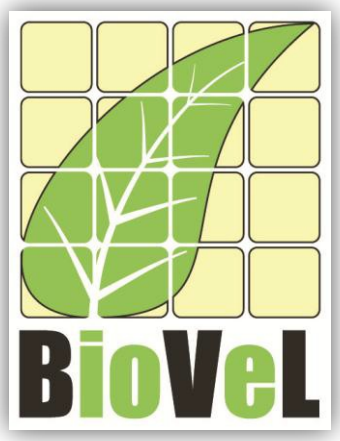


BioVeL – Biodiversity Virtual e-Laboratory

Workflow Documentation

Age specific survival workflow for local execution

June 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

Age specific survival workflow for local execution

1 Description

The age-specific survival analysis workflow gives the basic information on the mean, variance and coefficient of variation (cv) of the time spent in each stage class and the mean and variance of the time to death.

Analyses:

- Fundamental matrix (N)
- Variance (var)
- Coefficient of variation (CV)
- Meaneta
- Vareta

2 General

2.1 Name of the workflow in myExperiment.

Name: Age specific survival workflow for local execution.

Download: The pack can be downloaded at myExperiment under the following address: http://www.myexperiment.org/packs/631.html or the workflow at: http://www.myexperiment.org/workflows/3286.html

2.2 Date, version and licensing

Last updated: 2th November 2012

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 Gerard Oostermeijer unpublished results and publication: Oostermeijer, J.G.B. M.L. Brugman, E.R. de Boer; H.C.M. Den Nijs. 1996.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). Age specific survival 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, Age specific survival, Fundamental matrix (N), Variance (var), Coefficient of variation (cv), Meaneta, Vareta.

3.2 Scientific workflow description

The age-specific survival analysis workflow gives the basic information on age-specific survival; this includes the mean, variance and coefficient of variation (CV) of the time spent in each stage class and the mean and variance of the time to death.

The aim of the age-specific survival analysis workflow is to provide a connected environment to calculate the mean, variance and coefficient of variation (CV) of the time spent in each stage class and the mean and variance of the time to death. The workflow accepts input data in a .txt format. The output is provided as a set of R results.

a) Fundamental matrix (N): is the mean of the time spent in each stage class.
b) Variance (var): is the variance in the amount of time spent in each stage class.
c) Coefficient of variation (CV): is the coefficient of variation of the time spent in each class (SD/mean- the ratio of the standard deviation to the mean).
d) Meaneta: is the mean of time to death, of life expectancy of each stage.
e) Vareta: is the variance of time to death.

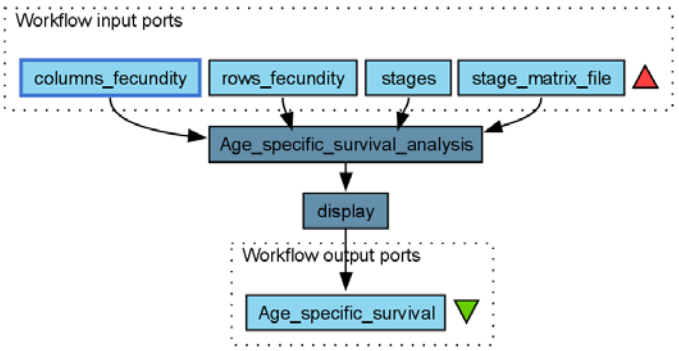


Figure 1. Age-specific survival analysis workflow in Taverna workbench.

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")
- Install package popbio
    - install.packages("popbio")
- Local R Server: (Rserve) running at port 6311. See https://wiki.biovel.eu/x/3ICD for additional information.

4.2.3 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 the age specific survival 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. The example matrix for the tutorial is available in: http://www.myexperiment.org/files/1135.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

Example from:
J. Gerard B. Oostermeijer; 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. *Journal of Ecology*, Vol. 84(2): 153-166.

Please note that the workflow as well as the tutorial is a beta version, and may contain errors. We hope you will still find it interesting to work with this workflow, and look forward to your feedback.

