ORCA – Online Research @ Cardiff

Citation for final published version:

De Giovanni, Renato, Williams, Alan R, Hernández Ernst, Vera, Kulawik, Robert, Quevedo Fernandez, Francisco and Hardisty, Alex 2016. ENM Components: a new set of web service-based workflow components for ecological niche modelling. Ecography 39 (4) , pp. 376-383. 10.1111/ecog.01552

Publishers page: http://dx.doi.org/10.1111/ecog.01552

ENM Components: a new set of web service-based workflow components for ecological niche modelling

Renato De Giovanni, Alan R. Williams, Vera Hernández E., Robert Kulawik, Francisco Quevedo Fernandez, Alex R. Hardisty

R. Giovanni (renato@cria.org.br), Centro de Referência em Informação Ambiental, Av. Dr. Romeu Tórtima, 388, 13084-791, Campinas, SP, Brazil. – A. R. Williams, School of Computer Science, University of Manchester, Oxford Road, Manchester, M13 9PL, United Kingdom – V. H. Ernst and R. Kulawik, Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Postfach 20 07 33, 80007, München, Germany – F. Quevedo and A. Hardisty, Cardiff School of Computer Science & Informatics, 5, The Parade, Cardiff, CF24 3AA, United Kingdom.

Abstract

Ecological Niche Modelling (ENM) Components are a set of reusable workflow components specialized for performing ENM tasks within the Taverna workflow management system. Each component encapsulates specific functionality and can be combined with other components to facilitate the creation of larger and more complex workflows. One key distinguishing feature of ENM Components is that most tasks are performed remotely by calling web services, simplifying software setup and maintenance on the client side and allowing more powerful computing resources to be exploited. This paper presents the current set of ENM Components in the context of the Taverna family of tools for creating, publishing and sharing workflows. An example is included showing how the components can be used in a preliminary investigation of the effects of mixing different spatial resolutions in ENM experiments.

Introduction

By being able to predict and to understand species' distribution under different scenarios, ecological niche modelling (ENM) recently became one of the most popular techniques in biodiversity research, with direct impact in the number of published papers (Lobo et al. 2010) and related tools (see Peterson et al. 2011 for references). Most of the work done in this field uses the correlative approach (Soberón and Peterson 2005), in which species occurrence points are combined with environmental data, serving as inputs to a modelling algorithm. The resulting models can then be projected into different geographical regions under different environmental scenarios, producing potential distribution maps with a wide range of uses.

Although the typical ENM procedure is usually straightforward for a single species with some of the existing software, many experiments can be quite complex, requiring several steps, usually mixing different tools. In such cases, a workflow approach through workflow management systems may offer several benefits. Scientific workflows can specify a sequence of data retrieval, data manipulation and data storage/publication steps. When a scientific procedure or protocol is captured as a workflow, this allows the protocol to be easily shareable and re-runnable. In addition, provenance data of what happened during a workflow run allows for research to be, within certain limits, reproducible.

Considering the two most popular ENM software found by a recent survey (Ahmed et al. 2015), users seem to be divided between simplicity and flexibility, as if these two features would be irreconcilable in the same software. That is, if users are looking for an easy to use interface with a short learning curve, they must live with inflexible point-and-click software, whereas if they wish flexibility, they must develop programming skills to use syntax driven software. None of the tools found by the survey are based on workflow management systems, which actually have the potential to provide both a simple and flexible interface. The creation of scientific workflows is commonly carried out within a graphical user interface which may be desktop based, for example Taverna Workbench[1] (Wolstencroft et al. 2013) and Kepler (Altintas et al. 2004); or browser based, for example Taverna Online[2] and Galaxy (Giardine et al. 2005). Such interfaces allow users to visually build custom workflows, usually by means of adding boxes on a panel (each box representing a task) and connecting them through input/output parameters. This intuitive way to design and control personalised workflows is one of the main reasons for scientific workflows to be currently used in a large number of disparate domains, for example bioinformatics, astronomy and preservation of digital resources.

Most workflow systems allow different types of steps to be included within a workflow, such as running user-defined scripts, interacting with the user to display or get data, and calling external programs locally or remotely. In this last case, workflows may perform tasks by interacting with web services. Web services are software applications supporting dynamic interactions with other programs over the Internet through open standards. Using web services inside workflows may bring up issues related to the need of having an Internet connection and to the reliability and limitations of third-party service providers. However, web services also offer considerable advantages in terms of minimising the need for software installation and maintenance on the client side. There can also be more powerful computational resources behind web services, allowing workflows to outsource part of the processing requirements and not be strictly constrained by a desktop environment.

