DRIVER SLEEPINESS ASSESSED BY ELECTROENCEPHALOGRAPHY – DIFFERENT METHODS APPLIED TO ONE SINGLE DATA SET

Martin Golz[1], David Sommer[1], Jarek Krajewski[2]
[1]University of Applied Sciences Schmalkalden, Germany
[2]University of Applied Sciences Cologne, Germany

Summary: An overview of several methods of electroencephalography (EEG) analysis in order to assess driver sleepiness is presented. All methods were applied to one single data set obtained from overnight driving simulations in our lab. 10 young adults (age 22.4 ± 4.1 years) participated and drove on rural roads; time on task was 7 x 40 min and time since sleep ranged between 16 and 22 hours. Results show large inter-individual variability of all variables and moderate correlation coefficients to one subjective and one objective independent variable of driver drowsiness. Only one method, the detection of microsleep-like EEG patterns, provides a variable with strong increases immediately before sleepiness related crashes. It is concluded that EEG analysis should attach more importance to short-term patterns and should renounce the analysis of spectral power in four bands.

INTRODUCTION

The assessment of driver fatigue is still a challenging task. Many different types of measurements and signal analysis have been proposed by several authors. Commercially available devices for fatigue monitoring have been constructed which rely mostly on eye blink behavior via contactless measurements. I many cases their outcome is the percentage of eye lid closure (PER-CLOS). It has been reported that this variable is reliable at a temporal scale of several minutes, but is lacking in moments of extreme driver fatigue on a scale of seconds (Caterpillar 2008; Golz et al, 2010). Therefore and in general, reference measures of driver sleepiness are needed in order to validate fatigue monitoring devices even under conditions of extreme sleepiness which is accessible in driving simulation laboratories only.

Electroencephalography (EEG) is one of the most reliable measurements of sleepiness due to its relative direct access to cortical and subcortical processes, its good test and retest reliability, and its high reproducibility. EEG provides an objective, functional mapping of brain activity on the scale of seconds. Many authors investigated EEG power spectral densities (PSD) during driving and found increased PSD in the alpha, theta, and in the broad band with increasing sleepiness.

Alpha & theta bursts. Intrusions of short alpha and theta bursts into beta activity (15-30 Hz) in active subjects under sleepy condition have initially been reported by Torsvall & Åkerstedt (1987). This concept has been extended to quantitatively assess the *alpha burst percentage* (ABP) and was applied to the EEG of drivers on the real road (Simon et al, 2011).

Theta band PSD. Other authors followed an alternative hypothesis that sleepiness leads to replacement of alpha by theta activity, initially stated by Harrison & Horne (1996). This concept of a general frequency downshift was utilized by several other authors. In contradiction, Kecklund & Åkerstedt (1993) observed that the "EEG shows a significant but moderate increase of hourly mean power spectral density only in the alpha band, but not in the theta band".

Broad band PSD. Broad band increases in the theta-alpha range were observed by several authors, initially reported by Cajochen et al. (1995). Lal & Craig (2001) reported on relative PSD increases of 22 %, 26 %, 9 %, and 5 % in the delta, theta, alpha, and beta band, respectively, for sleepy states compared to alert ones.

Microsleep EEG patterns. Coming down from the extreme end of sleepiness, the so-called microsleep (MS) events, is a new and promising approach to assess driver sleepiness. It is assumed that behavioral MS is a subcortical "switch-off" event influencing many cortex areas. This results in characteristic EEG changes (Peiris et al., 2006) detectable with high reliability if machine learning algorithms are utilized (Golz et al., 2007). The number of detected MS patterns of the EEG is larger than the number of behavioral MS observed by video ratings. This might indicate that there are some sleepiness events, which are observable in the EEG but not in behavior. The average amount of MS patterns in the EEG depends on the number and the length of patterns in a given interval (e.g. 2 minutes) and determines a new quantity, the *microsleep percentage* (MSP) (Sommer et al., 2008).

Here we compare all EEG sleepiness measures mentioned above on one single data set. For this purpose, the correlation of each measure to two independent variables of sleepiness is estimated. Furthermore, we ask if each sleepiness measure increases immediately before sleepiness related crashes.

EXPERIMENTS

Participants. 14 university students were randomly selected out of 73 volunteers who indicated their interest. All owned their driving licence for at least one year. Subjects provided written informed consent and were compensated for their time and effort. Within the three days before the experimental night subjects were checked by wrist actometry if sleep wake demands were fulfilled (wake-up time: 6 – 9 a.m.; bed time: 10 p.m. – 1 a.m.). Furthermore, it was checked if subjects had no nap sleep during the day before the experimental night. Subjects underwent at least an one hour training to become comfortable with the simulator and to screen for possible simulator sickness. All participants were asked to abstain from alcohol and drugs influencing fitness to drive three days before the study. Ten subjects completed all driving sessions (two females, eight males, mean age = 22.4 years, SD = 4.1, range 19 - 32 years). One male and two females quitted driving because of simulator sickness, one male quitted because of back pain.

