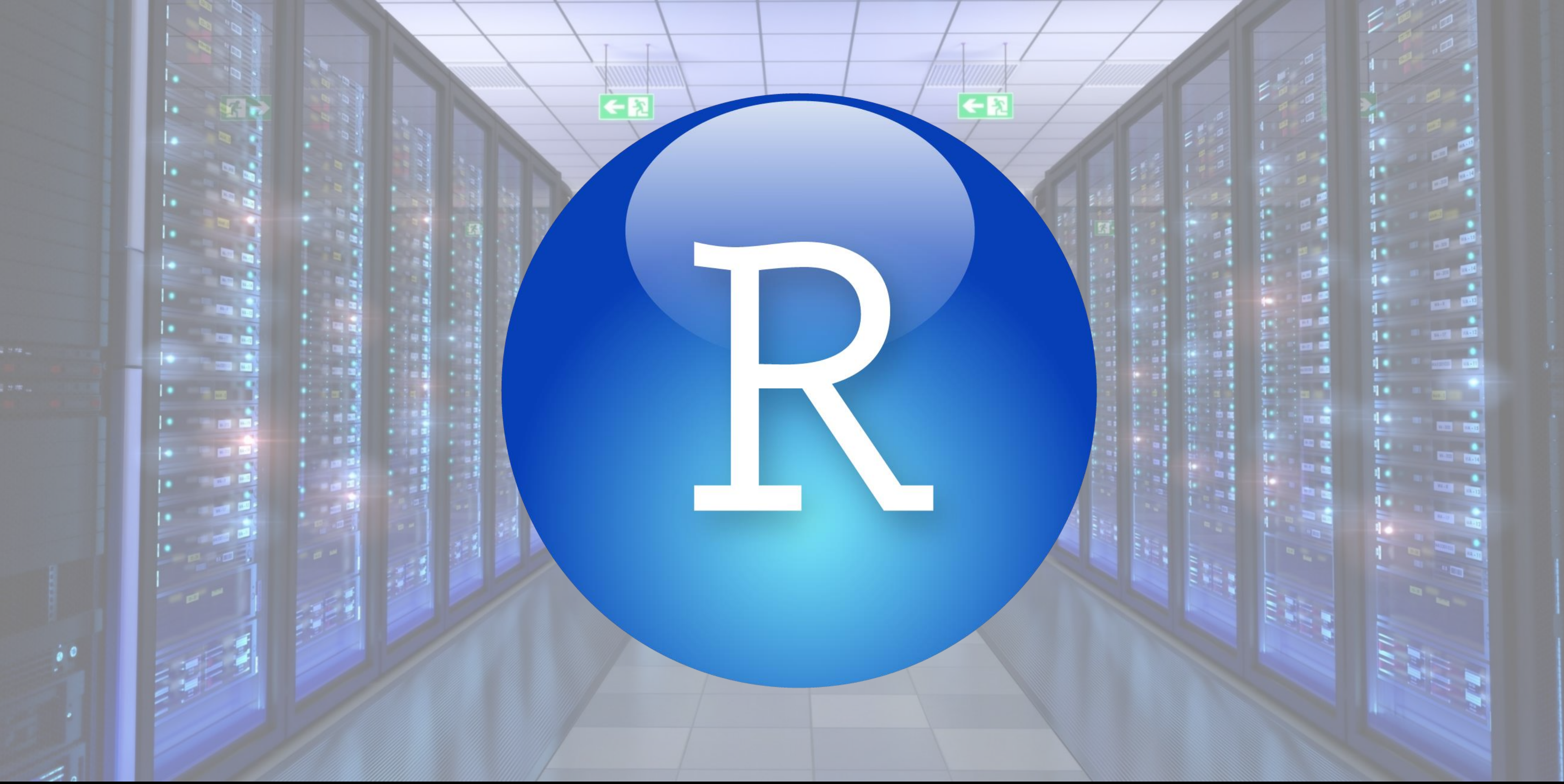
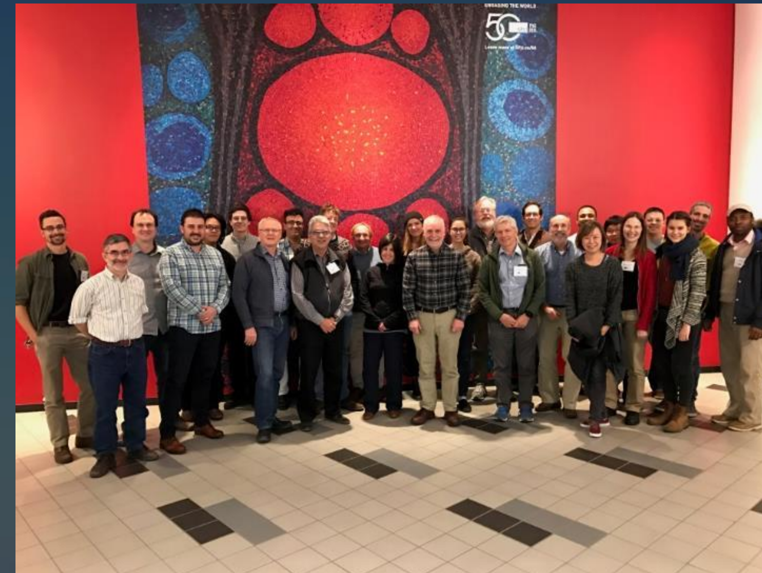


