

# EEG-Vorhersage zur Prognose von Mikroschlaf

Lisa Schneeweiß, Martin Patrick Pauli, Martin Golz

Hochschule Schmalkalden, Fakultät Informatik, Blechhammer 4, 98573 Schmalkalden

## Abstract

Elektroenzephalogramme (EEG) von 79 jungen Fahrzeugführer:innen, die bei Nachtfahrtsimulationen häufig Mikroschlafereignissen unterlagen, wurden mit dem Ziel analysiert, stets nur EEG vor dem Eintritt eines Mikroschlaf-Ereignisses zu verwenden. Zwei unterschiedliche Klassen wurden gebildet: (0) kein, (1) ein Mikroschlaf-Ereignis tritt nach dem Ende der EEG-Aufzeichnung ein. Mit fünf Methoden der Zeitreihenvorhersage wurden die EEG-Amplituden nach dem Ende der Aufzeichnung mit einem Prognosehorizont von bis zu fünf Sekunden geschätzt. Im Ergebnis zeigte sich, dass das ARIMA-Modell (auto-regressive integrated moving average) die geringsten Prognosefehler aufweist. Das EEG von zentralen und von occipitalen Lokalisationen war genauer zu prognostizieren als das von frontopolaren Lokalisationen. Die Methodik wurde zusätzlich bewertet mit der Genauigkeit der Diskrimination beider Klassen (0/1), wobei unterschiedliche Anteile von aufgezeichneten und vorhergesagten EEG-Amplituden in die Analysen eingingen. Es konnten mit Kreuzvalidierung mittlere Genauigkeiten von bis zu  $85,3 \pm 0,9$  % erreicht werden, wobei nur aufgezeichnetes EEG vor dem Mikroschlaf-Ereignis verwendet wurde. Durch das Hinzunehmen von vorhergesagtem und Weglassen von aufgezeichnetem EEG konnte eine Genauigkeitssteigerung von maximal 0,96 % erzielt werden. Allerdings traten Fluktuationen in den Ergebnissen auf, sodass die Ergebnisse nicht als statistisch sicher angesehen werden können.

## 1. Einleitung

In Deutschland wurden im Jahr 2021 1.507 Verkehrsunfälle mit Personenschaden durch Übermüdung registriert [1]. Allerdings wird diese Unfallursache meist angegeben, wenn keine andere in Frage kommt. Die tatsächliche Relevanz von Müdigkeit im Verkehrsunfallgeschehen wurde im Rahmen eines berufsgenossenschaftlich initiierten Projekts untersucht [2]. Akten zu schweren und tödlichen Wege- und Dienstwegeunfällen im Straßenverkehr (mit Pkw und Lkw) aus den Jahren 2014 bis 2019 wurden auf Hinweise für Müdigkeit als Unfallursache begutachtet. Dabei lagen in 24,2 Prozent der Unfälle Indizien dafür vor, dass der Fahrer eingeschlafen sein könnte. In weiteren 17,6 Prozent wurden Hinweise auf müdigkeitsbedingte Fahrfehler und Unaufmerksamkeit gefunden [2]. Damit wurden frühere Untersuchungen bestätigt, dass Müdigkeit in mindestens 15% bis 25% aller schweren Unfälle eine ursächliche Rolle spielt [3]. Die Untersuchung dieses Humanfaktors ist auch für andere Verkehrsmittel und Wirtschaftssektoren, wie beispielsweise Überwachungsaufgaben in Produktion und im Militär, von hoher Bedeutung. Das Einbrechen von schlafähnlichen Zuständen unter der Anforderung fortwährender Aufmerksamkeit bezeichnet man als Mikroschlaf. Er kann zu besonders schwerwiegenden und gefährlichen Situationen führen und sollte durch geeignete Gegenmaßnahmen verhindert werden. Ein Weg zu diesem Ziel ist die Fahrerüberwachung mit Detektion und Alarmierung. Kritisch ist allerdings, dass die Detektion erst dann einsetzt, wenn ein Mikroschlaf-Ereignis bereits eingetreten ist.

