Model-based multivariate decoding and model selection

Kay H Brodersen[1,2,†] · Florent Haiss[3] · Cheng Soon Ong[1] · Fabienne Jung[4] · Paul Allen[5] · Marc Tittgemeyer[4] · Joachim M Buhmann[1] · Philip McGuire[5] · Bruno Weber[3] · Klaas E Stephan[2,6]

[1] Department of Computer Science, ETH Zurich, Switzerland [2] Laboratory for Social and Neural Systems Research, Institute for Empirical Research in Economics, University of Zurich, Switzerland [3] Institute of Pharmacology and Toxicology, University of Zurich, Switzerland [4] Max Planck Institute for Neurological Research, Cologne, Germany [5] Division of Psychological Medicine, Institute of Psychiatry, King's College London, UK [6] Wellcome Trust Centre for Neuroimaging, University College London, UK
[†] Correspondence: kay.brodersen@inf.ethz.ch

1 Introduction

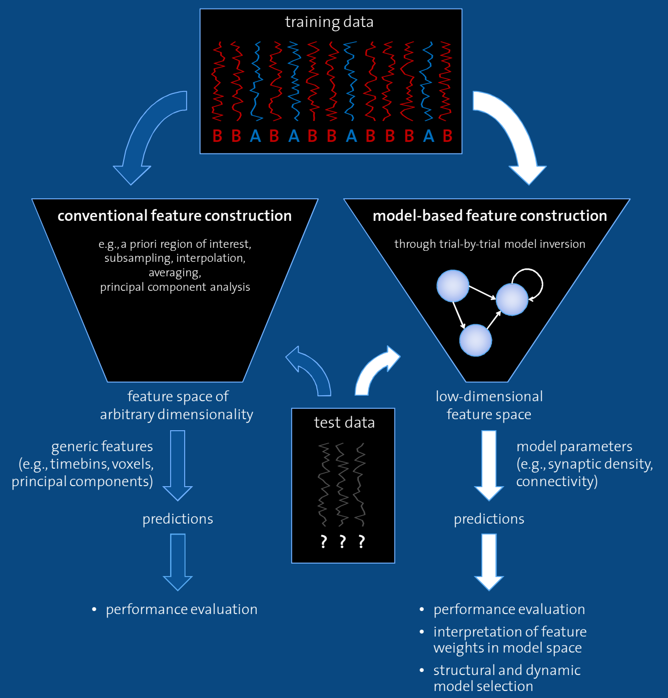
How much information can we decode from measurements of neural activity? Conventional approaches face two problems: feature spaces are high-dimensional; and results are difficult to interpret. Here, we propose a model-based decoding approach that addresses both challenges from a new angle.

Step 1: invert a dynamic causal model (DCM) of fMRI, EEG, MEG, or electrophysiological data in a trial-by-trial (or subject-by-subject) fashion;

Step 2: train a classifier on a strongly reduced feature space derived from the trial-wise (or subject-wise) model parameter estimates;

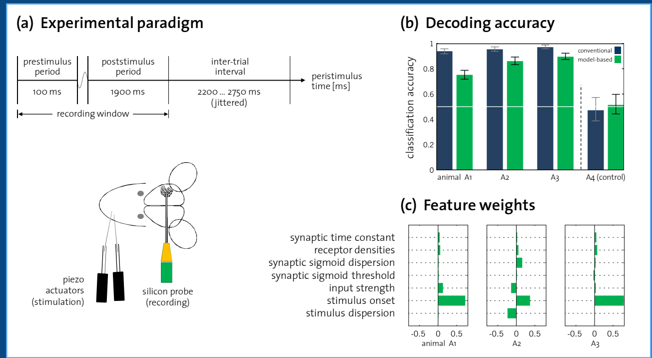
Step 3: test the classifier on new data, or reconstruct the separating hyperplane to assess which features were jointly informative.