2. Introduction to R for Digital Soil Mapping

Canadian Digital Soil Mapping Workshop, 2022

CDSM WG
CANADIAN SOCIETY OF SOIL SCIENCE
SOCIÉTÉ CANADIENNE DE LA SCIENCE DU SOL
Ontario
UNIVERSITY of GUELPH
DALHOUSIE UNIVERSITY FACULTY OF AGRICULTURE

About Us



Brandon Heung

Dalhousie University

Brandon.Heung@dal.ca

Daniel Saurette

Ontario Ministry of Agri-Food and Rural Affairs

University of Guelph

Daniel.Saurette@ontario.ca

What is R?

R is a free software environment for statistical computing and graphics.

https://www.r-project.org/

- **Free to use**
- Used by majority of statisticians
- R community provides technical support
- Automation of DSM processes
- Sharing of R scripts
- Open source

Installing R

R-3.6.0 for Windows (32/64 bit)

Download R 3.6.0 for Windows (80 megabytes, 32/64 bit)

Installation and other instructions

New features in this version

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the md5sum of the .exe to the fingerprint on the master server. You will need a version of md5sum for windows: both graphical and command line versions are available.

Frequently asked questions

- Does R run under my version of Windows?
- How do I update packages in my previous version of R?
- Should I run 32-bit or 64-bit R?

Please see the R FAQ for general information about R and the R Windows FAQ for Windows-specific information.

Other builds

- Patches to this release are incorporated in the r-patched snapshot build.
- A build of the development version (which will eventually become the next major release of R) is available in the r-devel snapshot build.
- Previous releases

Note to webmasters: A stable link which will redirect to the current Windows binary release is
<CRAN MIRROR>/bin/windows/base/release.htm.

Last change: 2019-04-26

Compatible with OS X, Windows, Linux
32-bit & 64-bit options

http://mirror.its.dal.ca/cran/

R Packages

Contributed Packages

Available Packages

Currently, the CRAN package repository features 14435 available packages.

Table of available packages, sorted by date of publication

Table of available packages, sorted by name

Installation of Packages

Please type `help("INSTALL")` or `help("install.packages")` in R for information on how to install packages from this repository. The manual R Installation and Administration (also contained in the R base sources) explains the process in detail.

CRAN Task Views allow you to browse packages by topic and provide tools to automatically install all packages for special areas of interest. Currently, 40 views are available.

Package Check Results

All packages are tested regularly on machines running Debian GNU/Linux, Fedora, OS X, Solaris and Windows.

The results are summarized in the check summary (some timings are also available). Additional details for Windows checking and building can be found in the Windows check summary.