Um diese Ereignisse zu detektieren, bevor sie eintreten, sollen im Weiteren Methoden der Zeitreihenprognose untersucht werden, wobei stets nur das EEG vor und nicht während eines Mikroschlaf-Ereignisses analysiert wird. Problematisch ist, dass das EEG eine hohe interindividuelle Variabilität und zudem eine komplizierte Signalcharakteristik aufweist. In diesem Beitrag sollen univariate Modelle eingesetzt werden und ihre freien Parameter empirisch optimiert werden, so dass verschiedene Fehlermaße (MAE, MSE, MAPE) an den vorliegenden Datensätzen minimal sind.

Die korrekte Prognose zukünftiger Amplitudenfolgen ist bei Signalen mit hohem Zufallsanteil nicht zwingend das ultimative Ziel. Nehmen wir an, dass eine erfolgreiche Vorhersage der wesentlichen dynamischen Signaleigenschaften möglich ist, auch, wenn die Amplituden und Phasenlagen nicht immer annähernd korrekt vorhergesagt werden können. Dann könnte eine Klassifikation mit lernfähigen Methoden, ob ein Mikroschlaf-Ereignis bevorsteht oder nicht, erfolgreich sein. So könnte der Fahrer vor dem bevorstehenden Mikroschlaf-Ereignis gewarnt werden, denn die komplette Signalanalyse wurde bereits vor dem Eintreten des Mikroschlaf-Ereignisses abgeschlossen.

## **2. Material**

Das EEG von 79 jungen, gesunden Personen aus verschiedenen Nachtfahrt-Studien mit siebenstündigen Aufzeichnungen pro Person wurde in die Datenanalysen eingeschlossen. Alle Studien wurden nach gleichem Konzept im Fahrsimulations-Labor der Hochschule Schmalkalden in authentischer Umgebung durchgeführt. Das Studienkonzept sicherte ab, dass folgende wesentliche Einflussfaktoren zu sehr hoher Fahrermüdigkeit führen: Wachzeit, Beanspruchungszeit, Tageszeit und Monotonie-Empfinden.

Hohe Wachzeiten (time since sleep) wurden erreicht durch das Starten der Fahrsimulation mindestens 14 Stunden nach Beendigung des Nachtschlafs und der Forderung an die Versuchsteilnehmer, dass zwischenzeitlich fortwährende Wachheit gewährleistet sein muss. Letzteres wurden mit Handaktometrie überprüft: die Beschleunigungen am Handgelenk, die einen definierten Schwellwert übersteigen, werden gezählt und als minütliches Histogramm gespeichert. Im Histogramm dürfen keine längeren Pausen vorliegen, anderenfalls wurde die Person nicht zur Studie zugelassen. Die Beanspruchungszeit (time on task) ergibt sich aus der Länge einer Fahrtsitzung (40 Minuten) und der Anzahl von sieben Sitzungen pro Person, sodass diese bei 280 Minuten lag. Eine höhere Beanspruchung ist kaum umsetzbar, unter anderem weil der dritte Einflussfaktor, die Tageszeit (time of day), mit 1 bis 7 Uhr gewählt wurde. In diesem Zeitraum sinkt der Stoffwechsel und somit die Körpertemperatur und es kommt zur vermehrten Ausschüttung des schlafbegleitenden Hormons Melatonin. Der wohl größte Einfluss ist nicht physischer sondern psychischer Natur: das nachhaltige Empfinden von Monotonie während der Tätigkeit. Auch dieser Faktor war in den Nachtfahrt-Studien beabsichtigt hoch. Man hätte das Monotonie-Empfinden beispielsweise durch intensive Kommunikation, eventuell sogar mit aggressiver Tendenz verringern und so ausreichend hohe Aktivierungen erreichen können, so würde der Schlafdruck eine untergeordnete Rolle bekommen. Weitere Maßnahmen zur Absicherung hoher Müdigkeit waren der angeordnete Verzicht auf psychotrope Substanzen (bspw. Koffein) am Tag der Studie, der allerdings nicht überprüft werden konnte.

Insgesamt stehen 9.608 Beispiele zur Verfügung, davon gehören 55,5 % der Klasse 0 (kein Mikroschlaf) und 44,5 % der Klasse 1 (Mikroschlaf) an. Die Anzahl beider Klassen

ist oftmals dadurch beschränkt, dass in einem relativ großen Zeitraum von 32 Sekunden kein weiteres Mikroschlaf-Ereignis und keine Artefakte im EEG vorliegen dürfen. Die interindividuelle Variabilität der Mächtigkeiten beider Klassen ist relativ hoch (Abbildung 1).