5.3 Select input data dialogue boxes.

The first step is to fill out the input ports:

5.3.1 INPUTPORTS

1) ***columns_fecundity*:** the column(s) in which the fecundity values are found, should be selected.

In the example of the *Gentiana* species (Oostermeijer et al. 1996. *Journal of Ecology*): The selected column is G (reproductive individuals), the number 4 will be used to

	S	J	V	G	D
S	0	0	0	7,666	0
J	0,0579	0,01	0	8,5238	0
V	0,4637	0,83	0,9009	0,2857	0,8604
G	0	0,04	0,009	0,619	0,1162
D	0	0,03	0,018	0	0,0232

identify the fecundity column (G) (see matrix next to text).

The numbers of the reproductive columns must be added one by one (in the case that there is more than one reproductive column. In this example there is only one). First press add value, fill the number of the column and press enter, then press add value and fill once again the next reproductive column, repeat the action until you have fill all the numbers of all reproductive columns (Fig 2).

e.g.: 4

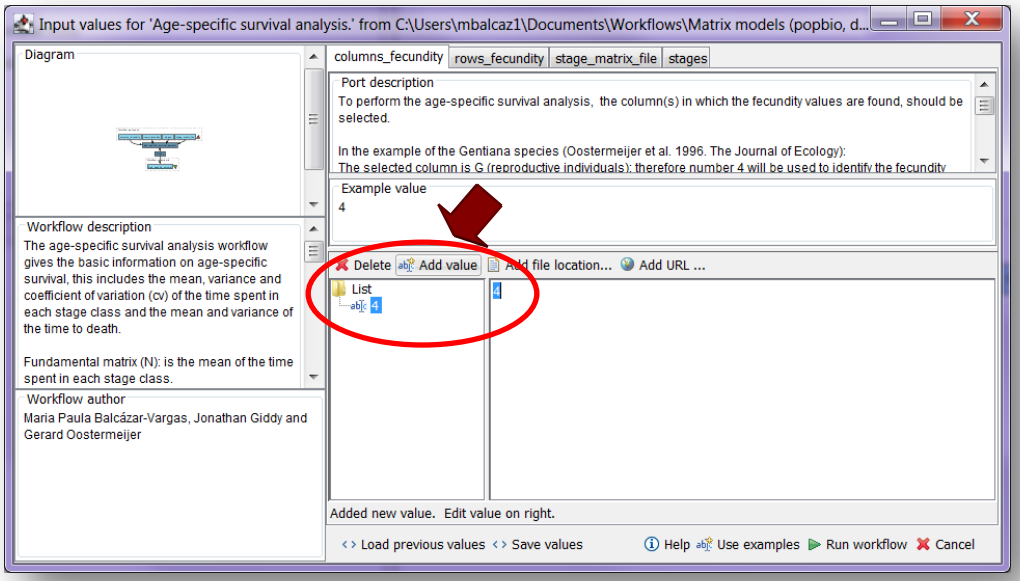


Figure 2. The number(s) of the column(s) to be filled in.

2) ***rows_fecundity:*** the row(s) in which the recruitment values are found, should be selected.

In the example of the *Gentiana* species (Oostermeijer et al. 1996. *Journal of Ecology*): Selected rows are S and J (seedlings and juveniles; these stages receive recruits each year from the stage G): therefore, the numbers 1 and 2 are used to identify these rows (see matrix next to the text).

	S	J	V	G	D
S	0	0	0	7,666	0
J	0,0579	0,01	0	8,5238	0
V	0,4637	0,83	0,9009	0,2857	0,8604
G	0	0,04	0,009	0,619	0,1162
D	0	0,03	0,018	0	0,0232

The numbers of the stages that receive recruits each year from the stage G must be added one by one (in the case that there are more than one recruit stage row). First press add value, fill the number of the row and press enter, then press add value and fill once again the next recruit stage row, repeat the action until you have fill all the numbers of all recruits rows (Fig. 3).

