

AUTOMATED ASSESSMENT OF VIGILANCE USING COMMITTEES OF RADIAL BASIS FUNCTION ANALYSERS

Stephen Roberts¹, Iead Rezek¹, Richard Everson², Helen Stone³, Sue Wilson⁴ & Chris Alford⁵

1. Department of Engineering Science, University of Oxford, UK.*
2. Department of Computer Science, University of Exeter, UK.
3. Sowerby Research Centre, BAE SYSTEMS, Bristol, UK.
4. University of Bristol, UK.
5. University of the West of England, Bristol, UK.

INTRODUCTION

It is clear that whenever people perform repetitive, boring or long-term tasks a loss in concentration can occur. These lapses in vigilance may have serious consequences under certain circumstances. Human cognitive processes are far from understood and the processes by which individuals have lapses in alertness are many. In this paper we consider the analysis of the human EEG during vigilance experiments as a case study in supervised data analysis. We address several issues which are found in many data analysis problems. The issue of finding informative signal parameterisations in multi-channel environments is approached using a feature selection process. Subsequent analysis uses Bayesian committees of Radial Basis Function analysers. A comparison is made of two analysis approaches, the first based on regression and the second setting the problem as a classification task with ‘extremal-label training’. Results are presented for a representative sample of subjects.

METHODOLOGY

Experimental details

A total of 12 healthy volunteers took part in this study. Each performed a set of four recording blocks. A simple tracking task was performed in each recording block along with distraction tasks based on reaction to external events. Each block lasted for 60 minutes and data was recorded continuously throughout that time. The tracking task consisted of keeping a simple ‘flight-simulator’ level. A total of 8 channels of EEG were recorded along with measures of eye movement activity (two EOGs) and muscle tone in the neck (two EMGs). All channels were recorded at a fixed sample rate of 256Hz. Observers were asked to maintain a simulated aircraft in a straight and level heading using a joystick to correct random errors on a head-up attitude display. Distraction tasks consisted of a reaction time task (responding to an intermittent stimulus) and a numerical task.

All observers tracked vertical deviations better than horizontal ones, however, there was no significant consistent worsening of tracking as the task proceeded. In a few cases performance on the tracking task noticeably improved over the first ten minutes. This seems to indicate that changes in the

*Email: sjrob@robots.ox.ac.uk, Fax: +44 (0)1865 273908.

task performance are due to a combination of not only drowsiness but also psychological effects as well e.g the subject’s ability to self-motivate. We would argue, however, that even in cases where drowsy subjects performed the task well, they had a greater propensity to poor performance, and that this is related to the underlying state of alertness. We hence chose to make predictions, based on the features extracted from the EEG and other channels, of a set of human-scored labels. This scoring was based on the EEG and scored by a human expert to a set of 6 wake-state stages where 6 represents the most ‘alert’ waking EEG and 1 corresponds to EEG indicative of non-alert drowsiness.

Feature encoding

A number of features derived from physiological data have been used to assess brain state. Recent studies of sleep EEG data [1, 2], over many observers, have found consistently that AR reflection coefficients are excellent indicators of the state of the cortex. All EEG signals were hence parameterised over successive 5s segments using the reflection coefficients of a 5-th order autoregressive (AR) model, as detailed, e.g. in [3]. The p -th reflection coefficient, k_p , defines the reduction in residual signal-model error, E , when the AR model increase its order from $p - 1$ to p ,

$$E_p = (1 - k_p^2) E_{p-1} \quad (1)$$

We use an elegant lattice-filter approach to the AR model which operates under a Bayesian framework [4] in which uncertainty in the AR model parameters is taken into account and integrated out.

Also calculated to a 5s resolution were the envelopes of the EMG channels (based on a simple smoothed rectification of the high-frequency activity) and the blink rate (the fraction of time eyes are closed) calculated from the EOG channels.

Feature selection

Having extracted AR model coefficients for the EEG channels (and differential channels, i.e. parameters for the EEG difference between two electrodes) and derived measures of muscle activity (from the EMG channels) and blink rate from the EOG we have several hundred combinations of possible features. Each of these combinations of features was used in turn as input to a regressor optimised so as to predict the smoothed human-scored labellings. For each combination of features the regressor was trained on one subject and one task after which the efficacy of the combination was assessed using the root mean square error between the predicted output and the true (smoothed) labels from another test task. A committee of RBF networks, each using thin-plate spline hidden units and trained under a Bayesian paradigm [5], was used for the regression (see next section).