Writing Your Own Packages

The manual Writing R Extensions (also contained in the R base sources) explains how to write new packages and how to contribute them to CRAN.

Repository Policies

The manual CRAN Repository Policy [PDF] describes the policies in place for the CRAN package repository.

Related Directories

Archive
: Previous versions of the packages listed above, and other packages formerly available.

Orphaned
: Packages with no active maintainer, see the corresponding README.

bin/windows/contrib
: Windows binaries of contributed packages

bin/macosx/el-capitan/contrib
: OS X El Capitan binaries of contributed packages

Each package consists of a suite of **tools** that are used for day-to-day operations in R.

Example:
caret: Machine-learning algorithms
sp: Spatial data analysis
RSAGA: Integration with SAGA-GIS

Each **package** includes a **reference manual** that describes the use of each **tool**.

https://cran.r-project.org/web/packages/

R Scripts

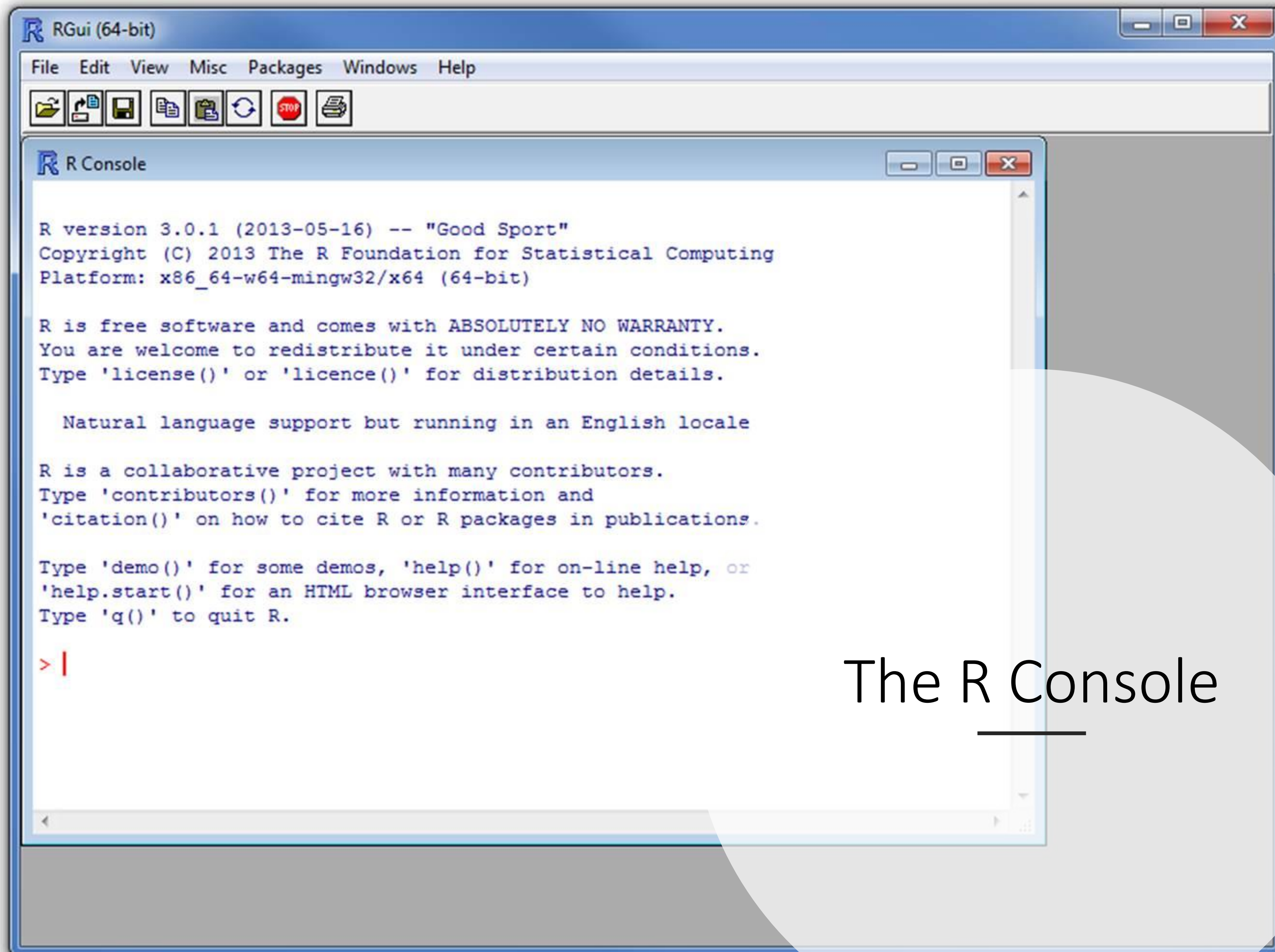
```r
##################################################################
### Conditioned Latin Hypercube Sampling of Principal Components ###
##################################################################

# STEP 1: Input working directory using setwd("x")

options( java.parameters = "-Xmx24g" )
setwd("E:/2018_Kamloops cLHS/Covariates")
setwd("E:/2018_Kamloops cLHS/Covariates/BEC Covariates/idfdk1_idfdk1")
setwd("E:/2018_Kamloops cLHS/Covariates/BEC Covariates/Reduced Variables")


dir.create("PCA")
dir.create("Sample Buffers")
dir.create("cLHS")


# STEP 2: Load necessary packages

library(clhs)
library(foreign)
library(ggplot2)
library(raster)
library(rgdal)
library(RStoolbox)
library(maptools)
library(sp)
library(spatstat)

rasterOptions(tmpdir="E:/2018_Kamloops cLHS/tmpdir")
```

