

BioVeL – Biodiversity Virtual e-Laboratory

Workflow Documentation

Integral Projection Model (IPM) construction and analyses Workflow for local execution

August 2012

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



Integral projection models construction and analyses Workflow for local execution

1 Description

The Integral Projection Model Workflow provides an environment for create an integral projection model and to calculate its population growth, structure, sensitivities and elasticities analyses, age-specific metrics, etc.

Create an IPM with no density dependence and perform several analyses on it.

Analyses:

- ✚ Lambda (λ).
- ✚ Sensitivity analysis.
- ✚ Elasticity analysis.
- ✚ Parameters elasticity and sensitivity of Lambda
- ✚ Parameters elasticity and sensitivity of R_0
- ✚ Parameters elasticity and sensitivity of Life Expectancy
- ✚ Survivorship
- ✚ Mean Life expectancy
- ✚ Passage Time

2. General

2.1 Name of the workflow and MyExperiment identifier

Name: Integral Projection Model (IPM) construction and analyses Workflow for local execution

Download: The workflow pack can be downloaded at myExperiment under the following address

<http://www.myexperiment.org/packs/482.html>

2.2 Date, version and licensing

Last updated: 6 August 2013

Version: 34

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 results reported in this publication (tutorial) come from data (Dr. Roberto Salguero-Gómez unpublished results and publication: Salguero-Gómez R. 2011. Physiological bases of plant shrinkage and its demographic implications. PhD thesis. University of Pennsylvania Press, USA. 3.). Through BioVeL workflows and services (www.biovel.eu). The Integral Projection Model construction and analyses

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

Integral Projection Models, Lambda (λ), Sensitivity analysis, Elasticity analysis, Parameters elasticity and sensitivity of Lambda (λ), Parameters elasticity and sensitivity of R_0 , Parameters elasticity and sensitivity of Life Expectancy, Survivorship, Mean Life expectancy, Passage Time.

3.2 Scientific workflow description

The aim of the Integral Projection Models (IPM) Workflow (Fig. 1) is to provide a connected environment for construct an IPM matrix and analyse it. The workflow accepts input data in a .csv 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:

1. **Identify Fields:** In the first section of the workflow the user must indicate the fields that provide fecundity information for the initial time period. And the user can add labels for fields to be used in output plots.
2. **Create Vital Rates:** In this section the functional forms are plotted, and the survival, growth and fecundity models are scored and selected.
3. **Create IPM Matrix:** Here the IPM is constructed, the diagnostics of the P matrix are run, and P and F matrices are plotted.
4. **Results & Analysis of IPM:** This section of the workflow shows the results:
 - a) Lambda or dominant eigenvalue (λ).
 - b) The numerical matrix and the IPM graph.
 - c) The numerical Sensitivity matrix and the Sensitivity graph.
 - d) The numerical Elasticity matrix and the Elasticity graph.
 - e) Vital rates of Sensitivity and Elasticity.
 - f) Survivorship
 - g) Mean Life Expectancy and Passage time.

Integral Projection Model (IPM) construction and analyses Workflow for local execution

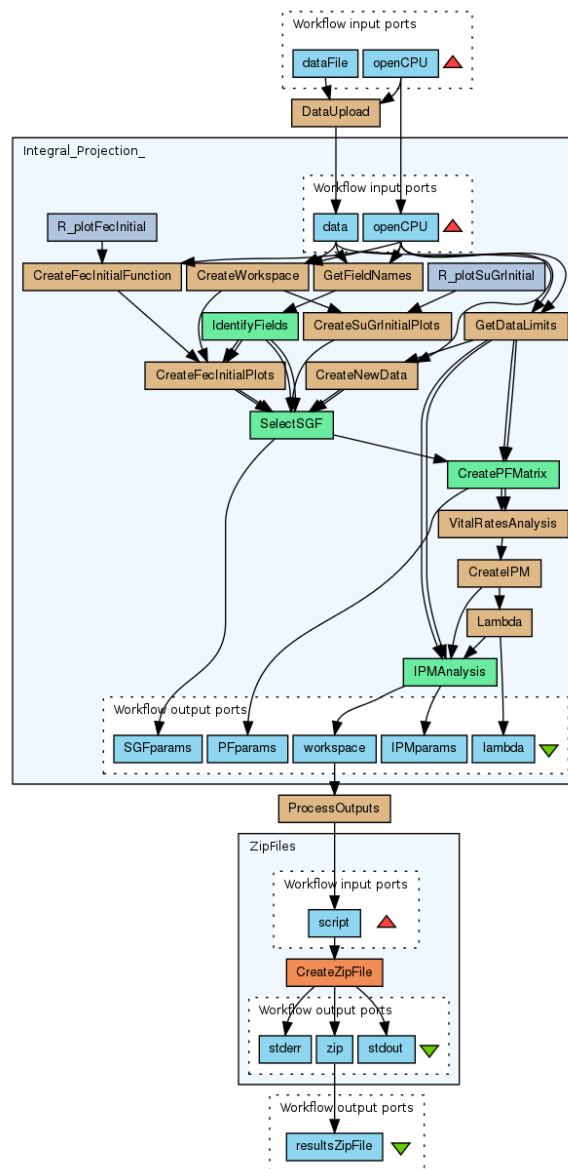


Figure 1. A) Graph simplification of the Integral Projection Model (IPM) workflow.

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

4. Technical specifications scientific 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 OpenCPU server with the IPMpack, Matrix, MASS, nlm, mvtnorm, methods, MCMCglmm, Truncnorm, lattice, graphics, grDevices packages installed.

4.2 Taverna installation, including updates and plugins

4.2.1 Taverna Workbench installation

You need to have the Taverna workbench (v2.4) installed on your computer. For installation files and instructions, please go to <http://www.taverna.org.uk/download/workbench/2-4/>

Please make sure you install the updates under Advanced -> Updates and plugins.

4.2.2 Interaction Plug-in

In addition, you need to install the Taverna Interaction plug-in, version 1.0.2 or higher.

Install the Interaction Service Plugin from Taverna Workbench menu Advanced > Updates and Plugins

4.2.3 Taverna Dependencies

The dependencies on the Taverna side include, [Interaction Plugin](#), which handles the interaction between browser applications and Taverna.

5. Tutorial

5.1 General

In this tutorial, you will use an input file Crypta.csv and other input data to construct an Integral Projection Model without density dependence and perform several analyses on it. Finally you will be able to export your results to different formats.

Please note that the workflow as well as the tutorial is a beta version, and is likely to have failures. Furthermore, we only explain a few analyses here, while many new analyses will be added and also made easier in the coming months. We hope you will still find it interesting to work with this workflow and not give up in despair, and looking forward to your feedback.

5.2 Annotation Construct an Integral Projection Model and analyse it.

The first step is to fill out the input ports:

5.2.1 INPUTPORTS

- 1) **dataFile**: The data must be organized in a table (in .csv format), where each row represents one observation of an organism in the population at one census time t with the following column names:

- ✦ **size**: size of individuals in census time t
- ✦ **sizeNext**: size of individuals in census time $t + 1$

- ✦ **surv**: survival of individuals from census time t to t +1 (contains: 0 for death or 1 for survival)
- ✦ **fec0, ...**: as many columns as desired relating size to sexual reproduction. For example, this might be:
 - ✦ { **fec0**: probability of reproduction (output: 0 for no reproductive or 1 for reproductive)
 - ✦ { **fec1**: number of reproductive structures (output: 1, 2, 3, ...) when individual is reproductive, that is, when fec0 = 1
 - ✦ { **fec2**: number of propagules (output: 1, 2, 3, ...) per reproductive structure (e.g. seeds per flower in reproductive plant individual)
 - ✦ { ...
- ✦ **stage**: stage of individuals in census time t, used to distinguish discrete and continuous stages, etc. This workflow only analyse population with continuous stages. For rows in the table where size is not an NA, then this must be the word “continuous”. Where size is NA, stage will be NA as well.
- ✦ **stageNext**: stage of individuals in census time t +1, in the simples case, “continuous” or dead”(which is redundant with “0” in the surv column.
- ✦ **offspringNext**: if the size contained in sizeNext corresponds to the size of an offspring, this column will contain either the value “sexual” or “clonal” (depending on whether sexual or clonal reproduction is being considered) In this workflow only the sexual reproduction is considered. If this column exists, rows that take these two values will be excluded from the growth analyses

Table 1. Input file in a .csv format to create an IPM.

No	size	fec0	fec1	stage	surv	ageNext	sizeNext	fec0Next	fec1Next	stageNext
100	NA	NA	NA	NA	NA	1	1	0	NA	continuous
72	1	0	NA	continuous	0	NA	NA	NA	NA	dead
87	43	1	6	continuous	1	7	25	1	8	continuous
142	NA	NA	NA	NA	NA	1	1	0	NA	continuous
107	1	0	NA	continuous	0	NA	NA	NA	NA	dead
108	1	0	NA	continuous	0	NA	NA	NA	NA	dead
111	1	0	NA	continuous	0	NA	NA	NA	NA	dead
112	1	0	NA	continuous	1	2	2	0	NA	continuous
113	2	0	NA	continuous	1	2	1	0	NA	continuous
118	23	1	8	continuous	1	7	19	1	5	continuous
119	71	1	26	continuous	1	6	79	1	41	continuous
138	1	0	NA	continuous	0	NA	NA	NA	NA	dead
139	1	0	NA	continuous	0	NA	NA	NA	NA	dead
144	NA	NA	NA	NA	NA	1	1	0	NA	continuous
509	NA	NA	NA	NA	NA	1	1	0	NA	continuous
148	1	0	NA	continuous	0	NA	NA	NA	NA	dead
591	NA	NA	NA	NA	NA	1	1	0	NA	continuous
168	50	1	19	continuous	0	NA	NA	NA	NA	dead
170	2	0	NA	continuous	1	2	9	1	2	continuous
592	NA	NA	NA	NA	NA	1	1	0	NA	continuous
177	18	1	1	continuous	1	3	20	0	NA	continuous
178	1	0	NA	continuous	0	NA	NA	NA	NA	dead
179	1	0	NA	continuous	0	NA	NA	NA	NA	dead
593	NA	NA	NA	NA	NA	1	1	0	NA	continuous
186	100	1	36	continuous	1	9	32	1	6	continuous
198	3	0	NA	continuous	1	2	11	0	NA	continuous
207	61	1	26	continuous	0	NA	NA	NA	NA	dead
211	14	1	2	continuous	1	7	11	1	8	continuous
212	30	1	2	continuous	0	NA	NA	NA	NA	dead
215	1	0	NA	continuous	0	NA	NA	NA	NA	dead
218	25	1	5	continuous	1	7	41	1	14	continuous
219	6	0	NA	continuous	1	2	11	1	4	continuous
237	16	0	NA	continuous	1	999	26	1	3	continuous
241	9	1	2	continuous	1	999	29	1	9	continuous
245	18	1	6	continuous	1	999	29	1	12	continuous

- 2) **openCPU**: The base URL of an OpenCPU server - see <http://opencpu.org/> for details about the OpenCPU framework.

Example: <http://opencpu1.biovel.eu/>

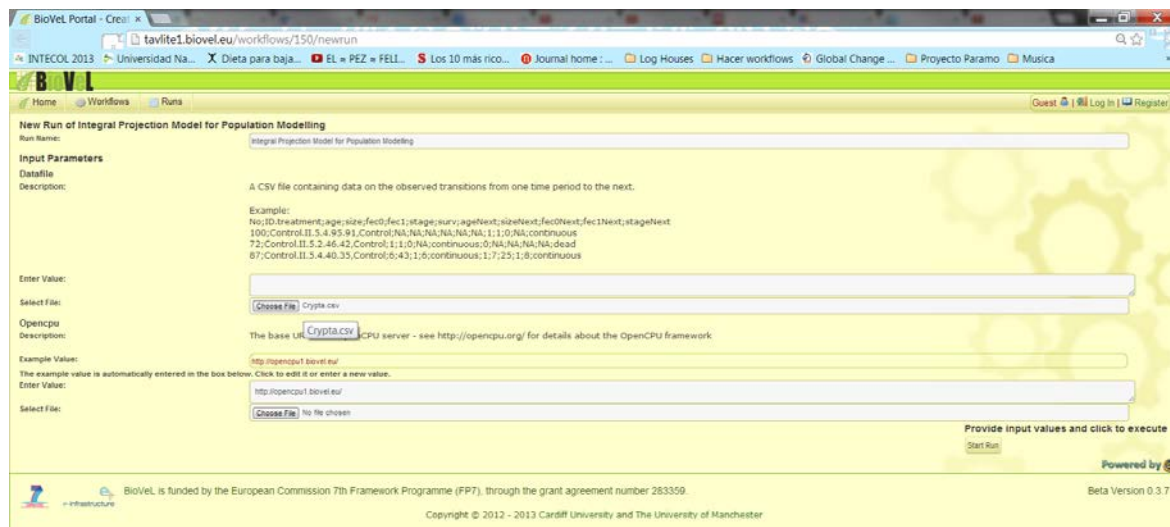


Figure 2. Inputs file and information to be filled in.