The Taverna suite of tools is a workflow management system allowing the creation, editing, sharing and running of workflows. Taverna workflows may be created and edited within the desktop Taverna Workbench or using the web-based Taverna Online. Workflows may be run: 1) directly within Taverna Workbench, 2) locally by the Taverna Command Line Tool or 3) remotely on the Taverna Server, which allows multiple simultaneous runs with secure user separation and offers a web service interface that can be remotely invoked by other programs. Additionally,

---

[1] http://www.taverna.org.uk

[2] http://onlinehpc.com

the running of a Taverna workflow can be included within a web application by using a self-contained software package called Taverna Player, similar to the way that videos are currently embedded within web applications. Taverna Player handles the marshalling of input data to and results from runs on a Taverna Server, also handling interaction requests from workflow runs. Taverna Player can be included in diverse web applications, such as IPython Notebook (Pérez & Granger 2007), Scratchpad (Smith et al. 2011) and web portals. Finally, any Taverna workflow can also be easily shared in the myExperiment platform[3].

Some of the recent developments in Taverna were carried out as part of the Biodiversity Virtual e-Laboratory (BioVeL) project[4]. BioVeL placed particular emphasis on setting up a robust web service infrastructure upon which scientific workflows can be built. This effort involved improving existing web services and creating new ones when necessary. All web services being used by BioVeL are registered in the Biodiversity Sciences Web Service Catalogue[5], including service endpoint, documentation, and monitoring information.

Workflows & ENM

Historically, ENM has been among case studies in many projects focused on scientific workflows. In 2004 the Biodiversity World project used the Triana workflow management system (Taylor et al. 2003) to build ENM workflows (Pahwa et al. 2006). Almost in parallel, the Science Environment for Ecological Knowledge (SEEK) project also created ENM workflows (Pennington et al. 2007), this time using the Kepler system. More recently, the ENM workflow approach was revived with the SAHM module in VisTrails (Morisette et al. 2013) and with the BioVeL project, where ENM is one of the major research areas (see Leidenberger et al. 2014 for an example).

---

[3] http://myexperiment.org
[4] http://biovel.eu
[5] http://biodiversitycatalogue.org

Even with all these initiatives, workflow management systems are still seen as a rather challenging environment for most researchers, traditionally requiring significant programming expertise to perform any different task that is typically needed when creating a custom workflow. Moreover, without sufficient specific analytical functions and features needed by ecologists and biodiversity researchers, the familiarisation effort required from researchers to start using workflow tools has not yet been perceived as sufficiently worthwhile. To overcome these challenges, one of the approaches explored at BioVeL has been to create families of workflow components specialized in common tasks for a certain area, such as ENM or phylogenetics. Each component is a sub-workflow representing a task-unit encapsulating implementation details. Components offer a high-level interface, allowing them to be more easily used and combined to create larger workflows.

BioVeL also created a web portal[6] where users can upload workflows or reuse workflows uploaded by other users. The portal allows users to start multiple workflow runs and retrieve results later, without needing an active Internet connection during the workflow run when there is no interaction involved. There are no additional requirements for a user to run a workflow through the portal except having an Internet browser.

Another major concern in BioVeL was to assure sustainability of assets beyond the project lifetime – especially considering that most of its workflows are strongly based on web services provided by different institutions. BioVeL's strategy to maintain a stable and persistent e-Infrastructure largely depends on institutional commitment, where each individual organisation takes responsibility to sustain various pieces of the e-Infrastrucure as part of its core business. A typical example is the ENM service provided by the Reference Center on Environmental Information (CRIA), which is currently used by ENM Components. The service has been running for many years at CRIA, well before the BioVeL project started, and will continue to run, as it is considered an important asset fully aligned to the institutional mission. Still regarding

---

[6] http://portal.biovel.eu

sustainability, BioVeL satisfies two pre-requisites pointed out by Henfridsson and Bygstad (2013) as being important factors for the adoption, spreading and evolution of a digital e-Infrastrucure: 1) loosely-coupled, service-oriented architecture and 2) decentralised management. All these factors contribute to the availability and improvement of ENM Components over time.