Contains a series of command lines that are used to make R do things

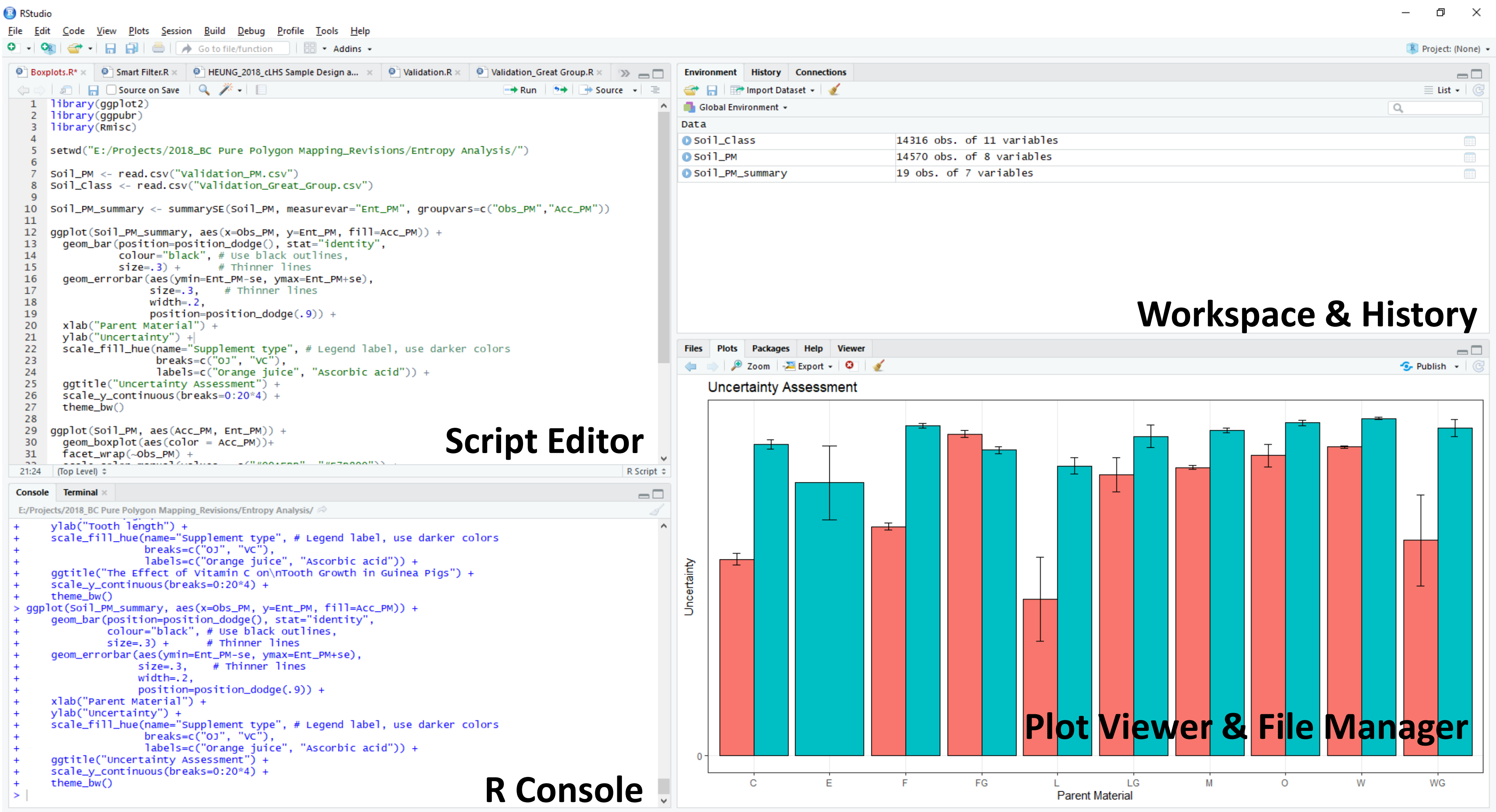
Useful for automation and reproducing your workflow

