

# [Re] Chaos in a long-term experiment with a plankton community

Owen Petchey<sup>1</sup>, Marco Plebani<sup>1</sup>, and Frank Pennekamp<sup>1</sup>

<sup>1</sup> Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

[owen.petchey@ieu.uzh.ch](mailto:owen.petchey@ieu.uzh.ch)

## Editor

Name Surname

## Reviewers

Name Surname

Name Surname

Received Sep, 1, 2015

Accepted Sep, 1, 2015

Published Sep, 1, 2015

Licence [CC-BY](#)

## Competing Interests:

The authors have declared that no competing interests exist.

 [Article repository](#)

 [Code repository](#)

 [Data repository](#)

## A reference implementation of

→ Benincà, E., Huisman, J., Heerkloss, R., Jöhnk, K.D., Branco, P., Van Nes, E.H., Scheffer, M. & Ellner, S.P. (2008) Chaos in a long-term experiment with a plankton community. *Nature*, 451, 822–825 DOI: 10.1038/nature06512

## Draft article

### Introduction

The original paper describes analyses of fluctuations in the abundance of organisms in a plankton community derived from the Baltic Sea, housed in a laboratory environment. The length of the time series (samples every few days for 2,300 days) allowed for analyses revealing that the observed dynamics exhibited characteristics consistent with chaos produced by non-linear species interactions. The article concludes that stability is not required for persistence of complex food webs, and that long-term prediction of abundances may be fundamentally impossible. The demonstration of chaotic dynamics and limited forecast horizons (*sensu* Petchey et al. [4]) is important in the field of ecology, since the ability to predict dynamics is an open question with considerable applied importance (Petchey et al. [4] Mouquet et al. [3].)

### Methods

This reproduction started with the raw data (source given below) and used information from the original paper, [the supplement to the original paper](#), and communications with Elisa Benincà, who provided for comparison digitised data from the original article, and Stephen Ellner, who provided code and data used to produce results in the original paper. Use of this data and code in this reproduction is indicated below.

### Scope of the reproduction

An attempt was made to reproduce the majority of the results in the original article. Instances where we did not attempt to reproduce a result are detailed below.

## The data

The data are available as an Excel file supplement to [an Ecology Letters publication](#) (Benincà et al. [1]). The Excel file contains several datasheets. Two were particularly important, as they were the source of the raw data (one contains original species abundances, the other the nutrient concentrations). We saved these two datasheets, without any alteration, as comma separated value (csv) text files. In the code associated with this reproduction, these data files are read from the associated github repository. Another datasheet contained transformed data (we also saved this as csv file, in order to use it in this reproduction). We also received a dataset direct from Steve Ellner, see below for details.

The original abundance dataset contains a duplicated date: 28/10/1996 (row 709 and 710 in excel sheet). We changed this (in R) to 26/10/1996, making the observation occur half way between the two surrounding dates. The Protozoa variable in the original dataset contains some numbers with a comma (instead of period) as the decimal separator; these were changed to periods.

## Reproduction environment

The R language and environment for statistical computing and graphics (R Core Team [5]) was used to make the reproduction. Additional R packages required are specified in the code associated with this reproduction. The code for making the analyses and figures presented in this reproduction resides in an R markdown document, as well as a source file containing some required functions. Some of the code takes several minutes to run, so an intermediate data file is provided with results from this code.

## Results

### Population dynamics

The reproduced populations dynamics (figure 1) were very similar to those in figure 1b-g of the original publication. As most analyses were performed on fourth root transformed values, we plot these, rather than raw abundances with a y-axis break as in the original article.

