Classics never die: a novel multilevel hidden Markov model to uncover mood states in patients with bipolar disorder

Keywords: Bayesian statistics, hidden Markov model, dynamics of behaviour, interpretable machine learning, bipolar disorder

Extended Abstract

Due to technological advances such as smartphones and smartwatches, sensors, and automatic coding of video recordings, it has become relatively easy and affordable to collect data on groups of individuals with a high temporal resolution (e.g., [1]). This novel type of data obtained using methods such as experience sampling, or ecological momentary assessment (EMA), receives the name of intensive longitudinal data (ILD). The steep increase in use of ILD reflects that researchers recognize the unique value these data have for studying behavioural phenomena that unfold over time (e.g., [2]). Although larger than traditional data sets in social and behavioural research (i.e., panel studies, longitudinal studies) ILD is often much smaller than current day Big Data, with individuals in the range of tens to hundreds, and a similar number of observations per individual. Given these characteristics, fully exploiting the wealth of information contained within ILD requires developing methods complex enough to offer a good representation of social and behavioural dynamics, yet simple enough such that they can be used to extract interpretable insights for applied researchers and clinicians from relatively small data sets.
Here we present a Bayesian multilevel extension of the classic hidden Markov model (HMM), a well-known machine learning method with a long history of success in time-series modelling (e.g., [3]), with an improved fit for social and behavioural data. By extending the model to the mixed-effect framework, the multilevel HMM offers three advantages over conventional HMMs: 1) it provides individuals with individual specific transition and emission parameters, along with individual specific trajectories over time, while ensuring that the meaning of the states remains the same across individuals, 2) it allows for the quantification of variability between individuals with random effects, and 3) it has a hierarchical structure with a regularizing effect over individual-specific parameters, which makes the model more robust to outliers and more reliable even with smaller data sets. We show the usefulness of the model proposed with an application to an empirical data set consisting of EMA data collected from patients with bipolar disorder.

Empirical motivation: Bipolar disorder (BD) is a chronic psychiatric condition characterized by large episodic changes in mood and energy. Recently, BD has been proposed to be conceptualized as chronic cyclical mood instability, as opposed to the traditional view of alternating discrete episodes with stable periods in-between. The problem lies in the fact that weekly or monthly therapy sessions may not be frequent enough to detect oscillations in mood of a higher frequency. Recognizing this mood instability may improve care and calls for high-frequency measures coupled with advanced statistical models.
Data: Empirical data consisted in 4-month EMA data in twenty patients comprising 12-item self-report questionnaires (5x daily) measuring manic and depressive constructs (yielding

~9.820 observations in total). Manic and depressive symptoms were also assessed weekly using clinically validated questionnaires (Altman Self-Rating Mania Scale and the Quick Inventory for Depressive Symptomatology Self-Report questionnaires, respectively).

Methods: Fitting of the MHMM was performed with the statistical software R and the developer version of the R package *mHMMbayes*. Uncovering the sequence of momentary mood states was done for each patient with the Viterbi algorithm based on patient-specific parameters, which yields the relative probability of each state on each occasion based on the pattern of response of a patient over the twelve EMA items. Model selection was performed by comparing relative model fit with AIC and the ability of the model to reproduce the original EMA data via Bayesian posterior predictive checks. The empirical duration of the mood states and the relative switching frequency were computed for the uncovered state sequence of each patient and over the aggregated sample. Alignment between HMM-uncovered momentary mood states and weekly questionnaires was assessed with a multilevel linear model. The reliability of the mood states uncovered with the MHMM was assessed with a small Monte Carlo simulation, focused on balanced accuracy of state uncovering, and absolute percent bias and coverage of parameter estimation.

Results: Our multilevel HMM uncovered four mood states (*euthymic*, *manic*, *mixed*, and *depressive*) that generally aligned well with weekly symptom scores (Figure 1). The Monte Carlo simulation indicated a high level of accuracy on the uncovered mood states ($M=88.2$%, SD=2.4%) averaged over patients and simulation repetitions. The absolute percent bias in the self-transition and emission parameters were generally low, ≤10% threshold (Median=7.9% and Median=7.4%, respectively), and a parameter coverage of at least 89% and 87%, respectively. Mood states switched more frequently than weekly data suggested, indication of rapid-cycling dynamics that escapes detection under traditional retrospective data. On average, patients remained <24h in one state: average state duration was 17.1h (SD=78.0) for *euthymic*, 16.3h (SD=70.8) for *depressive*, 9.1h (SD=11.6) for *manic*, and 8.8h (SD=10.9) for *mixed*. In almost half of the patients, significant mood instability was observed. State duration and switching varied strongly between individuals.

