BioVeL – Biodiversity Virtual e-Laboratory

Workflow Documentation

Cohen’s cumulative distance Workflow for local execution

July 2014

Capacities Programme of Framework 7: EC e-Infrastructure Programme –
e-Science Environments - INFRA-2011-1.2.1

Grant Agreement No: 283359
Project Co-ordinator: Mr Alex Hardisty
Project Homepage: http://www.biovel.eu
Duration of Project: 36 months

Cohen’s cumulative distance Workflow for local execution

1 Description

The Cohen's cumulative distance Workflow provides an environment to calculate the Cohen's cumulative distance**.** The Cohen's cumulative distance measures the difference between observed and expected vectors along the matrix path that the population would take to reach the expected population vector. It is a function of both the observed stage distribution (***n0***) and the structure of the matrix (**A**) (Williams et al 2011). Cohen's cumulative distance will not work for reducible matrices or imprimitive matrices with nonzero imaginary components in the dominant eigenpair, and returns a warning for other imprimitive matrices (Caswell 2001).

2 General

2.1 Name of the workflow in myExperiment.

Name: Cohen's cumulative distance Workflow for local execution

Download: The workflow pack can be downloaded at myExperiment under the following address: http://www.myexperiment.org/packs/651.html or only the workflow: http://www.myexperiment.org/workflows/4346.html

2.2 Date, version and licensing

Last updated: 23th July 2014

Version: 1

Licensing: CC-BY-SA

2.3 How to cite this workflow

To report work that has made use of this workflow, please add the following credit acknowledgement to your research publication:

The input data and results reported in this publication (tutorial) come from data (Dr. Gerard Oostermeijer unpublished results and publication: Oostermeijer, J.G.B. M.L. Brugman, E.R. de Boer; H.C.M. Den Nijs. 1996. Temporal and Spatial Variation in the Demography of *Gentiana pneumonanthe*, a Rare Perennial Herb. *The Journal of Ecology*, 84: 153-166.) using BioVeL workflows and services (www.biovel.eu). Cohen's cumulative distance workflow was run on *<date of the workflow run>*. BioVeL is funded by the EU's Seventh Framework Program, grant no. 283359.

3. Scientific specifications

3.1 Keywords

Matrix Population Models, BioVeL, Cohen's cumulative distance demography, Gentiana pneumonanthe, matrix, matrix population models, population, stable stage distribution (w), stage matrix.

3.2 Scientific workflow description

The Cohen's cumulative distance Workflow provides an environment to calculate the Cohen's cumulative distance. The workflow accepts input data in a .txt format. The output is provided as a set of R results.

The Cohen's cumulative distance measures the difference between observed and expected vectors along the matrix path that the population would take to reach the expected population vector.

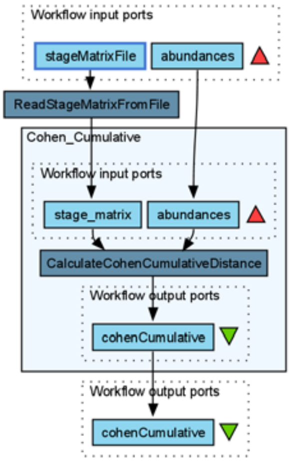


Figure 1. Cohen's cumulative distance workflow in Taverna workbench.

For more detailed description of the functions, please visit the Tutorial section.

4 Technical specifications

4.1 Execution environment and installation requirements

The Workflow requires a Taverna Engine. The simplest way to install a Taverna Engine is to install Taverna Workbench. The workflow also requires an Rserve installation with popdemo package installed. It is possible to setup the workflow to use a remote Rserve. However, instructions for installing a local Rserve are provided below.

4.2 Taverna installation, including updates and plugins

4.2.1 Taverna installations

- Taverna Workbench: Version 2.4 or 2.5. For installation files and instructions, please go to http://www.taverna.org.uk/download/workbench

4.2.2 Taverna Dependencies

- Install R software in your computer. See: http://www.r-project.org/
- Start R, and install package Rserve:
    - install.packages ("Rserve")
- For package popdemo, as it is archived in CRAN, use the package devtools to install it
    - install.packages ("devtools")
    - require(de2vtools)
    - install_url("http://cran.r-project.org/src/contrib/Archive/popdemo/popdemo_0.1-3.tar.gz")
- Local R Server: (Rserve) running at port 6311. See https://wiki.biovel.eu/x/3ICD for additional information.