Tons of resources available for you to copy & paste R scripts and modify for your own purposes.

The R Console



R Studio

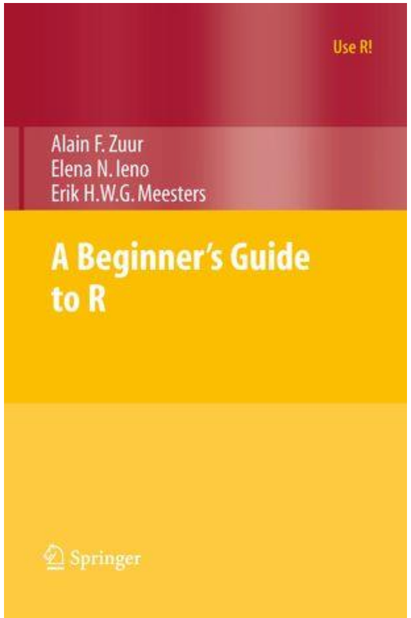


Scripting software that is specifically designed for R

Using R in Digital Soil Mapping

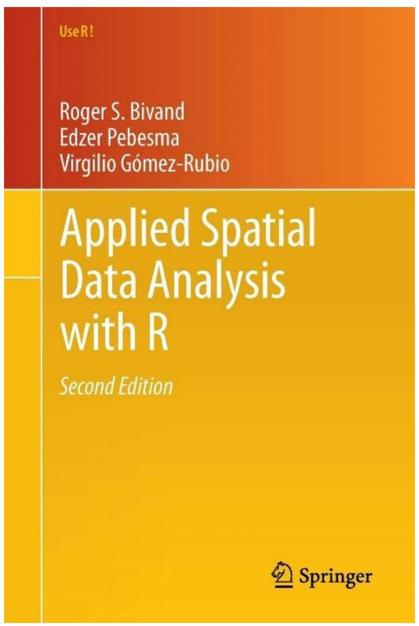
- Management of digital soil information and geospatial data
- Performs (geo-)statistical analyses
    - Conventional statistics
    - Spatial statistics
    - Machine-learning and artificial intelligence
- Integrates with GIS software (SAGA, GRASS GIS, QGIS)
- “Big Data” analytics