After the user has filled out the input ports and has clicked in run the workflow.

5.2.2 Dialogue

- 1) **Identify field**: Indicate the fields that provide fecundity information (fec; e.g. fec0, fec1) for the initial time period. You can add labels for fields to be used in output plots (Fig 3).

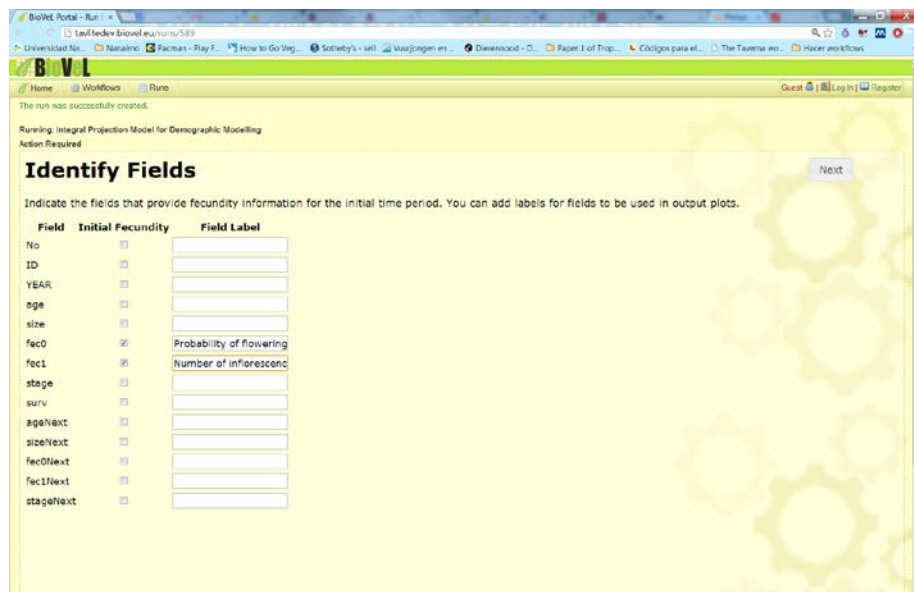


Figure 3. Identify fields. The user can add labels for fields to be used in output plots. Then the user clicks in Next.

- 2) **Create Vital Rates:** Here the user plots functional forms, scores and selects models. There are 5 tabs: Data, Survival, Growth, Fecundity and Submit (Fig 4). In the tab named Survival, first the user finds a graph that shows the relation between survival and size, dots are observed values of survival. (Fig 5). In this tab, below the graph, the user can perform the survival analysis. Survival analysis predicts the probability of surviving at a given size. The user can compare different linear models based on: [Bayesian Information criterion \(BIC\)](#) or the [Akaike Information Criterion \(AIC\)](#). Please select first the criteria and then click in compare (Fig 6). Then select the best fit for the survival probability based on biological criteria and / or the AIC or BIC values. Once you have selected your preferred linear model, click in submit (Fig 7).Once the user have submit the survival model, the tab Submit appears showing the graph and the linear model chosen by the user (Fig. 8).

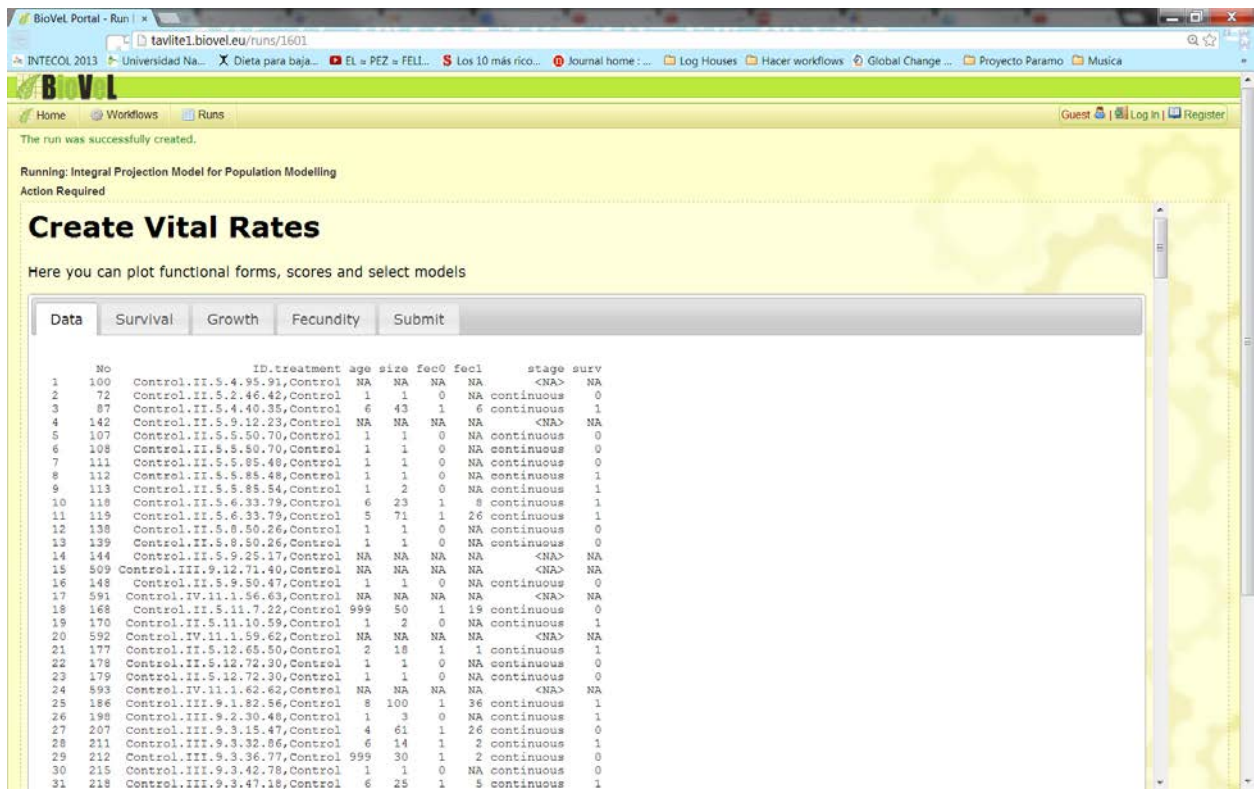


Figure 4. Create Vital Rates dialogue. Tab Data. There are 5 tabs: Data, Survival, Growth, Fecundity and Submit.

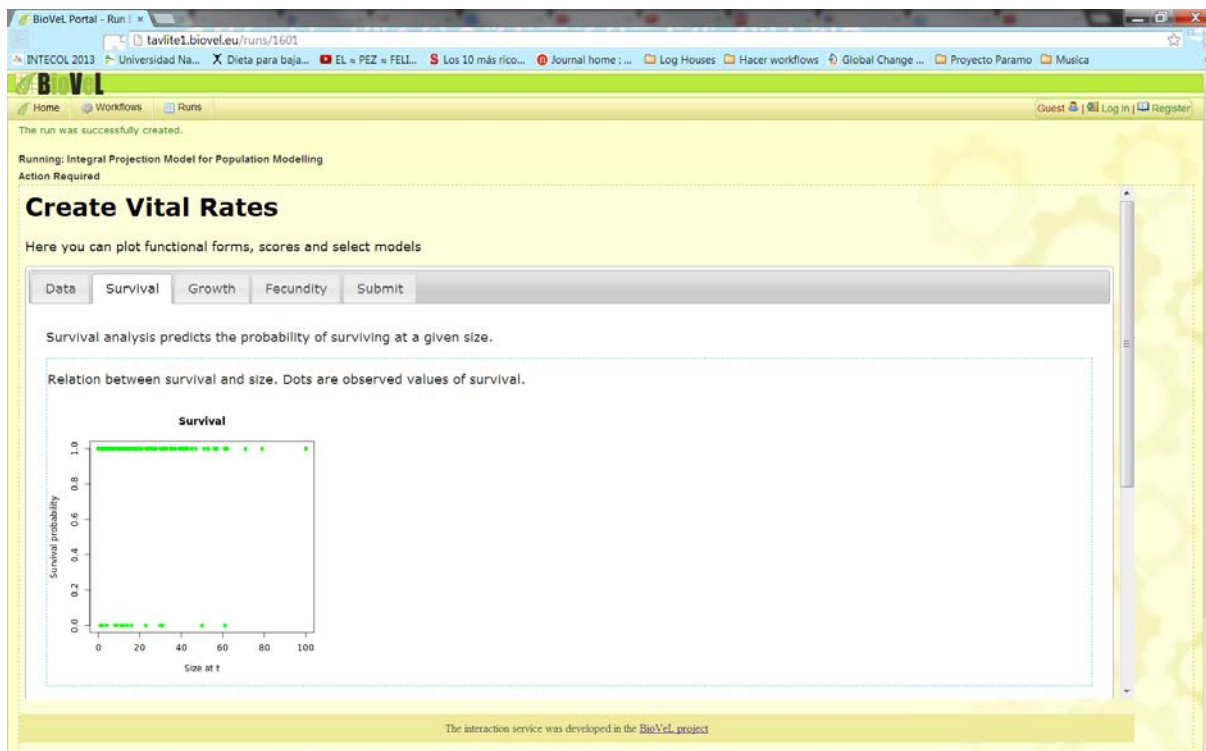


Figure 5. Create Vital Rates dialogue. Tab Survival. Showing the graph of survival probability.

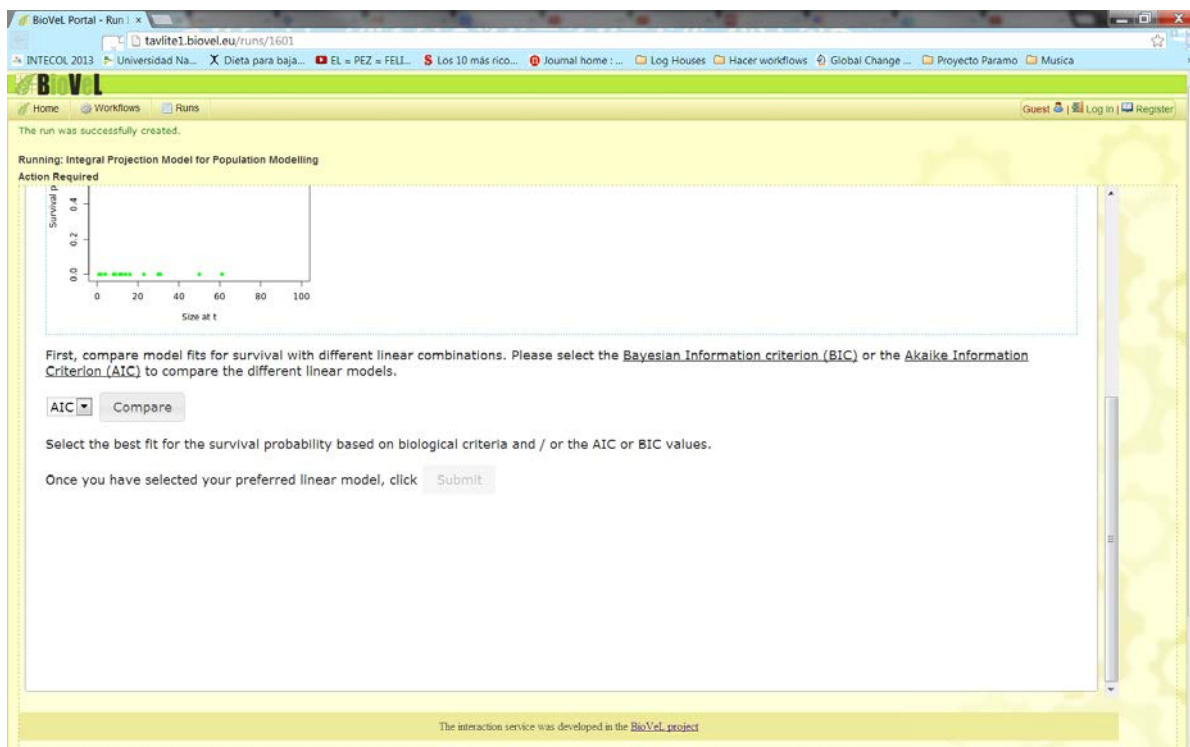


Figure 6. Create Vital Rates dialogue. Tab Survival. The user chooses the [Bayesian Information criterion \(BIC\)](#) or the [Akaike Information Criterion \(AIC\)](#) to compare the different linear models, and then click in compare.