ENM Components

The ENM Components were created with the Taverna workflow management system as part of BioVeL to simplify the existing ENM workflows produced by the project and to facilitate the creation of new workflows. Since Taverna components are special workflows themselves, they enjoy the same benefits of the Taverna suite: they can be designed and run using the same tools, they can be reused by other workflows and even shared in myExperiment, where ENM Components are all publicly available under a specific pack with the same name[7] (note: to use them it is not necessary to manually download the pack, as Taverna Workbench can dynamically interact with myExperiment to fetch remote components).

A main aspect of providing reusable components is to document how they can be used. In this respect, each ENM component has a short description of its functionality and of each input/output parameter (also called ports). Being workflows, all components can be opened with Taverna workbench and run independently (all ports provide example values that can be used for testing). Using ENM Components to build new workflows within the workbench is only a matter of dragging the desired component from the service panel into the workflow being designed. To facilitate the connecting of different components, most ports with equal interfaces (same parameters and data types) are assigned the same name (fig. 1 shows how the main components can be combined). More information about how to use ENM Components can be found in the corresponding pack description in myExperiment. All ENM workflows developed in

---

[7] http://www.myexperiment.org/packs/563

BioVeL are based on ENM Components, providing many examples of their usage (see the generic “Ecological niche modelling workflow”[8] and the “Bioclim workflow”[9]).

The web service currently used by ENM Components was developed on top of openModeller (Muñoz et al. 2011). OpenModeller is a toolbox mainly comprised by an ENM framework with a comprehensive list of functions that can be called by other programs. The framework has many algorithms available and makes use of other software libraries to handle different data formats and spatial reference systems. The openModeller toolbox also contains a set of command-line tools and the web service itself, both making use of the framework and sharing most data structures for input/output parameters. Since ENM Components are strongly based on the openModeller web service, sometimes it may be necessary to refer to the web service documentation[10] when designing new workflows. For example, many ports of the ENM Components return or expect data according to openModeller serialization rules. The three-tiered structure currently used by ENM Components (component/web service/server software) actually allows for alternative implementations in the future, provided the same input/output ports remain the same for each component. For instance, a different web service implementation could be used (not necessarily associated with openModeller tools), or even all web service calls could be replaced with interactions to locally installed software. At the moment, the implementation of ENM Components takes advantage of all algorithms available in openModeller and of its capabilities to handle different data formats and spatial reference systems, interacting with a remote web service provided by CRIA.

Using remote web services in ENM tasks brings a few changes in the way researchers are used to working with traditional stand-alone tools. For example, the service needs to be queried to know which algorithms can be used. Over time, new or enhanced algorithms may become available on the service being called (information about the currently available algorithms can

---

[8] http://www.myexperiment.org/workflows/3355

[9] http://www.myexperiment.org/workflows/3725

[10] http://openmodeller.sf.net/web_service_2.html

also be found in the openModeller web site[11]). Frequently used environmental layers are available on the server for convenience, and again the service can be queried to return this information. Alternatively, additional layers or masks can be provided to the service, as the modelling engine can access other geospatial web services such as WCS[12] or remote files available over the web. In this case, layers need to be uploaded somewhere, for example a BioSTIF[13] server. BioSTIF provides a set of standardized services for spatial data visualization, transformation and storage. Some of the ENM Components rely on BioSTIF to visualize points and maps.

Example: the effect of mixing different spatial resolutions

During the BioVeL project, one of the case studies faced a common situation in ENM: environmental layers came from different sources in different spatial resolutions, i.e., having different cell sizes (see Leidenberger et al. 2014 for more details). Although the sensitivity of models to spatial resolution has already been studied before (Guisan et al. 2007), we could not find specific references about mixing layers with different resolutions. Probably the main reason is that most of the existing ENM software actually forces researchers to use layers with exactly the same resolution, spatial reference system and extent – even when differences are negligible. Since openModeller does not have this constraint – and consequently also the service used by the ENM Components – users are left with the decision about what to do when there are such differences between layers.