Conclusion: This work shows how the novel Bayesian multilevel HMM can be used to extract interpretable insights about the dynamics of behaviour in social and behavioural data of a relatively small size by today standards. The results obtained with the multilevel HMM on the empirical data expand on current comprehension of the mood dynamics in bipolar disorder and suggest new implications for its conceptualization, as well as motivate the use of novel techniques such as ours and a higher frequency of mood assessment.

References

[1] S. Ariens, E. Ceulemans, and J. K. Adolf, 'Time series analysis of intensive longitudinal data in psychosomatic research: A methodological overview', *Journal of Psychosomatic Research*, vol. 137, p. 110191, Oct. 2020, doi: 10.1016/j.jpsychores.2020.110191.

[2] E. L. Hamaker and M. Wichers, 'No Time Like the Present: Discovering the Hidden Dynamics in Intensive Longitudinal Data', *Current Directions in Psychological Science*, vol. 26, no. 1, pp. 10–15, Feb. 2017, doi: 10.1177/0963721416666518.

[3] L. R. Rabiner, 'A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition', *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989, doi: 10.1109/5.18626.

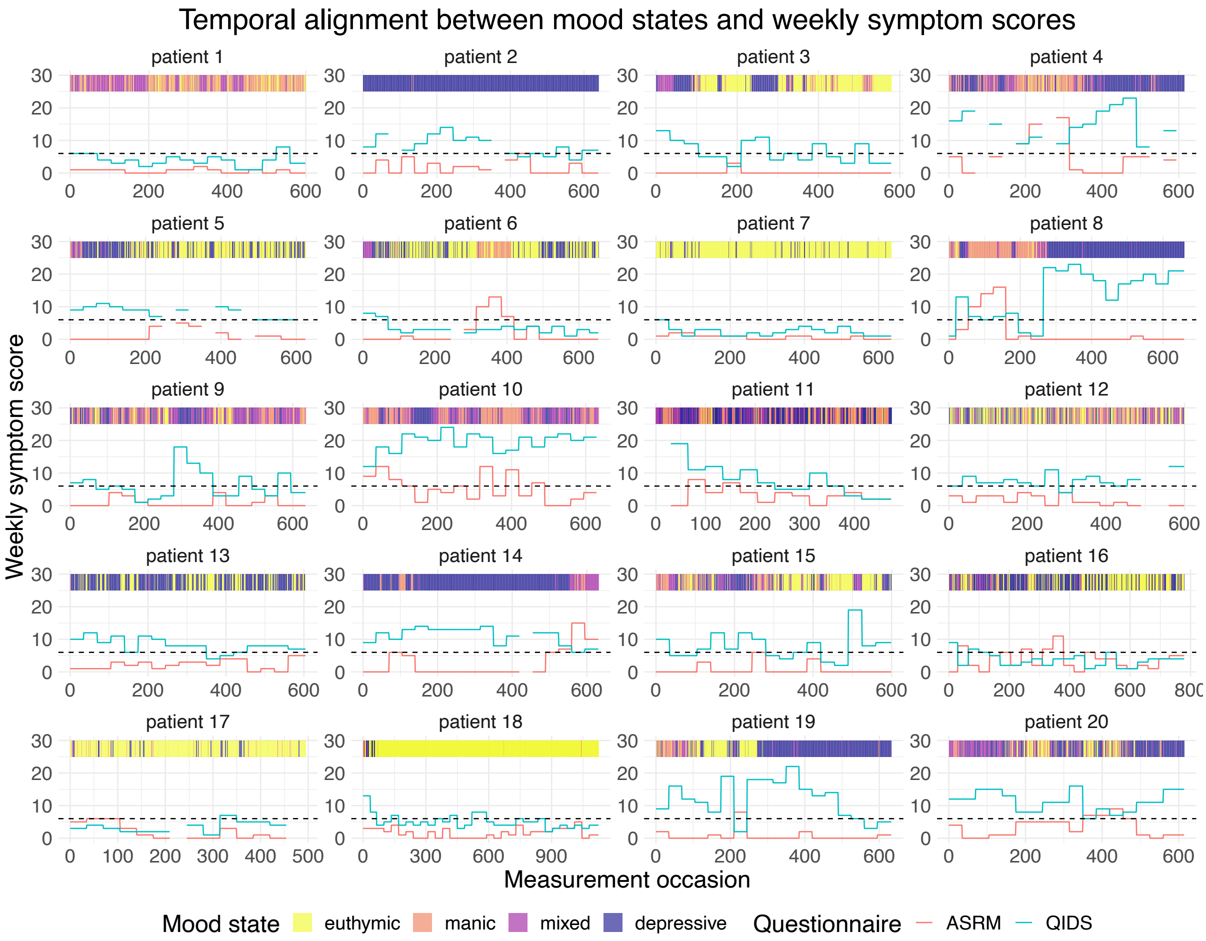

Figure 1. Temporal correspondence between uncovered mood states based on ecological momentary assessment data and manic and depressive weekly symptom-scores.