RESCIENCE C

Replication / Computational Neuroscience

[Re] Spatial constraints underlying the retinal mosaics of two types of horizontal cells in cat and macaque

Stephen J. Eglen[1, ID]

[1]Department of Applied Mathematics and Theoretical Physics, Cambridge, UK

Edited by
Nicolas P. Rougier ID

Reviewed by
Stephan Grein ID

Received
30 April 2020

Published
31 August 2021

DOI
10.5281/zenodo.5347786

Summary

This paper reports the successful reproduction of the results in a paper I co-authored in 2008[1]. A postprint of this paper is available at https://arxiv.org/abs/1208.0986. (Open access was not mandated by the funders of this research at the time of publication.) All the code has been uploaded to https://github.com/sje30/rescience-hor along with the Rmarkdown sources for this report.

1 Historical context

1.1 Scientific question

To investigate a problem in neurobiology concerning the relative spatial positioning of two types of neuron in the retina. We used three example datasets (pretty much all that was available at the time) to investigate this problem using spatial statistics and simulations of spatial patterning. These datasets were not available in digital format, but the images[2,3] were scanned in and digitised (each cell represented as an x,y location) by me as part of the work for the original paper. These datasets were not publicly shared before.

1.2 Computational context

Most of the code required for the 2008 paper had been developed by myself since about 2000 working on other related problems. The major innovation here was in the biological problem.

One of the key reasons I switched to R as a programming environment in the late 1990s was because of the strong support it had for spatial statistics. In particular the SPLANCS package[4] in R has been a key package in most of my research.

1.3 Original source code

The original source code was not published or publicly archived before. I keep good backups though, and in this case the projects were still in my home directory (copied from machine to machine over the years). No license was written systematically to cover the code I wrote, although the R packages are usually released under GPL or MIT.

Correspondence should be addressed to Stephen J. Eglen (sje30@cam.ac.uk)
The authors have declared that no competing interests exist.
Code is available at https://github.com/sje30/rescience-hor. – SWH swh:1:dir:6088b7312865b39a4a5f5af78556bf6da5fa6409.
Open peer review is available at https://github.com/ReScience/submissions/issues/34.

As a side-effect of producing this article, the following R packages of mine were made available:

1. sjedmin: dmin model for spatial point patterns. https://github.com/sje30/sjedmin
2. sjedist: unified calculation of spatial statistics. https://github.com/sje30/sjedist
3. sjevor: Wrapper around public domain code for Voronoi tessellation https://github.com/sje30/sjevor

The Voronoi tessellation was public domain C code provided by Steve Fortune[5]. (This code was not required for the current project, but was implicitly required for the parent package, sjedist.)

1.4 Retrieval of the software

This was relatively straightforward, as the software was all on my hard drive. One complication was that until a few years ago, I had been using RCS as my primary version control software. As I wished to release the packages onto github, I converted the history from RCS into Git using the tool `https://github.com/Oblomov/rcs-fast-export`.
I anticipated that loading the data might be tricky, as I had used R's internal Rda format to efficiently store the data files (they are not large, so there was no need to be concerned about efficiency). Given that this format could in principle change over time, this might have been a problem. However, R 4.0.0 was able to read in the file `bivariate_mosaics.Rda`.

2 Execution

The code ran fine in today's version of R (4.0.0); see end of document for the full list of R dependencies. I did not need to load any older versions of libraries. The code was tested on both my laptop (ThinkPad T480s running Manjaro linux) and on https://mybinder.org (see the end of this document for mybinder details.) The simulations themselves only take a couple of minutes to run. If recompiling on binder, it takes a few more minutes to install all the packages and initialise the tex distribution.
However, although the code worked fine, it took me about 20 minutes to read through the old scripts and work out what to do, as I had not provided a Makefile or similar.
Once the code was running, I only wished to make cosmetic updates to the project to modernise it. The most significant change was converting the format for figures from postscript to pdf as there is more support for the latter. The three figures that were generated are reproduced in this paper and can be compared against the figures from the original paper.
A Makefile is provided to help regenerate the simulation results and the figures. *make clean* will return the repository to a clean state. *make all* will first run the simulations and store the results in several *.Rda* files. There are then several R scripts to generate the figures from the corresponding *.Rda* files. Finally, after the figure pdfs have been generated, this document is compiled from its Rmarkdown source.