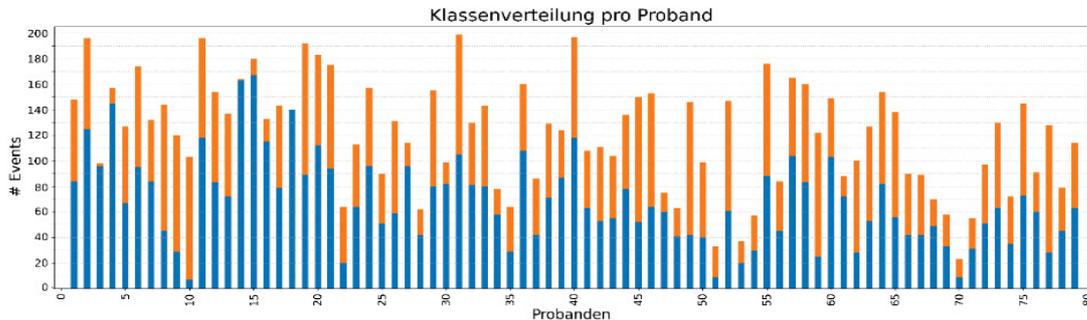


Abbildung 1: Mächtigkeiten der Klassen 0 (blau/unten) und 1 (orange/oben) über der ID-Nummer der beteiligten Personen.

### 3. Methoden und Zwischenergebnisse

Folgende fünf Methoden wurden aus Gründen des numerischen Aufwands an einem verkleinerten Datensatz auf ihre Tauglichkeit überprüft: simple exponential smoothing (SES), Holt-Methode ohne/mit linearem Trend (Holt1/2), autoregressives Modell (AR) und kombiniertes Modell der Autoregression und des gleitenden Mittelwerts über der Differenzen-Zeitreihe (ARIMA, auto-regressive integrated moving average) [4]. Validiert wurden die Modelle, indem folgende Fehlermaße aus dem vorhergesagten Amplituden und den beobachteten Amplituden gebildet wurden: mittlerer absoluter Fehler (MAE), mittlerer absoluter prozentualer Fehler (MAPE) und mittlerer quadratischer Fehler (MSE). Im Vergleich der Perzentile der Fehlerverteilung (Abbildung 2) zeigte sich, dass das ARIMA-Modell zu kleineren Prognosefehlern führte. Die Kastengrafiken zeigen, dass nur wenige Ausreißer vorliegen und der Großteil der Fehler im Intervall der Antennen (whisker) liegen. Das ARIMA-Modell ist in jedem Fehlermaß und in allen Perzentilen der Verteilung am niedrigsten. Daher wurde für die anschließenden Analysen aus Gründen des numerischen Aufwands stets die ARIMA-Vorhersage berücksichtigt.

Zu optimierende Parameter des ARIMA-Modells sind die Ordnungen des AR- und des MA-Modells, sowie die Differenzierungsordnung. Es wurden Parameter im Intervall 0 bis 9 zugelassen. Diese drei Parameter gelten stets einheitlich für alle in die Analysen eingeschlossenen 79 Personen. Anschließend erfolgte die Optimierung der Modellparameter ( $\alpha_i$ ,  $\beta_i$ ) für jede der Zielfunktionen (MAE, MAPE, MSE).

Es zeigte sich, dass für jeden EEG-Kanal andere Parameterkombinationen optimal waren. In beiden frontopolaren Kanälen (FP1, FP2) und in einem occipitalen Kanal (O1) waren die Häufigkeitsunterschiede gering, in den zentralen Kanälen (C3, C4) und im Kanal O2 waren drei Kombinationen dominant. Deshalb wurden im Weiteren für jeden EEG-Kanal separat die Parameterkombination ausgewählt, die am häufigsten optimal war.

