

# scanstatistics: space-time anomaly detection using scan statistics

Benjamin Allévius<sup>1</sup>

<sup>1</sup> Department of Mathematics, Stockholm University

DOI: [10.21105/joss.00515](https://doi.org/10.21105/joss.00515)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 22 September 2017

Published: 09 May 2018

## Licence

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

## Summary

The R package `scanstatistics` enables the detection of anomalous space-time clusters using the scan statistics methodology. Scan statistics are commonly applied in disease surveillance, where they are used to detect disease outbreaks as they emerge locally. In this setting, cases of a given disease are recorded continuously across a country, and are then aggregated spatially to (say) district level, and temporally to (say) weekly counts. Scan statistics accomplish the detection task by searching the recent records of clusters of neighboring districts for patterns that seem anomalous given either past counts or the counts outside the cluster currently searched.

The `scanstatistics` package implements several scan statistics, making it a partially overlapping complement to existing scan statistic software such as `SaTScan`. For example, the conditional Poisson (Kulldorff 2001) and space-time permutation (Kulldorff et al. 2005) scan statistics are available in both `SaTScan` and `scanstatistics`, while only the latter implements scan statistics for zero-inflated data (Allévius and Höhle 2018, in press), count data with overdispersion (Tango, Takahashi, and Kohriyama 2011), an unconditional (expectation-based) Poisson scan statistic (Neill et al. 2005), and a Bayesian scan statistic (Neill, Moore, and Cooper 2006).

The R package `scanstatistics` is available on [CRAN](#) and its source code is available on [GitHub](#).

## References

- Allévius, Benjamin, and Michael Höhle. 2018, in press. “An unconditional space-time scan statistic for ZIP-distributed data.” *Scandinavian Journal of Statistics*.
- Kulldorff, Martin. 2001. “Prospective time periodic geographical disease surveillance using a scan statistic.” *Journal of the Royal Statistical Society Series a-Statistics in Society* 164:61–72. <https://doi.org/10.1111/1467-985X.00186>.
- Kulldorff, Martin, Richard Heffernan, Jessica Hartman, Renato M. Assunção, and Farzad Mostashari. 2005. “A space-time permutation scan statistic for disease outbreak detection.” *PLoS Medicine* 2 (3):0216–24. <https://doi.org/10.1371/journal.pmed.0020059>.
- Neill, Daniel B., Andrew W. Moore, and Gregory F. Cooper. 2006. “A Bayesian Spatial Scan Statistic.” In *Advances in Neural Information Processing Systems*, edited by Y. Weiss, P. B. Schönkopf, and J. C. Platt, 18:1003–10. MIT Press.
- Neill, Daniel B., Andrew W. Moore, Maheshkumar Sabhnani, and Kenny Daniel. 2005. “Detection of emerging space-time clusters.” In *Proceeding of the Eleventh Acm Sigkdd International Conference on Knowledge Discovery in Data Mining - Kdd '05*, 218. New York, New York, USA: ACM Press. <https://doi.org/10.1145/1081870.1081897>.

Tango, Toshiro, Kunihiro Takahashi, and Kazuaki Kohriyama. 2011. “A Space-Time Scan Statistic for Detecting Emerging Outbreaks.” *Biometrics* 67 (1):106–15. <https://doi.org/10.1111/j.1541-0420.2010.01412.x>.