

**BioVeL – Biodiversity Virtual e-Laboratory** 

# **Workflow Documentation**

# Matrix Population Models construction and analyses (Integrated) 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: Project Co-ordinator: Project Homepage: Duration of Project:



283359 Mr Alex Hardisty http://www.biovel.eu 36 months



# **1** Description

The Matrix Population Models construction and analyses Workflow provides an environment for create a stage-matrix with no density dependence and perform several analyses on it.

Analyses:

- Eigen analysis.
- Age specific survival.
- Generation time (*T*).
- Net reproductive rate (*Ro*).
- Transient Dynamics.
- Bootstrap of observed census transitions (Confidence intervals of  $\lambda$ ).
- Survival curve
- Keyfitz's ∆
- Cohen's cumulative distance

# 2 General

# 2.1 Name of the workflow in myExperiment.

Name: Matrix Population Models construction and analyses (Integrated) Workflow for local execution

Download: The workflow pack can be downloaded at myExperiment under the following address:http://www.myexperiment.org/packs/483.htmloronlytheworkflow:http://www.myexperiment.org/workflows/3684.html

To see this documentation online:

https://wiki.biovel.eu/display/doc/Matrix+Population+Models+Workflow

# 2.2 Date, version and licensing

Last updated: 4th July 2013

Version: 20

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 Gerard Oostermeijer unpublished results and publication: Oostermeijer, J.G.B. M.L. Brugman, E.R. de Boer; H.C.M. Den Nijs. 1996.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). Matrix Population Models construction and analyses (Integrated) 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, stage matrix with not density dependence, Lambda ( $\lambda$ ), Sensitivity analysis, Elasticity analysis, Age specific survival, Generation time (*T*), Net reproductive rate (*Ro*), Transient Dynamic, Damping ratio, Bootstrap of observed census transitions (Confidence intervals of  $\lambda$ ), Survival curve, Keyfitz delta, Cohen's cumulative distance.

# **3.2** Scientific workflow description

The aim of the Matrix Population Models Workflow is to provide a connected environment for construct a stage matrix and analyse it. The workflow accepts input data in a .txt format. The output is provided as a set of R results and graphic plots. Currently, the workflow is composed of seven distinct parts and/or analyses (Fig 1a and B):

**1.** Construction of a stage matrix model from demographic monitoring of individuals in an animal or plant population with not density dependence. This section of the workflow create a stage matrix based on two years census data that contains transitions probabilities from each stage to the next, of tagged or mapped individuals of an animal or plant population. The resulting stage matrix is then used to perform the other components of the workflow: the Eigen analyses, age specific survival, generation time (T), net reproductive rate (Ro) and transient dynamics.

**2. Eigen analysis of a stage matrix model.** This component of the workflow performs the Eigen analysis. This analysis results are a set of demographic statistics:

- a) Lambda or dominant eigenvalue ( $\lambda$ ).
- b) The stable stage distribution.

- c) The sensitivity matrix
- d) The Elasticity matrix.
- e) Reproductive value.
- f) The damping ratio.

**3.** Age specific survival. This component of the workflow calculates the age-specific survival: that 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.

- 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.

**4. Generation time (***T***):** This component of the workflow calculates the generation time. The time *T* required for the population to increase by a factor of *Ro* (net reproductive rate). In other words, *T* calculates how much time takes to a plant/animal to replace itself by a factor of *Ro*.

**5.** Net reproductive rate (*Ro*): This section of the workflow calculates the net reproductive rate (*Ro*). Ro is the mean number of offspring by which a new-born individual will be replaced by the end of its life, and thus the rate by which the population increases from one generation to the next.

**6. Transient Dynamics:** This workflow produces plots of the short-term dynamics and convergence to stable stage distribution using stage vector projections.

**7.** Bootstrap of observed census transitions: This workflow calculates bootstrap distributions of population growth rates ( $\lambda$ ), stage vectors, and projection matrix elements by randomly sampling with replacement from a stage-fate data frame of observed transitions. The goal of a demographic analysis is very often to estimate lambda, because lambda is estimated from imperfect data, such estimation are uncertain. Therefore, when the results have policy implications it is important to quantify that uncertainty. Confidence interval is one of the traditional tools to doing so.

**8. Survival curve**: Plots the survival curve.

**9. Keyfitz's**  $\Delta$ : This workflow produces the Keyfitz's delta and the Cohen's cumulative distance, which are measurements of the distance to the stable stage distribution (SSD). Keyfitz's delta is a measure of the distance between any two probability vectors (Keyfitz 1968). In this case is a measure of the distance between n (Observed Stage Distribution) and w (Stable Stage Distribution). The maximum value is 1 and the minimum value is 0 (when the vectors are identical).

**10. 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).



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 including the Interaction Service plugin. The simplest way to install a Taverna Engine is to install Taverna Workbench, and then install the Interaction Service plugin.