We find that reflection coefficients from the differential channels between electrodes O1-O2 & T3-T4 and electrode Fz-GND (based on the standard 10/20 electrode positions) are picked in the best feature combinations, as is the blink rate. We note that signals from the Fz electrode have been used in other vigilance studies [6]. Interestingly, higher-order reflection coefficients (greater than 3) were found to be uninformative, indicating that EEG changes in vigilance may indeed be dominated by only one or two changes in the spectrum. All results presented in this paper were obtained using the first three reflection coefficients from electrodes T3-T4 and the blink rate measure.

Analysis

Figure 1 shows a typical human-scored labelling, scored into a set of stages on a 15s resolution. Stage 6 is the most ‘awake’ and stage 1 the least. We see that the labelling is characterised by regions of relative stability and regions of high variability. Regression or classification using this ‘raw’ scoring

fails. We have two choices: firstly we may regard the levels of the human labelling as quantisations of a continuous, smoother, measure and hence represent the time course of the labelling using e.g. a (crude) running mean and variance measure. We refer to this approach as the *regression-based method*. The second option is to take the approach advocated in [7] in which only *extremal* labels are used to determine the training set. This assumes that the extremes of e.g. wakefulness and drowsiness are more reliably assessed than the intermediate states. In this case we take sections of data with labels 5 or 6 to represent alert wakefulness and sections of data with label 1 to represent a drowsy state. This allows a standard classification approach to the problem in which $P(\text{alert} \mid \text{data})$ may be estimated. We refer to this approach as the *extremal-training method*.

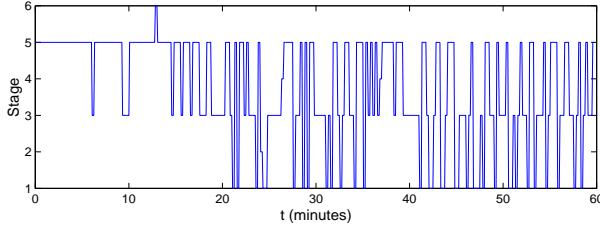


Figure 1: Typical human-scored labelling.

Radial Basis Function Analysers

Both regression and classification analysis was performed using committees of Bayesian Radial-Basis Function (RBF) analysers, as detailed in [5]. The basic theory and methodology may be found in [8] and is only briefly reviewed here.

We consider the analysis of a datum (feature vector) \mathbf{x} associated with a target t . The RBF operates by calculating the response of a set of simple non-linear functions to \mathbf{x} . The functions are often taken to be Gaussian but in this paper we utilise *thin-plate spline* functions. The response $\phi_k(\mathbf{x})$ of the k -th such function is written as

$$\phi_k(\mathbf{x}) = r^2 \ln r \quad (2)$$

where $r = \|\mathbf{x} - \boldsymbol{\mu}_k\|$ in which $\boldsymbol{\mu}_k$ is the location of the spline function in the feature space. The resultant output of the RBF, conditional on an input and a given set of weights, is thence:

$$y(\mathbf{x}) = g \left(\sum_k w_k \phi_k(\mathbf{x}) + w_{bias} \right) \quad (3)$$

in which w_k are a set of coupling weights, w_{bias} is a bias weight and $g(\cdot)$ a transfer function which is linear for regression and sigmoidal for classification approaches (see [8] for details). We may write Equation 3 in terms of a set of *latent*, or hidden, variables a such that

$$y(\mathbf{x}) = g(a) \quad (4)$$

This is a useful conceptually as we may look easily at the density over a as it is linear in the parameters of the model. Following the approach taken in [5] we utilise a committee of thin-plate spline networks. Committees of analysers are provably better, on average, than the mean performance of any one analyser [8] and with randomly located spline positions (i.e. the set of $\boldsymbol{\mu}_k$ are located at random within the bounds of the data set) the resultant systems become very rapid to train as the only free parameters of the system are the weights w_k . When $g(\cdot)$ is linear these may be set via a matrix pseudo-inverse approach. When $g(\cdot)$ is sigmoidal, however, a non-linear optimiser must be employed; in all the results presented here we used a quasi-Newton method based on the BFGS approximations to the *Hessian* matrix [9].