Experimental Procedure. Data were drawn from an experimental study conducted between February and May 2007 in our driving simulation lab. The study was designed to investigate driving performance and subjects' behaviour under high level of monotony and at higher levels of sleepiness. Subjects were instructed to track the lane as best as possible and to avoid falling asleep. After returning from MS subjects were reminded that if driving performance becomes too bad or signs of behavioral MS returned too fast, the experiment would be terminated. When crashes appeared an extensive soundscape as well as video scene was played back to increase the emotional importance of this event, to set a short break in the run of the driving simulation and to caution the driver.

Each subject had to perform overnight driving simulations within 7 sessions of 40 min length each (Fig. 1). Before and after each driving session several questionnaires and vigilance tasks were requested (not reported here). Breaks between each driving sessions proved to be necessary

to motivate the subjects to proceed with the experiment and to resist increasing sleep pressure. Time since sleep was at least 16 hours at the beginning of the first session.
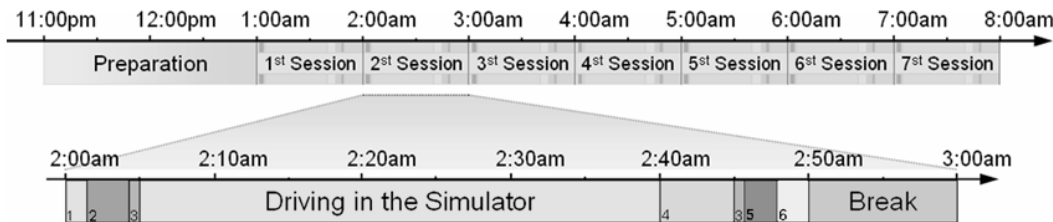

Figure 1. Schedule: driving simulations (40 min length) started hourly with preceding and subsequent questionnaires and vigilance tasks

Measurements. EEG (F1, F2, C3, Cz, C4, O1, O2, A1, A2, com.av.ref.), EOG (vertical, horizontal), ECG, and EMG (m. submentalis) and video recordings of driver's head & pose, driver's eyes, the driving scene were recorded. Video recordings were used for visual inspection if behavioral signs of MS occurred. Input variables of the simulator, e.g. steering angle, pedal position, as well as output variables, e.g. velocity, lateral position in lane, were recorded.

Sleepiness measures independent from the EEG. Every 4 minutes during driving subjects rated their subjectively experienced sleepiness on the Karolinska Sleepiness Scale (KSS) (Åkerstedt & Gillberg, 1990). The standard deviation of the lateral position in lane (SDLat) was utilized as an objective measure of driver sleepiness.

METHODS

Theta, alpha, broad band PSD. Artefacts within EEG recordings were eliminated based on amplitude (< 300 µV) and drift criteria (< 150 µV per 10 s). Ocular artefacts which superpose EEG recordings mainly at frontopolar electrodes (Fp1, Fp2) were not removed. PSD was estimated in de-trended segments ($L = 10$ s, 75 % overlap, Hann taper) utilizing the modified periodogram method. PSD values were added up within the common EEG spectral bands (θ-band: $4 \leq f < 8$ Hz, α-band: $8 \leq f < 12$ Hz) and in the broad band (δ-θ-α-β-band: $0.5 \leq f \leq 25$ Hz).

Microsleep percentage (MSP). All clear-cut behavioral MS events, found in the video recordings, and the same amount of non-MS examples were utilized for supervised machine learning in order to establish the MS detector (Golz et al., 2007). Afterwards, this detector was applied to the complete EEG recordings in order to search for MS-like patterns. Within 2-minute intervals the length of each MS pattern was accumulated in order to calculate the MSP.

Alpha burst percentage. ABP was estimated in the same manner as MSP. The alpha burst detector consists of fixed rules and is not an outcome of machine learning (Simon et al., 2011). The central part of this method is the search for such EEG activity which has maximum power in the α-band and which has a spectral bandwidth small enough to be a burst activity. Within 2-minute intervals the length of each AB pattern was accumulated in order to calculate the ABP.

Linear correlation model. Pairwise correlations between five EEG measures (α-, θ-, δ-θ-α-β-PSD, MSP, ABP) and two non-EEG measures (KSS, SDLat) were estimated utilizing a standard linear model for interval scale variables resulting in the Pearson product-moment correlation coefficient ρ and p-values for testing the hypothesis of no correlation against the alternative of nonzero correlation.