The workflow also requires an Rserve installation with the popbio and popdemo packages 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 <a href="http://www.taverna.org.uk/download/workbench">http://www.taverna.org.uk/download/workbench</a>

#### 4.2.2 Taverna Plug-ins

• Please install the Interaction plugin: version 1.0.2+

### 4.2.3 Taverna Dependencies

- Install R software in your computer. See: <u>http://www.r-project.org/</u>
- Start R, and install package Rserve:
  - install.packages("Rserve")
- Install package popbio
  - o install.packages("popbio")
- For package popdemo, as it is archived in CRAN, use the package devtools to install it
  - o install.packages("devtools")
  - o require(de2vtools)
  - install\_url("http://cran.r-project.org/src/contrib/Archive/popdemo\_0.1-3.tar.gz")
- Local R Server: (Rserve) running at port 6311. See <u>https://wiki.biovel.eu/x/3ICD</u> 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 Gentiana87-91.txt, called in myexperiment.org: Demographic data *Gentiana pneumonanthe* 1987-1991 from Terschelling File to construct a stage matrix without density dependence and 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 input data (a .txt-file) has to have the format of a table containing the demographic data on a series of individuals in two or more years. Each individual has a table row for each year and is identified by a number (plant column in example). For the year specified in the column 'year', each individual has a certain life stage ('stage'). The stage codes can be chosen by the user, but they have to be specified with a string. For all individuals, the number of offspring is specified in a chosen way (in the example, 'repstr' gives the number of flowers/fruits for each plant. In the example, it can be seen that only generative

adults (stage=G) had flowers. Individuals without offspring have to be indicated by filling in '0' in this column, which cannot have empty cells. The column 'recruitment' specifies those individuals that are new in the second year census to the population by means of a code. In the example, 'RS' is a new seedling, and 'RJ' a new juvenile plant. As can be seen in the example, no code is needed in this column for individuals that were already present (Table 1).

plant	year	stage	reprstr	recruitment
2	1987	D	0	
4	1987	J	0	
5	1987	D	0	
31	1987	V	0	
36	1987	D	0	
37	1987	J	0	
41	1987	J	0	
43	1987	S	0	
46	1987	V	0	
47	1987	s	0	
2	1988	v	0	
4	1988	D	0	
5	1988	V	0	
9	1988	J	0	RJ
14	1988	J	0	RJ
15	1988	J	0	RJ
16	1988	J	0	RJ
17	1988	J	0	RJ
18	1988	J	0	RJ
24	1988	S	0	RS
25	1988	s	0	RS
31	1988	v	0	
32	1988	s	0	RS
33	1988	J	0	RJ
35	1988	S	0	RS
36	1988	v	0	

**Table 1.** Input file in a .txt format to create a stage matrix.

The workflow accepts input data (Table 1) in a .txt format, All examples data input for the tutorial are available in the PACK: <u>http://www.myexperiment.org/packs/483.html</u> The example data Gentiana87-91.txt for the tutorial is available as well at: <u>http://www.myexperiment.org/files/850.html</u> or here below. The second data Dt1.txt is available at: <u>http://www.myexperiment.org/files/853.html</u> or here below.

#### 5.2.2 Input data

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

#### **Terschelling matrices**

- Gentiana87-91.txt (This table contain data from 87 to 91)
- Dt1.txt (This table only contain data from 87 to 88)

For this tutorial we use the <u>Gentiana87-91.txt</u> file.

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) <u>CensusDataFile</u>: The input data (a .txt-file) has to have the format of a table containing the demographic data on a series of individuals in two or more years. Each individual has a table row for each year and is identified by a number (plant column in example). For the year specified in the column 'year', each individual has a certain life stage ('stage'). The stage codes can be chosen by the user, but they have to be specified with a string. For all individuals, the number of offspring is specified in a chosen way (in the example, 'repstr' gives the number of flowers/fruits for each plant. In the example, it can be seen that only generative adults (stage=G) had flowers. Individuals without offspring have to be indicated by filling in '0' in this column, which cannot have empty cells. The column 'recruitment' specifies those individuals that are new in the second year census to the population by means of a code. In the example, 'RS' is a new seedling, and 'RJ' a new juvenile plant. As can be seen in the example, no code is needed in this column for individuals that were already present (Table 1).

Click in Set file location (red arrow in Fig 2), a window dialog appears and the user selects the file (e.g. <u>Gentiana87-91.txt</u>) and then clicks the Open button. Then the link to the file will appear on the Taverna window.

Choose files or directory	censusDataFile iterations label longTermYears shortTermYears     Port description     The input data file must have the format of a table containing the demographic data on a series of individuals     over several years. Each individual has a table row for each year and is identified by a number (plant column in     pecified in the column year, each individual has a certain life stage (stage'). The stage     the user, but they cannot be longer than 8 characters.	*
Look in: PACK-version2  Recent Annotations-MPM_update 2013-08-01  Annotations-MPM_update 2013-09-27  Annotations-MPM_update 2014-02-05  Annotations-MPM_update 2014-02-05  Annotations-MPM_update 2014-02-05  Computer  Wy Documents Computer  My Decumentation_MPM_C&A_V19_update 2014-04-16  Documentation_MPM_C&A_V19_update 2014-04-16  Documentation_MPM_C&A_V19_update 2014-04-16  Computer  My Decuments Computer  File game: Gentiana87-91.bt  File game: Gentiana87-91.bt	/19.t2flow  /19.t2flow  /19.t2flow  Drag to re-arrange, or drag files, URLs, or text to add  (> Load previous values (> Save values	el

Example of Input page:

Figure 2. Inputs file to be filled in.