### Data transformations

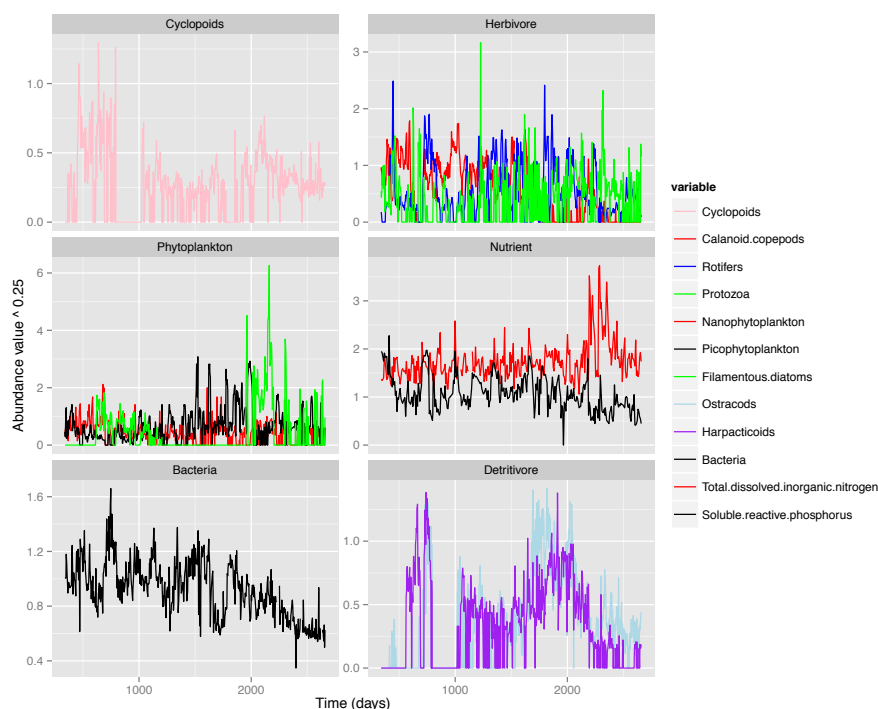
The original publication states that long sequences of zeros were removed from time series prior to transformation and further analysis. We removed zeros by matching the removals to those in the transformed version of the data in the published Excel file mentioned above. Subsequent transformation steps were:

1. Interpolation to create equally spaced observations in time series.
2. Fourth root transformation.
3. Detrending of five of the time series.
4. Rescaling to zero mean and unit standard deviation.

The reproduced transformed data closely matched the published transformed data (figure 2).

### Correlations among species abundances

Correlations among species abundances presented in Table 1 of the original article closely matched our reproduced correlations, calculated from the transformed data with zeros removed (figure 3). Deviations between the original and reproduced correlations are relatively small and infrequent.



**Figure 1:** Observed population dynamics.

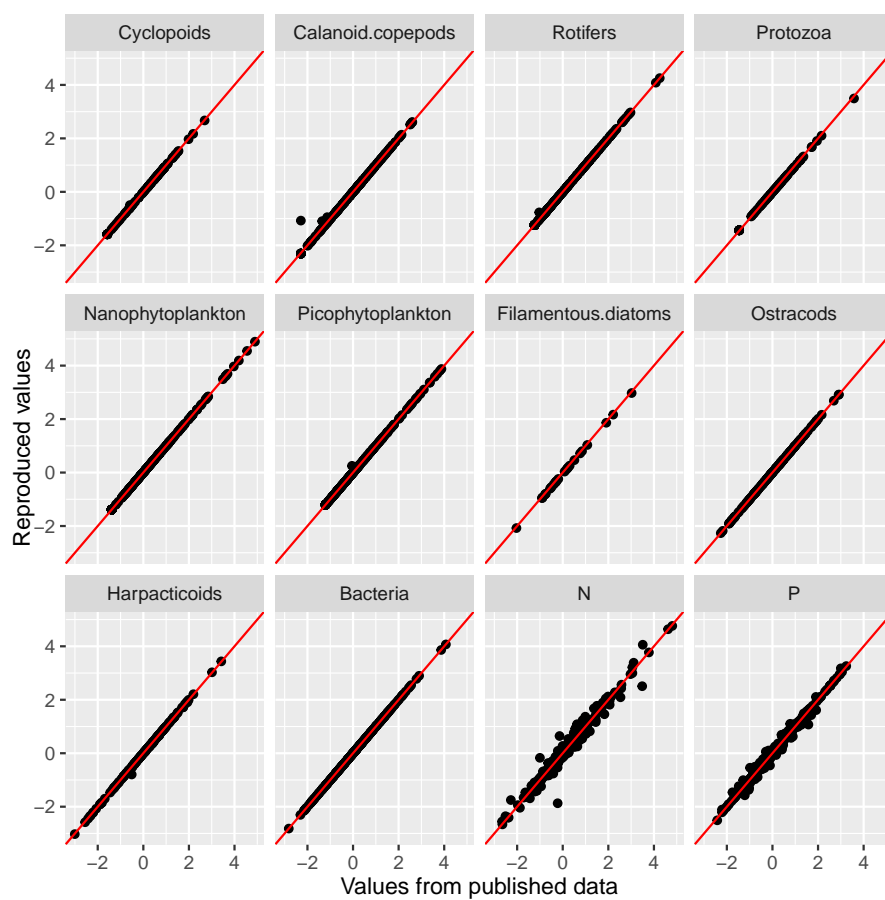
Highlighted in the text of the original paper were: negative correlations of picophytoplankton with protozoa, and of nanophytoplankton both with rotifers and calanoid copepods, positive correlation of picophytoplankton with calanoid copepods, negative correlation between bacteria and ostracods, and positive correlation between bacteria and phosphorus. All of these correlations were at least qualitatively reproduced.