4.2.4 How it works

- First, open R, once R is opened, type `library (Rserve)` and press enter; then type `Rserve()` and press enter again. You will see then the following message: Starting Rserve...

```
"C:\PROGRA~1\R\R-30~1.1\library\Rserve\libs\x64\Rserve.exe"
```

After this operation you can open Taverna and run the workflow.

5 Tutorial

5.1 Introduction

In this tutorial, you will use an input file MTers87_88.txt, called in myexperiment.org: Stage Matrix of Gentiana pneumonanthe 1987-88 File. You will use this file to perform several analyses on it. Finally you will be able to export your results to different formats.

5.2 Input data.

5.2.1 Data preparation/format

The workflow accepts input data (matrices) in a .txt format, all decimal numbers in each matrix must be indicated by dots e.g.: 0.578. All examples matrices for the tutorial are available in the PACK: http://www.myexperiment.org/packs/651.html or here below.

5.2.2 Input data

The input files are in a .txt format: to download click here in each file:

Terschelling

- MTers87_88.txt

5.3 Select input data dialogue boxes.

The first step is to fill out the input ports:

5.3.1 INPUTPORTS

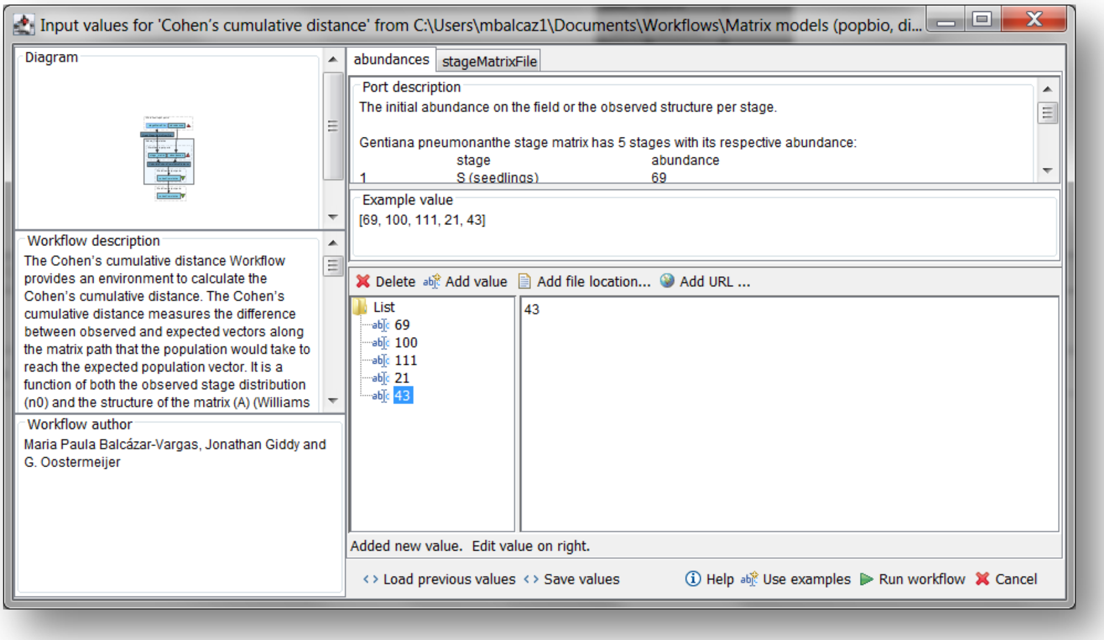
1) ***Abundance:*** here appear the fields with the initial abundance per stage observed in the field (Fig 2). As an example *Gentiana pneumonanthe* has 5 stages with its respective abundance.

The abundance of the stages or categories must be added one by one. First press Add value, fill the abundance of the first stage and press enter; then press Add value and fill once again the next stage abundance; repeat the action until you have fill all the abundance of all stages (Fig 2).