2) **BootstrapIterations:** Number of iterations for calculation of bootstrap distributions. Press add value, fill a number and press enter.

e.g.: 1000

3) *Label*: Descriptive title for labelling generated outputs (graphs). Press add value, fill a title and press enter.

e.g.: *Gentiana pneumonanthe*, **Terschelling** (*species taxonomic name*, **place where the research was conducted**)

4) *LongTermYears:* It is the maximal number of iteration or repetitions in the transient dynamic analysis. Press add value, fill a number and press enter.

e.g.: 50

This number of years (50) will be use as well in two type's outputs graphs (see output ports):

1) StageVectorPlotLongTermProportional: the proportion of individuals per stage in the long term.

2) StageVectorPlotLongTermLogarithmic: the number of individuals per stage in the long term.

5) **<u>ShortTermYears</u>**: This value will be used to plot a graph that shows a simulation of the number of individuals per stage in a few years period after the study. Press add value, fill a number and press enter.

e.g.: 10

This value represents the years of axis X of the output graph: StageVectorPlotShortTerm (see output port):

1) StageVectorPlotShortTerm.

After the user has filled out the input ports and has clicked the **Run Workflow button**, a dialogue pages appear in the user's internet browser. See the next section:

#### **5.3.2 INTERACTION**

1) **Stages:** With this dialogue automatically appears the names of the stages or categories of the census data file. When the dialogue appears, the stages are in disorder (Fig 3), so the user drags and organizes the stages according to the order in the life cycle. Then, the author chooses if the stage belongs to the recruited, reproductive category or it should be excluded (Fig 4 and 5). Recruited means that new individuals can be recruited to this (these) stage(s). Reproductive stages are those that reproduce (produce offspring) (in this example the stage G). In the census data file Gentiana87-91.txt, x is use to denote when a plant has died in the second year, so the user must selected in the excluded column.

Then the user clicks in confirm and you will read stages submitted (Fig 3, 4 and 5).

In the following example, the life cycle of *Gentiana pneumonanthe* has 5 stages or categories:

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

Stage	Recruited	Reproductive	Exclude	
D			-	
G				
1				
5				
v				
irm				
îrm				
irm	en click Confirm			
firm stages to sort, th	en click Confirm			
irm	en click Confirm			
firm	en click Confirm			
îrm stages to sort, th	en click Confirm			
îirm stages to sort, th	en click Confirm			
irm	n click Confirm			
irm	n click Confirm			
irm	n click Confirm			

Figure 3. Stages dialogue. When the dialogue appears, the stages are in disorder.

Stage	es			
Stage				
Stage				
	Recruited	Reproductive	Exclude	
S				
J		•	•	
v				
G				
b				
×				
irm			-	
irm	-			
rm			-	
m	click Confirm			
rm tages to sort, then c	click Confirm			
rm tages to sort, then c	click Confirm			
rm tages to sort, then c	click Confirm	•		
rm tages to sort, then c	click Confirm	•		
rm tages to sort, then c	click Confirm	•		
irm	click Confirm	•		
irm	click Confirm	•		
irm	click Confirm	•		
irm	click Confirm	•		
irm	click Confirm	•		
irm	click Confirm	•		
rm	click Confirm	•		
irm	click Confirm	•		
irm	click Confirm	•		
rm	click Confirm	•		
rm	click Confirm	•		
rm	click Confirm	•		

Figure 4. Stages dialogue. Drag each stage to the proper order. The stages should be organized as in the

life cycle by the user (e.g. S, J, V, G, D, x).

Channe	Described	Denneductive	Evoludo	
Stage	Recruited	Reproductive	Exclude	
j j	2			
v				
G		N.		
D				
×	-	· · · · · · · · · · · · · · · · · · ·	2	

**Figure 5.** Stages dialogue. Choose if the stage belongs to the recruited or reproductive category, or should be excluded.

2) <u>Year</u>: In this dialogue view the years of the census data file (Fig 6, example file from 1987 to 1991). Select one or multiple census years. In this example the years 1987 - 1988 and 1988 – 1989 have been chosen (Fig 7 and 8). That means that the analysis will create two matrices, one for the interval 1987 - 1988 and a second one for 1988 - 1989. After this step, the user should confirm the selection (Fig 8).

Taverna interaction	Taverna interaction	× +		Charles and
localhost 3080/interaction/interactionc427ea68d2	26461a8478cd22c3b89799.html		Y v C Google	P 🖬• 🟦 💿 •
ect Transitions				2
enerate Stage Matrices for:				
1987 - 1988				
1988 - 1989 1989 - 1990				
1990 - 1991				
ofirm				
ct year range to generate stage	matrices for transition C			
	matrices for transition. Ci	lick Confirm when done.		
	matrices for transition. Cl	lick Confirm when done.		
	matrices for transition. Cl	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		,
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		
	matrices for transition. C	lick Confirm when done.		

**Figure 6**. Year's dialogue. When the dialogue appears, the user selects the year interval of one or more year matrices.