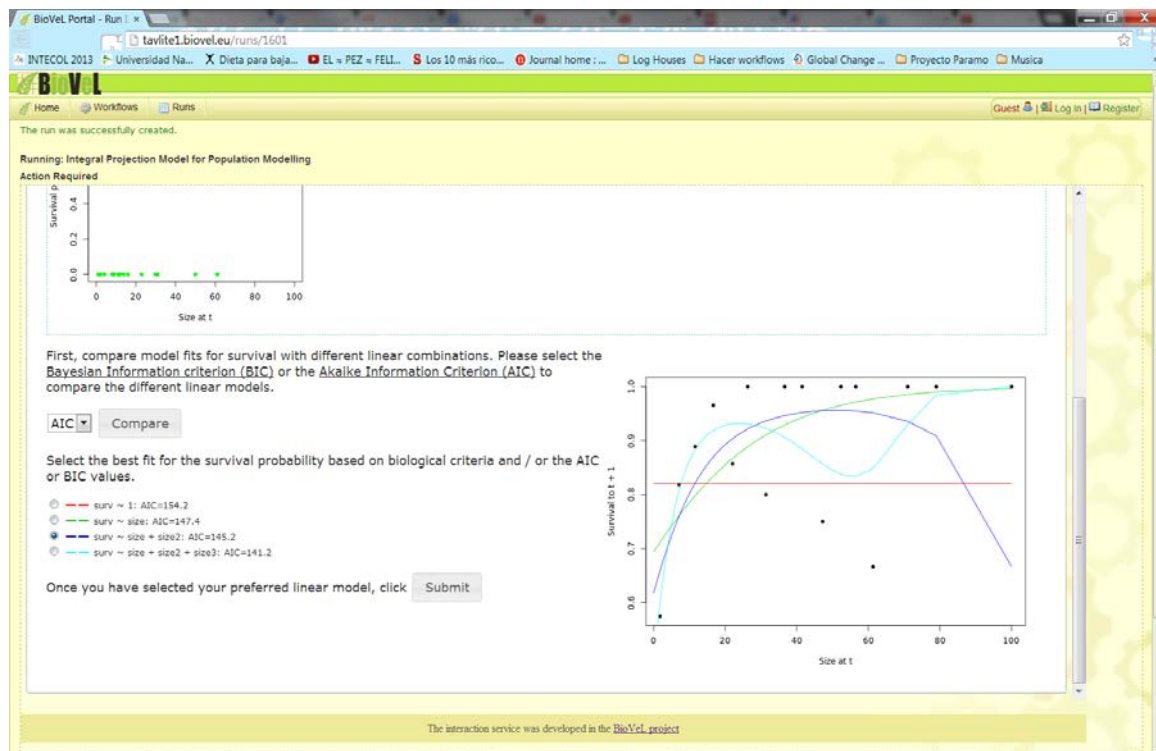


Figure 7. Create Vital Rates dialogue. Tab Survival. Then select the best fit for the survival probability based on biological criteria and / or the AIC or BIC values. Once you have selected your preferred linear model, click in submit.

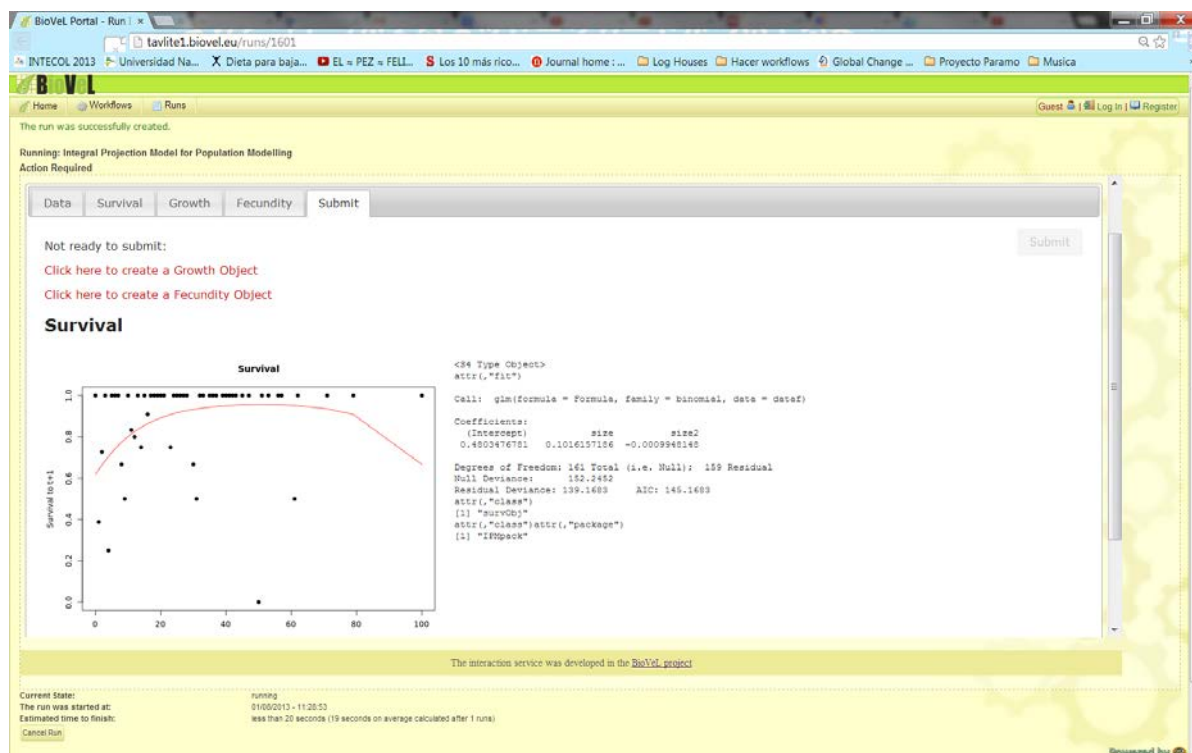


Figure 8. Create Vital Rates dialogue. Tab Submit. The survival model and the graph appear in the submit tab.

After the survival model is submitted, in the tab named growth, the user finds a graph that shows the relation between size at t vs. size at $t+1$ (Fig 9). In this tab, below the graph, the user can perform the growth analysis, which predicts the probability of growing at a given size. The user compares different linear models based on: [Bayesian Information criterion \(BIC\)](#) or the [Akaike Information Criterion \(AIC\)](#). Select first the criteria and then click in compare (Fig 10). Then selects the best fit for the growth probability based on biological criteria and / or the AIC or BIC values. Once the preferred lineal model is selected, click in submit (Fig 11).Once the growth model have been submitted, the tab Submit appears showing the graph and the linear model chosen by the user (Fig. 12).

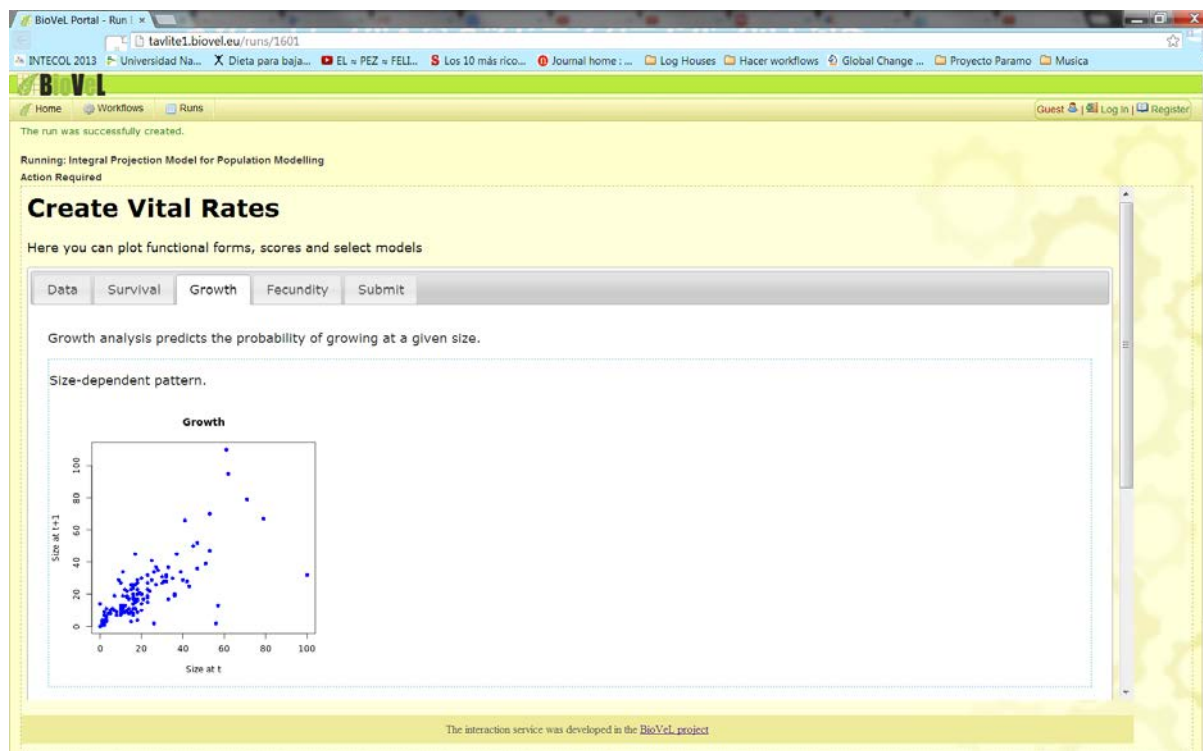


Figure 9. Create Vital Rates dialogue. Tab Growth. Showing the graph of size-dependent pattern, size at t vs. size at $t+1$.

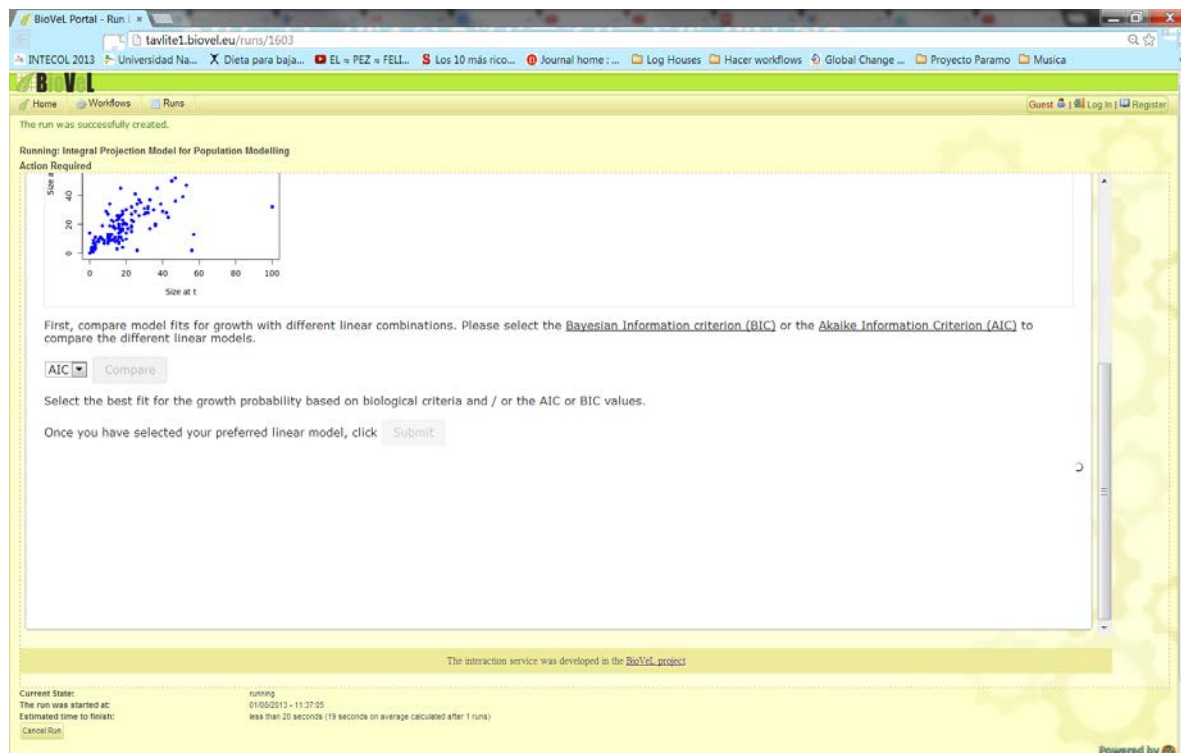


Figure 10. Create Vital Rates dialogue. Tab Growth. The user chooses the [Bayesian Information criterion \(BIC\)](#) or the [Akaike Information Criterion \(AIC\)](#) to compare the different linear models, and then click in compare.

