

The Retirement of R Packages with Many Reverse Dependencies

Edzer Pebesma, Roger Bivand

Overview

- ▶ some history of OSGEO and r-spatial
- ▶ why packages retire
- ▶ how we did it
- ▶ lessons learned / future recommendations

Some history of OSGEO

This community has settled on one standard and three key C++ libraries:

- ▶ *Simple Feature Access*, an open standard to define point-, line- and polygon-geometries and multi-forms of those
- ▶ **PROJ** a C++ library for cartographic projections and datum transformations
- ▶ **GEOS** for geometry measures, predicates and transformers of simple feature geometries
- ▶ **GDAL** a library to read and write vector or raster data to/from more than 300 different file formats, databases, or web services

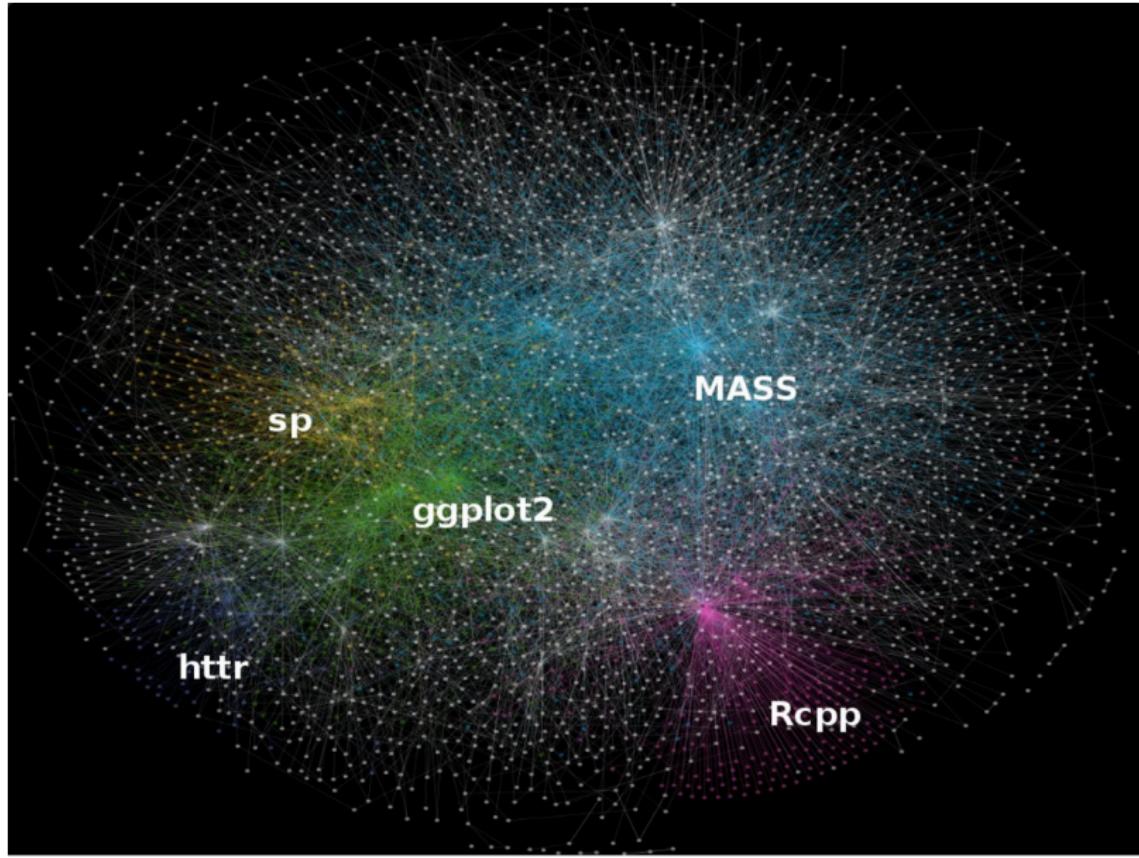
R, Python, Julia, QGIS, ArcGIS, PostGIS, etc all use this.

Some history of r-spatial (1)

- ▶ Spatial R packages have been there from the start, and in S-Plus (e.g., MASS had `spatial`)
- ▶ Since 2003 a community converged on using a set of classes for vector and raster data, using packages `sp` and `raster`
- ▶ `rgdal` (2003) interfaced GDAL and PROJ
- ▶ `rgeos` (2011) interfaced GEOS
- ▶ `maptools` was a collection of various useful things
- ▶ older, legacy packages include e.g. `maps` and `mapproj`
- ▶ `sp` predates simple features, and mirrors shapefiles

sp: Classes and Methods for Spatial Data

Classes and methods for spatial data; the classes document where the spatial location information resides, for 2D or 3D data. Utility functions are provided, e.g. for plotting data as maps, spatial selection, as well as methods for retrieving coordinates, for subsetting, print, summary, etc. From this section, 'sgdal', 'maptools', and 'rgdal' are no longer used at all, see <https://r-spatial.org/reps/2023/05/15/evolution-4.html> for details.