Equation 3, however, is dependent on the set of weights and in a Bayesian framework these are integrated out (the Bayesian framework, in essence, aims to remove unknown parameters by integrating over them - the interested reader is pointed to [10]). Furthermore, a probability density, rather than a point value, may be obtained. For computational simplicity (and as we have a large training set available) we choose to take the well-known *evidence* approach, popularised by MacKay

[11, 12] and applied to committees of thin-plate spline RBFs in [5]. Denoting the output y of the analyser to be generated by a mapping from the latent variable, a (Equation 4), the standard evidence framework gives a Gaussian density over a with mean $a^*(\mathbf{x}) = a(\mathbf{x}; \mathbf{w}^*)$ (where \mathbf{w}^* is the weight vector of most-probable free parameters i.e. those obtained after the training process) and variance of:

$$\sigma^2(\mathbf{x}) = \frac{1}{\beta(\mathbf{x})} + \mathbf{g}^T(\mathbf{x})\mathbf{H}^{-1}\mathbf{g}(\mathbf{x}) \quad (5)$$

in which $\frac{1}{\beta}$ is the estimated noise variance of the targets (zero in the case of classification), $\mathbf{g} = \partial y / \partial \mathbf{w}$ evaluated at \mathbf{w}^* and \mathbf{H} is the Hessian matrix (the matrix of second derivatives of the error function with respect to the weights - the error functional is least squares for regression and cross-entropy for classification, see [8] for details).

If we take a committee of such analysers, with weighting coefficients γ_i then the probability density on a will be a Gaussian mixture model. We choose to parameterise this model by a moment-matched Gaussian whose mean is just the weighted average over committee members, i.e.

$$\bar{a} = \sum_i \gamma_i a_i^* \quad (6)$$

and variance given as [5, 13]:

$$\sigma_{comm}^2 = \sigma_c^2 + \sigma_e^2 + \sigma_{wu}^2 \quad (7)$$

in which σ_c^2 represents the committee *variance*,

$$\sigma_c^2 = \text{var}[a_i^*] \quad (8)$$

σ_e^2 the target noise (and residual bias - though this is assumed to be small compared to the target noise for well-formed models) estimated by:

$$\sigma_e^2 = \sum_i \frac{\gamma_i}{\beta_i} \quad (9)$$

and σ_{wu}^2 the parameter uncertainty of the committee,

$$\sigma_{wu}^2 = \sum_i \gamma_i \mathbf{g}_i^T(\mathbf{x}) \mathbf{H}_i^{-1} \mathbf{g}_i(\mathbf{x}) \quad (10)$$

This is an intuitively pleasing result as the total error may be regarded in a common-sense manner as arising from three distinct causes.

- The first term penalises variant decisions *between* committee members,
- the second penalises the committee *as a whole* if the output is associated with a region of input space with high target noise levels
- and the third penalises solutions which have poorly set parameters.

In the case of the regression approach, $y = a$. As we regress, however, onto the mean (smoothed) label set we do not take into account the intrinsic variance in the human labelling process. The latter may itself be predicted. We hence regress onto the two-dimensional target space in which the first target is the mean label value and the second is the (log) variance in the labelling at that time. We consider the total predictive variance of our output to be:

$$\tilde{\sigma}^2(\mathbf{x}) = \sigma_{comm}^2(\mathbf{x}) + s^2(\mathbf{x}) \quad (11)$$

in which s^2 the predicted variance of the label set.

Although the networks are trained under a Bayesian paradigm, for the classification case we do not allow high uncertainty (variance) in the latent space to *moderate* (see [12, 8], for example) the resultant output probability as this results in the movement of $P(\text{alert} | \mathbf{x})$ for example, to fall towards the class prior of 1/2. Given that such values are thence interpreted as periods of loss of vigilance we regard this as undesirable.