The usual practice when environmental datasets come in different resolutions is to previously downscale the low resolution layers by subdividing their cells, or upscale the high resolution ones by coarsening their cell size. It is also important to note that two main factors should be taken into account when dealing with spatial resolution in an ENM experiment: 1) the resolution

---

[11] http://openmodeller.sf.net/documentation.html
[12] http://www.opengeospatial.org/standards/wcs
[13] http://www.biodiversitycatalogue.org/services/7

of the biological phenomenon being studied, since each species may respond to environmental signals at different scales (Peterson et. al 2011) and 2) the spatial uncertainty of the occurrence points being used. Ideally, this uncertainty should not be greater than the environmental cell size, otherwise models will be generated with mistakenly precise environmental data, which tends to degrade model performance (Graham et al. 2008). Thus, when both factors are compatible with the finest environmental resolution at hand, which approach – downscaling or upscaling environmental layers – produces better models? In this example, different features of the ENM Components and the workflow approach are demonstrated, showing a possible way to investigate the subject. In particular, we demonstrate the flexibility and modularity of ENM Components combining them in a workflow that contains user interaction, loop, custom code and more than one tool. The workflow also explores some of the capabilities of openModeller, such as generating virtual niches and handling environmental layers in different resolutions.

OpenModeller handles differences in spatial resolution and reference systems by treating each layer independently and simply fetching the corresponding environmental data at each point (presence, absence or background). Therefore, mixing layers with different spatial resolutions in openModeller is essentially equivalent to downscaling the low resolution layers with the nearest-neighbor method, which retains the same original cell value in the new smaller cells. The only difference with other ENM software is that there is no raster downscaling pre-processing step inside or outside openModeller – it uses the original layers without modifications. In this example, we simulate the situation of having environmental layers in different spatial resolutions and compare the results of models generated with the original layers (mixed resolutions, equivalent to downscaling the low resolution layers) with models generated after upscaling the high resolution layers. The workflow created can be summarized in eleven steps (fig. 2), with the first step involving user interaction to choose the environmental layers and study area (mask), followed by a loop containing most steps, including virtual niche generation, point sampling, model creation and model testing. A final step after the loop compares the results and generates a graph using another tool.

The workflow can be downloaded from myExperiment[14] and requires: Taverna 2.5; R (R Development Core Team 2008) with the Rserve package installed and running as a localhost service in the default port (6311); an active Internet connection so that the workflow can communicate with the external ENM service currently hosted at CRIA[15], and a web browser to handle user interactions. The R version used was 3.1.2 and the Web browser was Firefox 34.0.5. When running this workflow using the Taverna Workbench with the default values (10000 background points and 30 iterations), it is highly recommended to disable provenance capture and in-memory storage in the system preferences. The workflow run takes about an hour to complete with the current resources on the web service, but it may take longer depending on connection and service load. A simplified version of the workflow with a single iteration and including model projections is also available[16].

The basic idea of the workflow is to compare models generated with mixed resolution layers (downscaling scenario) with models generated only with low resolution layers (upscaling scenario), testing them against the same set of points extracted from a virtual species niche. The workflow initially retrieves all available layers on the server and asks the user to choose a set of environmental layers and then a mask delimiting the study area. This initial step is performed by a nested workflow labeled “choose layers and mask”, containing only a few interconnected ENM Components and constant values used as input parameters. Each kind of workflow element in Taverna has a different background color and any workflow element can be renamed. Components are displayed with a pink background, and most ENM Components used by this workflow were renamed to better indicate their purpose (original names can always be found in the details of each component). There are currently a few mask options offered by the

---

[14] http://www.myexperiment.org/workflows/4535
[15] http://modeller.cria.org.br/ws2/om
[16] http://www.myexperiment.org/workflows/4536

ENM service, all of them based on political boundaries, which does not affect an arbitrary mask choice for this study. For simplicity, it can even be assumed for any chosen mask that the whole area has been historically accessible to the virtual species that will be created in one of the next steps, so that presence points can be undoubtedly interpreted as being suitable for the species, and absence points unsuitable. After choosing a mask, the user is then asked to select a set of high resolution environmental layers, and in the next step to pick the corresponding low resolution ones. The choice of environmental layers is also arbitrary, and we can also assume that the chosen layers are the main variables that determine the virtual species' niche. For the purpose of this experiment, the only constraint when choosing layers is to select variables that are available at least in two different resolutions. Worldclim bioclimatic variables (Hijmans et al. 2005) are available on the ENM service in 30 arc-seconds and 10 arc-minutes resolutions, making them a convenient choice for the demonstration. Additionally, Worldclim layers were originally produced in 30 arc-seconds, with all other low resolution versions – including the 10 arc-minutes one – being obtained by upscaling (Hijmans et al. 2005). At the end of this initial step, the workflow has two sets of environmental layers with exactly the same variables, each set with a different spatial resolution.