e.g.: 1 and 2

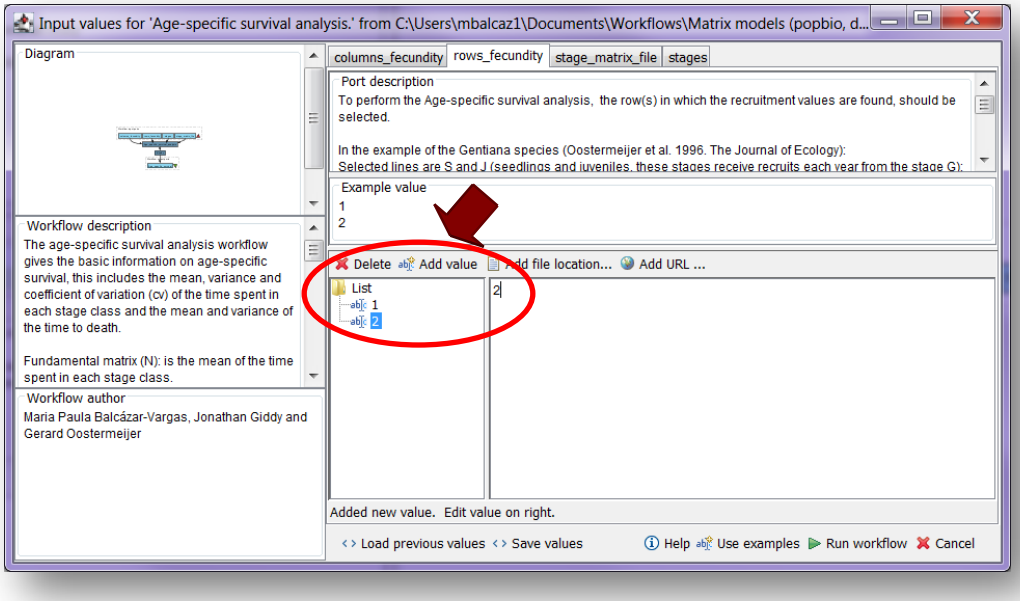


Figure 3. The number(s) of the recruit(s) row(s) to be filled in.

3) ***stage_matrix_file:*** Stage matrix in a .txt file format. Click in Set file location (red arrow in Fig 4), a window dialog appears and the user selects the file (e.g. MTers87_88.txt, Fig 5) and then clicks the Open button. Then the link to the file will appear on the Taverna window.

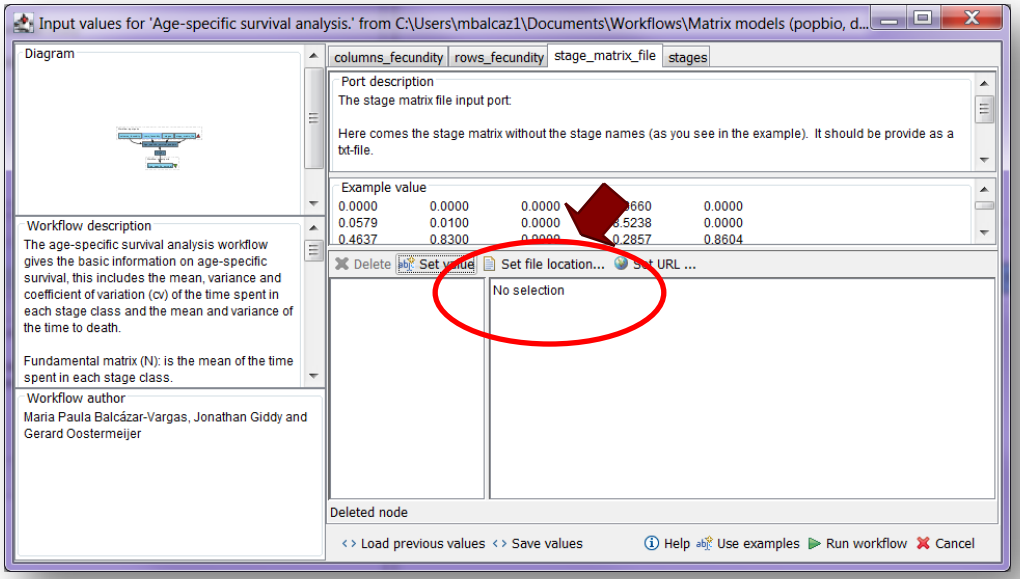


Figure 4. Set the Location of the matrix file.