We note that, as discussed in [5], thin-plate splines have the pragmatic advantage over Gaussians in an RBF analyser in that their response increases with distance away from the training data set. This means that the committee variance escalates rapidly in regions outside the training data set hence giving rise to wide (uncertain) distributions in the latent

space. This naturally screens against estimates which are derived from input data which is inconsistent with that of the training set.

RESULTS

In all the results presented below a set of ten RBF networks were used in each committee. The results are not highly sensitive to the number of committee members and this number was chosen on the grounds of computational pragmatism. The number of spline functions in each network was set using the evidence of the model given the data [11, 8, 5]. For the regression approach this gave rise to networks with ten splines. For classification eight were sufficient, however. All data are shown smoothed using an adaptive moving-average filter of length 1 minute (12 feature samples), based on an eigen decomposition of successive data windows [14]. In the regression case this filtering takes place at the output of the analyser and in the latent space for classification [15]. For ease of comparison the target label set is also shown smoothed to the same resolution. The training set in all cases consisted of two hours of data from two subjects who were thence excluded from further analysis.

We present results for both regression and ‘extremal-trained’ classification for three representative subjects. The upper plot of each of Figures 2, 3 & 4 shows the regression-based approach with the thick solid line being the predicted mean and the thin lines at ± 1 S.D. from this. The thick dashed line (often hidden behind the prediction) is the (smoothed) human-scored labelling. The lower plot of each figure shows the results of the classification approach, the thick line being the probability of alert wakefulness (scaled so that a probability of unity maps to label 6 and probability zero to label 0. This enables a direct comparison with the label set). The label set itself is shown in these plots as the thin solid line. Each plot consists of four subplots which are the four hour-long recording blocks for each subject.

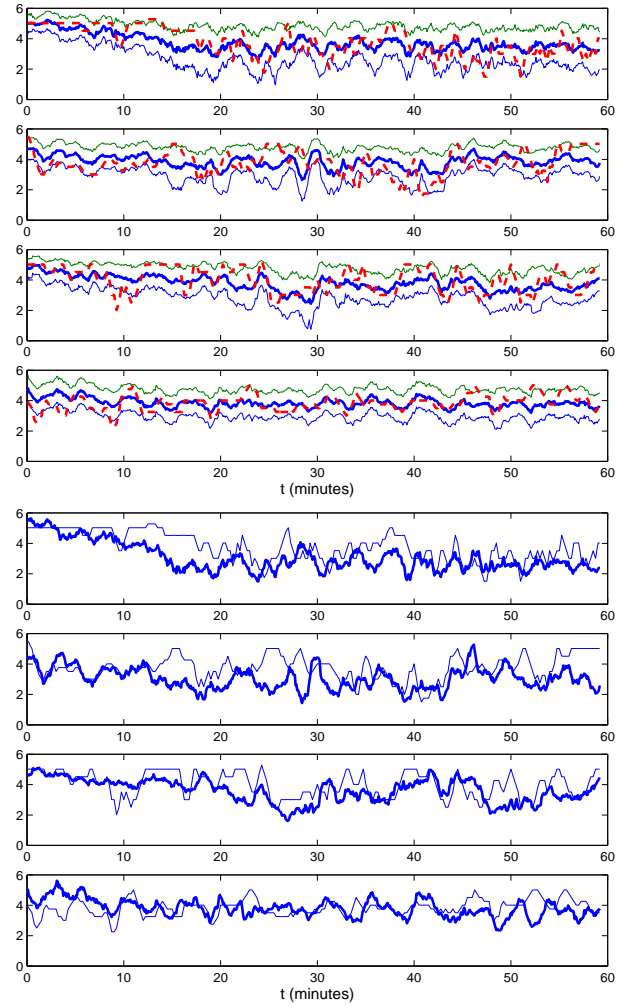


Figure 2: *Subject A. Analysis based on regression (top four plots) and classification (bottom four plots).*

We see clearly that both approaches, although not perfect, do track changes in alertness (as defined via the human-scored labels) reasonably well indicating that alertness may be, in principle, estimated from a small number of parameters from a single EEG channel and eye-movements. We would argue that although the classification approach gives reasonably good results there are a number of factors which are not in its favour:

1. As posterior (class) probabilities are being estimated, so point estimates, rather than distributions, are generated.
2. As only extreme-labelled data are