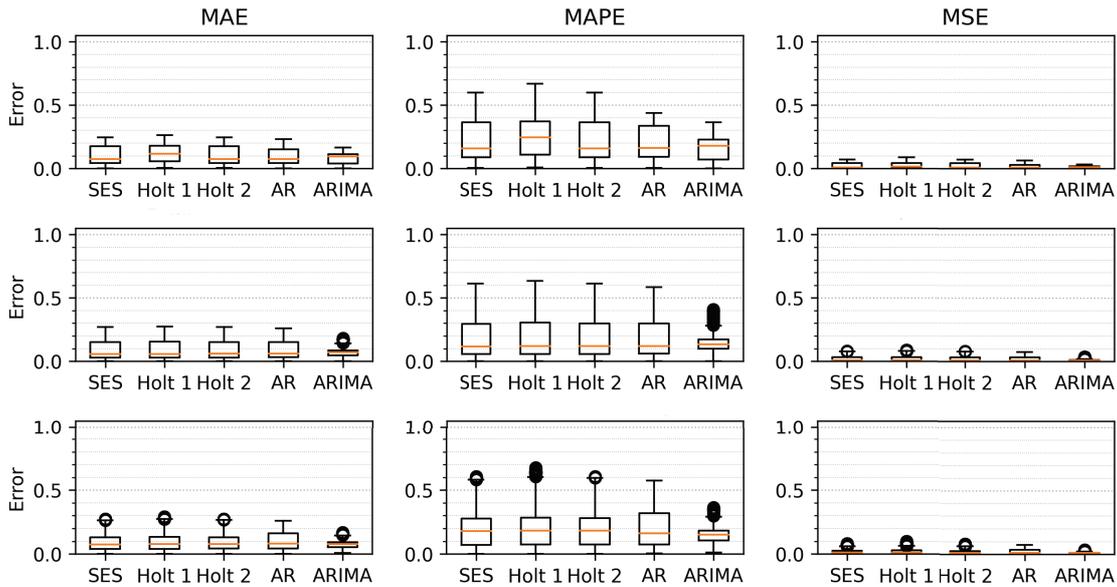


Abbildung 2: Perzentile der Fehlerverteilung für fünf Vorhersage-Modelle (SES, Holt1, Holt2, AR, ARIMA). Als Fehlermaße wurden der mittlere absolute, mittlere relative und mittlere quadratische Fehler (MAE, MAPE, MSE) verwendet. Es wurden drei Signalsegmente zufällig ausgewählt (horizontale Reihen).

