemporal cortical activation patterns. *Sci. Rep.* 8 (1), 1–9.
- Krauss, Patrick, Schilling, Achim, Bauer, Judith, Tziridis, Konstantin, Metzner, Claus, Schulze, Holger, Traxdorf, Maximilian, 2018b. Analysis of multichannel eeg patterns during human sleep: a novel approach. *Front. Hum. Neurosci.* 12, 121.
- Krauss, Patrick, Zankl, Alexandra, Schilling, Achim, Schulze, Holger, Metzner, Claus, 2019a. Analysis of structure and dynamics in three-neuron motifs. *Front. Comput. Neurosci.* 13 (5).
- Krauss, Patrick, Schuster, Marc, Dietrich, Verena, Schilling, Achim, Schulze, Holger, Metzner, Claus, 2019b. Weight statistics controls dynamics in recurrent neural networks. *PLoS One* 14 (4), e0214541.
- Kriegeskorte, Nikolaus, Pamela, K Douglas, 2018. Cognitive computational neuroscience. *Nat. Neurosci.* 21 (9), 1148–1160.
- Lazebnik, Yuri, 2002. Can a biologist fix a radio?—or, what i learned while studying apoptosis. *Canc. Cell* 2 (3), 179–182.
- LeCun, Yann, Bengio, Yoshua, et al., 1995. Convolutional networks for images, speech, and time series. *Handb. Brain Theor. Neural Network*. 3361 (10), 1995.
- LeCun, Yann, Bengio, Yoshua, Hinton, Geoffrey, 2015. Deep learning. *Nature* 521 (7553), 436–444.
- Martin, Wattenberg, Viégas, Fernanda, Johnson, Ian, 2016. How to Use t-Sne Effectively. *Distill*.
- Metzner, Claus, Mark, Christoph, Steinwachs, Julian, Lautscham, Lena, Stadler, Franz, Fabry, Ben, 2015. Superstatistical analysis and modelling of heterogeneous random walks. *Nat. Commun.* 6 (1), 1–8.
- Pedregosa, Fabian, Varoquaux, Gaël, Gramfort, Alexandre, Michel, Vincent, Bertrand, Thirion, Grisel, Olivier, Blondel, Mathieu, Peter, Prettenhofer, Weiss, Ron, Vincent, Dubourg, et al., 2011. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Phan, Huy, Andreotti, Fernando, Cooray, Navin, Chén, Oliver Y., De Vos, Maarten, 2018. Automatic sleep stage classification using single-channel eeg: learning sequential features with attention-based recurrent neural networks. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 1452–1455.
- Phan, Huy, Quan, Do, The-Luan, Do, Vu, Duc-Lung, 2013. Metric learning for automatic sleep stage classification. In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 5025–5028.
- Schilling, Achim, Metzner, Claus, Rietsch, Jonas, Gerum, Richard, Schulze, Holger, Krauss, Patrick, 2018. How Deep is Deep Enough?—Quantifying Class Separability in the Hidden Layers of Deep Neural Networks *arXiv preprint arXiv:1811.01753*.
- Schilling, Achim, Gerum, Richard, Zankl, Alexandra, Schulze, Holger, Metzner, Claus, Krauss, Patrick, 2020a. Intrinsic Noise Improves Speech Recognition in a Computational Model of the Auditory Pathway. *bioRxiv*.
- Schilling, Achim, Tomasello, Rosario, Henningsen-Schomers, Malte R., Surendra, Kishore, Haller, Martin, Karl, Valerie, Peter, Uhrig, Maier, Andreas, Krauss, Patrick, 2020b. Analysis of Ongoing Neuronal Activity Evoked by Continuous Speech with Computational Corpus Linguistics Methods *bioRxiv*.
- Stéfan van der Walt, S Chris Colbert, Varoquaux, Gael, 2011. The numpy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* 13 (2), 22–30.
- Tatum, William O., Dworetzky, Barbara A., Schomer, Donald L., 2011. Artifact and recording concepts in eeg. *J. Clin. Neurophysiol.* 28 (3), 252–263.
- Traxdorf, Maximilian, Krauss, Patrick, Schilling, Achim, Schulze, Holger, Tziridis, Konstantin, 2019. Microstructure of cortical activity during sleep reflects respiratory events and state of daytime vigilance. *Somnologie* 23 (2), 72–79.
- van der Maaten, Laurens, Hinton, Geoffrey, 2008. Visualizing data using t-sne. *J. Mach. Learn. Res.* 9 (Nov), 2579–2605.
- Wendt, Sabrina L., Welinder, Peter, Sorensen, Helge BD., Paul, E Peppard, Jennum, Poul, Perona, Pietro, Mignot, Emmanuel, Warby, Simon C., 2015. Inter-expert and intra-expert reliability in sleep spindle scoring. *Clin. Neurophysiol.* 126 (8), 1548–1556.
- Zhang, Junming, Wu, Yan, 2017. A new method for automatic sleep stage classification. *IEEE Trans. Biomed. Circuits Syst.* 11 (5), 1097–1110.
- Zhang, Junming, Wu, Yan, Bai, Jing, Chen, Fuqiang, 2016. Automatic sleep stage classification based on sparse deep belief net and combination of multiple classifiers. *Trans. Inst. Meas. Contr.* 38 (4), 435–451.