## Spectral analyses

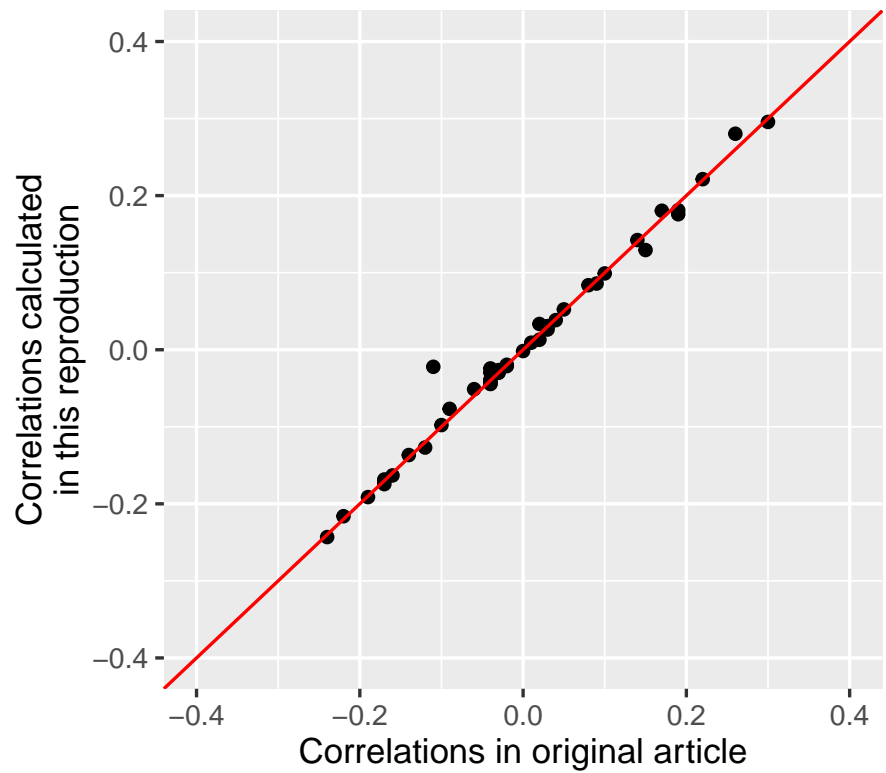
Spectral analyses in the original paper were presented graphically in figures S3 (raw spectrograms) and S4 (Welch periodograms). The original article states: “fluctuations covered a range of different periodicities”, and “picophytoplankton, rotifers and calanoid copepods seemed to fluctuate predominantly with a periodicity of about 30 days.” Our reproduced spectra (not shown here, but code provided) were a reasonably close match to the spectra in the original article.

