

## Referee Report for JRSS B

This paper provides a mathematically rigorous alternative to SiZer for time series. The main contribution is in proposing several ingenious modifications, proving that (under suitable conditions) the multiple testing quantiles computed for the independent case are still valid in the dependent case, and providing a new variance estimator. I believe that after a suitable revision this paper should be published somewhere.

In the following comments the page and line numbers are taken from the pdf sent for review.

1. **Section 3.2:** The authors recommend computing the quantiles for the independent Gaussian case by simulation. This suggestion is already in the original SiZer paper (Chaudhuri and Marron, 1999). However in the late 1990s computing power was not sufficient to make this suggestion feasible. This led to the use of approximation such as in Hannig and Marron (2006). I would like to ask how does the simulation based quantile compare to the approximation in Hannig and Marron (2006).
2. **Page 12, line 15:** What is random here? After a spending some time I believe that it is the  $\Pi_T$  but on first reading I thought  $E_T$ s are non-random. Please explain these various objects better.
3. **Page 18, line 52:** Please remove the speculative statements about what can be shown unless you actually show it in this paper.
4. **Section 4.1:** This section does not contain any truly new material and should be removed.
5. **Section 5.2:** I understand that you are doing comparisons to SiZer out of the box. However, some of the comparison might not be quite fair. SiZer is adjusting multiplicity "row-wise" while the proposed method is attempting a global multiple control. What would happen if your  $G_T$  only focused on one scale?
6. **Page 25, line 1-26:** I do not quite understand this figure. Would it be possible to rather reproduce the colorful SiZer figures that show the results of the test at various scales and locations? Also you should use several different signals. I believe that a single relatively large bump is not sufficient test bed. A good collection of signals can be found in

Donoho and Johnstone (1995). Also, would Hannig *et al.* (2013) be helpful in comparing the results?

7. **Page 31, line 1-39:** Can you plot the SiZer results on this data?

## References

- Chaudhuri, P. and Marron, J. S. (1999) SiZer for exploration of structures in curves. *Journal of the American Statistical Association*, **94**, 807–823.
- Donoho, D. L. and Johnstone, I. M. (1995) Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, **90**, 1200–1224.
- Hannig, J., Lee, T. C. and Park, C. (2013) Metrics for SiZer map comparison. *Stat*, **2**, 49–60.
- Hannig, J. and Marron, J. S. (2006) Advanced Distribution Theory for SiZer. *Journal of the American Statistical Association*, **101**, 484–499.