



Une task view sur épidémio(surveillance) ?

Le projet d'un groupe ?

Fabrice Dessaint,
Inrae, Agroécologie,
Dijon
fabrice.dessaint@inrae.fr

24 novembre 2020

Cati IMOTEP ?

Informations, **MO**dèles et **T**raitement des données en **É**pidémiologie et dynamique des **P**opulations.

Structure regroupant une quarantaine d'ingénieurs et techniciens exerçant au sein d'équipes de recherche en épidémiologie et dynamique des populations dans un objectif de partage des savoirs et savoir-faire.

- Support pour les plateformes d'épidémiosurveillance animale et végétale.
- Acquisition et gestion de données
- Traitement, analyse des données et modélisation
- Développement d'outils issus des travaux de modélisation

Organisation de webinaires.

Site Web: imotep.inrae.fr/

Projet : Origine

Le langage R: pourquoi ?

(1) Utilisation de R dans les travaux académiques: en 2020⁽¹⁾

- Science Direct: 8 000 références
- Springer : 4 750 références
- Wiley: 5 700 références

(2) Un grand nombre de packages. **Pour rappel:** en septembre 2020, il y avait

- 16 100 packages sur le site du CRAN
- 1 900 packages sur le site Bioconductor

(1) *Recherche de l'expression:* R Core Team

Univers des packages R

CONTRIBUTED RESEARCH ARTICLES

558

Navigating the R Package Universe

by Julia Silge, John C. Nash, and Spencer Graves

Abstract Today, the enormous number of contributed packages available to R users outstrips any given user's ability to understand how these packages work, their relative merits, or how they are related to each other. We organized a plenary session at useR!2017 in Brussels for the R community to think through these issues and ways forward. This session considered three key points of discussion. Users can navigate the universe of R packages with (1) capabilities for directly searching for R packages, (2) guidance for which packages to use, e.g., from CRAN Task Views and other sources, and (3) access to common interfaces for alternative approaches to essentially the same problem.

The R Journal Vol. 10/2, December 2018

Une *task view* sur Épidémiologie ?

An R View into Epidemiology

📅 2020-05-20

by Joseph Rickert

If you have been tracking the numbers for the COVID-19 pandemic, you must have looked at dozens of models and tried to make some comparisons. Even under the best of situations it is difficult to compare models, and this is especially true if you don't have sufficient domain knowledge. Experts tend to leave out assumptions and background material that they know other experts will take for granted. This leaves newcomers pretty much on their own.

It has been my experience that a good way for an R literate person to begin to acquire knowledge in a new field is to find some appropriate packages, study the vignettes, work through the examples, and read whatever source material they may reference. So, this post shows how one might go about finding those appropriate packages. Also, I thought it would be interesting to see what kind of special resources are available to epidemiologists working in R beyond the basic statistical infrastructure and packages for data manipulation and visualization.

Because there is no epidemiology task view, a good place to start is to search CRAN directly. (Note that there are task views on [differential equations](#), [spatial statistics](#), [time series](#) and other tools used by epidemiologists, so I confined my search to the basics.)

Une *task view* - c'est quoi ?

Une réponse (partielle) à l'explosion du nombre de packages R déposés sur le site du **CRAN**.

Le **CRAN** la définit comme

CRAN task views aim to provide some guidance which packages on CRAN are relevant for tasks related to a certain topic. They give a brief overview of the included packages ... The views are intended to have a sharp focus so that it is sufficiently clear which packages should be included (or excluded) - and they are not meant to endorse the 'best' packages for a given task.