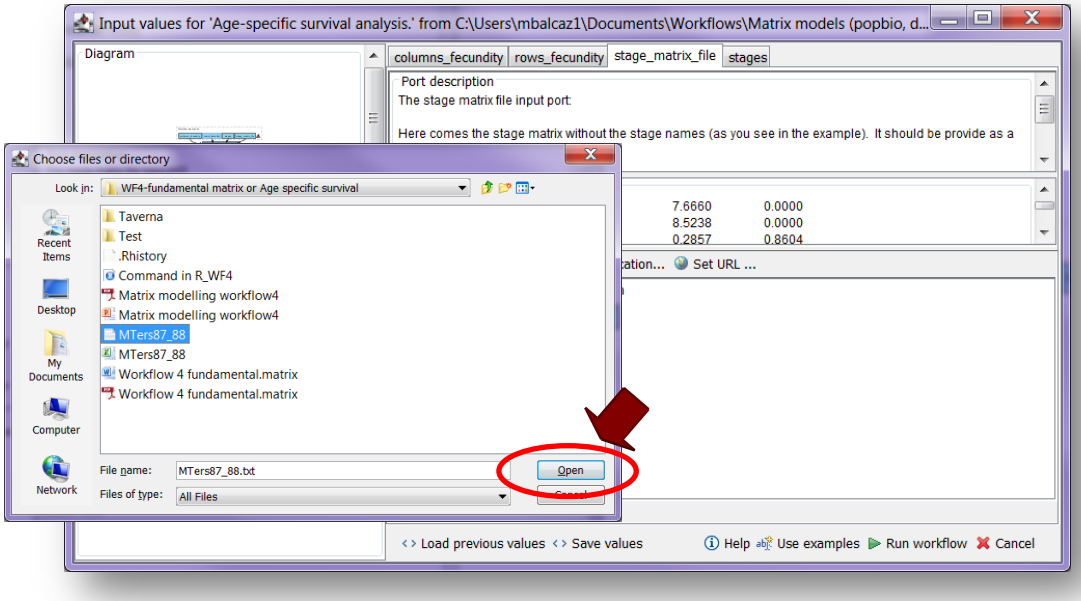


Figure 5. Choose the matrix file and click in Open.

4) ***stages*:** Here come the names of the stages or categories of the input matrix. No longer than 8 characters.

	S	J	V	G	D
S	0	0	0	7,666	0
J	0,0579	0,01	0	8,5238	0
V	0,4637	0,83	0,9009	0,2857	0,8604
G	0	0,04	0,009	0,619	0,1162
D	0	0,03	0,018	0	0,0232

In the following example, the matrix has 5 stages or categories:

e.g. the stages of this matrix are called:

1) Seedlings	S
2) Juveniles	J
3) Vegetative	V
4) Reproductive individuals	G
5) Dormant plants	D

The name of stages or categories must be added one by one. First press add value, fill the name of the stage no longer than 8 characters and press enter; then press add value and fill once again the next name of the stage; repeat the action until you have fill all the names of all stages (Fig 6).