2 Conventional vs. model-based decoding



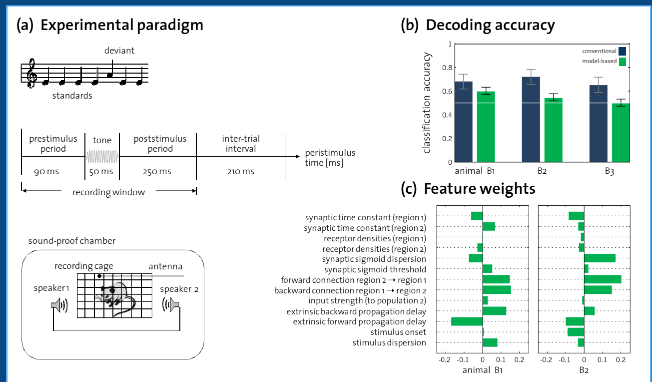
3 Example: decoding a sensory stimulus

Electrophysiological recordings were acquired from rat barrel cortex during a simple whisker-stimulation experiment (a). While both conventional and model-based decoding succeeded in predicting which whisker had been stimulated on a given trial (b), the model-based scheme also revealed which biophysical parameters provided most discriminative power.



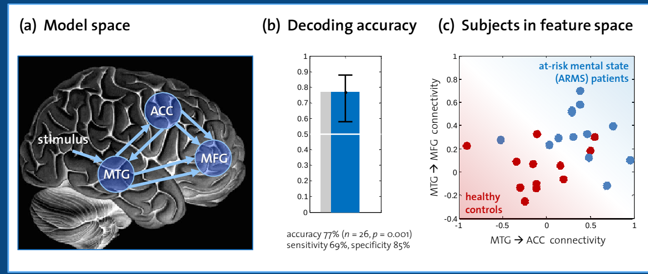
4 Example: decoding auditory mismatch

Electrophsyiological recordings were acquired from 2 electrodes in rat auditory cortex during an auditory-mismatch paradigm (a). In decoding which tone had been played on a given trial, model-based decoding performed significantly above chance in two out of three cases (b) and revealed a consistent pattern of influential parameters across both animals (c).



5 Example: decoding at-risk mental states

FMRI data were acquired from 13 at-risk mental state (ARMS) patients and 13 healthy controls. We considered a large set of dynamic causal models with different patterns of temporo-prefrontal connectivity. While Bayesian model selection failed to discriminate between patients and controls, model-based decoding provided strong prediction performance (b) and revealed which functional connections supported correct diagnoses (c).



6 Discussion and conclusions

Decoding with model-based feature construction offers three advantages over conventional decoding algorithms:

- The scheme provides biologically informed dimensionality reduction. This renders generic heuristics for feature selection obsolete.
- Decoding results can be interpreted in the context of a mechanistic model, by assessing which set of biophysical parameters underlie prediction performance. Such a mechanistic interpretation might prove particularly useful in clinical studies.
- Competing models can be compared to one another by evaluating how much information is preserved by each of them. The scheme therefore allows for decoding even when discriminability is not afforded by differences in model structure but only by patterns of parameter estimates under the same model structure; and it enables structural model selection in cases where Bayesian model selection is not applicable.

Acknolwedgements
This study was funded by the NEUROCHOICE project of SystemsX.ch (FH, BW, KES), the University Research Priority Program 'Foundations of Human Social Behaviour' at the University of Zurich (KHB, KES), the NCCR 'Neural Plasticity' (KES), and the Max Planck Society (FJ, MT).

In order to allow for cross-comparison of decoding methods, we have made our data available online at http://people.inf.ethz.ch/bkay/downloads.

References

- Brodersen, K.H. et al., 2010. Model-based feature construction for multivariate decoding. *NeuroImage* (in press).
- David, O. et al., 2006. Dynamic causal modeling of evoked responses in EEG and MEG. *NeuroImage*, 30(4), 1255-1272.
- Friston, K.J., Harrison, L. & Penny, W., 2003. Dynamic causal modelling. *NeuroImage*, 19(4), 1273-1302.
- Haynes, J. & Rees, G., 2006. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523-534.
- Kiebel, S.J., Garrido, M.I. & Friston, K.J., 2007. Dynamic causal modelling of evoked responses: The role of intrinsic connections. *NeuroImage*, 36(2), 332-345.