General Resources for R



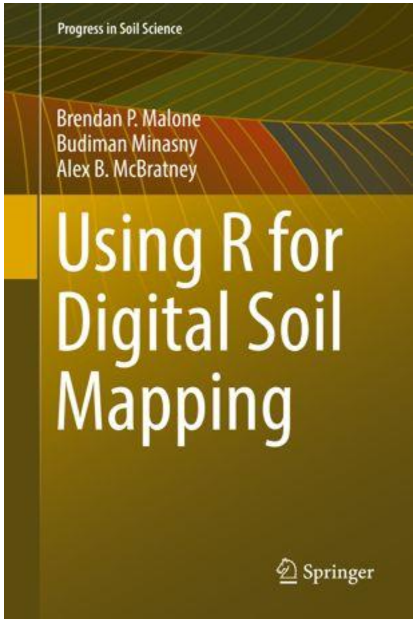
- Many resources for using R at the beginner's level.
- Includes R scripts and detailed walkthroughs

Spatial Data Analysis in R



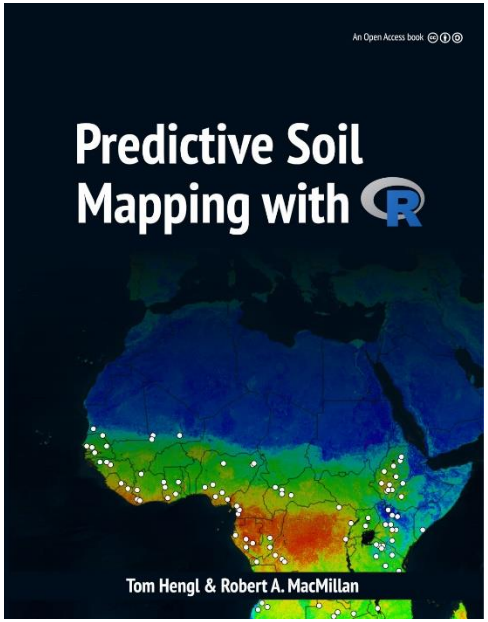
- Handling spatial data in R
- Visualizing spatial data
- Spatial data import and export
- Spatio-temporal data
- Spatial point pattern analysis
- Interpolation & geostatistics
- Modelling areal data

Digital Soil Mapping in R



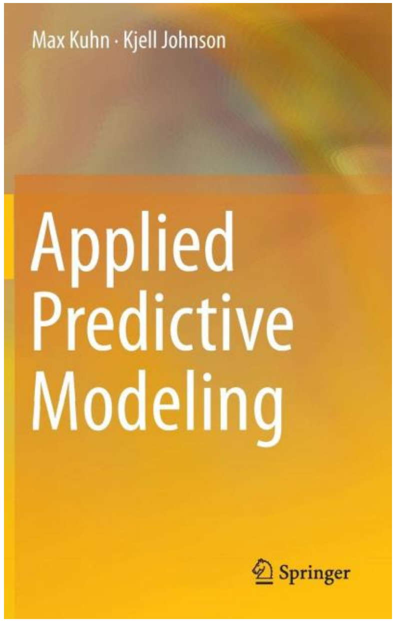
- R literacy for digital soil mapping
- Exploratory data analysis
- Modelling continuous and categorical soil data
- Quantification of prediction uncertainty
- Updating, harmonizing, and disaggregating legacy soil maps
- Digital soil assessments

Digital Soil Mapping in R



- Soil resource inventories
- Software installation
- Soil observations & variables
- Preparation of geospatial layers
- Statistical & machine-learning theory

