

Examples of how to publish datasets/code

Workshop FAIR data and data reuse for ESG researchers – Module 5

October 18, 2022, by Cindy Quik and Luc Steinbuch



Four ESG-related examples of publishing datasets/code

3 examples of datasets and 1 example of code

Short introduction to each example (brief content)

Focus on the way the datasets/code were published

Examples differ in access (open, restricted, closed)

Four ESG-related examples of publishing datasets/code

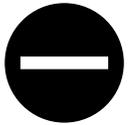
1. Dataset accompanying scientific paper,
published in a data repository with open access



2. Dataset accompanying a project report,
published in a data repository with restricted access



3. Dataset,
published internally (W-drive, closed access)



4. Code accompanying scientific paper,
published in a data repository with open access



1) Dataset accompanying scientific paper, published in a data repository with open access



- Paper that resulted from a PhD project
(Department of Environmental Sciences)
- Methodological study with the goal to improve the accuracy of age determinations of peat deposits
- Data from Dutch peat cores,
including plant species, organic matter content, age

1) Dataset accompanying scientific paper, published in a data repository with open access



- Considerations with publishing the paper and dataset of this project:
 - Funder (NWO) prefers open access publications
 - Access restrictions were not necessary
 - The dataset is small (278 kB)

1) Dataset accompanying scientific paper, published in a data repository with open access



- **Publication:**

 - Open Access at Elsevier

- **Dataset:**

 - Open Access to dataset description and ReadMe file at 4TU.ResearchData

 - Open Access to datafiles at 4TU.ResearchData



1) Dataset accompanying scientific paper, published in a data repository with open access



■ **Publication:**

Quik C, Palstra SWL, Van Beek R, Van der Velde Y, Candel JHJ, Van der Linden M, Kubiak-Martens L, Swindles GT, Makaske B, Wallinga J (2022).

Dating basal peat: The geochronology of peat initiation revisited.

Quaternary Geochronology 72:1-22. <https://doi.org/10.1016/j.quageo.2022.101278>

■ **Dataset:**

Quik C, Palstra SWL, Van Beek R, Van der Velde Y, Candel JHJ, Van der Linden M, Kubiak-Martens L, Swindles GT, Makaske B, Koudijs R, Wallinga J (2022).

Data from: Dating basal peat: The geochronology of peat initiation revisited.

4TU.ResearchData. <https://doi.org/10.4121/16923358>

1) Dataset accompanying scientific paper, published in a data repository with open access



- Publication:

Quik C, Palstra SWL, Van Beek R, Van der Velde Y, Candel JHJ, Van der Linden M, Kubiak-Martens L, Swindles GT, Makaske B, Wallinga J (2022).

Dating basal peat: The geochronology of peat initiation revisited.

Quaternary Geochronology 72:1-22. <https://doi.org/10.1016/j.quageo.2022.101278>

- Dataset:

Quik C, Palstra SWL, Van Beek R, Van der Velde Y, Candel JHJ, Van der Linden M, Kubiak-Martens L, Swindles GT, Makaske B, Koudijs R, Wallinga J (2022).

Data from: Dating basal peat: The geochronology of peat initiation revisited.

4TU.ResearchData. <https://doi.org/10.4121/16923358>

1) Dataset accompanying scientific paper, published in a data repository with open access



- Publication:

Quik C, Palstra SWL, Van Beek R, Van der Velde Y, Candel JHJ, Van der Linden M, Kubiak-Martens L, Swindles GT, Makaske B, Wallinga J (2022).

Dating basal peat: The geochronology of peat initiation revisited.

Quaternary Geochronology 72:1-22. <https://doi.org/10.1016/j.quageo.2022.101278>

- Dataset:

Quik C, Palstra SWL, Van Beek R, Van der Velde Y, Candel JHJ, Van der Linden M, Kubiak-Martens L, Swindles GT, Makaske B, **Koudijs R**, Wallinga J (2022).

Data from: Dating basal peat: The geochronology of peat initiation revisited.

4TU.ResearchData. <https://doi.org/10.4121/16923358>

Authors may not be equal!