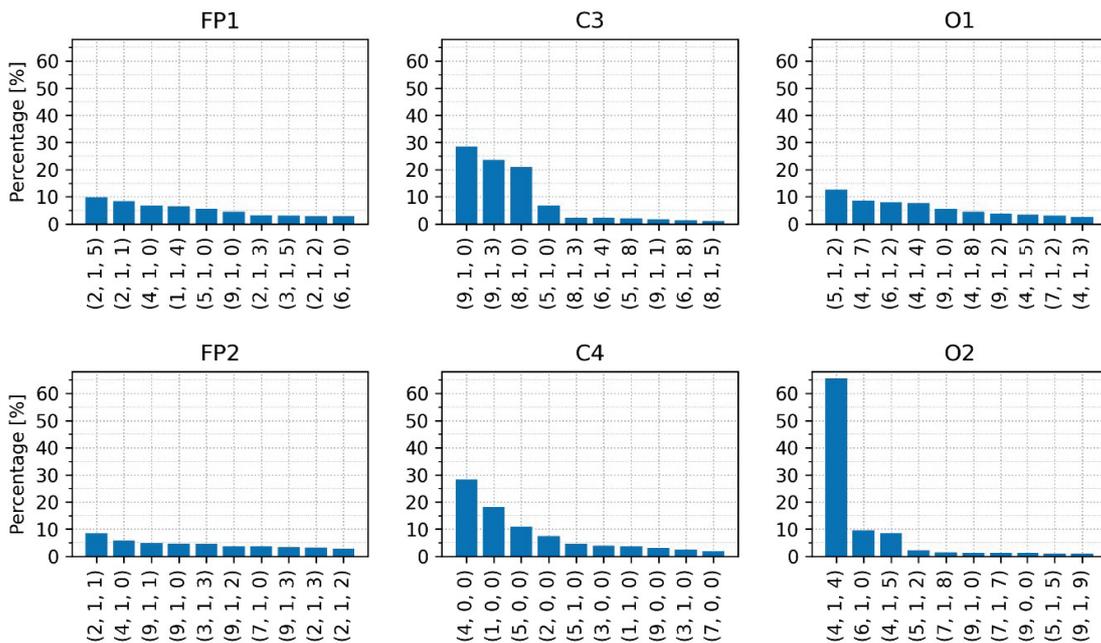
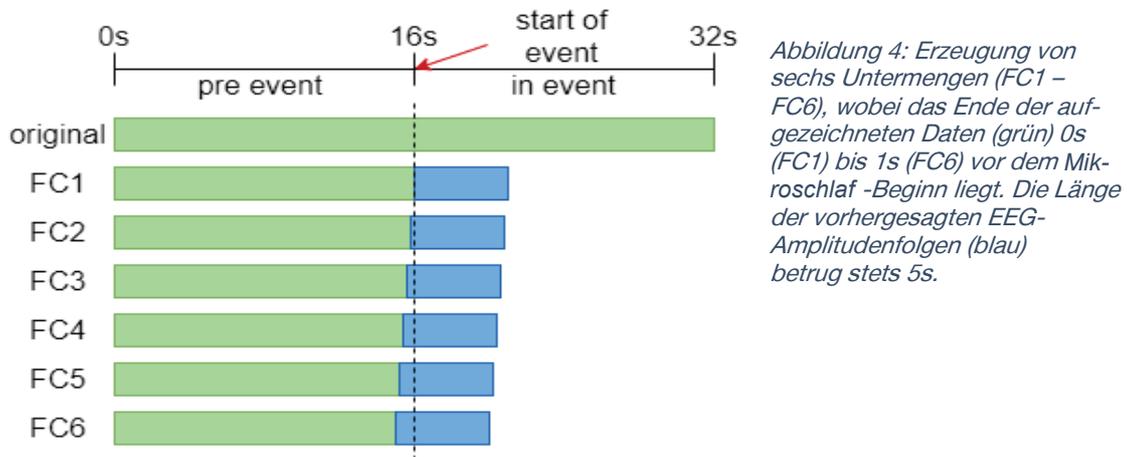


Abbildung 3: Relative Häufigkeit der 10 optimalen Parameterkombinationen des ARIMA-Modells getrennt für jeden der sechs EEG-Kanäle (FP – frontopolar, C – zentral, O – occipital).

Die anschließenden Klassifikationsanalysen haben die Genauigkeitssteigerung der Diskriminierung beider Klassen (0/1) als Ziel. Hierfür wurden sechs Untermengen des Datensatzes wie folgt erzeugt (Abbildung 4). Der original gegebene Datensatz hat eine Länge von 32 Sekunden pro Beispiel. Bei Klasse 1 beginnt das Mikroschlaf-Ereignis bei  $t = 16s$ . Bei Klasse 0 liegt kein Mikroschlaf vor, sodass zur Intervallmitte kein besonderes Verhaltensmerkmal vorliegt und dennoch eine Vorhersage der Amplitudenfolgen

durchgeführt wird. Die Untermenge FC1 enthält 16s aufgezeichnetes EEG und 5s vorhergesagtes EEG. Bei der Menge FC2 verringert sich der erste Anteil auf 15,8s; der zweite Anteil blieb fixiert auf 5s. Mit gleicher Schrittweite (0,2s) wurden die Untermengen FC3 - FC6 erzeugt, sodass in FC6 15s aufgezeichnetes und 5s vorhergesagtes EEG vorhanden sind.



Die mittleren Fehler für die Vorhersage der EEG-Amplitudenfolgen zeigen einen nicht signifikanten Anstieg von Untermenge FC1 bis FC6 (Abbildung 5). Deutlicher tritt die stärker werdende Streuung der mittleren Fehler hervor, die sowohl in der Höhe des Interquartilbereichs (Kastenhöhe) als auch in der Länge der Antennen (whisker) zu erkennen ist.