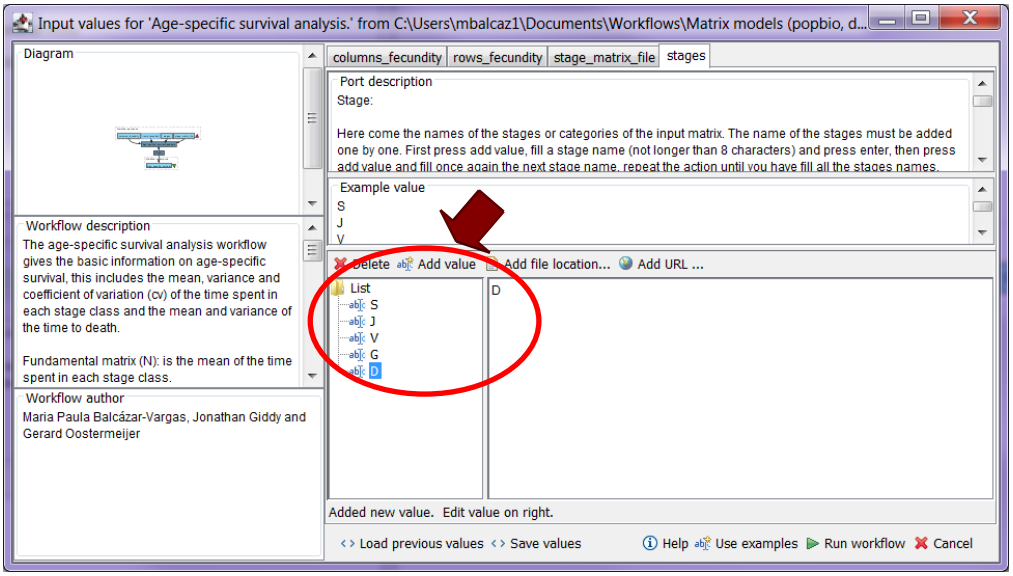


Figure 6. The name of the stages to be filled in.

After the user has filled out the input ports and has clicked the ***Run Workflow button***, the workflow performs the analysis. To complete all the analysis may take few seconds.

When the analyses are completed, they appear on a different window under results in Taverna, the user have to save each output separately.

5.4 Save data/results

5.4.1 OUTPUTS

First, click in the selected result window **Age_specific_survival** (Fig 7). Second, click in the left window on value 1 (blue oval), Third, on the right window click on the save value (red oval). Fourth, name the file and determinate the extension file. For a text file: e.g. .csv or .txt. (Fig 8), Finale, save the file in the chosen map.

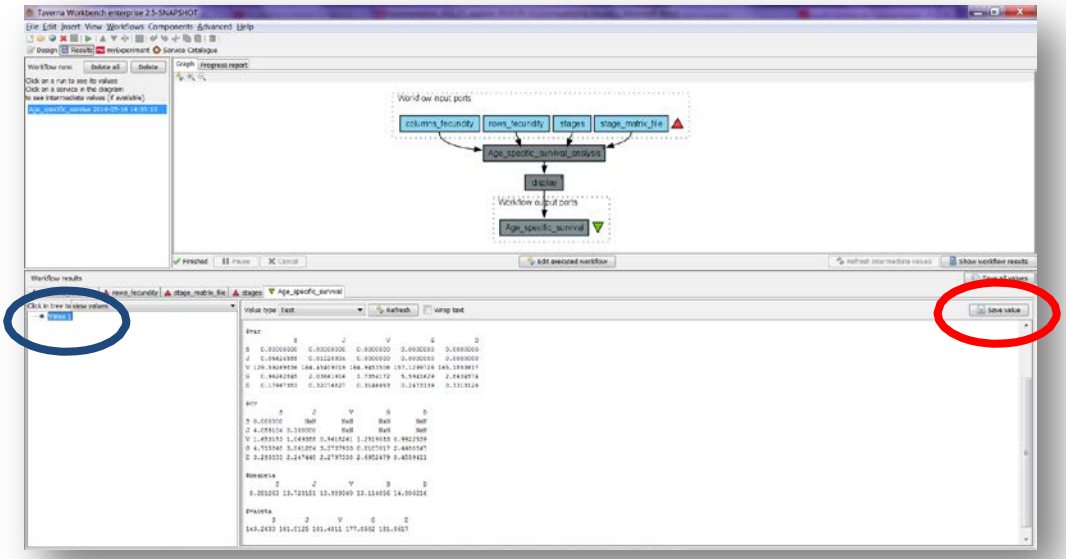


Figure 7. Age_specific_survival output window. Save value on the right window of the Age_specific_survival output results.