L lavena nonación	Taverna interaction	* +			E	20102	-
Callocathost 3080/Interaction/Interactionc427ea0	58d226461a8478cd22c5b89799.html		습 = C 🔂 * Google	P	₽.	ŧ	1
lect Transitions							-
enerate Stage Matrices for:							
1987 - 1988							
1988 - 1989							
1989 - 1990							
onfirm							
act year range to generate sta	age matrices for transition.	Click Confirm when done.					
ice year range to generate sta	age maches for cransicon.	click commit when done.					
							11
							.#*
							. H
							. #
							ц
							- # -:
							# ^ ·
							#
							4.
	The interaction	service was developed in the <u>BioVel. pr</u>	oject				#

Figure 7. Year dialogue. The user selects the year interval, in this case 1987-1988, for a year matrix 1987-1988.



**Figure 8**. Year dialogue. The user can select other year, in this case 1988 -1989, to generate one more matrix, a matrix from the year 1988-1989. The user confirms the choice.

3) <u>Initial abundance</u>: In this dialogue appears the fields with the initial abundance per stage observed in the field (Fig 9). As an example *Gentiana pneumonanthe* has 5 stages with its respective abundance. Abundances of the year 1987:

Stage abundance:

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

Taverna interaction	Taverna interaction	Taverna interaction	* +		*
S localhost 8080/interaction/interaction8p3e	P2da1764b36aa0bba176ec9d2bc.html		לְיָי ל 🕄 * Google	P 🖬 🕈	21.
et Abundanco	e				
tial year: 1987					
69					
100					
111					
21					
43					
onfirm					
20000					
	and the second se				

**Figure 9.** Abundance dialogue. The abundance per stage appears. This dialogue will appear for each selected year. The user confirms the field.

After the user confirms the field, the workflow performs the analysis. To complete all the analysis may take several minutes.

When the analyses are completed, they appear on different windows 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 e.g. **stage\_matrix** (Fig 10). Second, click in the left window on value 1 (Fig 11), Third, on the right window click on the save value (Fig 12). Fourth, name the file and determinate the extension file. For a text file: e.g. .csv or .txt. For images: .jpg or png (Fig 13), Finale, save the file in the chosen map.

state of the state	NAPSHOT	Contraction of the second s		
le Edit Insert View Workflows Com	ponents Advanced Help			
Design 🔚 Results 🔤 rryCiperiment 🧔	ゆ 壬 聖道 (音) Service Catalogue			
Iorkfow runs Delete all Delete	Groph Progress report			
ick on a run to see its values	4		-	
inae merendete values (f. eventete) Grund Constanting (f. eventete)				
	·			+
	V Finaned II Provide 26 Cancel	Secure worklow	Kolivsk internedule values	Show workflow results
W airgananalous sansitivity matrix 1	V aineronabolis sensitivity matrix 2 V alestrity	matrix V alasticity nint V fundamental matrix V apparation time V historyay	m ei trioc of lambda 🛛 👻 instite dalta 🔍 kon term k	warthmic stans writer rint
vecesanalysis sensitivity matrix 1     ▲ consultativite     vensultativite     vector p     log_term_proportional_stage_vector_p     lok in their to view values	▼ elsensmahsis sensthuty_matrix_2   ▼ elasticity ▲ label ▲ longTermYears ▲ shortTermYe sot ▼ net_reproductive_rate ▼ population_proj ▼ value type   Ted	matrix (V elastich; pbt) (V luddreasti) matrix (V elevanios time (V haboara ar V ochen; cumulative, distance (V confidence, intervel, Kiec, of lambda cottom (V relacios, matrix) (V sensitive)_point (V short_term_st v) (A Bafrash (V vap text)	n ci 93oc of Jambda V unfiz delta V long term 1 V damping ratio V elecenanolysis V ele age_vector_plot V stable_stage_statroution V scaje_m	otarithmic states vector plot manalysis slasticity motive at b second of the states of the Save value

Figure 10. Stage-matrix output window.

Taverna Workbench enterprise 2.5-5	NAPSHOT	and the second s	regulation of the	
e Edit Insert View Workflows Con	nponents Advanced Help			
※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※ ※	や子物語(音)			
f Design 🔚 Results 🧮 myCoperiment 🔇	Service Catalogue			
Ioridfow runs Delete all Delete	Groph Progress report			
ick on a run to see its values	4 4 9			
ck os a service is the diagram				÷
atro, Population, Mo 2014-04-16-14-59-25				
			一座 月在主	
		file sected worklow	Artivisk interniedativ vid	+ + Show workflow results
Noriflow results	e [ m ef Finand   H Pinne   2K Cancel		Ar fait visit intermediate visit	+ + Show workflow results C Sove all velues
Workflow results ▼ eleesanalysis, sensibirite, matrix, 1 4 connuctusTate & benations 5 interes to wardshift 1 in there to two wardshift	V Preaked III Price X Cancel V Preaked III Price X Cancel A label A lossTerritors A sheet pot V net_reproductor_max V populate	An and a second	A fathrada meramentative value     A fathrada meramentative value     A fathrada meramentative	stan stan verkfor results Store verkfor results Store vit volues rm. loaarthmic, slastory, metu ge_math V sarved, oure, giet
Workflow results The idease adapts - Sensitivity, matters, 1 A consultativitie A becabese Version - Sensitivities - Sensitivities C is the two is provided is the web into accurate is the web into accurate	Y Presided III Prane: X Cancel     V Presided III Prane: X Cancel     V observations. A short      V occurrentiation and      V occurrentiations. A short      V occurrentiation. A short      V occurrentiation.      V occurrentiations. A short      V occurrentiation.      V	Add conclusion of the other sector of the	Ar fathrach mannedate val     Ar fathrach mannedate val     Ar fathrach mannedate val     Arbannong and     X datamong an	ses Show workfow results Show Youlase Importantial states were able results and show you want able results and show you want able results and show you want able in the show you