Stage abundance of the year 1987:

S (seedlings)	69
J (Juveniles)	100
V (vegetative)	111
G (reproductive individuals)	21
D (dormant plants)	43

e.g. [69, 100, 111, 21, 43]



Input values for 'Cohen's cumulative distance' from C:\Users\mbalcaz1\Documents\Workflows\Matrix models (popbio, di...

Workflow description
The Cohen's cumulative distance Workflow provides an environment to calculate the Cohen's cumulative distance. The Cohen's cumulative distance measures the difference between observed and expected vectors along the matrix path that the population would take to reach the expected population vector. It is a function of both the observed stage distribution (n0) and the structure of the matrix (A) (Williams

Workflow author
Maria Paula Balcázar-Vargas, Jonathan Giddy and G. Oostermeijer

Port description
The initial abundance on the field or the observed structure per stage.

Gentiana pneumonanthe stage matrix has 5 stages with its respective abundance:

```
                stage                          abundance
1               S (seedlings)                  69
```

Example value
[69, 100, 111, 21, 43]

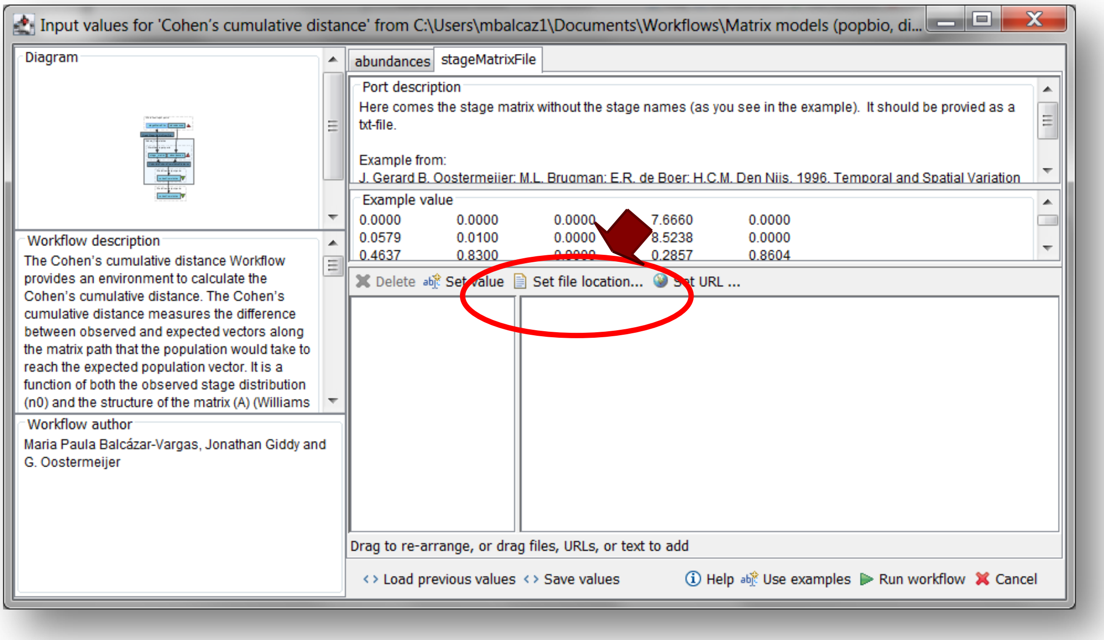
List
- 69
- 100
- 111
- 21
- 43

43

Added new value. Edit value on right.

Figure 2. Abundance input. Fill in the abundance per stage appears.

2) ***<u>StageMatrixFile</u>:*** Matrix in a .txt file format. Click in Set file location (red arrow in Fig 3), a window dialog appears and the user selects the file (e.g. MTers87_88.txt, Fig 4) and then clicks the Open button. Then the link to the file will appear on the Taverna window.



Input values for 'Cohen's cumulative distance' from C:\Users\mbalcaz1\Documents\Workflows\Matrix models (popbio, di...

Workflow description
The Cohen's cumulative distance Workflow provides an environment to calculate the Cohen's cumulative distance. The Cohen's cumulative distance measures the difference between observed and expected vectors along the matrix path that the population would take to reach the expected population vector. It is a function of both the observed stage distribution (n0) and the structure of the matrix (A) (Williams

Workflow author
Maria Paula Balcázar-Vargas, Jonathan Giddy and G. Oostermeijer

Port description
Here comes the stage matrix without the stage names (as you see in the example). It should be provied as a txt-file.