2.1 Retrieval of the software

This was relatively straightforward, as the software was all on my hard drive. One complication was that until a few years ago, I had been using RCS as my primary version control software. As I wished to release the packages onto github, I converted the history from RCS into Git using the tool `https://github.com/Oblomov/rcs-fast-export`.
I anticipated that loading the data might be tricky, as I had used R's internal Rda format to efficiently store the data files (they are not large, so there was no need to be

concerned about efficiency). Given that this format could in principle change over time, this might have been a problem. However, R 4.0.0 was able to read in the file `bivariate_mosaics.Rda`.

3 Execution

The code ran fine in today's version of R (4.0.0); see end of document for the full list of R dependencies. I did not need to load any older versions of libraries. The code was tested on both my laptop (ThinkPad T480s running Manjaro linux) and on https://mybinder.org (see the end of this document for mybinder details.) The simulations themselves only take a couple of minutes to run. If recompiling on binder, it takes a few more minutes to install all the packages and initialise the tex distribution.

However, although the code worked fine, it took me about 20 minutes to read through the old scripts and work out what to do, as I had not provided a Makefile or similar.

Once the code was running, I only wished to make cosmetic updates to the project to modernise it. The most significant change was converting the format for figures from postscript to pdf as there is more support for the latter. The three figures that were generated are reproduced in this paper and can be compared against the figures from the original paper.

A Makefile is provided to help regenerate the simulation results and the figures. *make clean* will return the repository to a clean state. *make all* will first run the simulations and store the results in several .*Rda* files. There are then several R scripts to generate the figures from the corresponding .*Rda* files. Finally, after the figure pdfs have been generated, this document is compiled from its Rmarkdown source.

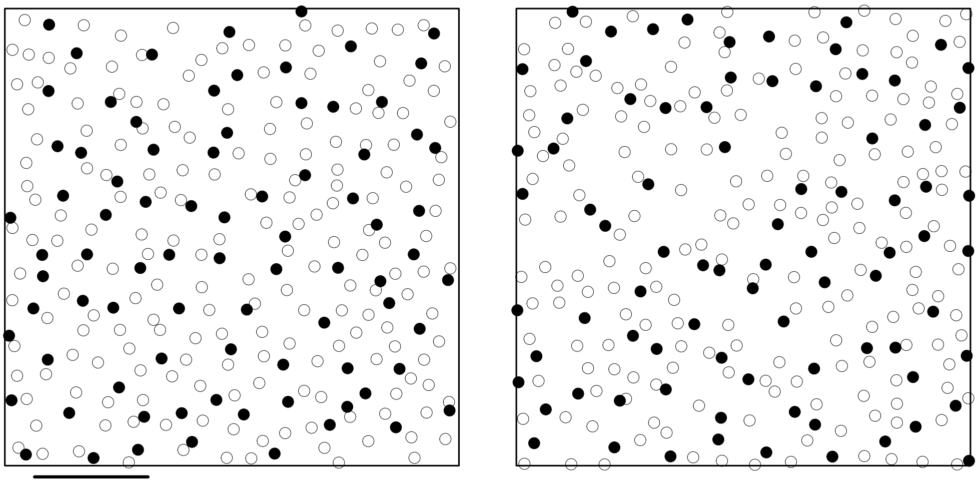
3.1 Reproduction of results

The results generated closely matched those that were presented in the original paper. As the simulation methods are stochastic (and no seed was set for reproducible random number generation), the P values reported in Figure 2 do not match those originally reported, nor does the simulation field (right hand side of Figure 1) match. However, the statistical properties of the models still match those of the experimental data.