RESULTS

Correlation coefficients. Correlation coefficients ρ between KSS and EEG measures were estimated (Table 1) for each of the 7 driving sessions (columns) and for the whole night (last column). In general, α-PSD offers higher correlations than θ-PSD and δ-θ-α-β-PSD. Therefore, results of α-PSD at 7 different electrode locations (rows) are depicted. Furthermore, correlations between KSS and MSP and KSS and ABP (two bottom rows) were estimated. MSP is extracted across all electrode sites, whereas ABP is extracted at occipital sites (O1, O2). Significance levels are $p < 0.05$ (*) and $p < 0.01$ (**).

Table 1. Correlation coefficients ρ between KSS and α-PSD, MSP, and ABP

	starting time of driving sessions							whole
	01:00	02:00	03:00	04:00	05:00	06:00	07:00	night
C3	0.59 **	0.51 **	0.47 **	0.43 **	0.32 **	0.31 **	0.30 **	0.75 **
C4	0.63 **	0.64 **	0.63 **	0.65 **	0.65 **	0.66 **	0.65 **	0.75 **
Cz	0.59 **	0.60 **	0.59 **	0.60 **	0.61 **	0.61 **	0.61 **	0.75 **
Fp1	0.67 **	0.66 **	0.65 **	0.64 **	0.62 **	0.62 **	0.62 **	0.64 **
Fp2	0.62 **	0.51 **	0.47 **	0.41 **	0.26 **	0.23 **	0.22 **	0.70 **
O1	0.63 **	0.56 **	0.52 **	0.48 **	0.37 **	0.35 **	0.34 **	0.76 **
O2	0.72 **	0.69 **	0.67 **	0.66 **	0.61 **	0.60 **	0.59 **	0.79 **
MSP	0.47 **	0.65 **	0.72 **	0.71 **	0.70 **	0.66 **	0.51 **	0.76 **
ABP	0.09	-0.08	-0.19 **	-0.20 **	-0.37 **	0.05	-0.02	-0.11

Table 2. Correlation coefficients ρ between SDLat and α-PSD, MSP, and ABP

	starting time of driving sessions							whole
	01:00	02:00	03:00	04:00	05:00	06:00	07:00	night
C3	0.31 **	0.52 **	0.54 **	0.58 **	0.54 **	0.55 **	0.50 **	0.59 **
C4	0.30 **	0.56 **	0.54 **	0.51 **	0.52 **	0.50 **	0.36 **	0.58 **
Cz	0.23 **	0.49 **	0.52 **	0.56 **	0.58 **	0.56 **	0.46 **	0.59 **
Fp1	0.09	0.11	0.25 **	0.37 **	0.31 **	0.32 **	-0.01	0.43 **
Fp2	0.30 **	0.39 **	0.46 **	0.57 **	0.34 **	0.36 **	-0.03	0.44 **
O1	0.39 **	0.60 **	0.57 **	0.60 **	0.60 **	0.55 **	0.45 **	0.57 **
O2	0.40 **	0.57 **	0.64 **	0.58 **	0.56 **	0.48 **	0.42 **	0.54 **
MSP	0.45 **	0.71 **	0.71 **	0.65 **	0.61 **	0.54 **	0.53 **	0.59 **
ABP	0.02	-0.16 *	-0.23 **	-0.22 **	-0.01	0.32 **	0.32 **	-0.02

In the same manner correlation coefficients ρ between SDLat and all EEG measures were estimated (Table 2). In general, correlations with SDLat are lower than with KSS.
The temporal development of KSS scores and all EEG measures were investigated utilizing 2-minute accumulation intervals (Fig. 1). Correlation coefficients ρ between KSS scores and EEG measures for each driving session (7 text boxes at the top) and for total time of driving (text box

right beneath) are depicted. Electrode localization for PSD and ABP estimation is O2. Significance levels are $p < 0.05$ (*), $p < 0.01$ (**). KSS, MSP, α-PSD, δ-θ-α-β-PSD have a time on task effect, but θ-PSD, ABP not. All measures, despite ABP, have a time since sleep effect. Each measure has large standard deviation, mainly caused by inter-individual variations.