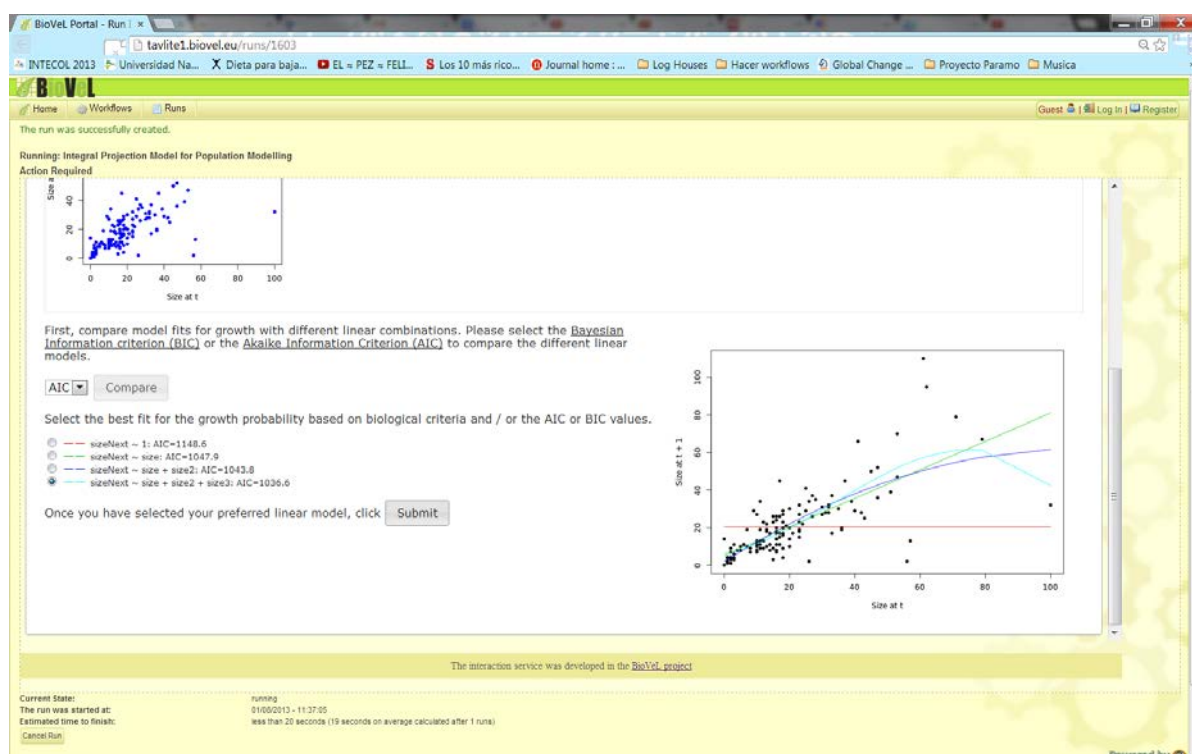


Figure 11. Create Vital Rates dialogue. Tab Growth. Then select the best fit for the growth probability based on biological criteria and / or the AIC or BIC values. Once you have selected your preferred linear model, click in submit.

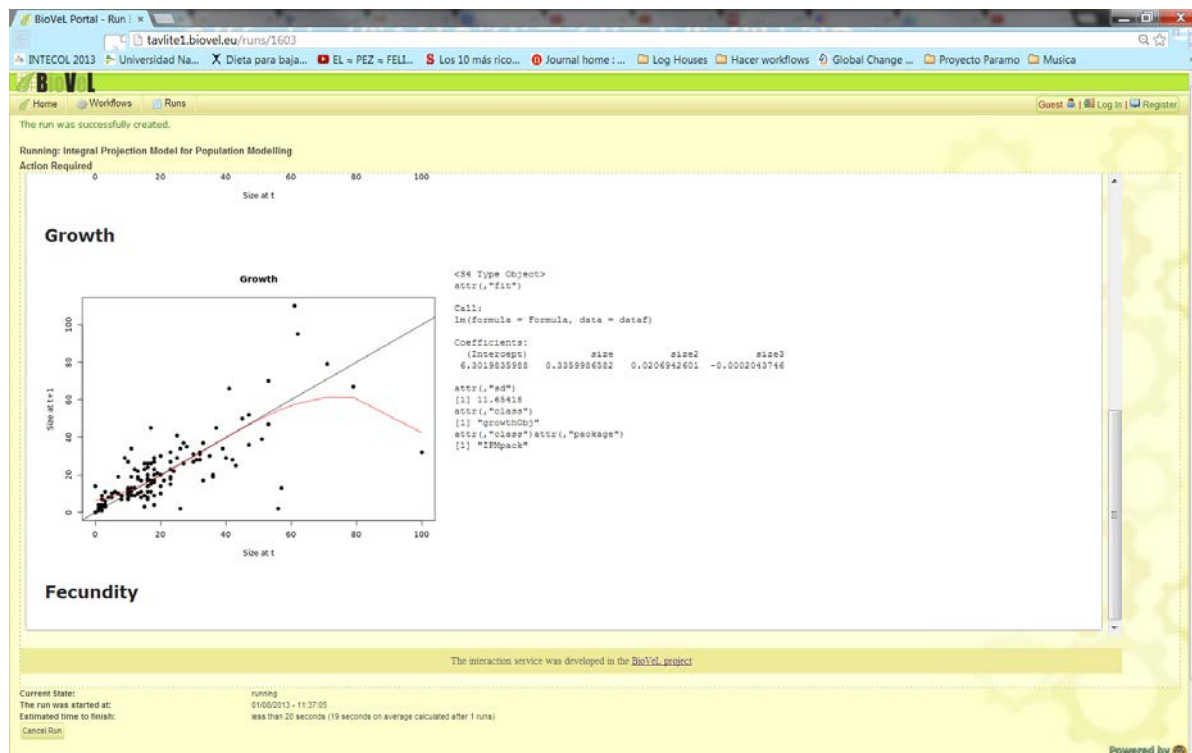


Figure 12. Create Vital Rates dialogue. Tab Submit. The growth model and the graph appear in the submit tab.

Once the user submitted the growth model. In the tab named fecundity, the user finds under fec0, two graphs that show the fecundity patterns, the left graph shows the relation between fecundity and size (e.g. fec0 is: probability of reproduction (output: 0 for no reproductive or 1 for reproductive)). The right graph illustrates the frequency of being reproductive or not (Fig 13). Below these figures are found two more graphs, the left graph shows the relation between size vs. the number of reproductive structures (e.g. fec1: number of reproductive structures (output: 1, 2, 3 ...) when individual is reproductive, that is, when fec0 = 1) and on the right side the frequency of observation vs. size (Fig 14).

Below the graphs, the user analyses and plots each fecundity data field with a suitable family and transform. These analyses predict the probability of flowering at a given size. The user selects the field (e.g. fec0), the family (e.g. binomial) and transform (e.g.: none), then click in plot (Fig 15). Below it appears three linear models compared with based on the [Akaike Information Criterion \(AIC\)](#). So the user selects the best fit for the fecundity probability. Once you have selected your preferred linear model, click in submit (Fig 16). Once the user have submitted the first fecundity model and if the data have more **fec** columns the user repeat once more the procedure (with fec1) and click in submit (Fig 17). The tab Submit appears showing the graphs and the models that were chosen by the user (Fig. 18). In the upper part of the Tab Submit click in submit (Fig 19).

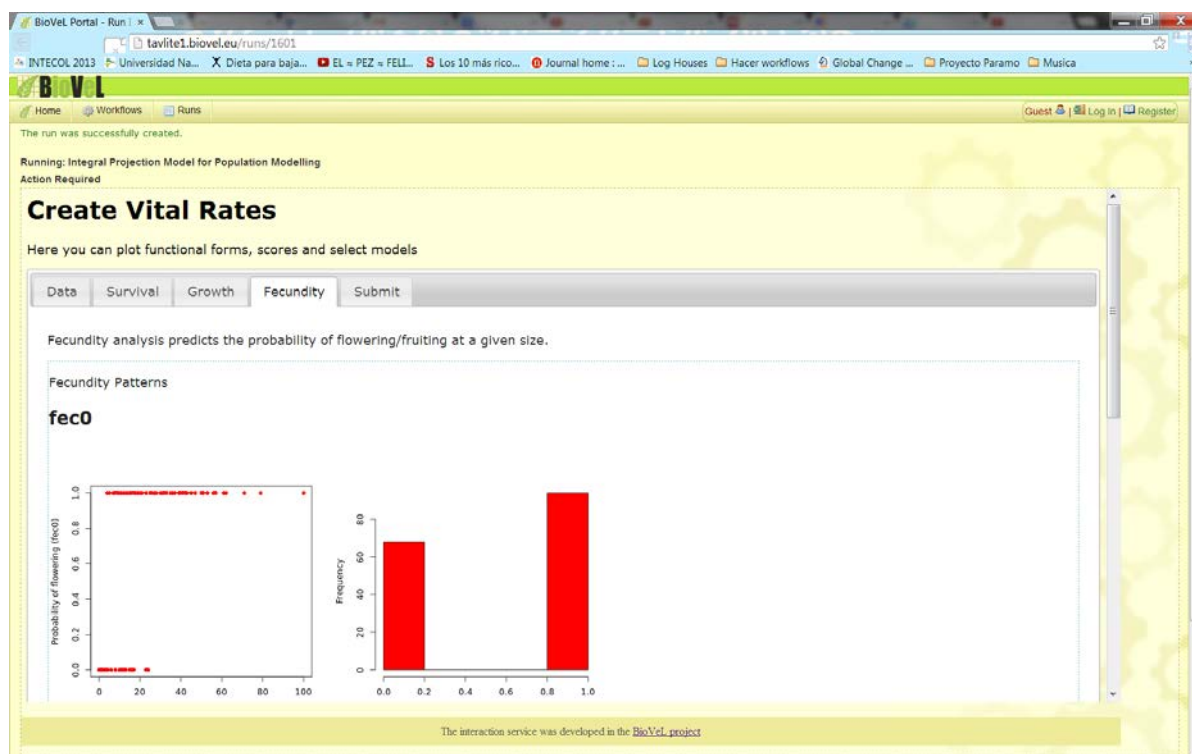


Figure 13. Create Vital Rates dialogue. Tab Fecundity. Left: Probability of flowering, Right: Frequency of flowering.

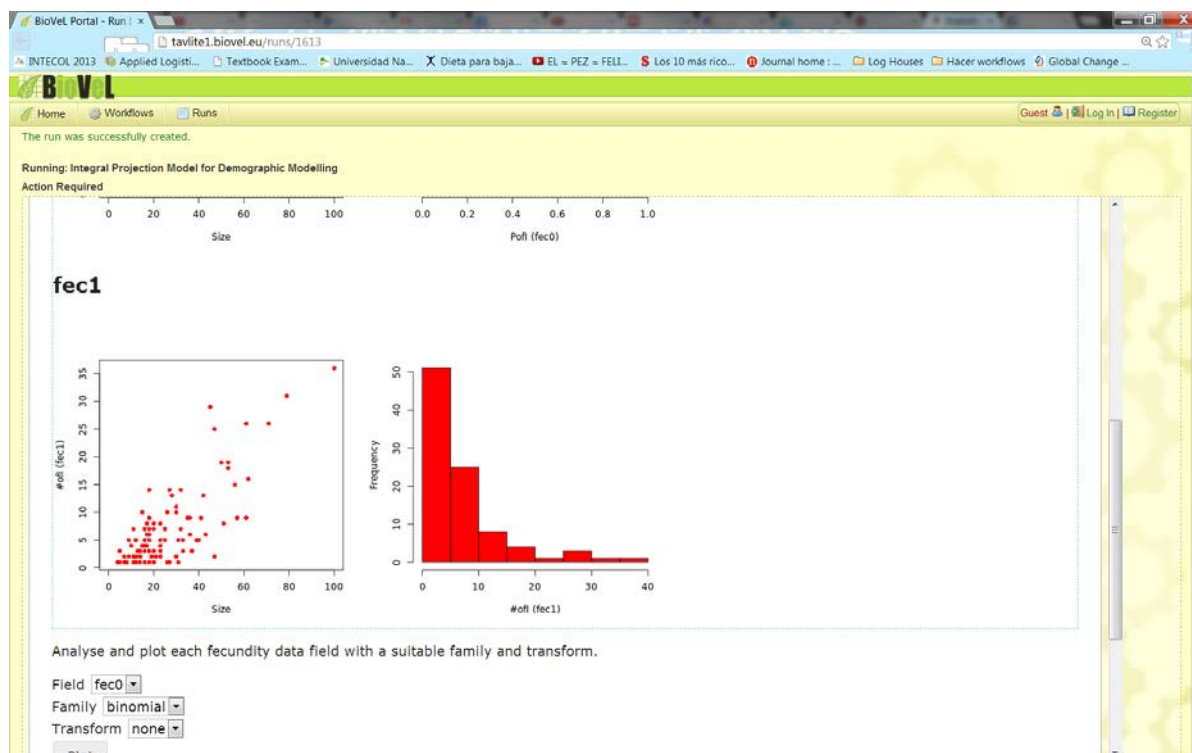


Figure 14. Create Vital Rates dialogue. Tab Fecundity. Left: the relation between sizes vs. the number of reproductive structures, Right: Frequency of observations vs. size

