Comment on "Multiscale change point inference" by K. Frick, A. Munk and H. Sieling.

Rebecca Killick
Mathematics & Statistics Department, Lancaster University

October 29, 2013

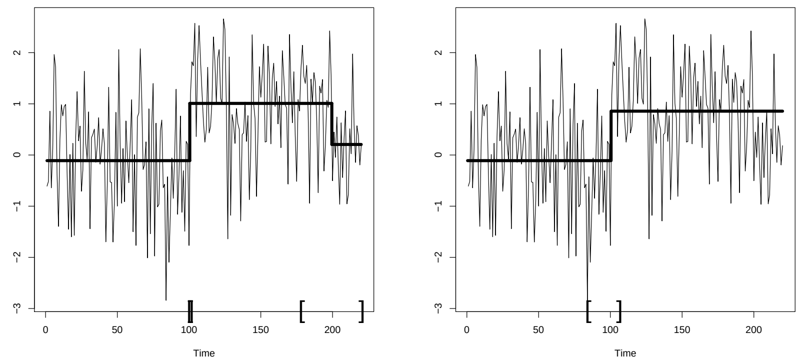
The authors are to be congratulated on a valuable contribution to the changepoint literature and commended for making their algorithm available within the *stepR* package (Hotz and Sieling, 2013). The automatic penalty selection and construction of confidence intervals for changepoint locations are of particular interest.

Firstly a question on the automatic penalty selection. In Section 3 the authors reframe SMUCE as a dynamic program for ease of computation. By Lemma 1 in Section 3 the solution of the dynamic program is also the solution of the original problem if γ is chosen larger than $[2(nq + n\sqrt{2\log(en)})^2]^{-1} + \ell(Y, m^{-1}(\overline{Y}))$. The choice of γ is not mentioned in the paper as the authors instead focus on the choice of q. Given $q = q_n^*$ as indicated in the paper, what γ was used in the simulations and more importantly, what γ is pre-programmed into the *stepR* package?

Secondly, whilst the presentation in the paper is restricted to SMUCE, I believe the concept of confidence intervals for the changepoint locations can be extended to other search algorithms, in particular PELT (Killick et al., 2012). The confidence intervals for the changepoint locations in SMUCE are constructed by considering all sets of solutions where the test statistic $T_n(Y, \vartheta)$ is less than the threshold q (equation 5). In contrast, the PELT algorithm keeps all changepoint locations that are within the penalty value of the maximum in order to prune the search. The same idea of confidence could be applied to these changepoint locations as their test statistics are close to the maximum and are thus also likely candidates for a changepoint. Obviously the key question is what theory is there to support this criteria as a way of constructing a confidence interval?

Initial simulations using this method show desired properties such as:

- as you increase the penalty (i.e. increase your expected confidence in a changepoint) you become more uncertain about the proposed locations;
- for a given penalty value, the larger the change, the smaller the confidence interval;
- the coverage doesn't depend on the size of the change.



(a) PELT (b) SMUCE

Figure 1: Simulated data with two changes in mean at 100 and 200 with (a) PELT and (b) SMUCE changepoints estimates and confidence intervals.

- the longer the interval between changes, the smaller the confidence interval;

Figure 1 gives an example of this last point using the *changepoint* package (Killick and Eckley, 2014). In the simulations the coverage of both SMUCE and PELT was larger than 99% using default values. The theory behind this conjecture needs to be thoroughly treated but at least empirically this seems promising.

References

Hotz, T. and Sieling, H. (2013). *stepR: Fit Step-Functions*. R package version 1.0.

Killick, R. and Eckley, I. A. (2014). changepoint: An R package for changepoint analysis. *Journal of Statistical Software*. To appear.

Killick, R., Fearnhead, P., and Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *JASA*, 107(500):1590 − 1598.