ce que l'on peut résumer par

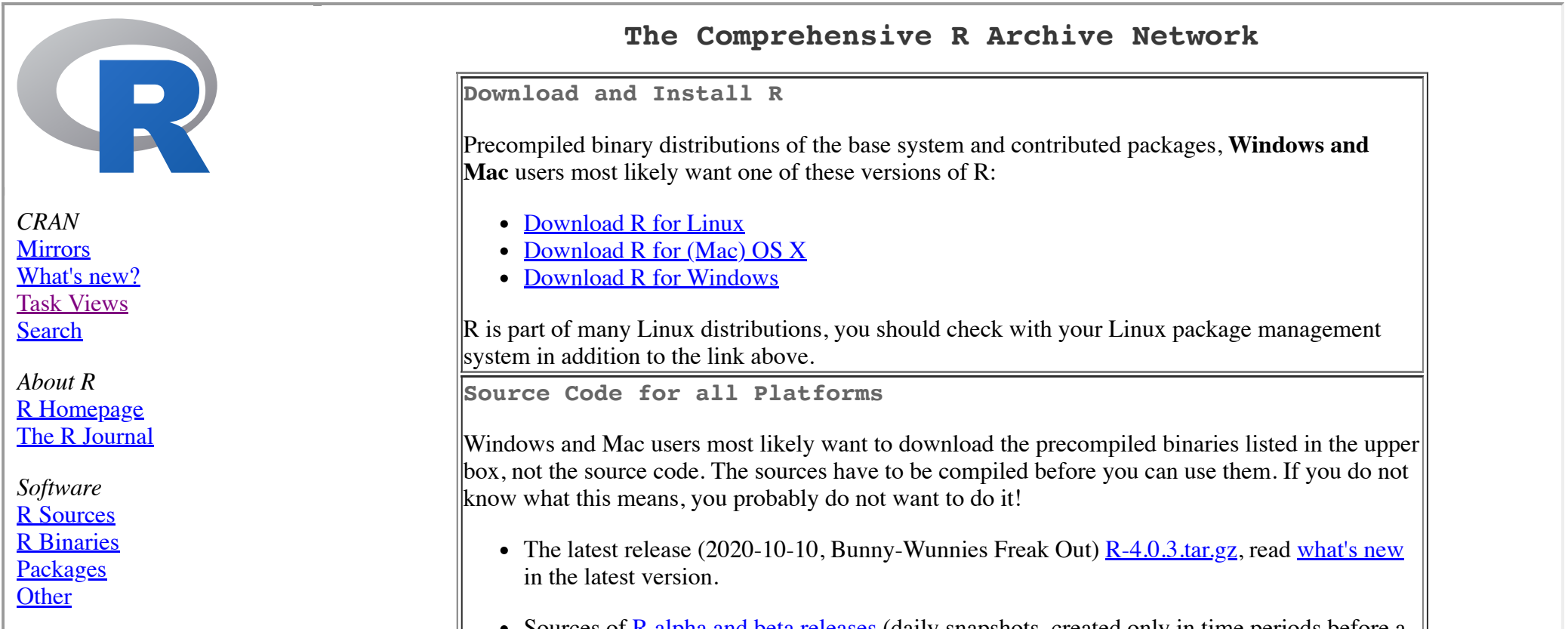
Une task view c'est une compilation, regroupant par grands thèmes ou disciplines, les packages R jugés les plus utiles.

Une *task view* - en pratique

- Une page web accessible via la rubrique Task Views sur le site du **CRAN**
- Proposée et mise à jour par un ou des volontaires (Maintainer)
- Sur une thématique plus ou moins large (Topic)
- Regroupant une liste annotée plus ou moins longue de packages R utiles pour la thématique
- Avec des informations sur des ouvrages ou articles, des cours, d'autres *task views*

Une *task view* - une page web

- Accessible via la rubrique Task Views du **CRAN**



The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2020-10-10, Bunny-Wunnies Freak Out) [R-4.0.3.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a

CRAN
[Mirrors](#)
[What's new?](#)
[Task Views](#)
[Search](#)

About R
[R Homepage](#)
[The R Journal](#)

Software
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Other](#)

Une *task view* - le(s) volontaire(s)

- Proposée et mise à jour par un ou des volontaires (Maintainer)

CRAN Task View: Missing Data

Maintainer: Julie Josse, Nicholas Tierney, Nathalie Vialaneix (r-miss-tastic team)

Contact: r-miss-tastic at clementine.wf

Version: 2020-10-09

URL: <https://CRAN.R-project.org/view=MissingData>

Missing data are very frequently found in datasets. Base R provides a few options to handle them using computations that involve only observed data (`na.rm = TRUE` in functions `mean`, `var`, ... or use `= complete.obs` | `na.or.complete` | `pairwise.complete.obs` in functions `cov`, `cor`, ...). The base package `stats` also contains the generic function `na.action` that extracts information of the NA action used to create an object.

These basic options are complemented by many packages on CRAN, which we structure into main topics:

- [Exploration of missing data](#)
- [Likelihood based approaches](#)
- [Single imputation](#)
- [Multiple imputation](#)
- [Weighting methods](#)
- [Specific types of data](#)
- [Specific application fields](#)

In addition to the present task view, this [reference website on missing data](#) might also be helpful.