## Lyapunov exponents by direct method

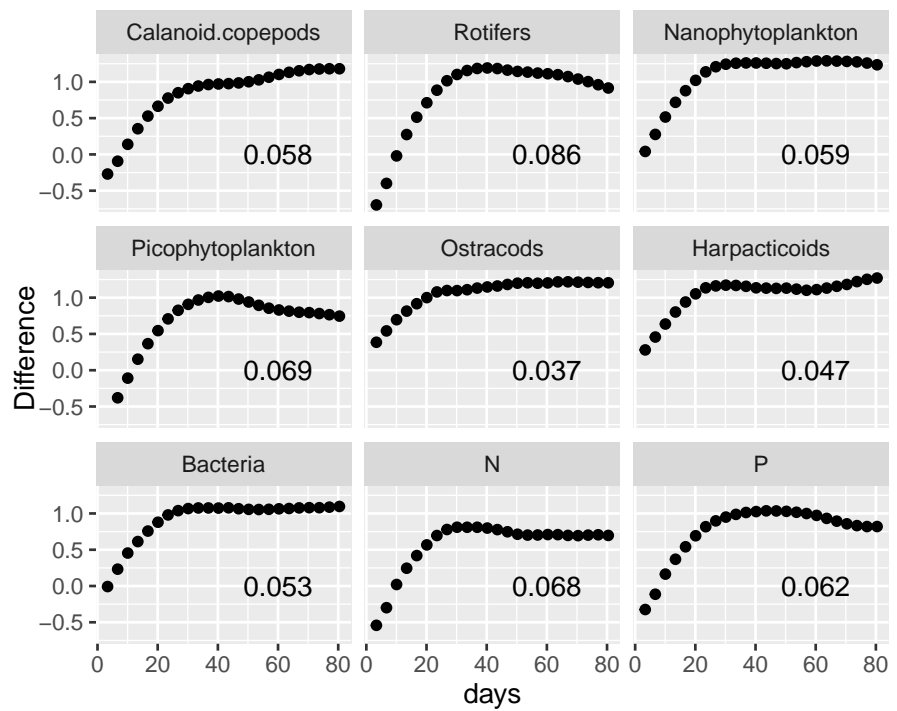
The original article states: “the distance between initially nearby trajectories increased over time, and reached a plateau after about 20–30 days”. The reproduced results (figure 4) are consistent with this statement. The original article also stated that the analyses “yielded significantly positive Lyapunov exponents of strikingly similar value for all species (Fig. 3; mean exponent = 0.057 per day, s.d. = 0.005 per day, n = 9)”. Reproduced exponents had very similar mean value, but had larger standard deviation (mean = 0.060 and s.d. = 0.014), resulting from quantitative differences in reproduced divergence rates (figure 4) and therefore Lyapunov exponents (figure 5).



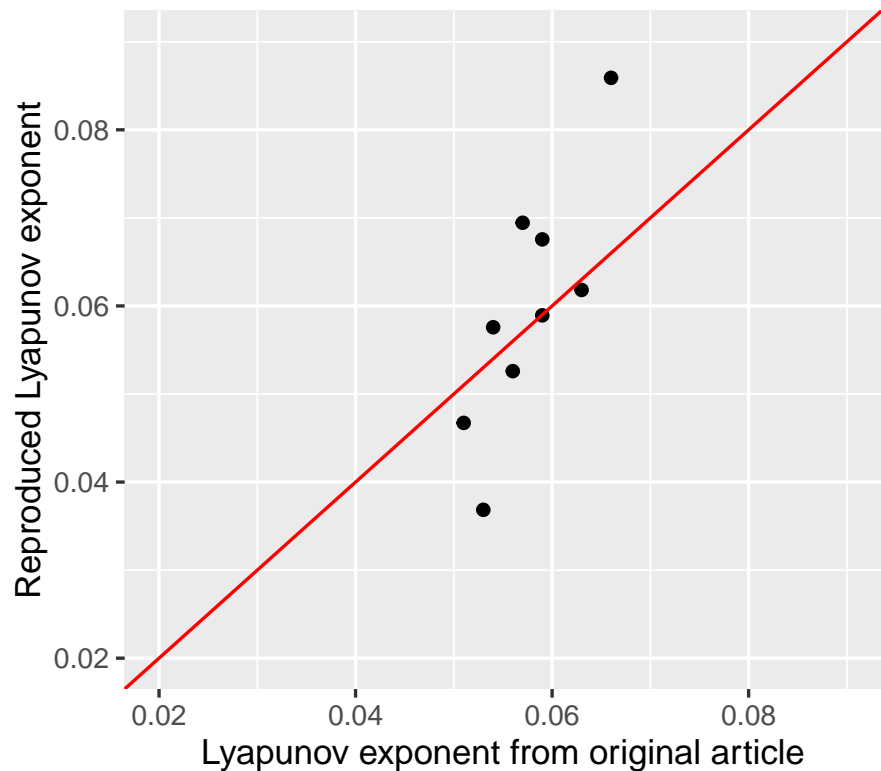
**Figure 2:** Comparison of published transformed data with reproduced transformed data. The red line is the 1:1 line.