Figure 11. Value 1 on the left window of the stage\_matrix output results.

ie Edit Insert View Workflows Compo 1 ● ④ ★ ■ ▶ ▲ ♥ ☆   ■   ♥ ☆ = F Design ■ Results ■ m/Ciperiment ۞ Serv	write Advanced Hain			
📴 🥥 🗶 🎆 🕨 🔺 🦞 🔅 📰 🖗 🤤 e	ents Brivancer Deb			
Design 🛅 Results 🧮 myCiperiment Q Serv	- No. (1)			
	see Catalogue			
Vorkfow runs Delete all Delete	Progress report			
ki de anno beelt vielet de la construction de la construction de la co				
-	m. Firished II Pause X Cancel	tdt. executed workflow	Autresh intermediate values	t Show workflow results
Workflow results				DID See all unline
A consuddrafile A constent A to V long term proportional stage_vector_plot lick in the to view values if is with 1 value is a vanisation	bel         A short reminer:         Y           Y net_reproducerele         Y populosrelection         Y           Y alies their face         Y         Y	orbin, considered, distance   V enforce, record, Sice, d. Jacobá version, construit (V encoder, particity), port (V siter, term, ubage Staffment)   V way too: S 0 S 0 S 1000000 S 100000 S 1000000 S 100000 S 10000000 S 100000 S 100000 S 100000 S 100000 S 100000 S 1000000 S 10000000 S 10000000 S 10000000 S 1000000 S 1000000 S 1000000 S 1000000 S 10000000 S 100000000 S 1000000000 S 1000000000000000000000000000000000000	▼ derepon_into   ▼ escenetidos   ▼ dosmanta yester_pot   ▼ dobe_stage_bitroutos   ▼ δογε_non	V survival, curve _plac

Figure 12. Save value on the right window of the stage\_matrix output results.

pre Edit Insert View Workflows Com T == ♥ ★ ■ I ► A ♥ ⊕ I ■ Ø ♥ F Cesign ■ Resulta ■ myCoperiment ♥ :	onnerts anvanced teep + Re R   B   ervice Cotalogue			
Norkfow runs Delete all Delete	Groph (Hogress report) Generation (Control of Control o	🛃 Save		
The and memory of a sched of you can be a set		Save jn: Save j	PACK-version2     Annotations-MPM_update 2013-08-01     Annotations-MPM_update 2013-09-27     Annotations-MPM_update 2013-09-27     Annotations-MPM_update 2014-02-05     Annotations-MPM_update 2014-02-05     Annotations-MPM_update 2014-02-05     Documentation_MPM_v19_update 2014-04-16     Dt1     Gentiana87-91     matrix_population_models_construction_and_analy     matrix_population_models_construction_and_analy     Workflow-MPM_update 2013-08-06	→
= i traviti   vice	Vela 5 pro (e4) 9 5.0000000 5.0 6.000000 7 5.1000000 5.0 6.000000 7 5.0000000 5.4 5.000000 8 5.0000000 5.4 5.0000000 8 5.0000000 5.4 5.0000000	• Urbanisa         • Urba laci           • 1.454488         0.0000000           • 1.454488         0.0000000           • 1.454488         0.000000           • 1.454488         0.1427492           • 1.455488         0.1427492	File pame: Stagematro.td	, Sove

**Figure 13**. Name the file and determinate the extention file. For a text file: e.g. .csv or .txt. For images: e.g.: .jpg or png.

#### Construction matrix

1) **Stage\_matrix**: a stage matrix contains transitions probabilities from each stage to the next. In this example the selected species is *Gentiana pneumonanthe* has a matrix of 5 x 5 stages (Fig. 14; Oostermeijer et al., 1996).

Htt	но	ine	insert Pa	ge Layour 🛛 Fo	rmulas Data	Review	otw Add-215	Acrobat		- Internet	NUMBER OF T	-				_	_	0	0=0	β - Σ
A) 155	From Web	From Text	From Other Sources -	Existing Corrections	Refresh All - Ecit	entico Losici M	Son Filter	The Chenr To Prescoly Ty Advanced at	Text to Remove Columns Duplicate	Data S validation	Consolidate W	tue d Nord	Group U	ngroup Subtra	*3 thow Det *3 Hide Deta #	4				
-	09	VILL	- (*	fr	60-1600					0.000				0.0.0						- 4
1	A		В	C	D	F	F	G	н	E.	i i	,	(	L.	M	N	0	р	0	Ċ
			s	1	v	G	D							-					~	
s				0 0	0 0	7.666667	0													
J			0.057971	0.01	0	8.285714	0													
V	0		0,463768	0.83	0,900901	0,285714	0,850465													
G	;		0	0,04	0,009009	0,619048	0,116279													
D	)		c	0,03	0,018018	0	0,023256													
																				4
i	* st	200	natrix /									4								ari.
	1.00	-															time of the lot of the	100 C	D	100

Figure 14. stage\_matrix output file open in excel.