Example from:
J. Gerard B. Oostermeijer; M.L. Brugman; E.R. de Boer; H.C.M. Den Nijs. 1996. Temporal and Spatial Variation

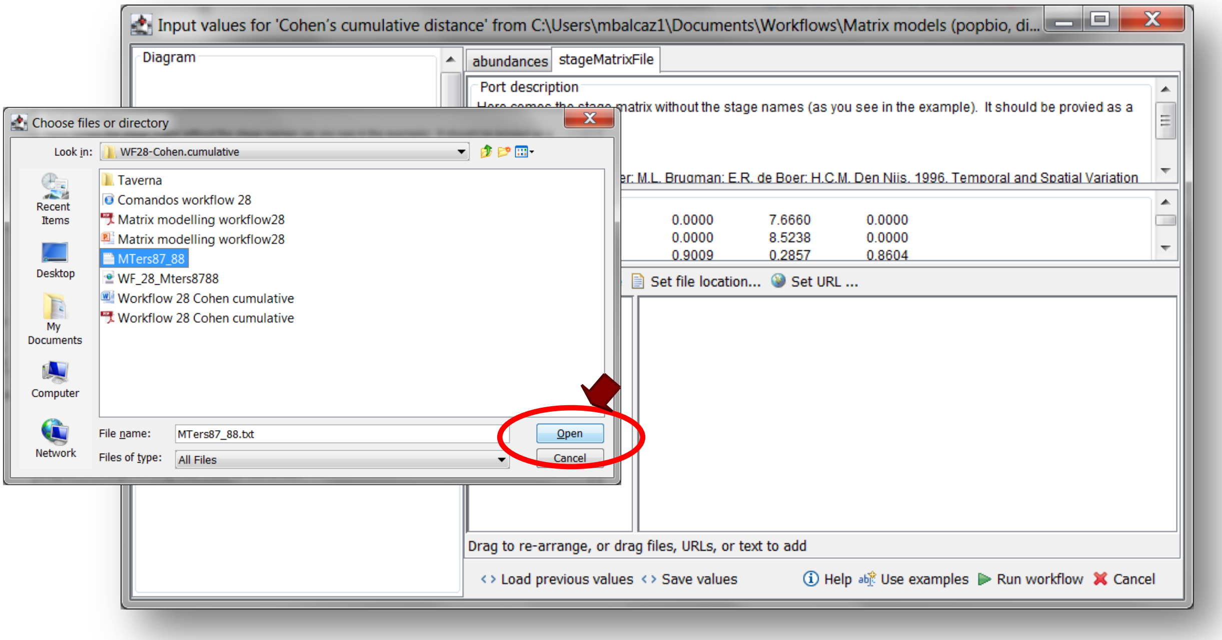
Example value

```
0.0000      0.0000      0.0000      7.6660      0.0000
0.0579      0.0100      0.0000      8.5238      0.0000
0.4637      0.8300      0.0000      0.2857      0.8604
```

Drag to re-arrange, or drag files, URLs, or text to add

Figure 3. Set the Location of the matrix file.

Cohen’s cumulative distance Workflow for local execution



Input values for 'Cohen's cumulative distance' from C:\Users\mbalcaz1\Documents\Workflows\Matrix models (popbio, di...

Port description

matrix without the stage names (as you see in the example). It should be provied as a

er: M.L. Brugman; E.R. de Boer; H.C.M. Den Nijs. 1996. Temporal and Spatial Variation

```
0.0000    7.6660    0.0000
0.0000    8.5238    0.0000
0.9009    0.2857    0.8604
```

Choose files or directory

Look in: WF28-Cohen.cumulative

- Taverna
- Comandos workflow 28
- Matrix modelling workflow28
- Matrix modelling workflow28
- MTers87_88
- WF_28_Mters8788
- Workflow 28 Cohen cumulative
- Workflow 28 Cohen cumulative

File name: MTers87_88.txt

Files of type: All Files

Figure 4. Choose the matrix file and click in Open.