1) Dataset accompanying scientific paper, published in a data repository with open access



- Let's see what it looks like online:

<https://doi.org/10.4121/16923358>

2) Dataset accompanying a project report, published in a data repository with restricted access



- Project report from the Dairy Business Information Service and Support (DairyBISS) Project (WUR Centre for Development Innovation)
- Survey among commercial farms, firms and advisors working in the dairy value chain in Ethiopia
- Data from 100 households,
including dairy herd structure, feeding and farm management

2) Dataset accompanying a project report, published in a data repository with restricted access



- Considerations with publishing the paper and dataset of this project:
 - No information on preferences of the funder
 - The type of data probably requires access restrictions
 - The dataset is small (718 kB)

2) Dataset accompanying a project report, published in a data repository with restricted access



- **Publication:**

- Open Access at WUR E-depot

- **Dataset:**

- Open Access to dataset description and ReadMe file at DANS-Easy

- Restricted access to datafiles (depositor needs to grant permission)

Data Archiving and Networked Services

DANS

2) Dataset accompanying a project report, published in a data repository with restricted access



- **Publication:**

Buizer NN, Berhanu T, Murutse G, Van Vugt SM (2015).

DairyBISS Baseline report.

Wageningen UR Centre for Development Innovation. <https://edepot.wur.nl/375521>

- **Dataset:**

Vernooij AG (Wageningen Livestock Research, WUR), Berhanu T (2021).

Farm practices.

DANS. <https://doi.org/10.17026/dans-xnh-9aha>

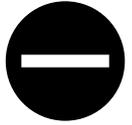
2) Dataset accompanying a project report, published in a data repository with restricted access



- Let's see what it looks like online:

<https://doi.org/10.17026/dans-xnh-9aha>

3) Dataset,
published internally (W-drive, closed access)

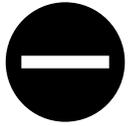


“Register every dataset at pure.library@wur.nl”

Two examples, from outsider viewpoint at research.wur.nl

(internally also the location of the dataset is saved)

3) Dataset, published internally (W-drive, closed access)



Dataset Fiscal Development under Colonial and Sovereign Rule

Ewout Frankema (Creator), Marlous van Waijenburg (Creator)

WASS, Rural and Environmental History

Dataset

Date made available

28 Jan 2022

Publisher

Wageningen University

Contact person

Ewout Frankema

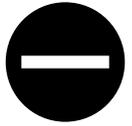
Related publication is a bookchapter:

[Fiscal Development under Colonial and Sovereign Rule](#)

Frankema, E. H. P. & van Waijenburg, M., 17 Feb 2022, *Global Taxation: How Modern Taxes conquered the World*.

Genschel, P. & Seelkopf, L. (eds.). Oxford University Press

3) Dataset, published internally (W-drive, closed access)



ISRIC-WISE revised soil property estimates for the soil types of the World

Niels Batjes (Creator)

ISRIC - World Soil Information

Dataset

Description

This dataset presents derived property values for 28 soil chemical and physical attributes identified as being useful for AEZ studies, crop growth simulation, and analyses of global environmental change.

It is based on the analysis of over 9600 globally distributed soil profiles, held in a working copy of WISE.

Criteria for clustering the data are in accordance with conventions developed at FAO, IIASA, and ISRIC for use with the Soil Map of the World.

The lists of derived soil properties can be linked to the 1:5 million scale Soil Map of the World through the 1974 Legend soil unit code, and to SOTER databases through the Revised Legend code. Map unit complexity should be considered when making such linkages (see documentation).

Date made available

17 Mar 2014

Publisher

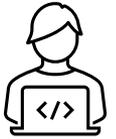
ISRIC - World Soil Information

Keywords

soil science

Related
publication is a
report (from
2002)

4) Code accompanying scientific paper, published in a data repository with open access



Source code in the R programming language, belonging with: Model based geostatistics from a Bayesian perspective: Investigating area-to-point kriging with small datasets

L. Steinbuch, T.G. Orton, D.J. Brus

Soil Geography and Landscape, Mathematical and Statistical Methods - Biometris, PE&RC

Research output: Non-textual form > Software > Other research output

4) Code accompanying scientific paper, published in a data repository with open access



ABOUT YOUR DATA ▾ ABOUT 4TU.RESEARCHDATA ▾ ABOUT OUR

Bayesian-area-to-point-kriging-master

export_tables

session_results_1D_sim

exportcolour

exportbw

session_results_2D_sim

data_burkina_faso

bayesian_areal_kriging

calculate_posterior_inference.R

calculate_validation_stats.R

calculate_validation_stats_case.R

collect_summarise_validation_and_Bay_inference.R

create_covariance_objects.R

create_covariance_objects_1D_sim.R

LICENSE

REFERENCES

- <https://doi.org/10.1007/s11004-019-09840-6> View PDF
- <https://git.wur.nl/stein012/Bayesian-area-to-point-kriging>

4) Code accompanying scientific paper, published in a data repository with open access



Bayesian-area-to-point-kriging

README.pdf

Source code in the R programming language, belonging with paper: *Model-Ba. Perspective: Investigating Area-to-Point Kriging with Small Data Sets*. Luc Steir Mathematical Geosciences (2019). <https://doi.org/10.1007/s11004-019-09841> world dataset is available on request; ask the first author (Luc Steinbuch).