Furthermore this workflow creates the:



2) *Projection\_matrix*: This creates a grid of coloured rectangles to display the *stage\_matrix* (Fig 15).

Figure 15. *Projection\_matrix* opens as a coloured graph.

## 3) Eigen\_analysis

- a) **Lambda or dominant eigenvalue**: This value describes the population growth rate of a stage matrix. The population will be stable, grow or decrease at a rate given by lambda: e.g.:  $\lambda = 1$  (population is stable),  $\lambda > 1$  (population is growing) and finally  $\lambda < 1$  (population is decreasing).
- b) **The stable stage distribution** (w): It is the proportion of the number of individuals per stage. It is given analytically by the right eigenvector (another property of the transition matrix) that corresponds to the dominant eigenvalue

*Elasticity and Sensitivity*: Sensitivity and elasticity analyses are prospective analyses.

c) **The sensitivity matrix**: The sensitivity gives the effect on  $\lambda$  of changes in any entry of the matrix, including those that may, at a given context, be regarded as fixed at zero or some other value. The derivative tells what would happened to  $\lambda$  if  $a_{ij}$  was to change, not whether, or in what direction, or how much, aij actually change. The hypothetical results of such impossible perturbations may or may not be of interest, but they are not zero. It is up to you to decide whether they are useful (Caswell 2001).

When comparing the  $\lambda$ -sensitivity values for all matrix elements one can find out in what element a certain increase has the biggest impact on  $\lambda$ . However, a 0.01 increase in a survival matrix element is hard to compare to a 0.01 increase in a reproduction matrix element, because

the latter is not bound between 0 and 1 and can sometimes take high values. Increasing matrix element a14 (number of S (seedlings) the next year produced by a G (Reproductive individuals)) with 0.01 from 7.666 to 7.676 does not have a noticeable effect on  $\lambda$ . For comparison between matrix elements it can therefore be more insightful to look at the impact of proportional changes in elements: by what percentage does  $\lambda$  change if a matrix element is changed by a certain percentage? This proportional sensitivity is termed elasticity (Description based on Oostermeijer data, based on Jongejans & de Kroon 2012).

- d) **The Elasticity matrix**: The elasticities of  $\lambda$  with respect to the stage are often interpreted as the "contributions" of each of the stages to  $\lambda$ . This interpretation relies on the demonstration, by de Kroon et al (1986) that the elasticities of the  $\lambda$  with respect to the stage, always sum to 1. For further information see: de Kroon, et al., 1986. and Caswell 2001.
- e) **Reproductive value** (v): scaled so v[1]=1. To what extent will a plant or animal of a determinate category or stage , contribute to the ancestry of future generation.
- f) The damping ratio: it can be considered as a measure of the intrinsic resilience of the population, describing how quickly transient dynamics decay following disturbance or perturbation regardless of population structure, the larger the p, the quicker the population converges.

4 L ()	the second second	-	-	-	And Personnel of Concession, name	eig	en analysis -	Microsoft Exc	0			-	-			- 0	*
Home Insert	Page Layour F	ormulas Data	Review	View Add	f-Ins Acroo	at										000	8-X
From From Fram Cth s Web Text Sources Get Extensil Data	er Existing Corrections	Retresh All - Ecit	ections 11 ection 11 unic: 11	Sort He	The Court So Realing So Advan	Ar Tee to Column	Remove s Duplicates vi	Data Com Ridation -	alder Anet Anitiss	Group Ungro	up Subtotal	thow Detail Hide Detail					
A1 - (*	for Sta	mbda1															4
A B	C D	) E	F	G	н	1	1	к	LA	A N	0	р	0	B	\$	Т	C
Slambda1 [1] 1.232338																	
S I V	G D																
0.14218294.0.16165957	0.05944861.0	02285525.0.0	383863														
Ssensitivities 5 J V 5 0.0000000 0.000000 1 0.14255579 0.1620878 V 0.08206359 0.093307 6 0.0000000 0.3355759	G D 0.00000 0.00 8 0.00000 0.00 4 0.38050 0.0 6 11.28901 0.3	06042076 0.000 (2914380 0.000 13190880 0.00 (91255815 0.23	000000 00000 498595 590160 846721														
0 0.00000000 0.002.007	0 1.55057 0.0	00000000000	PHOTES														
\$elasticities \$ 1 V \$ 0.00000000 0.00000	G D	00 0.03758918	7 0.000000	000													
1 0.006705037 0.001315	5287 0.000000	00 0.15406651	0.0000000	000													
V 0.030883150 0.06284	4075 0.278237	57 0.00305827	1 0.005578	5792													
G 0.00000000 0.08983	2455 0.082528	331 0.19654184	8 0.022355	\$201													
0 0.00000000000000000000000000000000000	00100.015655	555 0.00000000	0.000033	215													
Srepro.value 5 J V G 1.000000 3.792458 2.1	5 D 183170 64.755	197 7.781288															
Sdamping ratio [1] 2.0902																	
																	-1
																	- 6
+ eigen analysis	<b>1</b> 2								4					1/1C		1	1
de l															100% (-)		-116

Figure 16. *Eigen analysis* output file open in excel.