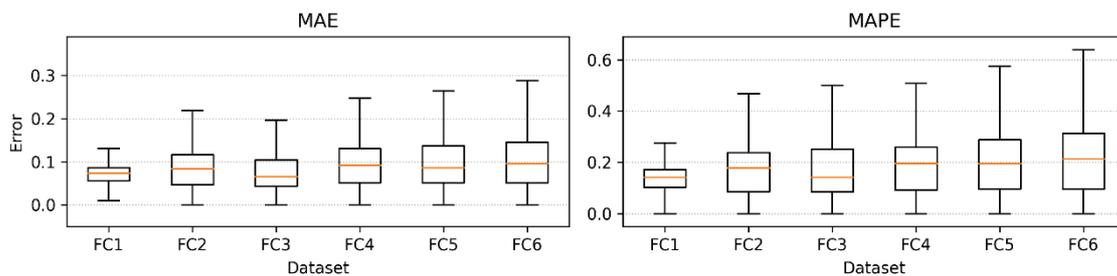


Abbildung 5: Mittlere absolute und relative Vorhersagefehler (MAE, MAPE) für 5s lange EEG-Amplitudenfolgen, die aus allen sechs Untermengen (FC1 – FC6) gewonnen wurden.

Die abschließende Klassifikationsanalyse mit computerbasierter Intelligenz hatte zum Ziel, innerhalb der Untermengen FC1 bis FC6 weitere Untermengen zu finden, sodass die mittlere Klassifikationsgenauigkeit maximiert wird. Frühere Analysen am gleichen Datensatz zeigten, dass eine Untermengenlänge von 4s bis 6s optimal ist und dass der zeitliche Versatz (offset) der Untermenge bzgl. des Mikroschlaf-Eintritts einen hohen Einfluss hat und deshalb empirisch zu optimieren ist [5]. Deshalb wurde mit einer Länge von 4s und mit einem variablen Versatz im Intervall 0s bis 2s variiert (Abbildung 6). Der Versatz ist definiert als Differenz zwischen Endzeitpunkt der Untermenge und Zeitpunkt des Eintritts des Mikroschlaf-Ereignisses. Die schrittweise Erhöhung des Versatzes verringert den Anteil des aufgezeichneten EEG in der Untermenge und erhöht den Anteil der vorhergesagten EEG-Amplituden.

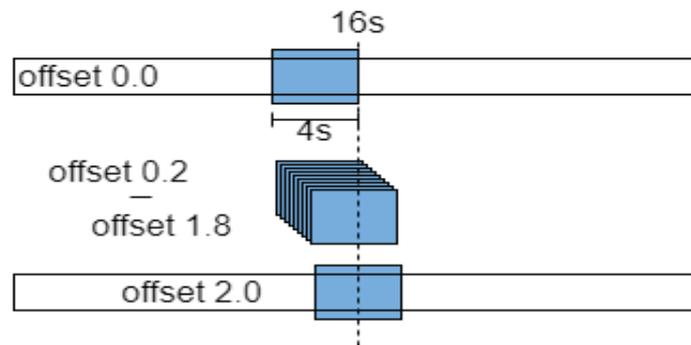


Abbildung 6: Generierung von Untermengen mit variablem Versatz (offset) zur empirischen Optimierung der mittleren Klassifikationsgenauigkeit zur Diskriminierung von auftretenden gegenüber nicht auftretenden Mikroschlaf-Ereignissen.

In gleicher Weise wurden Referenz-Datensätze erzeugt, die jedoch ausschließlich aufgezeichnete EEG-Amplitudenfolgen enthalten und keine vorhergesagten. Dadurch verkürzte sich die Untermengenlänge, sobald der Versatz größer als Null war.

Die weitere Verarbeitung erfolgte wie in [5]. Für jede Untermenge und für jeden EEG-Kanal wurden die logarithmierten spektralen Leistungsdichten mit der Periodogramm-Methode statistisch geschätzt und in 1Hz breiten Intervallen im Frequenzintervall von 0 bis 25Hz gemittelt. Die so entstandenen 25-dimensionalen Merkmalsvektoren sind die Eingangsvariablen für die Methoden der computerbasierten Intelligenz. Als Klassifikationsmethode wurde LightGBM von Microsoft Research Inc. verwendet [6]. Diese Gradient-Boosting-Methode (GBM) enthält ein Ensemble aus Entscheidungsbäumen.