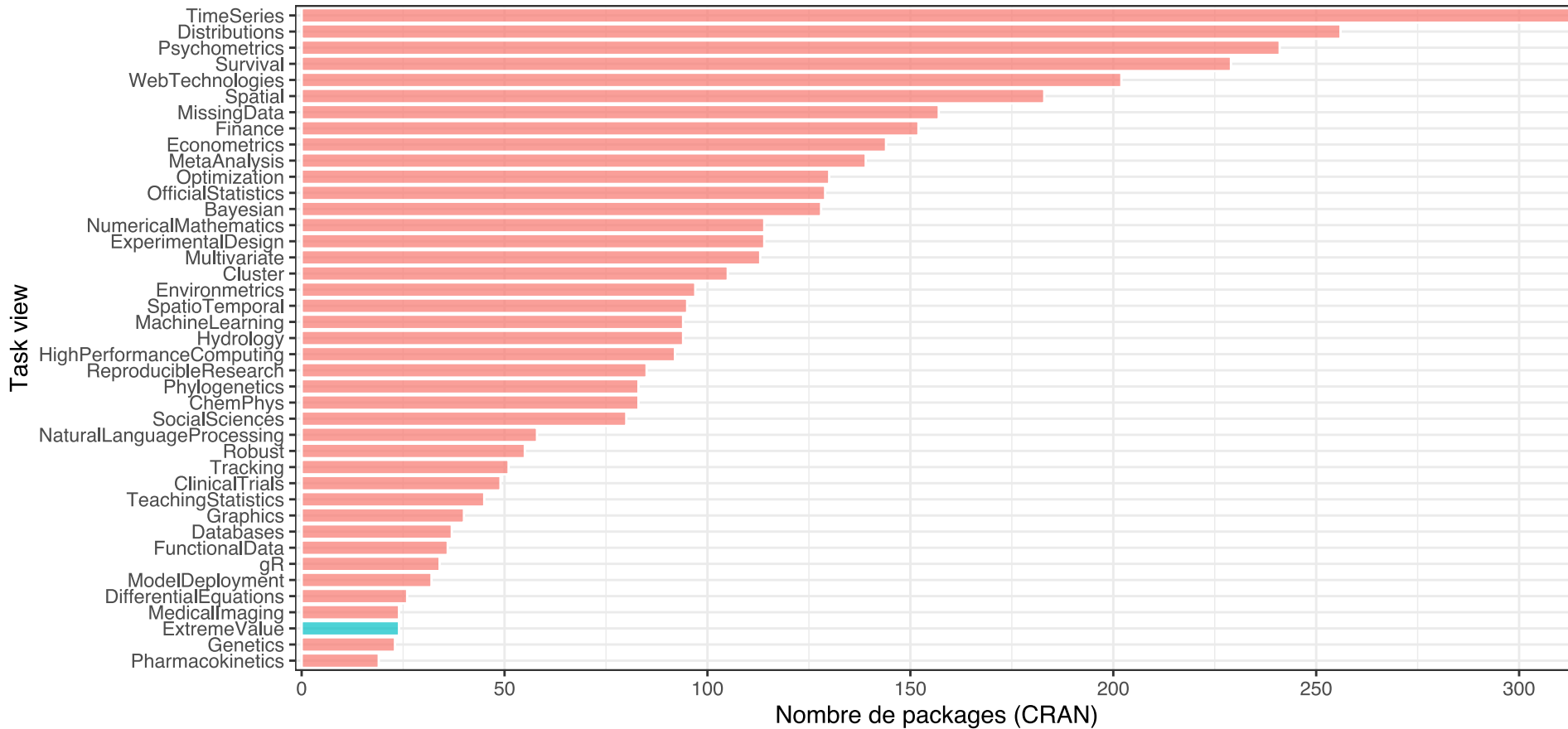
Une *task view* - une thématique

- Actuellement le **CRAN** propose 41 *task views*.

[1] "Bayesian"	"ChemPhys"	"ClinicalTrials"	"Cluster"
[5] "Databases"	"DifferentialEquations"	"Distributions"	"Econometrics"
[9] "Environmetrics"	"ExperimentalDesign"	"ExtremeValue"	"Finance"
[13] "FunctionalData"	"Genetics"	"Graphics"	"HighPerformanceComputing"
[17] "Hydrology"	"MachineLearning"	"MedicalImaging"	"MetaAnalysis"
[21] "MissingData"	"ModelDeployment"	"Multivariate"	"NaturalLanguageProcessing"
[25] "NumericalMathematics"	"OfficialStatistics"	"Optimization"	"Pharmacokinetics"
[29] "Phylogenetics"	"Psychometrics"	"ReproducibleResearch"	"Robust"
[33] "SocialSciences"	"Spatial"	"SpatioTemporal"	"Survival"
[37] "TeachingStatistics"	"TimeSeries"	"Tracking"	"WebTechnologies"
[41] "gR"			

Une *task view* - les packages

- Une liste annotée plus ou moins longue de packages R utiles pour la thématique.