4) **Bartplot**: A bar plot which shows the stable stage distribution (w) of the analysed matrix. In other words, it plots the proportion of individuals per stage (Fig 17).



Figure 17. Bartplot output file.

5) *Elasticity, Sensitivity 1 and Sensitivity 2 matrices:* Creates a grid of coloured rectangles to display the *elasticity\_matrix* (Fig 15), *sensitivity\_matrix 1* calculates the sensitivity values of non-zero elements and *sensitivity\_matrix 2* calculates the sensitivities of all elements in the matrix (Fig 18, 19 and 20).



Figure 18. *elasticity\_matrix* opens as a coloured graph.



Figure 19. sensitivity\_matrix1 opens as a colour graph.



Figure 20. sensitivity\_matrix2 opens as a coloured graph.

#### 6) Age specific survival

The age specific survival workflow gives the basic information on age-specific survival (Fig 21):

a) **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.

- 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.g. The mean age at death is the life expectancy; the life expectancy of a new individual seedling is 8 years.
- e) *Vareta:* is the variance of time to death.



Figure 21. Age specific survival opens in excel.

7) *Generation time (T)*: Generation time is the time *T* required for the population to increase by a factor of *Ro* (net reproductive rate).

e.g. 8.055186

8) **Net reproductive rate (Ro):** The net reproductive rate is mean number of offspring by which a newborn individual will be replaced by the end of its life, and thus the rate by which the population increases from one generation to the next.

e.g. for Generation time and Net reproductive rate: If T = 8.05 and Ro = 5.46.

The average plant of the species *Gentiana pneumonanthe* in Terschelling in the year 1987 replaced itself with five new plants and took approximately 8.05 years to do so.

## 9) Transient dynamic analysis

#### Population projection (Fig 22).

- a) Lambda or dominant eigenvalue: The population will be stable, grow or decrease at a rate given by lambda: e.g.:  $\lambda = 1$  (population is stable),  $\lambda > 1$  (population is growing) and finally  $\lambda < 1$  (population is decreasing). e.g. The projected population growth rate ( $\lambda$ ) is 1.237, meaning that the population is projected to increase with 23% per year if these model parameters remain unchanged.
- b) **The stable stage distribution (stable.stage)**: It is the proportion of the number of individuals per stage and it is given by (w).
- c) **Stage vector (stage.vectors)**: it is the projection of the number of individuals per category per year in the long term, the long term was stipulated in the input port (long term years) (e.g. 50 years).
- d) **Population sizes (pop.sizes)**: it is the total population size per year in the long- term (e.g. 50 years).
- e) **Population change (pop.changes):** are the lambda values per year in the long-term (e.g. 50 years).

10) **Damping Ratio:** The ratio between the dominant eigenvalue and the second highest eigenvalue of a transition matrix is called the damping ratio, and it can be considered as a measure of the intrinsic resilience of the population, describing how quickly transient dynamics decay following disturbance or perturbation regardless of population structure, (the larger the damping ratio, the quicker the population.

converges). High damping ratios tell you that the dominant stable stage distribution is reached fairly soon.

e.g. damping ratio= 2.0901.

1	<b>B</b> 1			Come	diors 11		W. Cour Ye Feat		81		B)=	12	•81 •81	翻 :	Show Detail Hide Detail					
s From ss Web	Text Source Set External Da	ther Eind is Corne ta	tions All	connections	- <sup>21</sup>	Sort Filte	y Adra Filter	column	s Duplicates	Validation Data Too	Consolidate	Analysis *	Group Ungroup	o subicial Outline	6					
H7	- (	f.																		¥
A	8	с	D	E	F	G	н	1	1	к	L	M	N	0	p	Q	R	5	т	-
Slambda																				- 6
1]	1,232338																			
5stable.st	age																			
	S	1	V	G	D															
0,142188	0,10107	0,059449	0,022855	0,013839		-														
Cotanan Lon	1					-	_													
harmRe.ve	0	,	2	3	đ	5	6	7												
	60	161	176 3333	187 2207	210.0905	261 924	318 2264	2233 085												
1	100	179	201.6948	214.5771	249,7801	298.2718	362.0884	442,9598												
/	111	258	467,4033	685,9864	903,753	1149,347	1437,187	1783,397												
G	21	23	24,42009	28,57702	34,164	41,50779	50,7972	62,39105												
D	43	6	10,15818	14,70876	19,13949	24,22235	30,22041	37,46071												
	8	9	10	11	12	13	14													-
\$	478,3314	588,5237	724,7047	892,7536	1099,981	1355,435	1670,287													
J	543,9605	669,2132	824,0306	1015,091	1250,704	1541,154	1899,142													
v	2204,992	2721,568	3356,41	4137,717	5099,941	6285,367	7746													
G	76,76396	94,5267	116,4461	143,4758	176,7958	217,8635	268,4762													
D	46,29325	57,12499	70,44214	86,83496	107,0256	131,9009	162,5519													
e.	2058 218	1636.53	1435 836	38	4742.058	5840.076	7300 146													
1	2340.337	2884.058	3554.118	4179.865	5197.468	6651.501	8196.895													
v	9545,87	11763.84	14497.09	17865,36	22016,18	27131,38	33435,04													
G	330,8504	407,7177	502,445	619,1814	763,0404	940,3234	1158,796													
D	200,3221	246,8664	304,224	374,9074	462,013	569,3564	701,6396													
	22	23	24	25	26	27	28													
5	8884,104	10948,22	13491,9	16626,59	20489,57	25250,08	31116,63													
() (i)	10101,34	12448,27	15340,47	18904,65	23296,92	28709,67	35380,02													
*	41203,28	50776,36	62573,64	77111,88	95027,89	117106.5	144314,8													
5	1428,028	1759,814	2168,685	2672,553	3293,488	4058,691	5001,679													
0	864,6572	1065,55	1313,118	1618,205	1994,175	2457,498	3028,468													
* * po	pulation pro	ection	2									4					No.		1	+
																		120% 💬		.(+)