Updates will be available at repository <https://git.wur.nl/stein012/Bayesian-a>

Contents of this README.pdf

Source code functionality	1
Directories.	1
Contact.	1
Coding conventions	1
Functions, .R files and dependencies.	2 (note: different page fo
License.	3
Session info	3

4) Code accompanying scientific paper, published in a data repository with open access



README.pdf:

Perspective: Investigating Area-to-Point Kriging with Small Data Sets. Luc Steinbuch, Thomas G. Orton, Dick Brus. *Mathematical Geosciences* (2019). <https://doi.org/10.1007/s11004-019-09840-6> The associated spatial real world dataset is available on request; ask the first author (Luc Steinbuch).

In the related paper:

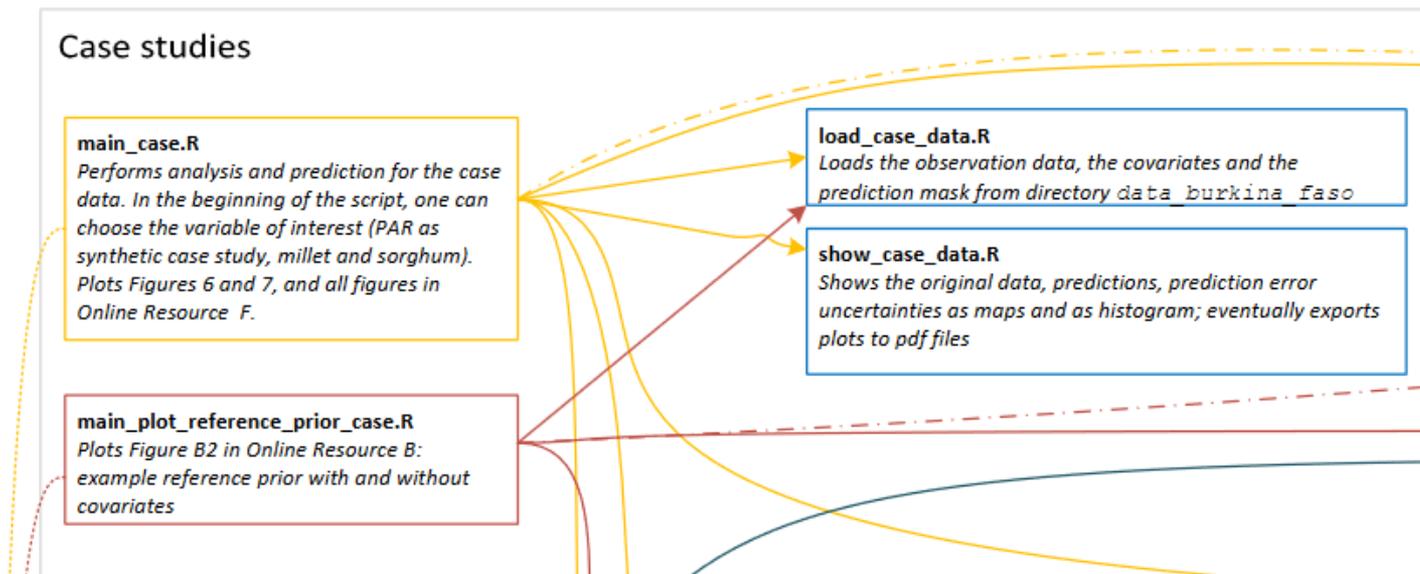
As a real-world case study, this paper predicts yields of sorghum and millet, both cereal staple foods, in Burkina Faso, West Africa. The observation areas are the 45 provinces, for which only the average yields are known (averaged over the years 2000–2013, and provided by AGRHYMET), as shown in Fig.

4) Code accompanying scientific paper, published in a data repository with open access



Functions, .R files and dependencies

README.pdf



Break!

14:40 -> The other way around: data reuse



Please be back
at 14:40