Structure d'une *task view*

Pour chaque *task view*, la page web détaille une liste **annotée** de packages regroupés

(1) par grands groupes de méthodes et/ou de fonctionnalités.

Exemple: la *task view* Extreme Value Analysis, proposée par Christophe Dutang et Kevin Jaunatre, affiche les thèmes:

- Univariate Extreme Value Theory
- Bivariate Extreme Value Theory
- Multivariate Extreme Value Theory
- Classical graphics

(2) par niveau avec des packages recommandés (core) marqués par une * et d'autres suggérés.

[1]	copula	evd*	evdbayes	evir*	evmix	extremefit	extRemes	extremeStat	fExtremes
[10]	in2extRemes	ismev	lmom	lmomco	lmomRFA	mev	POT	ptsuite	QRM
[19]	ReIns	Renext	revdbayes	RTDE	threshr	VGAM			

Exemple: Extreme Value Analysis

CRAN Task View: Extreme Value Analysis

Maintainer: Christophe Dutang, Kevin Jaunatre

Contact: Christophe.Dutang at ensimag.fr

Version: 2020-02-20

URL: <https://CRAN.R-project.org/view=ExtremeValue>

Extreme values modelling and estimation are an important challenge in various domains of application, such as environment, hydrology, finance, actuarial science, just to name a few. The restriction to the analysis of extreme values may be justified since the extreme part of a sample can be of a great importance. That is, it may exhibit a larger risk potential such as high concentration of air pollutants, flood, extreme claim sizes, price shocks in the four previous topics respectively. The statistical analysis of extreme may be spread out in many packages depending on the topic of application. In this task view, we present the packages from a methodological side.

Applications of extreme value theory can be found in other task views: for financial and actuarial analysis in the [Finance](#) task view, for environmental analysis in the [Environmetrics](#) task view. General implementation of probability distributions is studied in the [Distributions](#) task view.

The maintainers gratefully acknowledge E. Gilleland, M. Ribatet and A. Stephenson for their review for extreme value analysis packages (2013) and Achim Zeileis for his useful comments. If you think information is not accurate or if we have omitted a package or important information that should be mentioned here, please let us know.

Univariate Extreme Value Theory:

- **Block Maxima approach:**

Exemple: Missing data

CRAN Task View: Missing Data

Maintainer: Julie Josse, Nicholas Tierney, Nathalie Vialaneix (r-miss-tastic team)

Contact: r-miss-tastic at clementine.wf

Version: 2020-10-09

URL: <https://CRAN.R-project.org/view=MissingData>

Missing data are very frequently found in datasets. Base R provides a few options to handle them using computations that involve only observed data (`na.rm = TRUE` in functions `mean`, `var`, ... or `use = complete.obs | na.or.complete | pairwise.complete.obs` in functions `cov`, `cor`, ...). The base package `stats` also contains the generic function `na.action` that extracts information of the NA action used to create an object.

These basic options are complemented by many packages on CRAN, which we structure into main topics:

- [Exploration of missing data](#)
- [Likelihood based approaches](#)
- [Single imputation](#)
- [Multiple imputation](#)
- [Weighting methods](#)
- [Specific types of data](#)
- [Specific application fields](#)

In addition to the present task view, this [reference website on missing data](#) might also be helpful.

If you think that we missed some important packages in this list, please contact the maintainer.

Une task view Epidémio(surveillance) Comment faire ?

Pour le CRAN

Création d'un fichier .ctv de la forme:

```
<CRANTaskView>
  <name>Epidemiology</name>
  <topic>Epidemiology, epidemiosurveillance</topic>
  <maintainer email="X@inra.fr">X</maintainer>
  <version>aaaa-mm-dd</version>
  <url>https://CRAN.R-project.org/view=Epidemiology</url>

  <info>
    Emphasize something <b>important</b> with standard HTML tags, and reference
    the package <pkg>foo</pkg> and also the view <view>TimeSeries</view>.
  </info>

  <packagelist>
    <pkg>foo</pkg>
    <pkg priority="core">bar</pkg>
  </packagelist>

  <links>
    <a href="http://path/to/homepage/">The Title of a Relevant Homepage</a> <view>TimeSeries</view>
  </links>
</CRANTaskView>
```

Achim Zeileis, Writing CRAN Task Views ([disponible ici](#)).