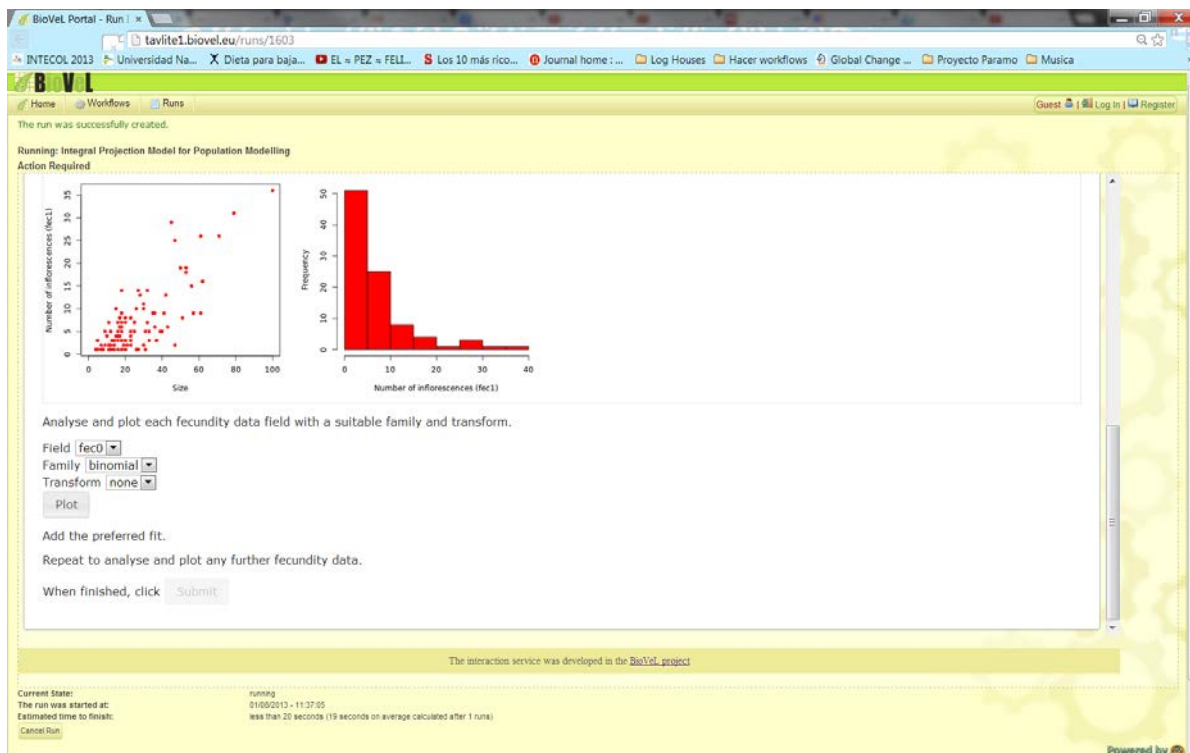


Figure 15. Create Vital Rates dialogue. Tab Fecundity. Analyse and plot each fecundity data field with a suitable family and transform, then click in plot.

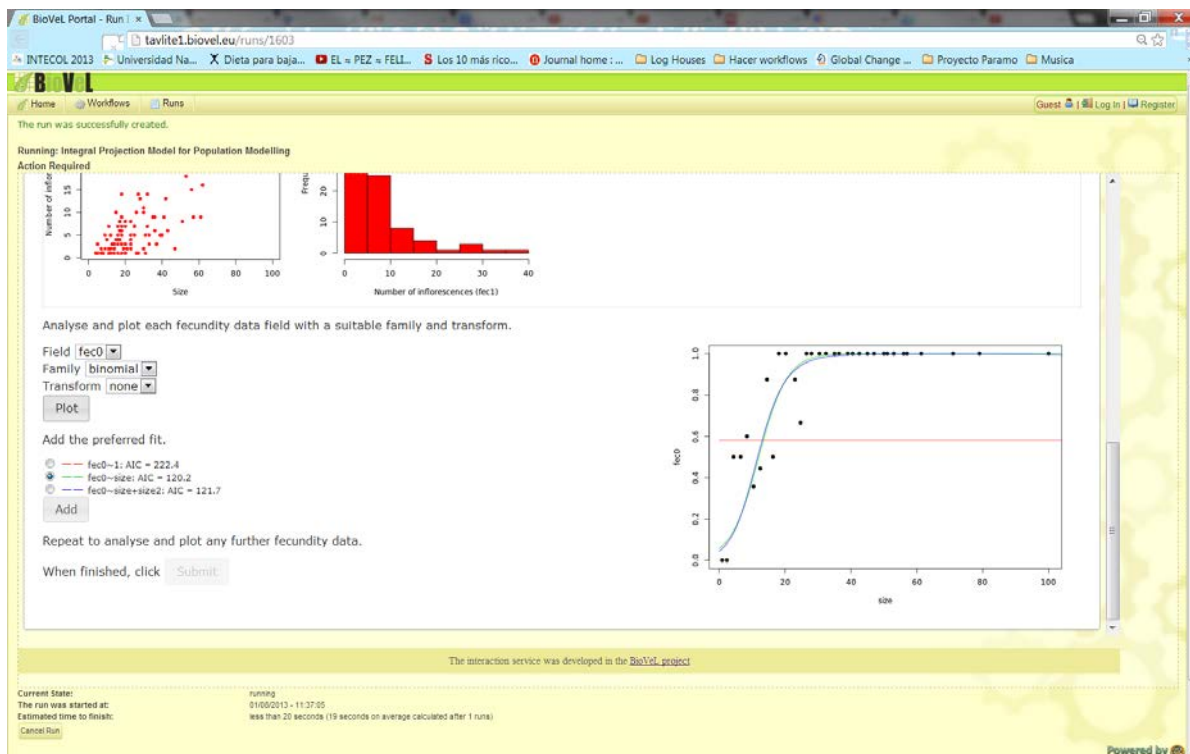


Figure 16. Create Vital Rates dialogue. Tab Fecundity. Three linear models appear. And the user selects the best fit and clicks add.

BioVel Documentation

Integral Projection Models Workflow

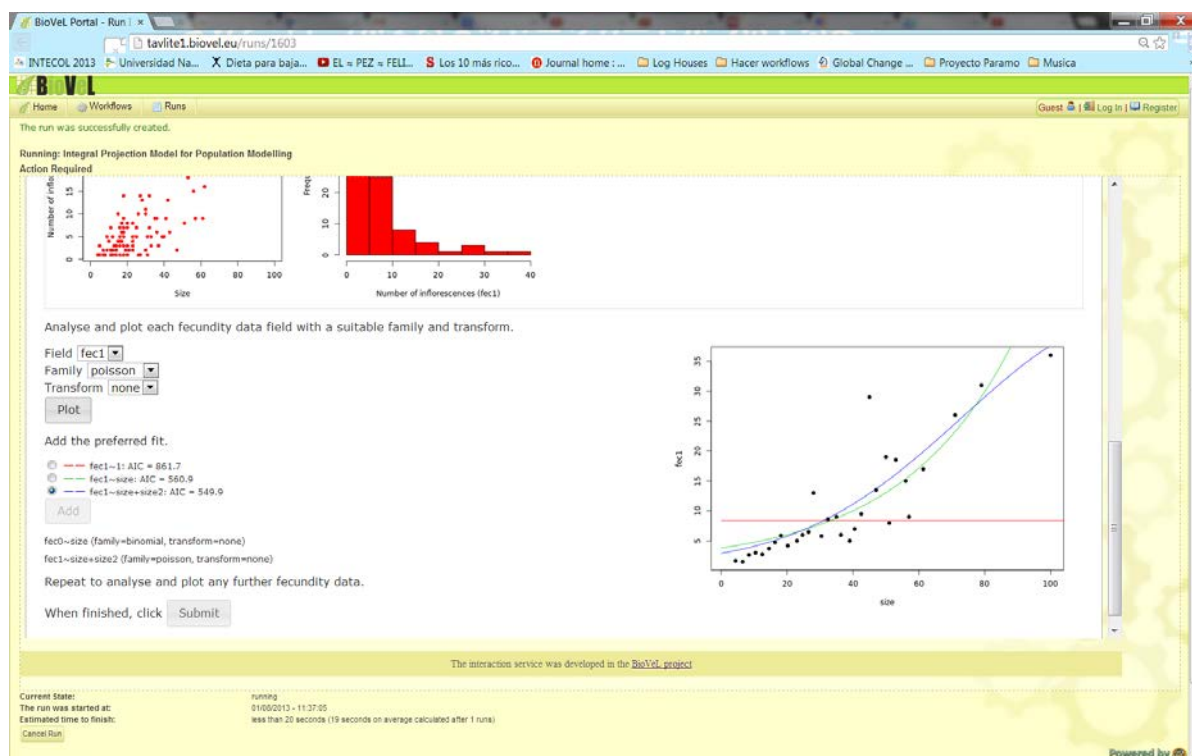


Figure 17. Create Vital Rates dialogue. Tab Fecundity. Repeat once more the procedure and click in submit.



Figure 18. Create Vital Rates dialogue. Tab Submit. The fecundity models and the graphs appeared in the submit tab.

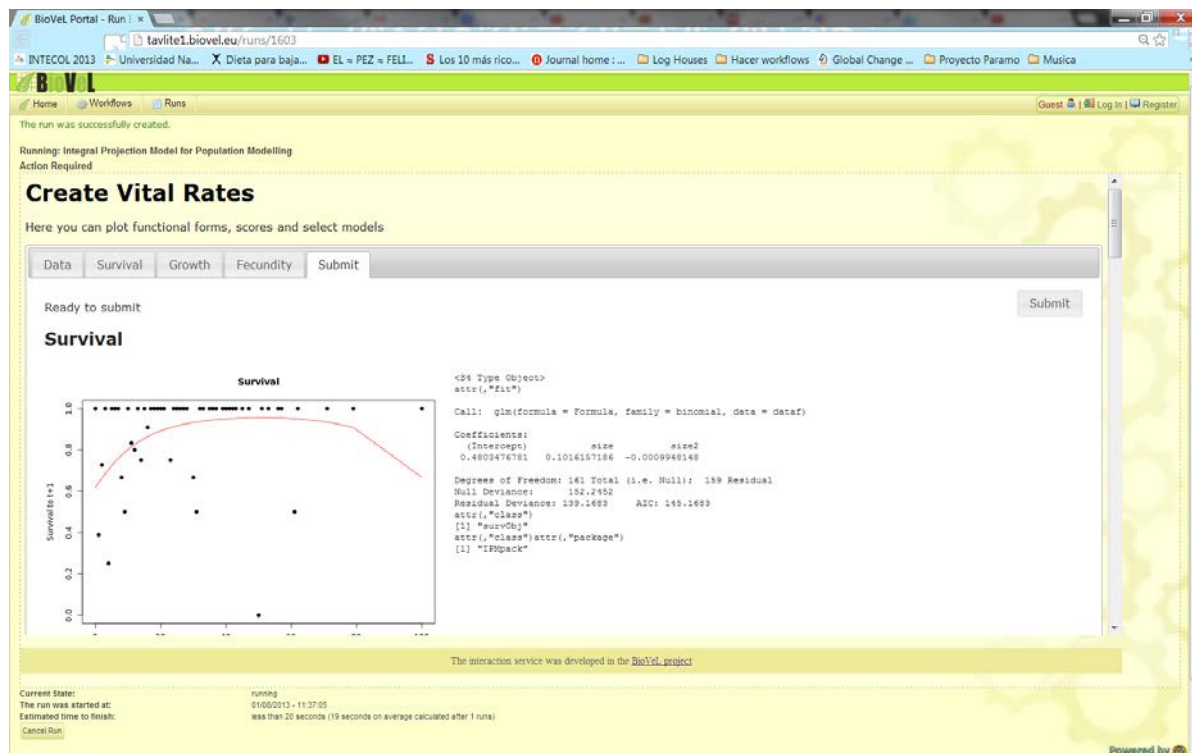


Figure 19. Create Vital Rates dialogue. Tab Submit. Upper part of the tab, click in submit.

- 3) **Create IPM Matrix:** Here the user constructs the IPM, runs diagnostics of the P matrix, and plots P and F matrices. Uses growth and survival objects (analysed and created in Tabs Survival and Growth) to construct a matrix defining probabilities for transitions between sizes due to growth and survival. In this section, the user finds the following fields to field in (Fig 20), once the user fielded in and choose the options, clicks Plot All.
 - a. Matrix Dimensions: number of size bins in the P and F matrix, defaults to 100.
 - b. Min Size: minimum size of the P matrix, defaults to the smallest size in the two years.
 - c. Max Size: maximum size of the P matrix, defaults to the largest size in the two years.
 - d. Correction: correction type, defaults to constant. This option will multiply every column of the IPM by a constant sufficient to adjust values to those predicted for survival at that size. The second option is discretizeExtremes which will place all transitions to sizes smaller than min size into the smallest n bin, and transitions to sizes larger than max size into the largest bin. The third option is no transformation.
 - e. Integrate Type: integration type, defaults to "midpoint" (which uses probability density function); other option is "cumul" (which uses the cumulative density function).
 - f. Pre-census (F matrix only): logical (TRUE or FALSE), indicating whether the fecundity object should represent an interval between pre-breeding or a post-breeding censuses. Defaults to TRUE (pre-breeding census), meaning that all reproduction and offspring rates required for the F matrix are embedded in Fecundity model. Alternatively, an F matrix based on post-breeding census (preCensus=FALSE) requires additional Survival models and growth models, to cover the survival and growth of the parents until the reproduction event.