Temporal relation to crashes. In addition to correlations, sleepiness measures should also be reliable markers immediately before sleepiness related crashes. Such crashes occurred only in the second half of the night (5:00 p.m. and later) and were seldom. MS and AB occurred more frequently by factor 6 and 9, respectively. Despite MSP, all other EEG sleepiness measures had no effect immediately before crashes (Fig. 2).
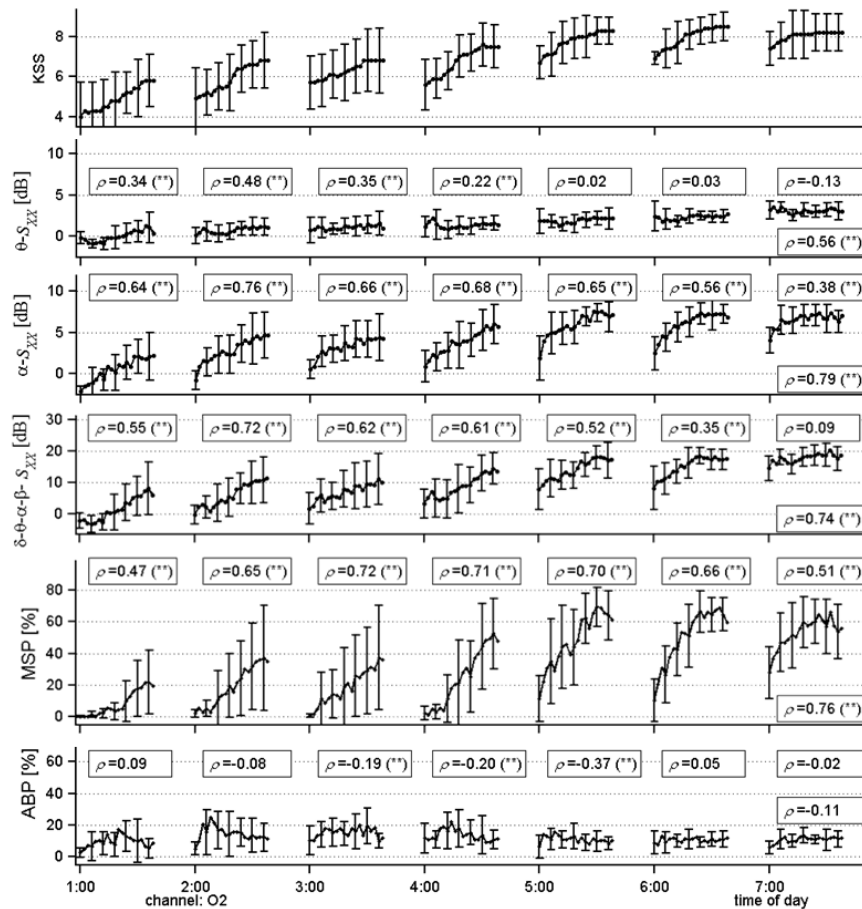

Figure 1. Mean and standard deviation of KSS scores and the 5 EEG sleepiness measures versus time of day

Mean α-PSD and even more θ-PSD and δ-θ-α-β-PSD are nearly unaffected by upcoming sleepiness related crashes (Fig. 2). It's no wonder that ABP, which has no time on task and no time since sleep effect, is also unaffected by upcoming crashes. MSP values between 60 and 70 % indicate high sleepiness level in the second half of the night. Twenty seconds before crash mean MSP raises up to nearly 90 %. MSP and also ABP have low standard deviations within this kind of analysis.
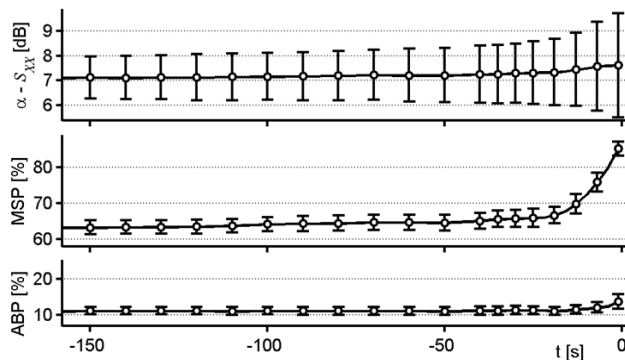
Figure 2. Mean and standard deviation of α- PSD, MSP, and ABP versus time before crash

DISCUSSION

High correlations were found between KSS and α-PSD as well as δ-θ-α-β-PSD at electrode localization O2 and also at Fp1, Cz, and C4. In contrast to ABP, MSP was highly correlated to KSS. Correlations between SDLat and all EEG measures were generally lower, presumably because of low time on task and low time since sleep effects on SDLat in the first 4 driving sessions (1 - 4 a.m.). During the last 3 sessions (5 - 7 a.m.) strong SDLat increases with time on task were observed, but with larger standard deviations, compared to KSS. Even during the second half of the night, no EEG measure had higher correlations to SDLat.