**Figure 3:** Comparison of calculated correlations among species abundances in the original article and this reproduction.



**Figure 4:** Reproduced divergence rates and Lyapunov exponents (equivalent to figure 3 in the original article).



**Figure 5:** Comparison of original and reproduced directly estimated Lyapunov exponents.

### Lyapunov exponents by indirect method

The data received directly from Stephen Ellner (file ‘interp\_short\_allsystem\_newnames.csv’ in the reproduction repository) was interpolated, but without zeros removed. The reproduced interpolated data, without zeros removed, matched closely this data (figure 6).

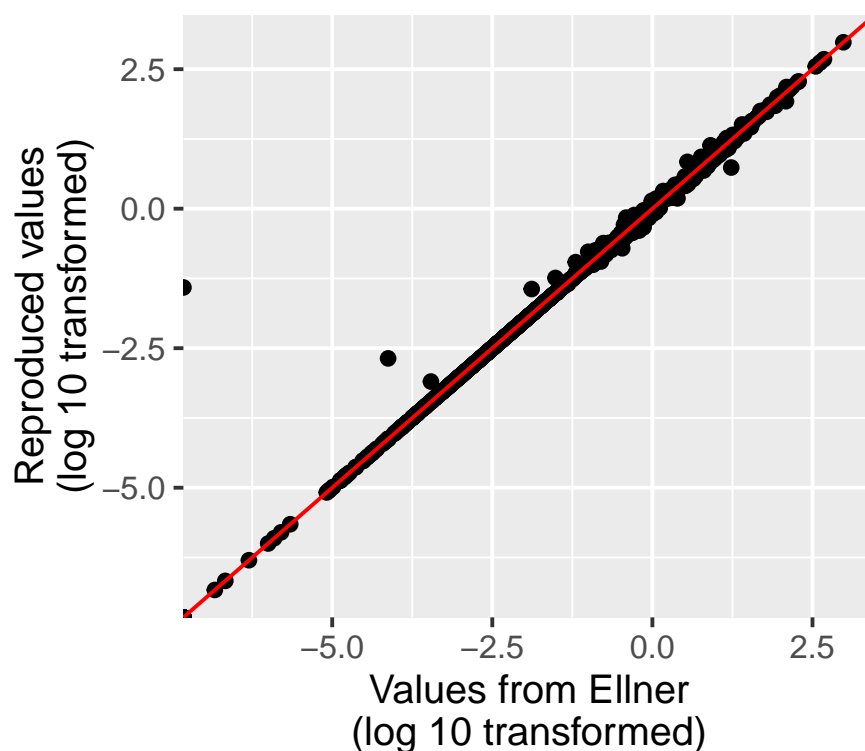
The original paper reported a global Lyapunov exponent calculated via two modelling approaches (neural network and generalised additive models [GAMs]). Only the GAM approach was reproduced, with the assistance of code donated by Stephen Ellner. The original article obtained a global Lyapunov exponent of 0.08 per day. The reproduced value was 0.04. We did not reproduce the bootstrapping used to give confidence intervals around this estimate.