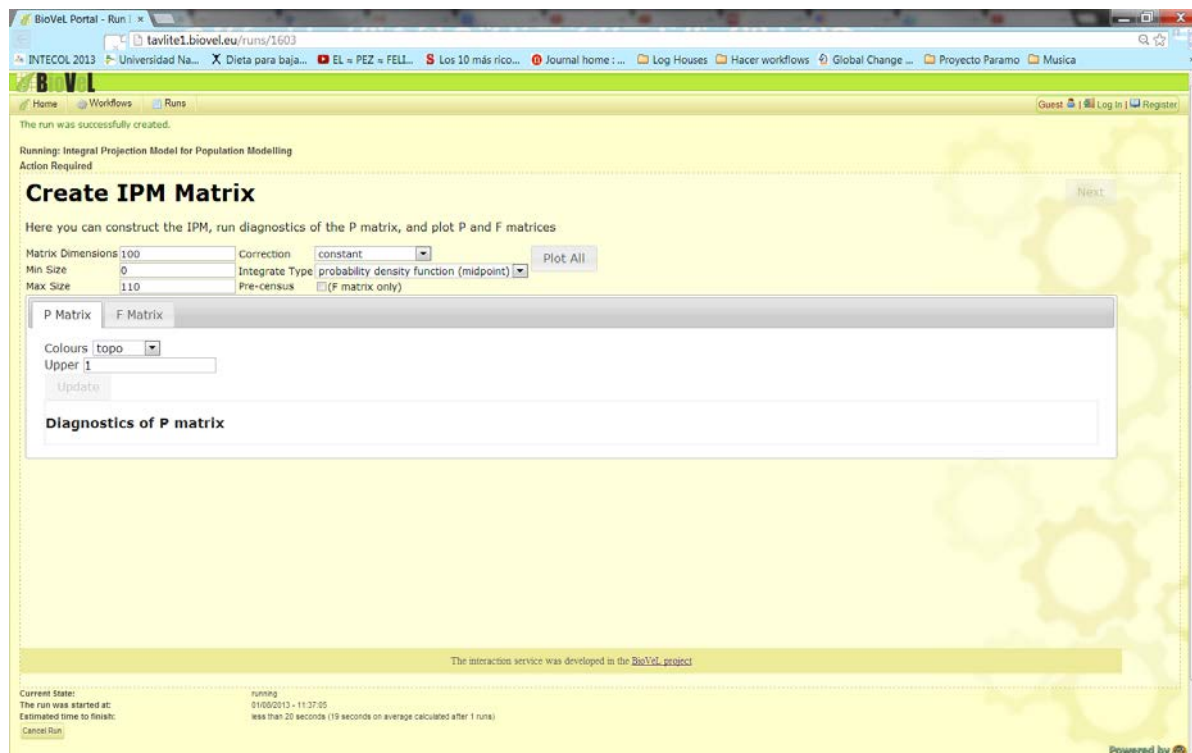


Figure 20. Create IPM matrix. Here the user constructs the IPM, runs diagnostics of the P matrix, and plots P and F matrices.

The results appears in the P and F matrix tab. In P matrix tab appears a 2D plot, a 3D plot (Fig. 21), the diagnostics of the P matrix (Fig. 22) and the numerical P matrix. Next to, in the tab F Matrix a 2D plot, a 3D plot (Fig. 23) and the numerical F matrix appears. If the user agrees with the results, then click in Next if not repeats the procedure and clicks in Plot All.

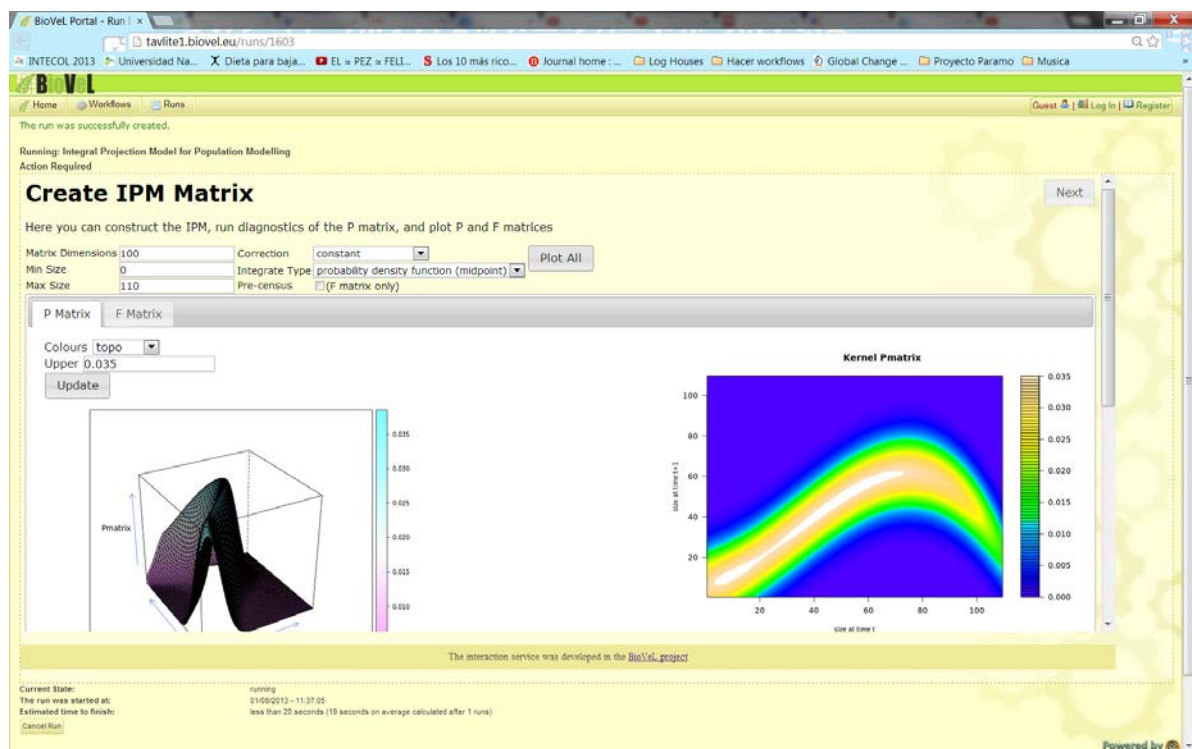


Figure 21. Create IPM matrix. Tab P Matrix. 2D plot and 3D plots of the P Matrix.

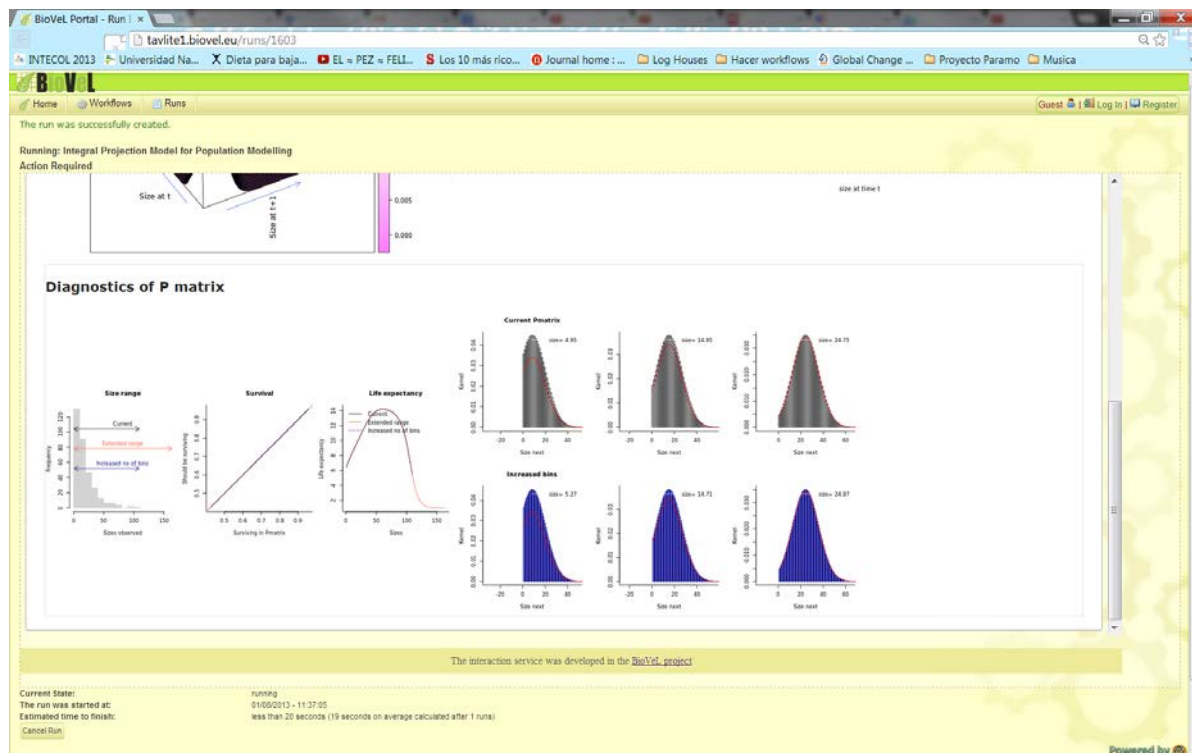


Figure 22. Create IPM matrix. Tab P Matrix. Shows the diagnostics of the P matrix

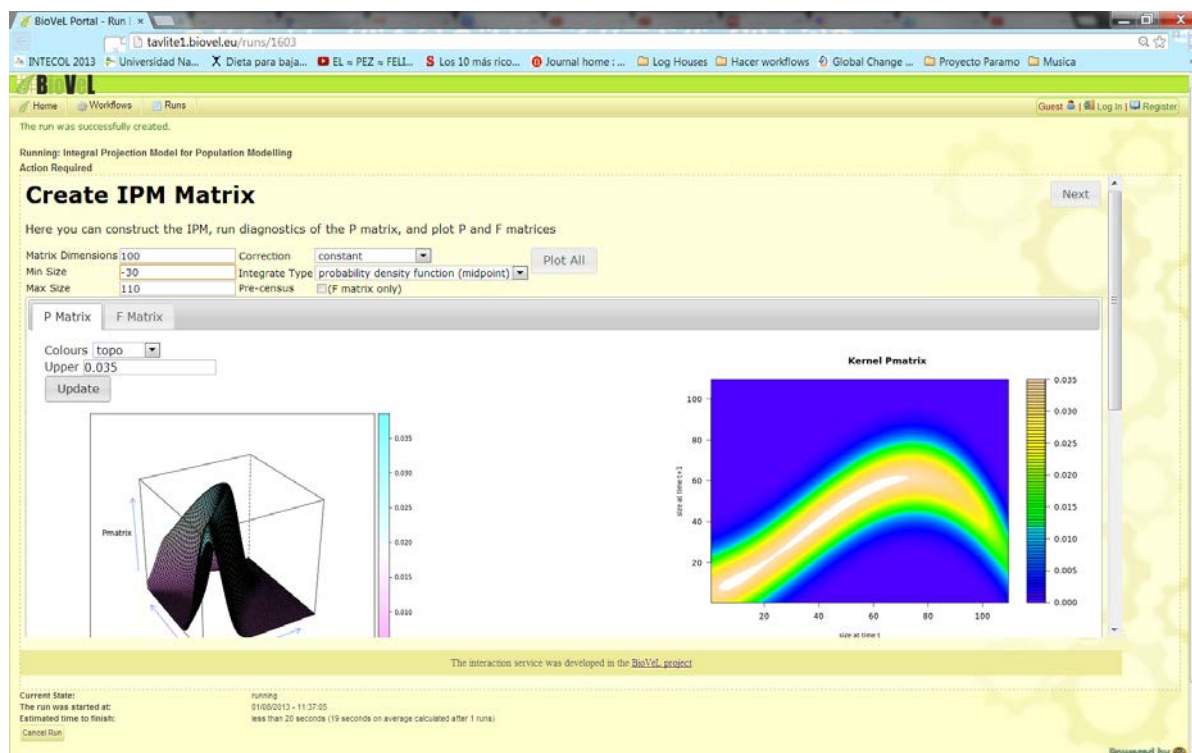


Figure 23. Create IPM matrix. Tab F Matrix. 2D plot and 3D plots of the F Matrix.