After the user has filled out the input ports and has clicked the ***Run Workflow button***, the analysis will few seconds.

When the analysis is completed, it appears on different windows under results in Taverna, then the user can save the result.

5.4 Save data/results

5.4.1 OUTPUTS

First, click in the selected result window e.g. **cohenCumulative** (Fig 5, green oval). Second, click in the left window on value 1 (Fig 5, blue oval), Third, on the right window click on the save value (Fig 5, red oval). Fourth, name the file and determinate the extension file. For a text file: e.g. .csv or .txt. Finale, save the file in the chosen map (Fig 6).

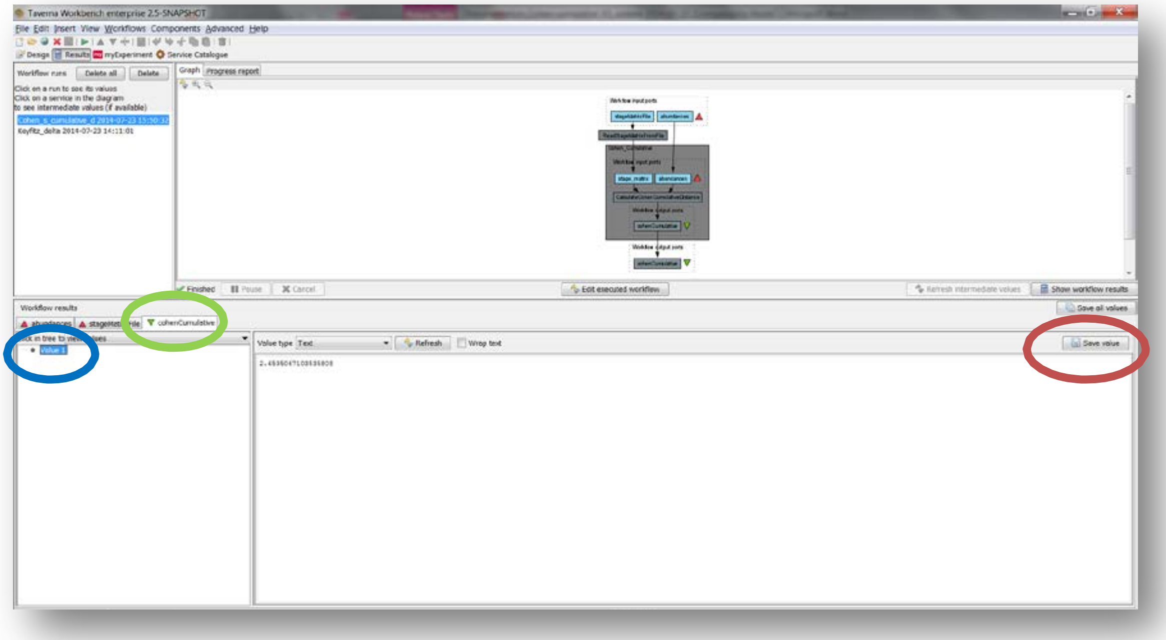
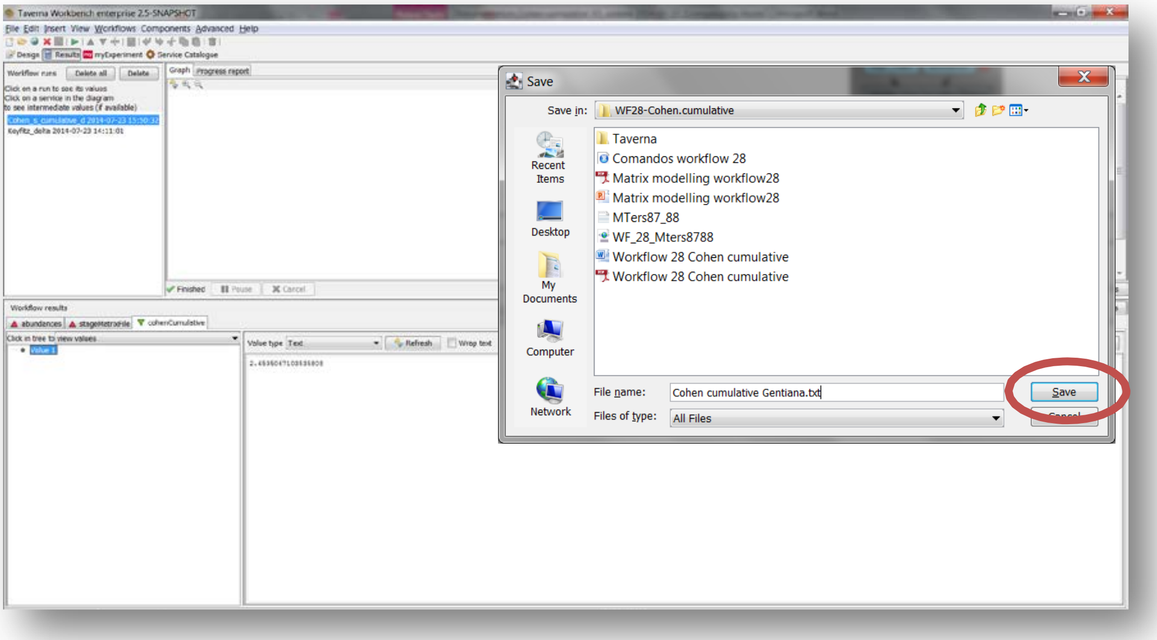