Pour la construction

1. Définir le thème de la *task view* ✓
2. Vérifier que c'est un thème non en cours de traitement pour le CRAN ✓
3. Définir le périmètre/étendue (sous-thèmes): ce qui sera traité
4. Construire la liste des packages à inclure, les présenter
5. Proposer la *task view* au CRAN

Construire la liste des packages

Pour construire cette liste

(1) Tester: Moissonnage sur des mots clés

- sur les dépôts de packages : CRAN, Bioconductor, GitHub
- dans les publications

(2) À faire : Enquête sur les pratiques des chercheurs/ingénieurs en épidémiologie

- localement au sein du CATI et par contagion aux membres de leurs unités
- nationalement, via les départements,
- les réseaux thématiques, ...

... ou d'autres options ?

Recherche "automatique"

Recherche les packages du CRAN contenant la chaîne `epidemi*` dans leur description, on a 86 packages:

```
[1] "EpiModel"          "biostat3"          "Epi"               "epifitter"        "POUMM"            "msce"
[7] "MKclass"          "EpiILM"           "HDMT"             "mem"              "memapp"          "Haplin"
[13] "EpiCurve"         "landsepi"         "RSurveillance"    "epiR"             "EpiStats"        "EpiReport"
[19] "bets.covid19"     "argo"             "ICD10gm"          "FMCCSD"           "bhpm"            "table1"
[25] "ems"              "surveillance"     "packMBPLSDA"     "pts2polys"        "outbreaker2"     "epimdr"
[31] "EpiILMCT"         "DSAIDE"           "popEpi"           "TreeSim"          "bootLR"          "epibasix"
[37] "epinet"           "powerSurvEpi"     "StrainRanking"   "epitools"         "wktmo"           "MOQA"
[43] "eesim"            "epicontacts"      "pop"              "R0"               "TreePar"         "CARE1"
[49] "sac"              "allestimates"     "asymptor"         "BayesSenMC"       "card"            "chest"
[55] "contact"          "contactdata"      "EpiContactTrace" "EpiEstim"         "epiflows"        "epiphy"
[61] "epitrix"          "expose"           "i2extras"         "incidence"        "incidence2"      "matchmaker"
[67] "MedDietCalc"      "MGDrive2"         "micar"            "mStats"           "NobBS"           "nosoi"
[73] "OCNet"            "predkmeans"       "projections"      "R2019nCoV"        "recmap"          "riskCommunicator"
[79] "SimInf"           "SPPcomb"          "STraTUS"         "tipr"             "trendeval"       "ungroup"
[85] "varrank"          "weibull4"
```

Exemple: le package DSAIDE

Exploration of simulation models (apps) of various infectious disease transmission dynamics scenarios. The purpose of the package is to help individuals learn about infectious disease epidemiology (ecology/evolution) from a dynamical systems perspective. All apps include explanations of the underlying models and instructions on what to do with the models.

DSAIDE Package Tutorial

[Andreas Handel](#)

2020-01-08

Summary

Dynamical Systems Approaches to Infectious Disease Epidemiology/Ecology/Evolution (DSAIDE) is an R package that allows you to explore and study concepts of infectious disease epidemiology (ecology/evolution) using dynamical systems models. The main way of interaction with the models is through a graphical user interface. It is not necessary - though possible - to read or write computer code.



The package also provides a path toward using and building your own models by relatively seamlessly moving

Exemple: le package `epifitter`

Analysis and visualization of plant disease progress curve data. Functions for fitting two-parameter population dynamics models (exponential, monomolecular, logistic and Gompertz) to proportion data for single or multiple epidemics using either linear or no-linear regression. Statistical and visual outputs are provided to aid in model selection. Synthetic curves can be simulated for any of the models given the parameters.

`epifitter`: Analysis and Simulation of Plant Disease Progress Curves

Analysis and visualization of plant disease progress curve data. Functions for fitting two-parameter population dynamics models (exponential, monomolecular, logistic and Gompertz) to proportion data for single or multiple epidemics using either linear or no-linear regression. Statistical and visual outputs are provided to aid in model selection. Synthetic curves can be simulated for any of the models given the parameters. See Laurence V. Madden, Gareth Hughes, and Frank van den Bosch (2007) <[doi:10.1094/9780890545058](https://doi.org/10.1094/9780890545058)> for further information on the methods.