Figure 22. Population projection opens in excel.

11) <u>stageVectorPlotLongTermProportional</u>: Is the proportion of individuals per stage in the long term (e.g.: 50 years) (Fig 23)



Figure 23. Stage vector plot long term proportional opens as a graph.

12) stageVectorPlotLongTermLogarithmic: Is the number of individuals per stage in the long term (e.g.: 50 years) (Fig 24).



Figure 24. Stage vector plot long term logarithmic opens as a graph.

13) **<u>stageVectorPlotShortTerm</u>**: It is a plot that charts the number of individuals per stage vs. years in the short-term (e.g. 5, 10 years). This value (5 or 10 years) is related to the short-term years input value. (Fig 25).



Figure 25. <u>Stage vector plot short term</u> opens as a graph.

## Bootstrap analysis

14) **Confidence\_interval**: 95% Confidence interval of Lambda. In statistics, a confidence interval (CI) is a type of interval estimate of a population parameter and is used to indicate the reliability of an estimate. It is an observed interval (i.e. it is calculated from the observations), in principle different from sample to sample, that frequently includes the parameter of interest if the experiment is repeated. How frequently the observed interval contains the parameter is determined by the confidence level or confidence coefficient

e.g.:	2.5%	97.5%
	0.9566	1.4648

15) *CI\_histogram:* Histogram plotting the frequencies of the lambda values and the 95% confidence intervals resulting from the bootstrap analysis (Fig 26).



Figure 26. The confidence interval histogram opens as a graph.

**16**) *Survival curve:* A plot of survival curve, where the probability of survival is plotted per stage. (Fig. 27)



Figure 27. Survival curve.

**17)** Keyfitz's  $\Delta$ : Keyfitz's delta is a measure of the distance between n (Observed Stage Distribution) and w (Stable Stage Distribution). Its maximum value is 1 and its minimum is 0 (when the vectors are identical).

e.g. 0.3367

**18).** Cohen's cumulative distance 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.

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: <u>https://wiki.biovel.eu/display/BioVeL/Glossary</u>

# 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. 1986. Life cycle models for plants. Lectures on Mathematics in the Life Sciences 18: 171-233.
- **Caswell**, H. 2001. Matrix population models: Construction, analysis and interpretation, 2<sup>nd</sup> Edition. Sinauer Associates, Sunderland, Massachusetts.
- **Cohen**, J.E. 1979. The cumulative distance from an observed to a stable age structure. *SIAM Journal on Applied Mathematics*, 36:169–175.
- de Kroon, H. J., A. Plaiser, J. van Groenendael, and H. Caswell. 1986. Elasticity: The relative contribution of demographic parameters to population growth rate. Ecology 67: 1427-1431.
- **Horvitz, C., D.W. Schemske, and Hal Caswell.** 1997. The relative "importance" of life-history stages to population growth: Prospective and retrospective analyses. In S. Tuljapurkar and H. Caswell. Structured population models in terrestrial and freshwater systems. Chapman and Hall, New York.
- Jongejans E. & H. de Kroon. 2012. Matrix models. Chapter in Encyclopaedia of Theoretical Ecology (eds. Hastings A & Gross L) University of California, p415-423
- Keyfitz, N. 1968. Introduction to the Mathematics of Populations. Addison-Wesley, Reading, MA, USA.
- Mesterton-Gibbons, M. 1993. Why demographic elasticities sum to one: A postscript to de Kroon et al. Ecology 74: 2467-2468.
- **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., S. Townley and D.J. Hodgson** 2011. A framework for studying transient dynamics of population projection matrix models. Ecology Letters 14: 959–970

- **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
- van Groenendael, J., H. de Kroon, S. Kalisz, and S. Tuljapurkar. 1994. Loop analysis: Evaluating life history pathways in population projection matrices. Ecology 75: 2410-2415.
- 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