- 4) **Results & Analysis of IPM:** Here the user sees the multiples results of the IPM (Fig. 24 – 29). Then clicks in Next.
- IPM:** Uses survival/growth and fecundity kernels to make an IPM kernel. Shows the value of Lambda, a 2D and a 3D graphs of the constructed IPM and the numerical IPM (Fig 24)
 - Sensitivity:** Estimates sensitivities of each element of a discretized IPM. Shows a 2D and a 3D graphs of the Sensitivity matrix and the numerical Sensitivity matrix (Fig. 25).
 - Elasticity:** Estimates elasticities of each element of a discretized IPM. Shows a 2D and a 3D graphs of the Elasticity matrix and the numerical Elasticity matrix (Fig. 26).
 - Vital rates Sensitivity & Elasticity:** Uses perturbation to estimate the sensitivity and elasticity of all the parameters underlying an IPM. Shows the histograms of the sensitivity and elasticity of lambda, R0 and Life expectancy to parameters underlying an IPM. (Fig. 27).
 - Survivorship:** Calculates the fraction of the cohort surviving across age for a chosen starting continuous stage value (e.g. 1 year with a maximum age of 50). Shows a graphs of survivorship across age (Fig. 28).
 - Mean life expectancy and Passage time:** **Mean life expectancy:** Based on the P matrix is defined the life expectancy in units of the time step used. **Passage time:** Defines passage time to a chosen continuous stage. Estimates the time in units of the chosen time-steps that it will take to reach a chosen continuous (e.g. size) stage for the first time conditional on surviving from each of the meshpoints of the IPM (Fig. 29).

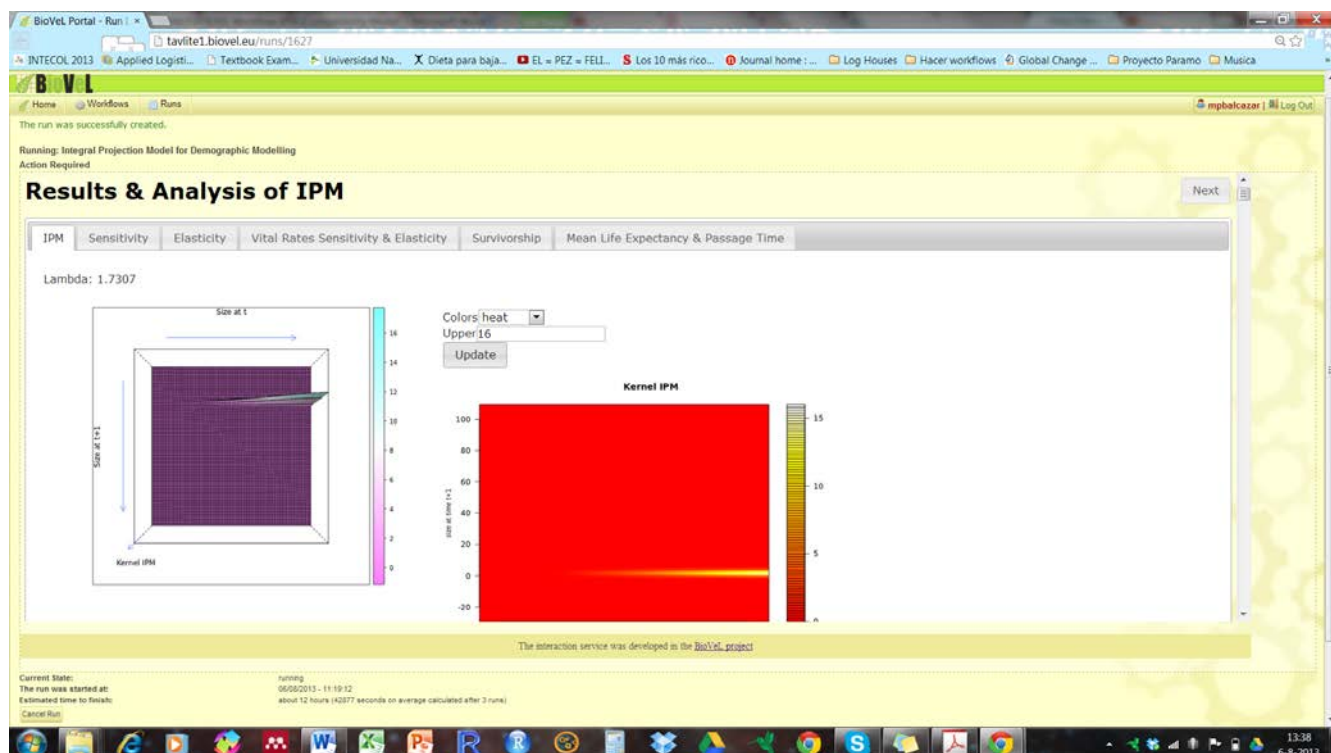


Figure 24. Results & Analysis of the IPM. Tab IPM. 2D plot and 3D plots of the IPM and the numerical IPM matrix.

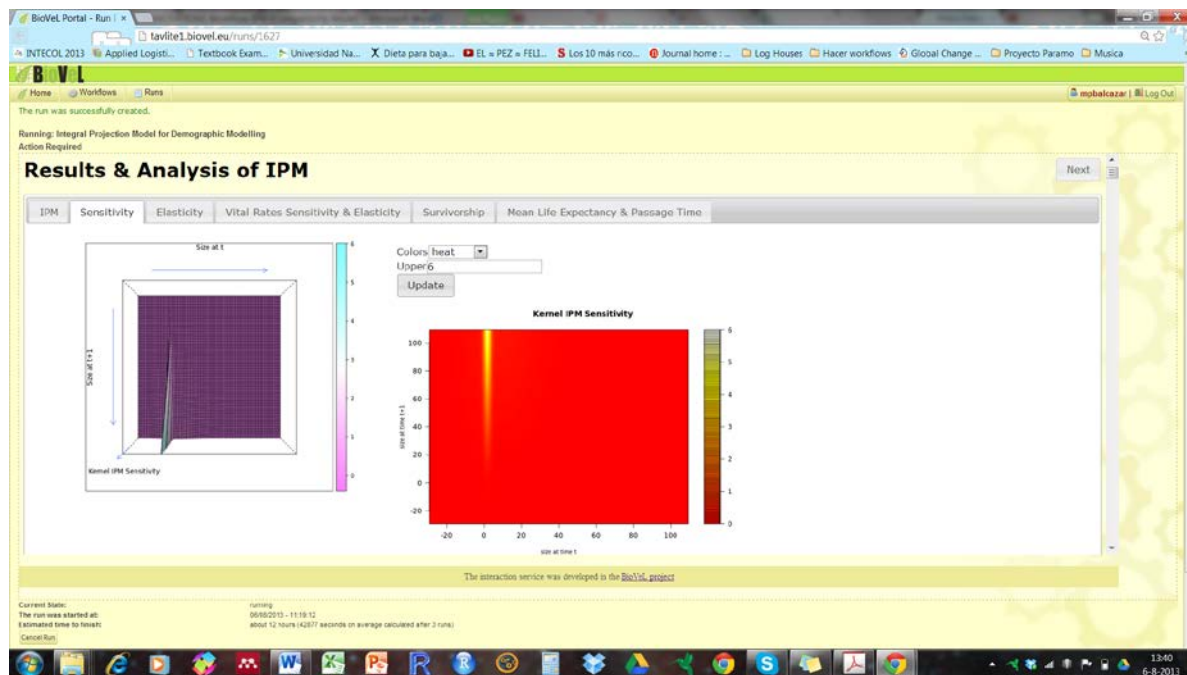


Figure 25. Results & Analysis of the IPM. Tab Sensitivity. 2D plot and 3D plots of the Sensitivity matrix and the numerical Sensitivity matrix.

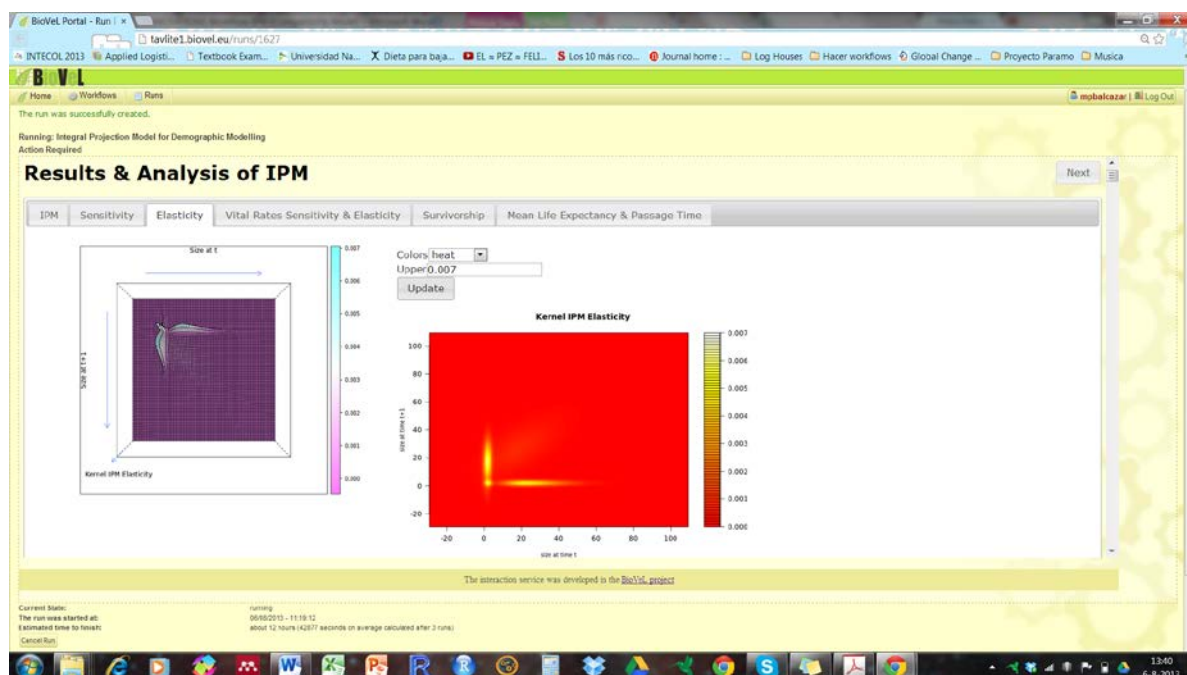


Figure 26. Results & Analysis of the IPM. Tab Elasticity. 2D plot and 3D plots of the Elasticity matrix and the numerical Elasticity matrix.

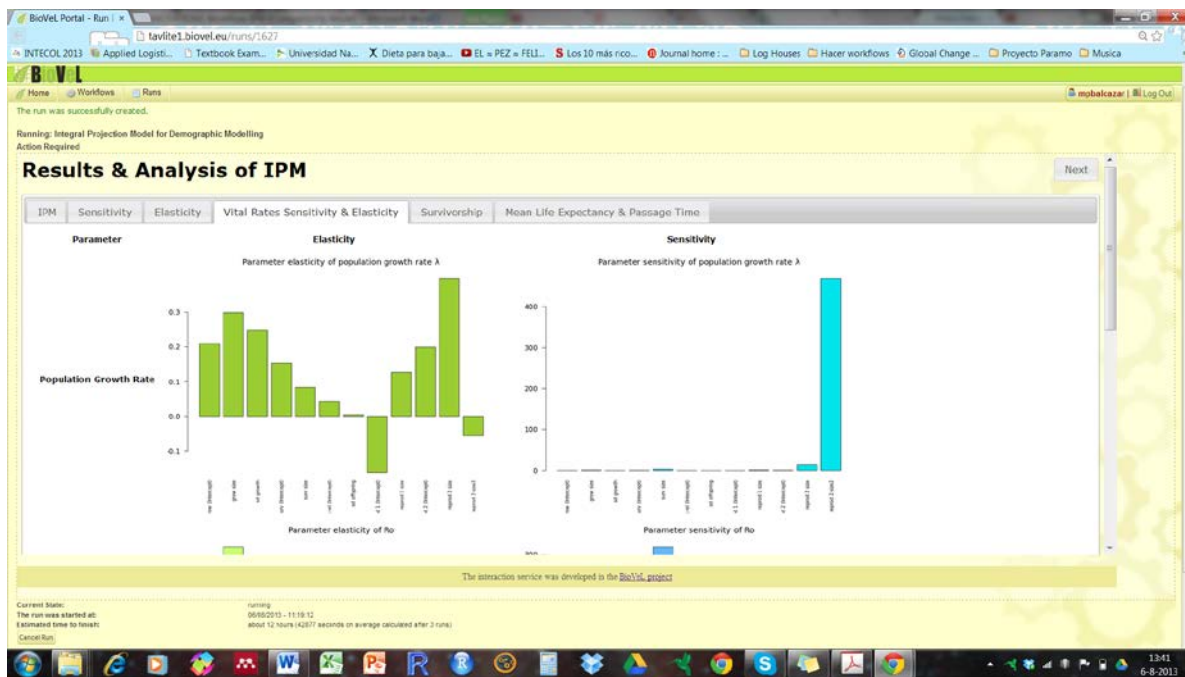


Figure 27. Results & Analysis of the IPM. Tab Vital rates Sensitivity & Elasticity. Histograms of the sensitivity and elasticity of lambda, R0 and Life expectancy to parameters underlying an IPM.