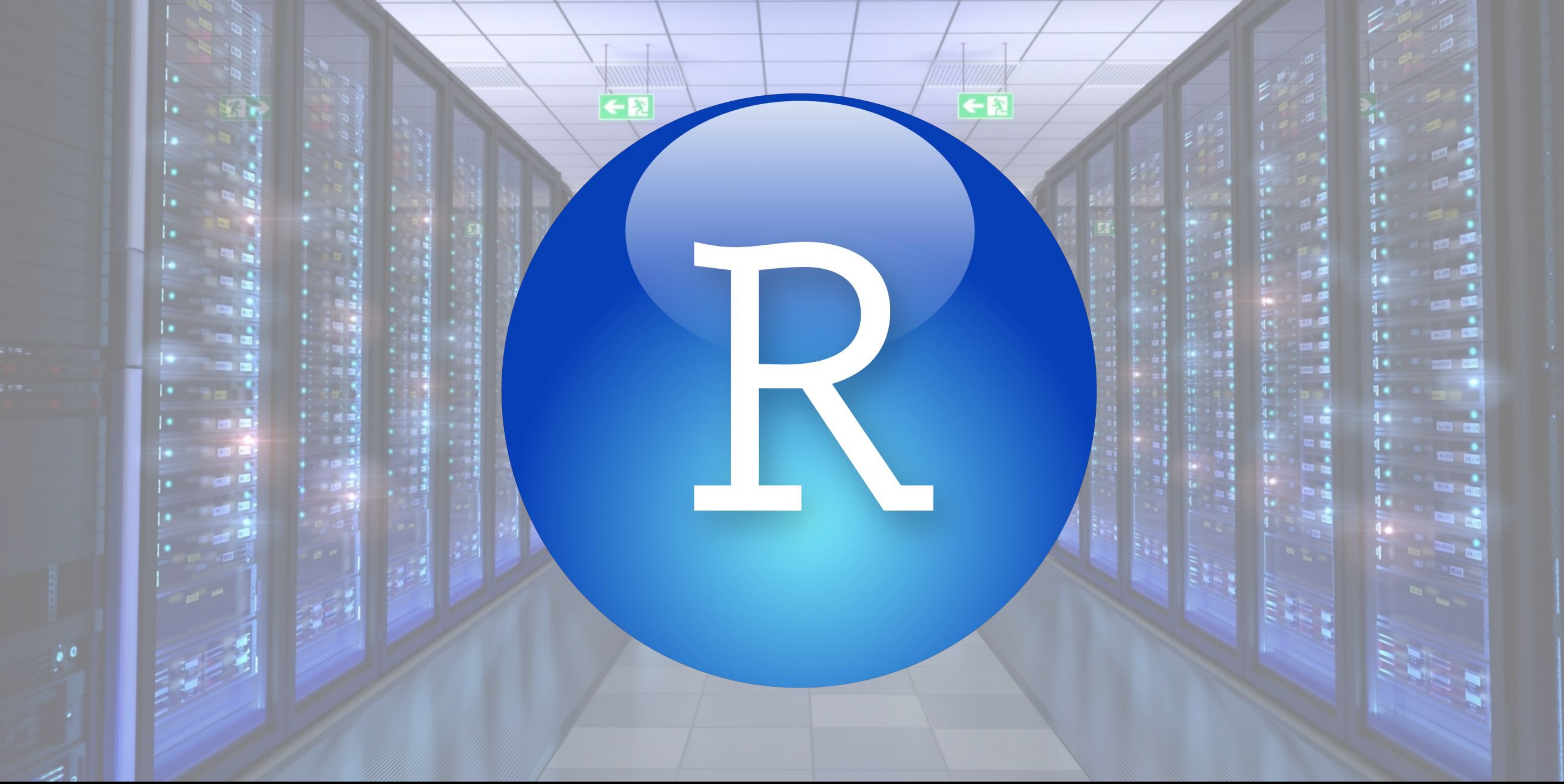
Predictive Modelling in R



- Data pre-processing
- Model fitting and overfitting
- Regression modelling
- Classification modelling

Module 2: Starting R

- Installation & loading libraries in R
- Setting up your working drive & accessing data
- Basic data exploration



2. Introduction to R for Digital Soil Mapping

Canadian Digital Soil Mapping Workshop, 2022

CDSM WG

CANADIAN SOCIETY OF SOIL SCIENCE
SOCIÉTÉ CANADIENNE DE LA SCIENCE DU SOL

Ontario

UNIVERSITY of GUELPH

DALHOUSIE UNIVERSITY
FACULTY OF AGRICULTURE