

Changepoint analysis with the changepoint package in *R*

R. Killick^{1,*}, I.A. Eckley¹

1. Department of Mathematics & Statistics, Lancaster University, Lancaster, U.K.

*Contact author: r.killick@lancs.ac.uk

There exists a rich literature exploring changepoint problems which dates back to the 1950s. To date there have been few implementations of changepoint methods within existing *R* packages. At the time of writing, those packages which are currently available implement novel methodology only. As a consequence it is more difficult (i) to compare the performance of new methods with those proposed in the established literature and (ii) for practitioners to test changepoint methods on their data.

The changepoint package attempts to resolve the above by providing both well-established and new methods within a single package. In doing so, we hope to make it easier for practitioners to implement existing methods and for researchers to compare the performance of new approaches against the established literature.. The package has therefore been designed to give users access to many techniques for changepoint analysis within a few easy to use functions. It includes functions that detect changes in mean, variance and regression under various distributional and distribution-free assumptions.

Given the capacity to sample time series at ever higher frequencies the need for computationally efficient search methods is increasing. This is currently an active area of research and so we include four search options in this initial version of the package. The search options which are available are At Most One Change (AMOC), Binary Segmentation ([Scott and Knott, 1974](#)), Segment Neighbourhoods ([Auger and Lawrence, 1989](#)) and PELT ([Killick et al., 2011](#)).

The presentation will be structured as an introduction to changepoint analysis followed by a demonstration of the methods within the **changepoint** package. The methods will be illustrated using datasets from a variety of application areas including genetics, oceanography, finance and linguistics.

References

- Auger, I. E. and Lawrence, C. E. (1989). Algorithms for the optimal identification of segment neighborhoods. *Bulletin of Mathematical Biology*, 51(1):39–54.
- Killick, R., Fearnhead, P., and Eckley, I. A. (2011). Optimal detection of changepoints with a linear computational cost. *In Submission*.
- Scott, A. J. and Knott, M. (1974). A cluster analysis method for grouping means in the analysis of variance. *Biometrics*, 30(3):507–512.