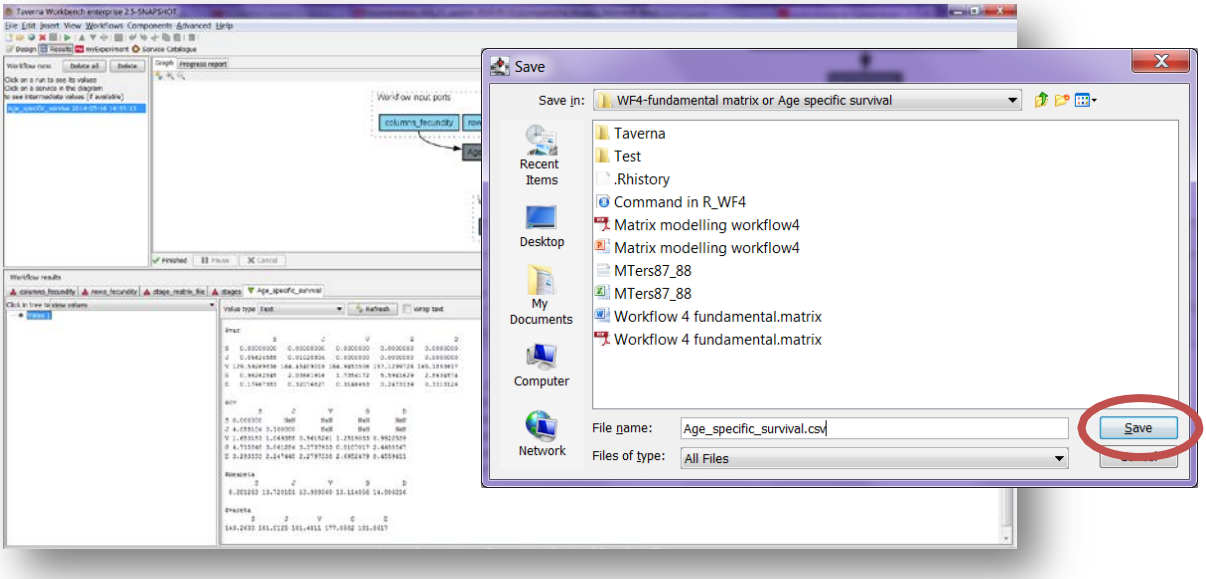


Figure 8. Name the file and determinate the extension file. For a text file: e.g. .csv or .txt.

Age specific survival

The age specific survival workflow gives the basic information on age-specific survival (Fig 9):