### Predictability decay

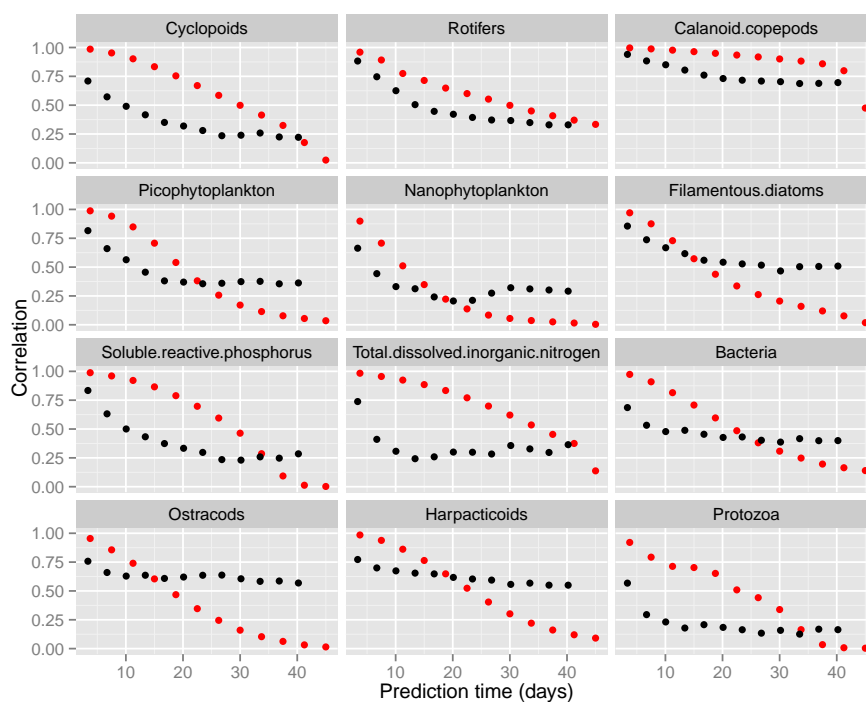
The article stated: “For short-term forecasts of only a few days, most species had a high predictability of  $R^2 = 0.70 - 0.90$  (Fig. 2). However, the predictability of the species was much reduced when prediction times were extended to 15–30days.” The reproduced predictabilities, which were calculated from the GAMs, were consistent with these qualitative statements, though were quantitatively different (7). We did not reproduce the predictability estimates for linear models.

## Conclusion

The reproduced results were qualitatively identical to those in the original article, and therefore support the conclusion of chaotic dynamics. For example, all Lyapunov



**Figure 6:** Comparison of interpolated data provided by Ellner with reproduced interpolated data (no removal of zeros). The red line is the 1:1 line.



**Figure 7:** Predictability (correlation between predicted and observed abundances) and prediction distance (days) (figure 2 in the original article). Reproduced data in red, and data from original publication in black.

exponents estimated by direct method were positive, as in the original article. Quantitative differences may have resulted from difference in algorithms used. For example, the original used the [Tisean software](#) to calculate Lyapunov exponents. As this was available from CRAN [until mid 2014](#) and since it is a bit less well integrated with R, we instead used the `tseriesChaos` package (Fabio Di Narzo [2]), which in any case was largely inspired by the TISEAN project. In addition, there may have been some difference in algorithm parameters, as not all parameters required by the function we used were reported in the original ms.

Considerable quantitative difference in predictability decay (figure 7) remain unexplained, though could result from original analyses using neural networks, and the reproduction using GAMs.

In conclusion, this reproduction supports the general scientific conclusions of the original article, but also shows how difficult can be an accurate quantitative reproduction, even in the presence of the extensive methodological details provided alongside the original article.

## Acknowledgements

This reproduction was made as part of the Reproducible Research in Ecology, Evolution, Behaviour, and Environmental Studies (RREEBES) Course, lead by Owen Petchey at the University of Zurich. More information about the course [here](#) on github. Many thanks to Elisa Benincà and Stephen Ellner for assistance, and to Stephen Ellner for providing unpublished data and code.

## References

- [1] Elisa Benincà et al. "Coupled predator-prey oscillations in a chaotic food web." In: *Ecology letters* 12.12 (Dec. 2009), pp. 1367–78. ISSN: 1461-0248. DOI: [10.1111/j.1461-0248.2009.01391.x](#). URL: <http://www.ncbi.nlm.nih.gov/pubmed/19845726>.
- [2] Antonio Fabio Di Narzo. "tseriesChaos: Analysis of nonlinear time series". In: (2013). R package version 0.1-13. URL: <http://CRAN.R-project.org/package=tseriesChaos>.
- [3] Nicolas Mouquet et al. "Predictive Ecology In A Changing World". In: *Journal of Applied Ecology* (2015), n/a–n/a. ISSN: 00218901. DOI: [10.1111/1365-2664.12482](#). URL: <http://doi.wiley.com/10.1111/1365-2664.12482>.
- [4] Owen L. Petchey et al. "The ecological forecast horizon, and examples of its uses and determinants". In: *Ecology Letters* 18 (2015), pp. 597–611. ISSN: 1461023X. DOI: [10.1111/ele.12443](#). URL: <http://doi.wiley.com/10.1111/ele.12443>.
- [5] R Core Team. "R: A Language and Environment for Statistical Computing". In: (2015). URL: <https://www.R-project.org/>.