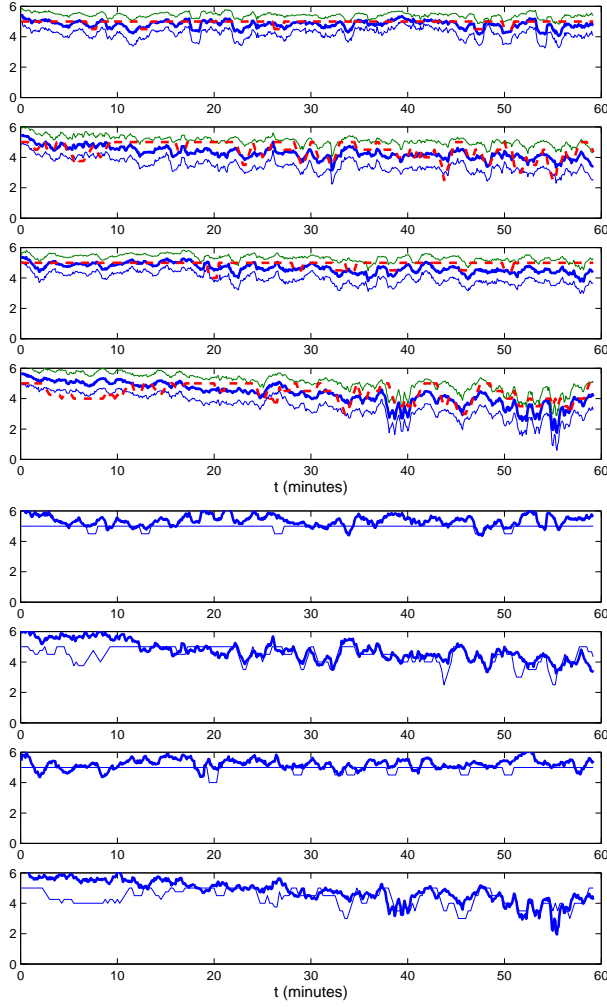


Figure 3: *Subject B. Analysis based on regression (top four plots) and classification (bottom four plots).*

utilised, the training set is reduced in size.

In the regression approach we see gradual changes in the predicted variance in the vigilance estimate such as starting around 10 minutes and 20 minutes into the first (uppermost) blocks of, respectively, subjects A and C. We also note the existence of a rhythmicity in the vigilance measures with period of order five minutes. This is most clearly seen in subject C and is not an artefact of the filtering or analysis process. The origin of this oscillation, which we note is present in many of the alertness indicators which we have analysed, is not currently known.