Colin Gillespie @csgillespie · Apr 18

Updated `#rstats` dependencies map of CRAN (original by @RevoAndrie see blog.revolutionanalytics.com/2015/07/the-new-dependencies-map-of-cran/)

pic.twitter.com/4hXpnubQ4A



7



23

Reverse dependencies Sep 2021

	rgdal	rgeos	maptools	sp*
Direct “strong”	213	140	93	528
Recursive “strong”	265	190	641	1552
Direct “most”	358	225	180	638
Recursive “most”	15702	15702	15702	15702

Some history of r-spatial (2)

- ▶ CRAN binary packages are static builds, meaning they include upstream libraries (GEOS, PROJ, GDAL)
- ▶ CRAN check and build systems contain these libraries, they are *not* vendored in the respective R packages
- ▶ the CRAN team needs to regularly update, or rebuild these
- ▶ this involves communication (both ways), attention, and problem solving skills

Some history of r-spatial (3)

- ▶ sf came out in 2016 and supports simple features in R
- ▶ it uses S3 rather than S4, integrates with tidyverse and ggplot2
- ▶ stars (2018) provided classes for the raster side and data cubes
- ▶ terra (2020) uses S4 classes for raster and vector, and succeeds raster

Why packages retire

- ▶ Maintaining packages can be a lot of work
- ▶ Packages can get superseded by better solutions
- ▶ Maintainers retire!

Why rgdal, rgeos and maptools retired

- ▶ Roger Bivand, their maintainer, retired
- ▶ The code is hard to read
 - ▶ uses `.Call()`, rather than `Rcpp` or similar
 - ▶ many conditional branches for upstream changes
- ▶ simple features in R (`sf`) simplifies life, a lot.
- ▶ raster data analysis moved to `stars` (2018) and `terra` (2020)
- ▶ R Consortium funded an *evolution* project (Sept 21 call)

sp evolution status: examples of migration from retiring packages

May 15, 2023 • Roger Bivand, Edzer Pebesma

- [sp evolution status](#)
 - Use of `sp` affected functions and methods
 - Possible difficulties
 - Why not `terra` ?

[\[view raw Rmd\]](#)

Summary:

This is the fourth report on the R-spatial evolution project and is addressed to maintainers of packages and workflows using `sp` classes, methods and functions. The project involves the retirement (archiving) of `rgdal`, `rgeos` and `maptools` during 2023. The [first report](#) set out the main goals of the project. The [second report](#) covered progress so far, steps already taken, and those remaining to be accomplished. The [third](#) gave detailed guidance for maintainers of packages using the retiring packages.

From June 2023, the internal evolution status setting of `sp` will be changed from "business as usual" to "use `sf` instead of `rgdal` and `rgeos`". Packages depending on `sp` may need to add `sf` to their weak dependencies, and to monitor any changes in output.

The final step will occur during October 2023, in five months, `rgdal`, `rgeos` and `maptools` will be archived on CRAN, and packages with strong dependencies on the retiring packages must be either upgraded to use `sf`, `terra` or other alternatives or work-arounds by or before that time. Making all required changes in the period from now to the June `sp` change will mean just one round of adaptations rather than two rounds.

This spreadsheet lists methods and functions in retiring packages and `sp` found by `pkgapi`. They are listed by function name as used by packages, and the analogous list by package may be found of functions is [here](#).

Keynote: R Spatial



KEYNOTE SPEAKER
EDZER PEBESMA
Professor, Institute for Geoinformatics,
University of Münster
R Spatial

Monday July 5, 2021
3:30pm – 4:30pm UTC
<https://www.r-spatial.org>

The logo for the R conference, featuring a stylized blue 'R' inside a circle.

1

Edzər Pebesma @edzervebesma · Jun 23

The abstract is now online too. It says "... when rgdal and rgeos retire in 2024." Anyone volunteering to take over maintenance from @RogerBivand ? <https://r-spatial.org/contributor/2021/index...>

3

 You Retweeted

Refining to [@edturnerema](#)

I'm not sure that taking over maintenance is a sensible use of effort (add maptools to that list). The main condition for this will be transition from raster to terra, as raster uses rgdal heavily. If sf/stars and terra become the main workflow bases, things get simplified.

when

RGDAL and RGI's retire in 2024.