Figure 5. Cohen cumulative output window.



Save in: WF28-Cohen.cumulative

File name: Cohen cumulative Gentiana.txt

Files of type: All Files

Value type: Text

Figure 6. Name the file and determinate the extention file. For a text file: e.g. .csv or .txt.

1) Cohen's cumulative distance: it measures the difference between observed and expected vectors along the matrix path that the population would take to reach the expected population vector.

e.g. 2.6732

For further details see: References

.

6 Support

For questions with using the workflow, please write ***support @biovel.eu***

For definitions of technical and biological terms, please visit the BioVeL glossary page:
https://wiki.biovel.eu/display/BioVeL/Glossary

7 References

This workflow was created using and based the package popdemo (Stott, Hodgson and Townley, 2013)

Caswell, H. 2001. Matrix population models: Construction, analysis and interpretation, 2nd Edition. Sinauer Associates, Sunderland, Massachusetts.

Jongejans E. & H. de Kroon. 2012. Matrix models. Chapter in Encyclopaedia of Theoretical Ecology (eds. Hastings A & Gross L) University of California, p415-423

Oostermeijer J.G.B., M.L. Brugman; E.R. de Boer; H.C.M. Den Nijs. 1996. Temporal and Spatial Variation in the Demography of Gentiana pneumonanthe, a Rare Perennial Herb. The Journal of Ecology, Vol. 84(2): 153-166.

Stott, I., D.J. Hodgson and S. Townley. 2013. popdemo: Provides Tools For Demographic Modelling Using Projection Matrices. Version 0.1-3.

Williams, J.L., M.M. Ellis, M.C. Bricker, J.F. Brodie and E.W. Parsons. 2011. Distance to stable stage distribution in plant populations and implications for near-term population projections. *Journal of Ecology*, 99, 1171–1178.

7.1 Acknowledgements

7.1.1 Authors

1. ***Maria Paula Balcazar-Vargas*** Instituut voor Biodiversiteit en Ecosysteem Dynamica (IBED), Universiteit van Amsterdam.
2. ***Jonathan Giddy*** Cardiff School of Computer Science and Informatics, Cardiff University, Cardiff CF24 3AA, United Kingdom.
3. ***J. Gerard B. Oostermeijer*** Instituut voor Biodiversiteit en Ecosysteem Dynamica (IBED), Universiteit van Amsterdam.

7.1.2 Project funding

The workflow described in this documentation has been designed and implemented as part of the BioVeL project.

BioVeL is funded by the European Commission 7th Framework Programme (FP7) as part of its e-Infrastructures activity. Under FP7, the e-Infrastructures activity is part of the Research Infrastructures programme, funded under the FP7 'Capacities' Specific Programme. It focuses on the further development and evolution of the high-capacity and high-performance communication network (GÉANT), distributed computing infrastructures (grids and clouds), supercomputer infrastructures, simulation software, scientific data infrastructures, e-Science services as well as on the adoption of e-Infrastructures by user communities.

7.2 Publications