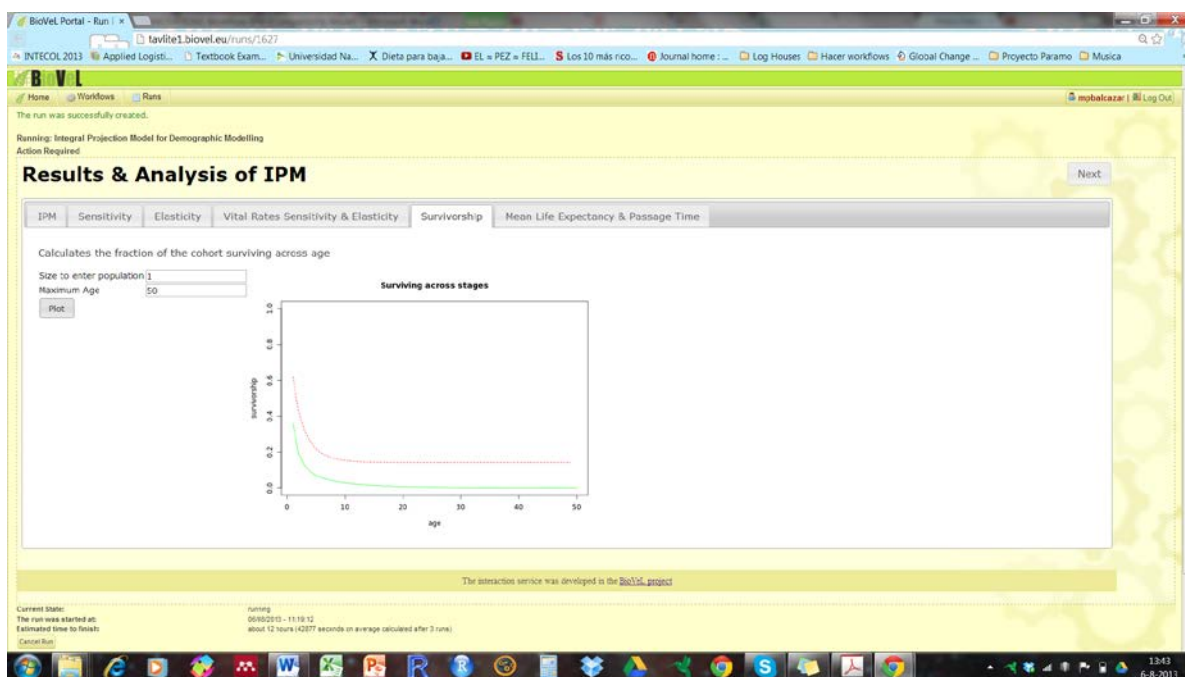


Figure 28. Results & Analysis of the IPM. Tab Survivorship. Shows a graphs of survivorship across age.

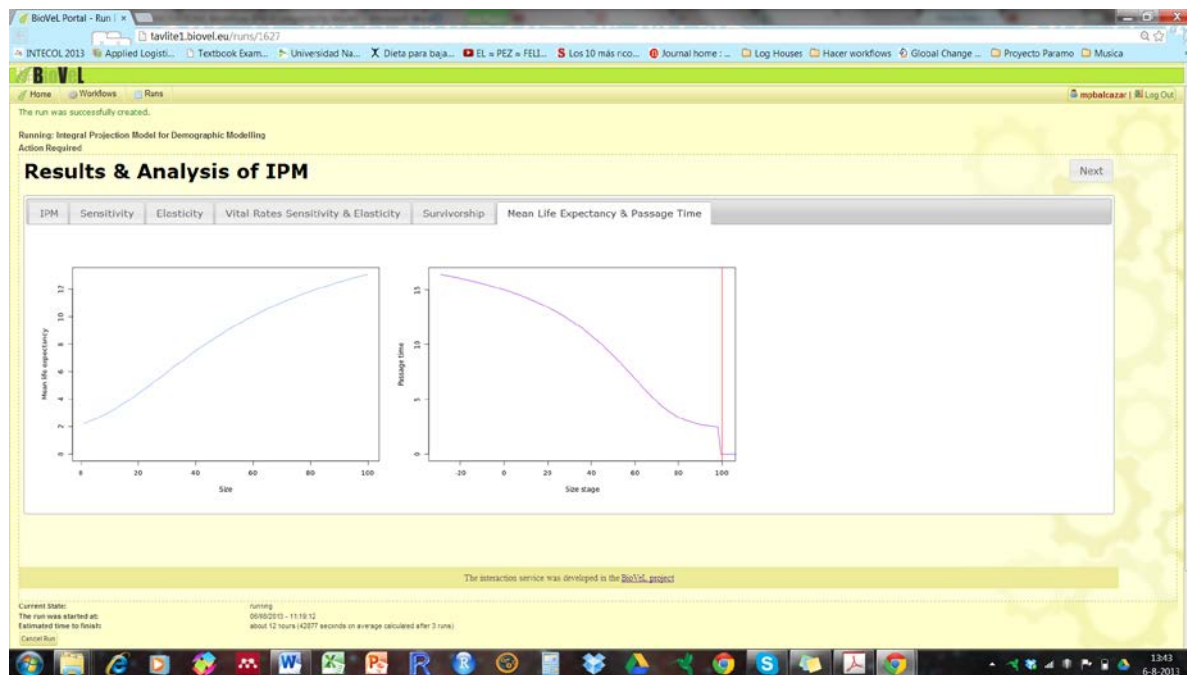


Figure 29. Results & Analysis of the IPM. Tab Mean Life Expectancy & Passage Time. Graphs of the Mean Life expectancy and Passage Time.

- 5) **Download zip files:** The user clicks in Dowload and all the the results (numerical and graph will be download as a zip file that can be save by the user (Fig 30 and 31).

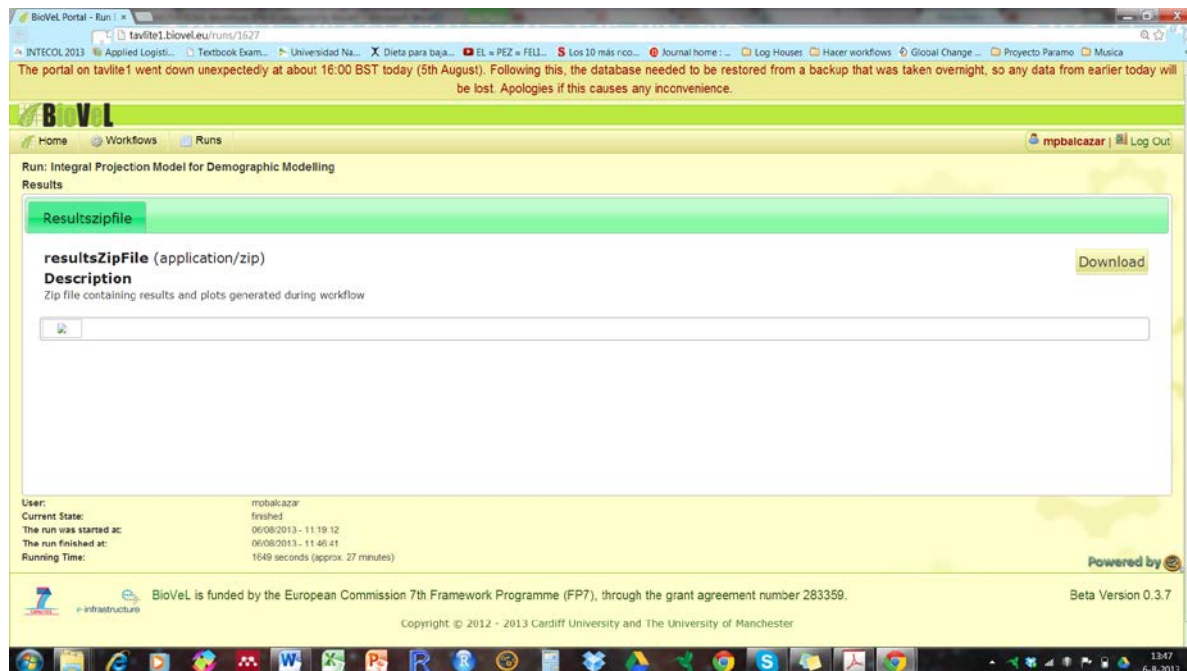


Figure 30. Results zip files. Clicks in Dowload.

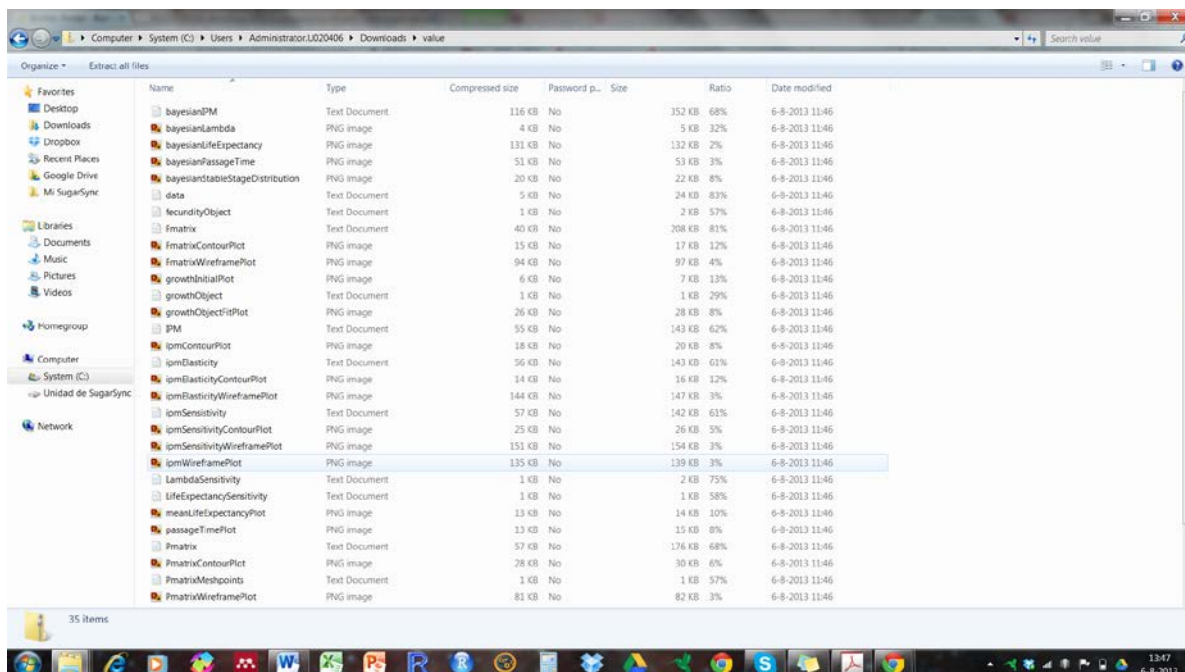


Figure 30. Zip files. Numerical and graph results download as a zip files.

For further details see: References page

5.3 Input file format

The input file is in a .csv format (see Table 1. Section 5.1.1): to download click here [Crypta.csv](#)

6 Support

For feedback, suggestions, and comments, please write to 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 Package ‘IPMpack’ (Stubben & Milligan 2007; Stubben, Milligan & Nantel 2011), Matrix, MASS, nlme, mvtnorm, methods, MCMCgImm, truncnorm, lattice, graphics, grDevices in R.

- 1) Metcalf, C.J.E., S.M. McMahon, R. Salguero-Gomez, E. Jongejans, and C. Merow. 2013. IPMpack: an R package for demographic modeling with Integral Projection Models (v. 2.0) May 19, 2013
- 2) Metcalf, C.J.E., S.M. McMahon, R. Salguero-Gomez, E. Jongejans, and C. Merow. 2012-7-4. Package 'IPMpack', Builds and analyses Integral Projection Models (IPMs). July 26, 2013, Version 2.0.
- 3) Metcalf C.J.E., S.M. McMahon, R. Salguero-Gomez and E. Jongejans. 2013. IPMpack: an R package for Integral Projection Models. *Meth Ecol Evol* 4:195-200.
- 4) Salguero-Gómez R, Siewert W, Casper BB & Tielbörger K. 2012. A demographic approach to study effects of climate change in desert plants. *Philosophical Transactions of the Royal Society, series B* 367, 3100-3111.

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