4:20 PM · Jun 23, 2021 · Twitter Web App



How?

- ▶ terra in 2018-20 took over rgdal dep from raster
- ▶ progress reported to r-spatial evolution blogs (for motivation)
- ▶ handling reverse dependencies (starting 2022 RC project)
 - ▶ package raster dropped dependency on rgdal and rgeos (Sept 2022), using terra instead
 - ▶ reverse dependency checks, emulating removal; monitor
 - ▶ **heavy** startup warnings in sp (worried people who were not affected!)
 - ▶ remove rgdal and rgeos dependencies of sp in favour of sf in june 2023

Activities Google Chrome

6 Jul 09:54

link.springer.com/article/10.1007/s10109-020-00336-0

SPINGER LINK

Find a journal Publish with us Track your research Search Log in Cart

Home > Journal of Geographical Systems > Article

Progress in the R Ecosystem for representing and handling spatial data

Original Article | Open access | Published: 16 October 2020
Volume 23, pages 515–546, (2021) Cite this article

Download PDF ↴ You have full access to this open access article

Roger S. Bivand 

9419 Accesses 101 Altmetric 1 Mention Explore all metrics →

Abstract

Twenty years have passed since Bivand and Gebhardt (*J Geogr Syst* 2(3):307–317, 2000. <https://doi.org/10.1007/PL00011460>) indicated that there was a good match between the then nascent open-source R programming language and environment and the needs of researchers analysing spatial data. Recalling the development of classes for spatial data presented in book form in Bivand et al. (*Applied spatial data analysis with R*, Springer, New York, 2008, *Applied spatial data analysis with R*, 2nd edn. Springer, New York, 2013), it is important to present the progress now occurring in representation of spatial data, and possible consequences for spatial data handling and the statistical analysis of spatial data. Beyond this, it is imperative to discuss the relationships between R-spatial software and

Journal of Geographical Systems

Aims and scope →
Submit manuscript →

Use our pre-submission checklist →
Avoid common mistakes on your manuscript.

Sections Abstract Introduction Spatial data classes in the sp package Spatial data classes in the sf and stars packages Upstream software dependencies of the R-spati... Outlook Notes References

<https://link.springer.com/article/10.1007/s10109-020-00336-0>

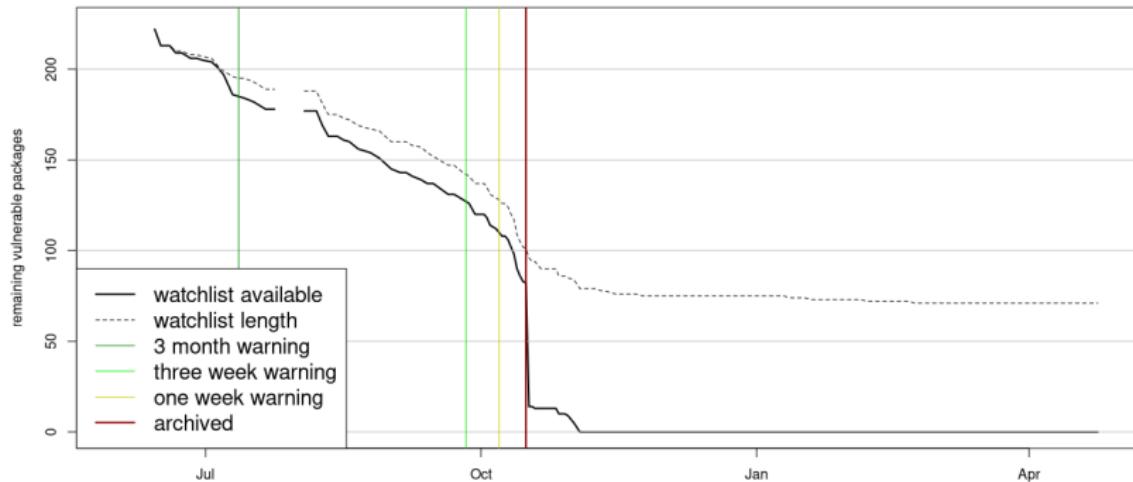
Harder dependencies

- ▶ fiesta (USDA) developed gdalraster, which provides a low-level GDAL interface for R
- ▶ URSA now imports sf, and suggests gdalraster, terra, and stars
- ▶ More response from GH issues than from emails to pkg maintainers
- ▶ it was necessary to be irritating, and provide patches for numerous pkgs (by RB);
- ▶ pkgapi was very useful: what is a package using? (S4 methods reporting patchy)