3.2 Could someone else reproduce the work?

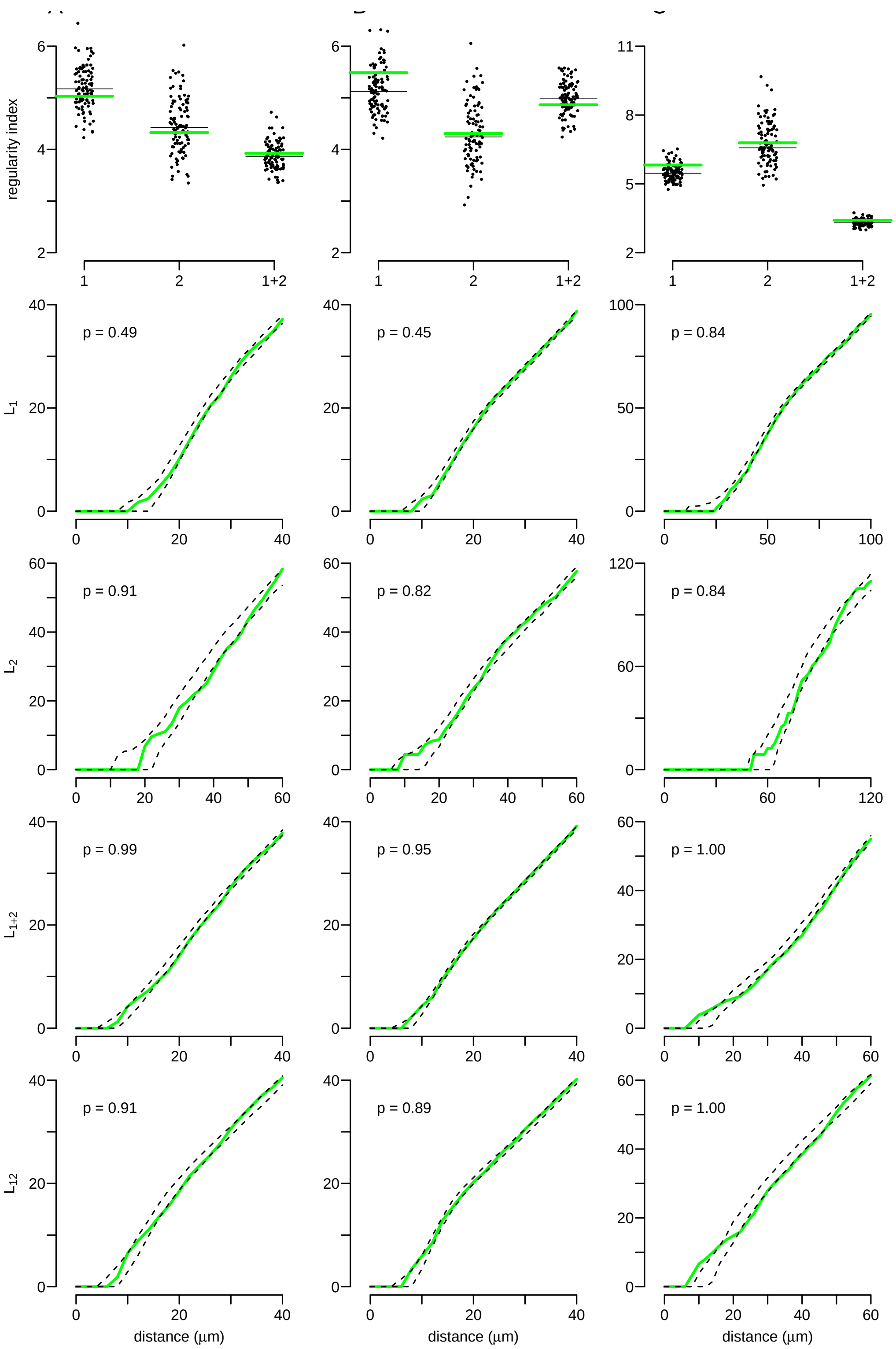
If I had simply given the R code and packages to someone else, the lack of documentation or Makefile would have meant that this person would have needed to study the code carefully to work out what files were generated. I would have hoped though that someone fluent with R would have made sense of the code itself, and all of the parameters for the simulations were included in the source files.