Besides time on task and time since sleep effects which are both accumulated measures on a temporal scale of several minutes up to some hours, also short-term effects should be reflected by an EEG sleepiness measure.

CONCLUSIONS

With regard to all three lines of analysis no convincing indications have been found that common EEG band powers are reliable, individual sleepiness measures. There are indeed time since sleep and time on task effects, but with large inter- and intra-individual variability. It must be considered that correlations in time series are oftentimes demonstrable but have no causal meaning as long as effects on reliable control variables are not ensured. Such variables are not available in sleepiness research.

This is aggravated by the large variability of correlation coefficients observed between different electrode localizations. It is difficult to understand why large correlation coefficients were observed in the right hemisphere at occipital and frontal sites, but not for central sites. To our knowledge, reliable inter-hemispheric differences have not been reported by other authors.

From our point of view, the strongest concept to validate sleepiness measures is their temporal change shortly before sleepiness related crashes. Results of this investigation indicate that all EEG measures completely fail, except MSP. This demonstrates that there are several sleepiness related changes in the EEG, but not assessable with a very limited number of fixed rules of signal analysis. Only comprehensive machine learning algorithms provide the capability to be both sensitive and specific to several EEG patterns of sleepiness. Future analysis has to investigate if it is possible to simplify the patterns found in order to give back general information from machine learning to human EEG experts.

One reason, why it is difficult to assess sleepiness by EEG measures is their very large physiological variability. Santamaria & Chiappa (1987) reported on different drowsiness related EEG patterns in their review paper. They visually inspected the EEG grapho-elements and stated: "There is a great deal of variability in the EEG of drowsiness among different subjects". Many reports on quantitative EEG analysis of drowsy subjects did not consider these patterns and the large variability between subjects. Other authors investigated EEG with modern nonlinear signal processing methods or localized time-frequency methods, but didn't search for distinct drowsiness related patterns within the whole recordings. To our knowledge, the ABP and MSP measures are the only ones which consider these concerns.

Future quantitative EEG analysis should meet these requirements. Even under the limitation of relatively small sample sizes, machine learning methodologies in the framework of pattern recognition have the potential to handle these issues.

CONFLICT OF INTEREST

The authors wish to certify that no actual or potential conflict of interest and financial conflicts in relation to this manuscript article exists.

REFERENCES

Åkerstedt, T. & Gillberg, M. (1990) Subjective and objective sleepiness in the active individual. *Int J Neurosci, 52*, 29-37.

Cajochen, C., Brunner, D., et al. (1995) Power density in theta/alpha frequencies of the waking EEG progressively increases during sustained wakefulness. *Sleep, 18*, 890-894.

Caterpillar Machine Research (2008). *Operator Fatigue Detection Technology Review.* (Project Report, November 2, 2008). Peoria, IL.

Golz, M., Sommer, D., Chen, M., Trutschel, U. & Mandic, D. (2007) Feature fusion for the detection of microsleep events. *J VLSI Signal Process*, *49*, 329-342.

Golz, M., Sommer, D., Trutschel, U., Sirois, B. & Edwards, D. (2010) Evaluation of fatigue monitoring devices. *Somnology (Berlin) 14*, 187-199.

Harrison, Y. & Horne, J. (1996) Occurrence of microsleeps during daytime sleep onset in normal subjects. *Electroencephal Clin Neurophysiol, 98*, 411-416.

Kecklund, G. & Åkerstedt, T. (1993) Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics, 36*, 1007-1017.

Lal, S. & Craig, A. (2001) Electroencephalography activity associated with driver fatigue: implications for a fatigue countermeasure device. *J Psychophysiol, 15*, 183-189.

Peiris, M.R, Jones, R.D., Davidson, P.R., Bones, P.J. (2006). Detecting behavioral microsleeps from EEG power spectra. *Proc 28ⁿᵈ Int Conf IEEE EMBS,* New York City, 5723-5726.

Santamaria, J. & Chiappa, K. (1987) The EEG of Drowsiness in Normal Adults. *J Clin Neurophysiol*, *4*, 327-382.

Simon, M., Schmidt, E., Kincses, W., Fritzsche, M., et al. (2011) EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clin Neurophysiol, 122*, 1168-1178.

Sommer, D., Golz, M. & Krajewski, J. (2008) Consecutive detection of extreme central fatigue. *Proc 4th Europ Conf Int Fed Med Biol Eng*, 243-246.

Torsvall, L. & Åkerstedt, T. (1987) Sleepiness on the job: continuously measured EEG changes in train drivers. *Electroencephal Clin Neurophysiol, 66*, 502-511.