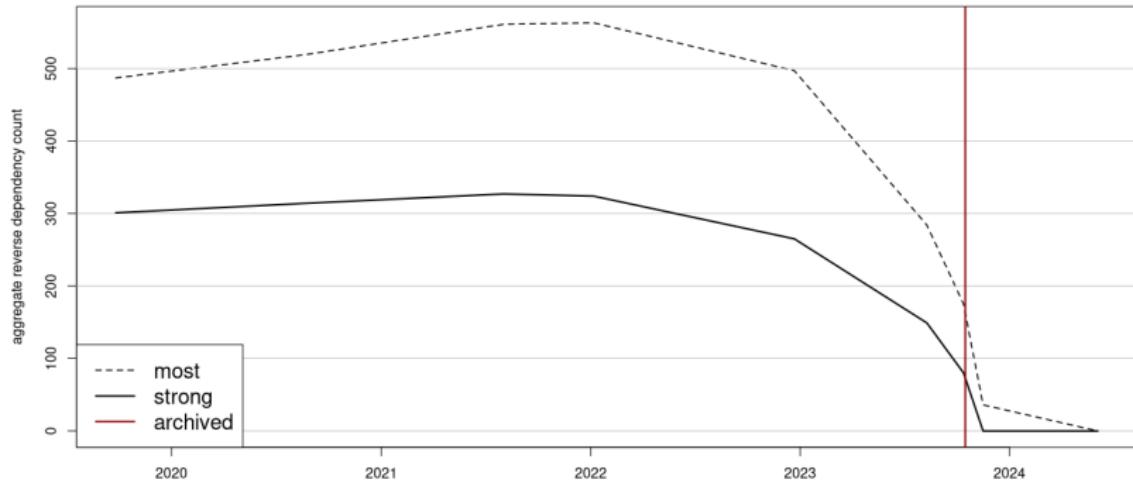
Watchlist: started jun 2023

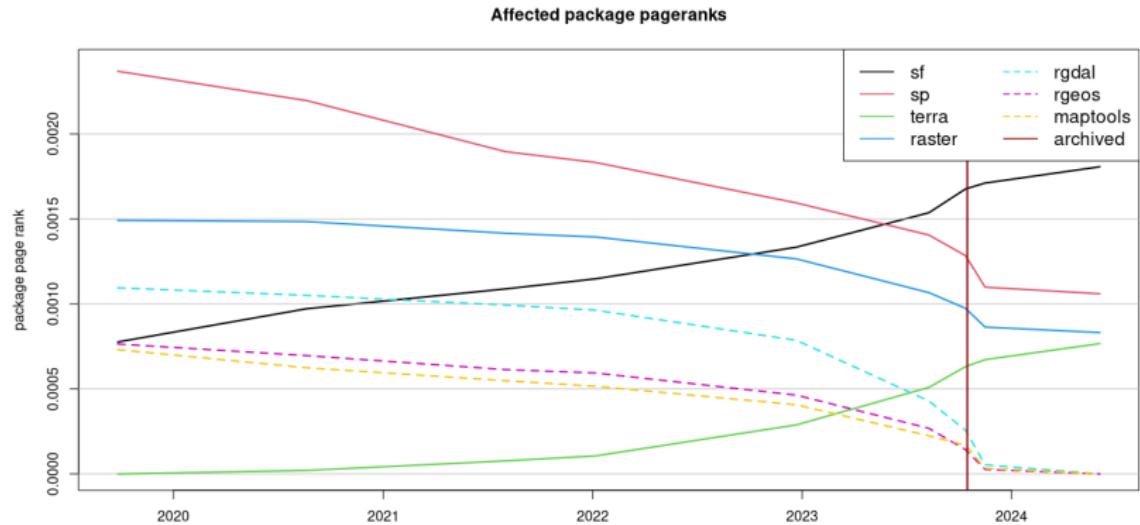
- ▶ ran nightly to see whether a revision had appeared
- ▶ check whether it fixed revdep issue, look into & report
- ▶ final removal at release date of Bioconductor 3.18 (Oct 2023)
- ▶ Watchlist output and code used is on the GH evolution repo.
- ▶ Data and code is on github.com/r-spatial/evolution

Watchlist attrition from June 2023



Reverse dependency counts for retiring packages





Recommendations

If you want to retire one or more heavily used CRAN packages and limit damage to CRAN:

- ▶ announce early, widely, loudly and often
- ▶ actively approach package maintainers, irritate them, help them
- ▶ closely watch the development of reverse dependencies
- ▶ communicate with the CRAN team
- ▶ ask for help and support from the community

Thanks to

- ▶ Roger Bivand for his tireless efforts
- ▶ R Consortium for supporting this project
- ▶ A reference group of key package developers and user community representatives who helped us giving feedback on planning and approach