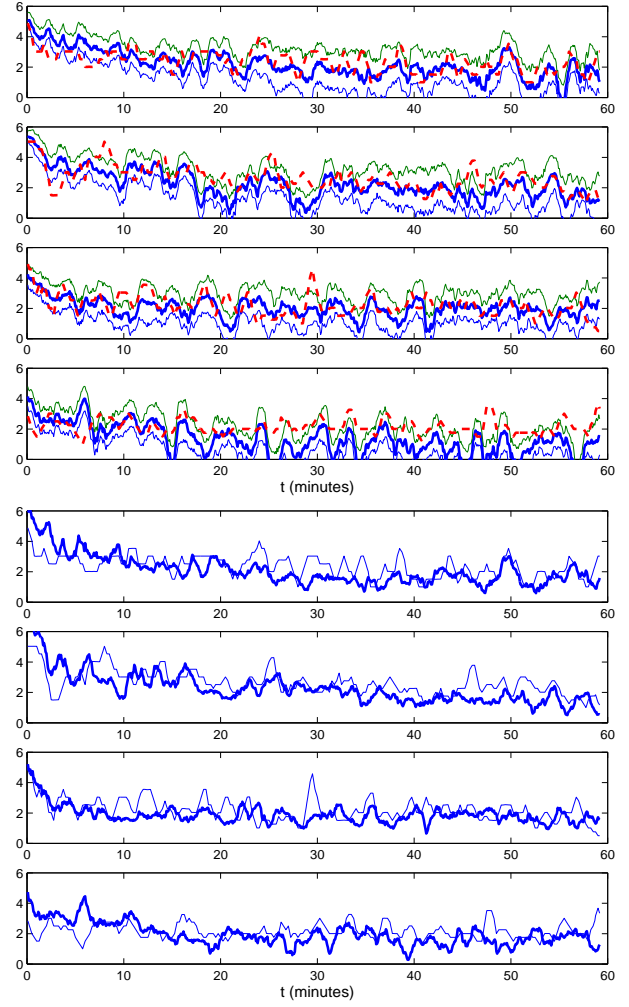


Figure 4: *Subject C. Analysis based on regression (top four plots) and classification (bottom four plots).*

CONCLUSIONS & FUTURE WORK

The assessment of vigilance from a small number of physiological measurements may be of importance in safety monitoring in a number of professions. We have shown that it is possible to make reasonable estimates of the state of alertness of a subject based on EEG and eye-movement information. The estimates appear to be fairly robust across subjects using appropriately-chosen features from the signals. It is noted, however, that in general neither the smoothed labels (human-scored) nor the resultant estimated measures correlate well with the corresponding tracking performance measure. Indeed as dis-

cussed earlier, the existing experimental data set shows, save for a couple of exceptions, no trends of deterioration in tracking performance as the one-hour recording proceeds. What we would ideally like then is to determine changes in the data which are due to changes in the subjects' *coping strategies* as well as true changes in alertness. We hope that new experimental data will shed light on this.

References

- [1] P. Sykacek, I. Rezek, S.J. Roberts, A. Flexer, and G. Dorffner. Bayesian wrappers versus conventional filters: feature subset selection in the SIESTA project. In *Proceedings of EMBEC-99*, Vienna, Austria, November 1999.
- [2] M. Pregenzer, S.J. Roberts, and G. Pfurtscheller. Feature selection in the SIESTA project using DSLVQ. In *Proceedings of EMBEC-99*, Vienna, Austria, November 1999.
- [3] J. Pardey, S. Roberts, and L. Tarassenko. A Review of Parametric Modelling Techniques for EEG Analysis. *Med. Eng. Phys.* 18(1):2–11, 1996.
- [4] P. Sykacek, I. Rezek, S.J. Roberts, A. Flexer, and G. Dorffner. Reliability in Preprocessing - Bayes rules. In *Proceedings of EMBEC-99*, Vienna, Austria, November 1999.
- [5] S.J. Roberts and W. Penny. A Maximum Certainty Approach to Feedforward Neural Networks. *Electronics Letters*, 33(4):306–307, 1997.
- [6] S. Makeig and M. Inlow. Lapses in alertness: coherence of fluctuations in performance and eeg spectrum. *Electroenceph. Clin. Neur.*, 86:23–35, 1993.
- [7] S. Roberts and L. Tarassenko. New Method of Automated Sleep Quantification. *Med. & Biol. Eng. & Comput.*, 30(5):509–517, 1992a.
- [8] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, 1995.
- [9] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1991.
- [10] J.M. Bernardo and A.F.M. Smith. *Bayesian Theory*. John Wiley, 1994.
- [11] D.J.C. MacKay. A Practical Bayesian Framework for Backpropagation Networks. *Neural Computation*, 4:448–472, 1992.
- [12] D.J.C. MacKay. The Evidence Framework applied to Classification Networks. *Neural Computation*, 4:720–736, 1992.
- [13] W.D. Penny, D. Husmeier, and S.J. Roberts. *The Bayesian Paradigm: second generation neural computing*. in: *Artificial Neural Networks in Biomedicine*, P. Lisboa, I. Ifeachor (Eds.). Springer, 1999.
- [14] J. B. Elsner and A. A. Tsonis. *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. Plenum Press, New-York, 1996.
- [15] S.J. Roberts and W.D. Penny. Real-time Brain Computer Interfacing: a preliminary study using Bayesian learning. *Medical and Biological Engineering & Computing*, 38(1):56–61, 2000.

Acknowledgements

The authors would like to thank Will Penny for helpful discussions. Thanks also to Kevin Hapeshi for the flight simulator software and help with the experimental design. This work was supported by a grant from BAE SYSTEMS to whom the authors are most grateful.