Version: 0.1.0
 Depends: R (≥ 3.2)
 Imports: [deSolve](#), [dplyr](#), stats, [ggplot2](#), [knitr](#), [tidyr](#), [DescTools](#), [minpack.lm](#), [magrittr](#), [tibble](#)
 Suggests: [rmarkdown](#), [ggridges](#), [cowplot](#)
 Published: 2020-10-29
 Author: Kaique dos S. Alves  [aut, cre], Emerson M. Del Ponte  [aut]
 Maintainer: Kaique dos S. Alves <kaiquedsalves at gmail.com>
 BugReports: <https://github.com/AlvesKS/epifitter/issues>

Exemple: le package `landsepi`

A stochastic, spatially-explicit, demo-genetic model simulating the spread and evolution of a plant pathogen in a heterogeneous landscape to assess resistance deployment strategies. It is based on a spatial geometry for describing the landscape and allocation of different cultivars, a dispersal kernel for the dissemination of the pathogen, and a SEIR ('Susceptible-Exposed-Infectious-Removed') structure with a discrete time step. It provides a useful tool to assess the performance of a wide range of deployment options with respect to their epidemiological, evolutionary and economic outcomes.

`landsepi`: Landscape Epidemiology and Evolution

A stochastic, spatially-explicit, demo-genetic model simulating the spread and evolution of a plant pathogen in a heterogeneous landscape to assess resistance deployment strategies. It is based on a spatial geometry for describing the landscape and allocation of different cultivars, a dispersal kernel for the dissemination of the pathogen, and a SEIR ('Susceptible-Exposed-Infectious-Removed') structure with a discrete time step. It provides a useful tool to assess the performance of a wide range of deployment options with respect to their epidemiological, evolutionary and economic outcomes. Loup Rimbaud, Julien Papaïx, Jean-François Rey, Luke G Barrett, Peter H Thrall (2018) <[doi:10.1371/journal.pcbi.1006067](https://doi.org/10.1371/journal.pcbi.1006067)>.

Version: 1.0.1
 Depends: methods, utils, grDevices (≥ 3.0.0), graphics (≥ 3.0.0), R (≥ 3.3.0), [sp](#) (≥ 1.0-17)
 Imports: [Rcpp](#) (≥ 0.9.0), stats (≥ 3.0.2), [Matrix](#), [mvtnorm](#), [fields](#), [splancs](#), [sf](#), [DBI](#), [RSQLite](#), [foreach](#), parallel, [doParallel](#)
 LinkingTo: [Rcpp](#), [testthat](#)
 Suggests: [testthat](#), [shiny](#), [shinyjs](#), [DT](#), [knitr](#), [rmarkdown](#)
 Published: 2020-07-02

Autre proposition: article

Les *task views* ne sont pas les seules solutions. On trouve de plus en plus régulièrement des articles/guide sur le choix des packages R.

Received: 19 October 2018 | Accepted: 23 September 2019

DOI: 10.1111/1365-2656.13116



BIOLOGGING: REVIEW

Journal of Animal Ecology 

Navigating through the R packages for movement

Rocío Joo¹  | Matthew E. Boone¹  | Thomas A. Clay²  | Samantha C. Patrick²  |
Susana Clusella-Trullas³  | Mathieu Basille¹ 

Autre proposition: site web

Le site web **R** Epidemics **C**onsortium propose aussi de recenser/développer des packages R utiles.



The logo for RECON (R Epidemics Consortium) features the word "RECON" in a large, blue, sans-serif font. The letters "R", "E", and "C" are solid blue, while the "O"s and "N" are composed of a dense cluster of small, grey, semi-transparent triangles, giving them a textured, 3D appearance.

The **R** Epidemics **C**onsortium (RECON) is an international not-for-profit, **non-governmental organisation** gathering experts in data science, modelling methodology, public health, and software development to create the next generation of analytics tools for informing the response to disease outbreaks, health emergencies and humanitarian crises using the R

Groupe *task view*

- CATI: Sandie Arnoux, Thierry Hoch, Aminah Keliet, Jean-François Rey, Hervé Richard, Emily Walker.
- Extérieur: Marie Grosdidier, Maude Jacquot, Anne Lehébel, Aurelien Madouasse, Julien Papaix, Patrizia Zamberletti,