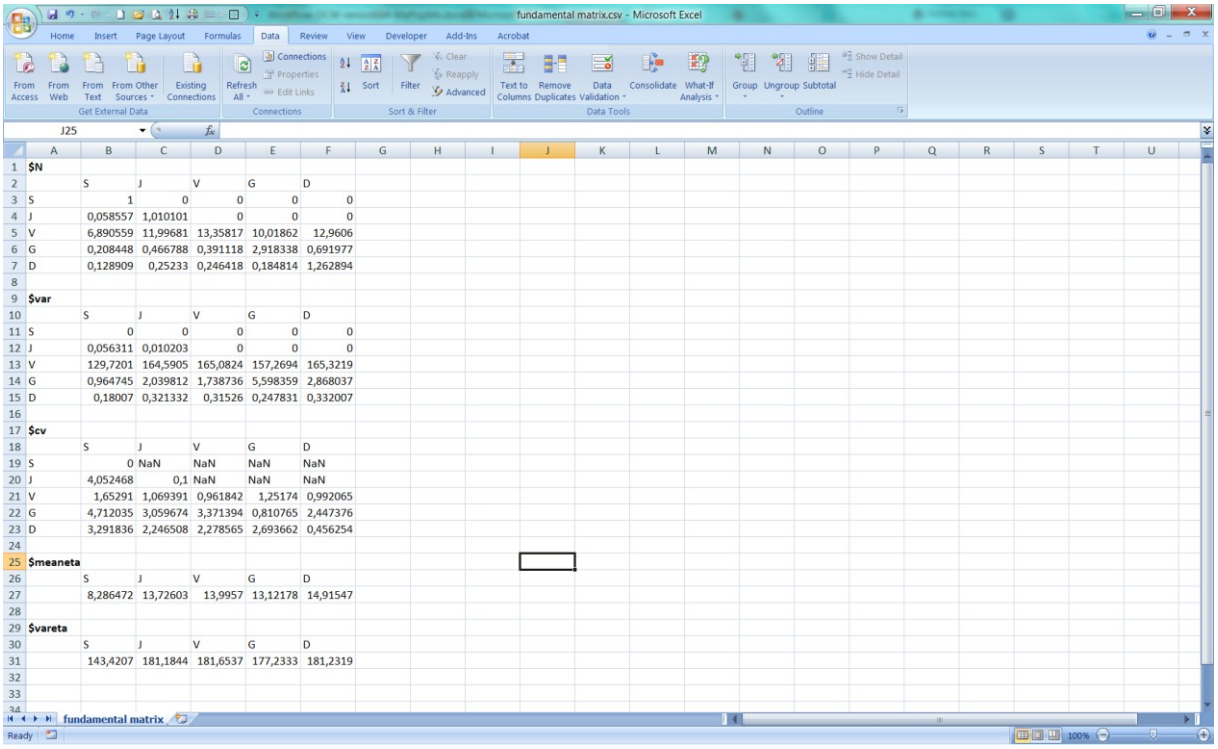
11) ***Fundamental matrix (N):*** is the mean of the time spent in each stage class. e.g.: For our *Gentiana* example means a J plants will spends, on average, about 1 year as a Juvenile, 11 as vegetative plant, less than a year a reproductive and dormant plant.

12) ***Variance (var):*** is the variance in the amount of time spent in each stage class.

13) ***Coefficient of variation (CV):*** is the coefficient of variation of the time spent in each class (SD/mean- the ratio of the standard deviation to the mean).

14) ***Meaneta:*** is the mean of time to death, of life expectancy of each stage. e.g. The mean age at death is the life expectancy; the life expectancy of a new individual seedling is
8 years.

15) ***Vareta:*** is the variance of time to death.



	A	B	C	D	E	F
1	$N					
2		S	J	V	G	D
3	S	1	0	0	0	0
4	J	0,058557	1,010101	0	0	0
5	V	6,890559	11,99681	13,35817	10,01862	12,9606
6	G	0,208448	0,466788	0,391118	2,918338	0,691977
7	D	0,128909	0,25233	0,246418	0,184814	1,262894
8						
9	$var					
10		S	J	V	G	D
11	S	0	0	0	0	0
12	J	0,056311	0,010203	0	0	0
13	V	129,7201	164,5905	165,0824	157,2694	165,3219
14	G	0,964745	2,039812	1,738736	5,598359	2,868037
15	D	0,18007	0,321332	0,31526	0,247831	0,332007
16						
17	$cv					
18		S	J	V	G	D
19	S	0	NaN	NaN	NaN	NaN
20	J	4,052468	0,1	NaN	NaN	NaN
21	V	1,65291	1,069391	0,961842	1,25174	0,992065
22	G	4,712035	3,059674	3,371394	0,810765	2,447376
23	D	3,291836	2,246508	2,278565	2,693662	0,456254
24						
25	$meaneta					
26		S	J	V	G	D
27		8,286472	13,72603	13,9957	13,12178	14,91547
28						
29	$vareta					
30		S	J	V	G	D
31		143,4207	181,1844	181,6537	177,2333	181,2319
32						
33						
34						

Figure 9. ***Age specific survival*** opens in excel.

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 on Packages 'popbio' in R. (Stubben & Milligan 2007; Stubben, Milligan & Nantel 2011) and 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.

Stubben, C & B. Milligan. 2007. Estimating and Analysing Demographic Models Using the popbio Package in R. Journal of Statistical Software 22 (11): 1-23

Stubben, C., B. Milligan, P. Nantel. 2011. Package 'popbio'. Construction and analysis of matrix population models. Version 2.3.1

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