Die Validierungsanalysen für LightGBM wurden mit wiederholter zufälliger Partitionierung in Trainings- und Validierungsmenge durchgeführt (repeated random subsampling), wobei 80% der 9.608 Datenbeispiele als Trainings- und 20% als Validierungsmenge verwendet wurden. Diese zufällige Partitionierung wurde 20mal wiederholt und aus den erreichten Klassifikationsgenauigkeiten wurden Mittelwert und Standardabweichung geschätzt.

#### 4. Ergebnisse

Die Versatzoptimierungen führten zum stets gleichen Ergebnis: bei einem Versatz von 0s wurden die maximalen mittleren Klassifikationsgenauigkeiten an Validierungsmengen gefunden (Tabelle 1, Spalte 2). Beim maximal gewählten Versatz von 2s wurden die minimalen Genauigkeiten gefunden. Die mittlere Differenz zwischen beiden Ergebnissen lag bei ca. 5%. Die Prognose hat die Reduktion an aufgezeichnetem EEG im Analyseintervall nicht adäquat ersetzen können. Die Ergebnisse sind am höchsten für die Untermenge FC1 und am geringsten für FC6.

Untermenge	$\bar{a}_V$ (Versatz: 0s)	$\bar{a}_V$ (Versatz: 2s)
FC1	85,3 ± 0,9 %	81,1 ± 0,7 %
FC2	84,9 ± 0,9 %	80,6 ± 1,3 %
FC3	84,2 ± 1,0 %	79,4 ± 0,8 %
FC4	83,1 ± 0,8 %	78,2 ± 0,4 %
FC5	82,5 ± 0,8 %	76,7 ± 0,4 %
FC6	81,5 ± 0,9 %	74,7 ± 0,6 %

Tabelle 1: Mittelwert und Standardabweichung der Klassifikationsgenauigkeiten  $\bar{a}_V$  an Validierungsmengen für sechs Untermengen (FC1 - FC6) gemäß Abbildung 4 und für zwei Versatzparameter.

Im Vergleich zur Klassifikation an Referenz-Untermengen zeigt sich, dass die Hinzunahme von vorhergesagten Amplitudenfolgen in der überwiegenden Zahl der Fälle einen positiven Effekt hat. Genauigkeitssteigerungen bis zu 0,96% konnten erreicht werden (Tabelle 2). Allerdings kann kein einheitlicher Trend festgestellt werden. Innerhalb einer Untermenge (Spalte) kann es zu rückläufigen Genauigkeitsänderungen (rot) kommen. Während Untermenge FC5, die durch erhebliche Reduktion der Original-Datenmenge gekennzeichnet ist, die höchsten Genauigkeitssteigerungen zeigt, ist die Untermenge FC6, die nur eine geringfügig größere Datenreduktion bzgl. FC5 hat, durch die durchschnittlich (letzte Zeile) geringste mittlere Genauigkeitssteigerung gekennzeichnet.

Versatz	FC1	FC2	FC3	FC4	FC5	FC6
0,0	± 0,00	+ 0,53	+ 0,47	+ 0,19	+ 0,29	+ 0,22
0,2	+ 0,14	+ 0,12	+ 0,69	+ 0,38	+ 0,38	+ 0,29
0,4	+ 0,15	+ 0,14	+ 0,93	+ 0,28	+ 0,43	+ 0,02
0,6	+ 0,33	+ 0,05	+ 0,51	+ 0,27	+ 0,24	- 0,16
0,8	+ 0,06	+ 0,25	+ 0,38	+ 0,05	+ 0,68	+ 0,26
1,0	- 0,05	- 0,18	- 0,04	+ 0,6	+ 0,24	- 0,74
1,2	+ 0,25	+ 0,12	- 0,24	+ 0,57	+ 0,96	+ 0,12
1,4	+ 0,48	+ 0,51	+ 0,52	+ 0,52	+ 0,81	+ 0,39
1,6	- 0,22	+ 0,29	+ 0,95	+ 0,57	+ 0,88	- 0,25
1,8	+ 0,21	+ 0,53	+ 0,43	+ 0,81	+ 0,90	+ 0,73
2,0	- 0,69	+ 0,26	+ 0,09	+ 0,37	+ 0,76	- 0,18
∅	+ 0,06	+ 0,24	+ 0,43	+ 0,42	+ 0,60	+ 0,06

Tabelle 2: Prozentuale Veränderung der mittleren Klassifikationsgenauigkeiten bzgl. der Referenz, die nur aus aufgezeichneten, jedoch nicht aus vorhergesagten EEG-Amplitudenfolgen besteht.