4 Reproduction of Figure 1

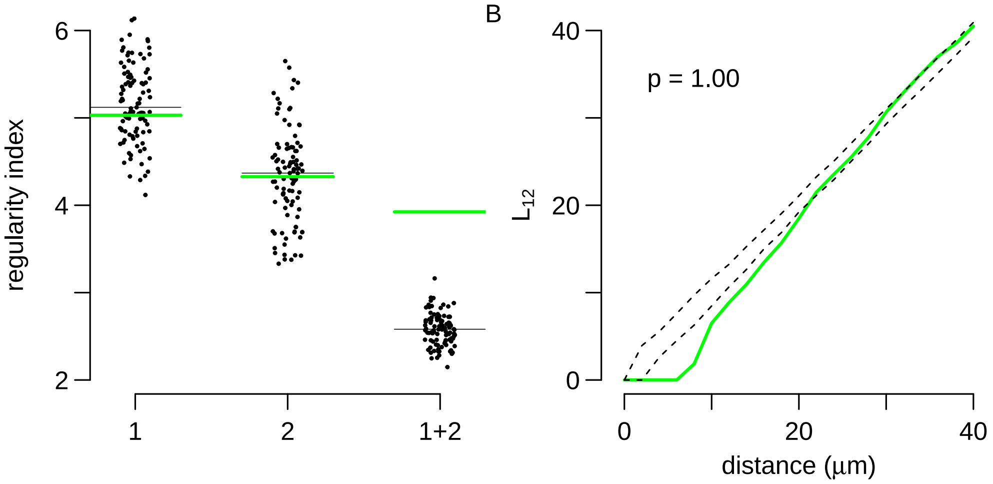


We also made a supplementary figure of all 3 fields.

5 Reproduction of Figure 2



6 Reproduction of Figure 3



7 Reproduction of Table 2

Key column from table 2 is the rejects column which is reproduced in Table 1. The other columns of this table were determined statically from the data sets, so were not recalculated here. (Had I known more about generating tables in R at the time, I think my R code should have created this entire table.)

field	rejects	n_{pairs}	rejects / n_{pairs}
C	242 ± 10	25500	0.01
A	937 ± 36	15334	0.06
B	7202 ± 300	17716	0.41

Table 1. Reproduction of table 2 results

8 Conclusions

It was relatively easy to recompute the results in this report, because of the stability of the R environment and its packaging system. Even its binary data format (.Rda) for storing objects was reusable over a 10 year timescale. The largest challenge was the user issue of remembering what I had done to run the analysis and generate the results. Using mybinder.org was useful to iron out minor issues that others might face trying to run the software on their machine. In its current state, the repository requires R and make – if make is unavailable, the scripts would need to be run by hand, which is not onerous, but writing them out explicitly might help those unfamiliar with Make syntax.

References

1. S. J. Eglen and J. C. T. Wong. "Spatial constraints underlying the retinal mosaics of two types of horizontal cells in cat and macaque." In: **Vis. Neurosci.** 25.2 (Mar. 2008), pp. 209–214. DOI: 10.1017/S0952523808080176.
2. H. Wässle, L. Peichl, and B. B. Boycott. "Topography of horizontal cells in the retina of the domestic cat." In: **Proc. R. Soc. Lond. B Biol. Sci.** 203.1152 (Dec. 1978), pp. 269–291.

3. H. Wässle, D. M. Dacey, T. Haun, S. Haverkamp, U. Grünert, and B. B. Boycott. "The mosaic of horizontal cells in the macaque monkey retina: with a comment on biplexiform ganglion cells." en. In: **Vis. Neurosci.** 17.4 (July 2000), pp. 591–608. DOI: 10.1017/s0952523800174097.
4. B. Rowlingson and P. Diggle. **splancs: Spatial and Space-Time Point Pattern Analysis**. R package version 2.01-40. 2017. URL: https://CRAN.R-project.org/package=splancs.
5. S. Fortune. "A sweepline algorithm for Voronoi diagrams." In: **Algorithmica** 2 (1987), pp. 153–172.