## 5. Diskussion

Die Vorhersage von EEG-Amplituden stellt eine hohe Herausforderung dar, da das EEG Ergebnis eines Multi-Prozesses ist und somit die Beschreibung mit einem fixierten ARIMA-Modell nur sehr fehlerhaft gelingt. Vorhersagen mit einem Horizont bis zu 5s sind sehr unsicher und von fragwürdiger Qualität. Allerdings bestand die Hoffnung, dass

die vorhergesagten Amplitudenfolgen wesentliche dynamische Charakteristika enthalten, sodass die adaptive Klassifikationsanalyse mit LightGBM erfolgreich sein könnte. An Trainingsmengen konnten Klassifikationsgenauigkeiten zwischen 99 % und 100 % gefunden werden. Leider konnten diese Ergebnisse nicht an Validierungsmengen beobachtet werden.

Ein Ausweg könnte die Verwendung von Merkmalszeitreihen anstelle der Amplitudenzeitreihen sein. Merkmale sind durch die integrale Analyse vieler Amplituden gekennzeichnet und haben somit tendenziell geringere Fluktuationen. Ein weiterer, jedoch aufwendigerer Weg könnte der Einsatz von Methoden der computerbasierten Intelligenz zur Zeitreihen-Vorhersage sein. Hier könnten rückgekoppelte Netze vom Typ LSTM (long short-term memory) zum Einsatz kommen, die bereits bahnbrechende Erfolge in der Spracherkennung, Musikklassifikation und der Steuerung von Computerspielen erreicht haben.

Das hier vorgestellte Konzept der gemischten Verwendung von aufgezeichneten und vorhergesagten Amplitudenfolgen und der anschließenden adaptiven Klassifikationsanalyse konnte bereits eine, wenn auch bescheidene Leistungsfähigkeit erreichen. Insgesamt lässt sich daher sagen, dass sich mit einer Vorhersage der EEG-Amplituden die Klassifikationsgenauigkeit von Sekundenschlafereignissen erhöhen lässt. Allerdings scheint keine der verwendeten fünf Methoden für die Vorhersage der komplexen EEG-Signale geeignet zu sein, um die komplexe Dynamik der Signale erfassen zu können. Dennoch ist jeder noch so kleine Vorhersage-Horizont nützlich, denn gerade im Straßenverkehr kann die Erkennung von bevorstehendem Sekundenschlaf und der umgehende Einsatz von technischen Gegenmaßnahmen Unfälle vermeiden und Leben retten.

## Quellen

- [1] Statistische Bundesamt (2022) *Anzahl der durch Übermüdung verursachten Straßenverkehrsunfälle mit Personenschaden in Deutschland von 1991 bis 2021*. [de.statista.com/ /statistik/daten/studie/317703/umfrage/unfaelle-mit-personenschaden-uebermuedung/](https://de.statista.com/statistik/daten/studie/317703/umfrage/unfaelle-mit-personenschaden-uebermuedung/)
- [2] Berufsgenossenschaft Energie Textil Elektro Medienerzeugnisse (2020), Erst müde, dann Krankenhaus. impuls Ausgabe 04/2021. [www.bgetem.de/medien-service/profi-zeitschrift-fuer-versicherte](http://www.bgetem.de/medien-service/profi-zeitschrift-fuer-versicherte)
- [3] Langwieder K, Sporer A, Hell W (1994) Struktur der Unfälle mit Getöteten im Freistaat Bayern im Jahr 1991. *München: HUK-Verband, Büro für Kfz-Technik*.
- [4] Box GEP, Jenkins GM (1970) *Time series analysis: Forecasting and control*. Holden-Day, San Francisco.
- [5] Pauli MP, Pohl C, Golz M (2022) Optimal EEG Segmentation for Microsleep Detection Based on Machine Learning. *Current Directions in Biomedical Engineering*, 8(2), 749-752.
- [6] Ke G, Meng Q, Finley T, Wang T, et al. (2017) LightGBM: A highly efficient gradient boosting decision tree. *Adv Neural Inf Proc Syst*, 30:3146-54.