Next, the workflow uses another ENM Component to randomly sample 10000 background points over the whole study area. At this stage, two other elements are used to demonstrate how to include custom code in a workflow. Depicted in brown background, they are known as Java BeanShells[17] in Taverna. One of them ("merge all layers") concatenates the identifiers of all selected layers in a single string list before sampling background points, to ensure that all sampled points have valid values across all different layers and resolutions. The other one ("for loop triggering") simply creates a list with the same size of the "replicates" workflow parameter, since workflow loops can be activated by lists. Although custom code may require programming skills, BeanShells can easily be transformed into new workflow components if necessary, and

---

[17] https://jcp.org/en/jsr/detail?id=274

then stored in specific components families to be used by other users. A few additional BeanShell examples can be found in other parts of the workflow.

Next, we use a workflow loop to repeat the same steps a specified number of times. These steps are inside the “create and test models” nested workflow, where a virtual niche is generated, training and testing points are sampled based on the virtual species distribution, and finally the two models for each set of layers are generated and tested. In the first part, the ENM Component for sampling points is used again to sample a single point to be passed as a parameter to the Virtual Niche algorithm in openModeller. This algorithm assumes that the corresponding environmental values for the point are the optimal conditions for the virtual species, randomly defining standard deviations for each variable to create a continuous niche across the study area. This is all performed with the high resolution environmental layers, producing a high resolution niche to be considered the truth for the virtual species. The corresponding niche is then evaluated over all background points to get the niche values, which are ordered and split based on a random threshold separating suitable from unsuitable conditions, ensuring a random arbitrary prevalence between 0.1 and 0.7. These two groups of points (suitable/unsuitable) are used to randomly sample presence points for model creation (a number between 30 and 100) and 100 points for independent model testing (50 presences and 50 absences). Finally, the workflow creates two models using one of the most popular ENM algorithms also available in openModeller: Maxent (Phillips et al. 2006). The first model is created with the corresponding low resolution environmental layers (upscaling scenario) and the other with a random balanced mix of high and low resolution layers (downscaling scenario). These models are tested with the same testing points by measuring the area under the Receiver Operating Characteristic curve (AUC) – a threshold-independent test suitable for algorithms that produce a continuous (non binary) output such as Maxent. AUC values range from 0 to 1, where 1 indicates perfect discrimination between the presence and absence points being tested, 0.5 indicates a predictive discrimination equivalent to a random guess, and values below 0.5 indicate discrimination worse than random. All steps from virtual niche generation until model tests are repeated 30 times in the workflow to generate enough variation in the virtual niche, training points, testing points and resolution mix with the selected layers. In the last part of the workflow, results are compared using an R script which also produces a graph plotting side by side AUC values for each set of layers in each iteration. This way, the example also demonstrates how to use different tools in different parts of the same workflow. The probability (p-values) of getting a better model when mixing resolutions (downscaling scenario) instead of using only low resolution layers (upscaling scenario) is estimated as the percentage of times that the former AUC is greater than the later one. This is a two-tailed test also used by Elith et al. (2006) to compare the performance of different algorithms. A value close to 1 means that mixing resolutions produces better models than using only low resolution layers, and vice-versa for a value close to 0.

In the first workflow run, we used Mexico as the mask and WorldClim bio2 (mean diurnal range), bio5 (maximum temperature of warmest month), bio6 (minimum temperature of coldest month), bio12 (annual precipitation) and bio14 (precipitation of driest month) as the environmental variables. Most models using mixed layer resolutions produced better AUCs, although the differences were small (fig. 3) and the result was not significant (p=0.73). We also used the simplified version of the workflow with the same parameters to project models, illustrating a virtual species distribution (fig. 4) and its corresponding projected model with mixed resolutions (fig. 5). Back to the complete workflow, an identical pattern was found in a subsequent run with different parameters: Finland as the mask and bio2, 3, 4, 6, 13 and 14 as the environmental variables (p=0.73). A third run using India as the mask and bio1, 4, 11, 15 and 16 as the environmental layers pointed to the same direction, but with less intensity (p=0.53).

Since the main purpose of the example was to demonstrate the use of ENM Components, we tried not to add more complexity to the workflow. For a more extensive investigation, future versions of the workflow could for example include automatic variation of mask, number of layers and proportion of mixed layers, also including more spatial resolutions. An additional step to produce biased training points could produce a wider and more realistic range of AUCs. Other modelling algorithms could be tested as well.

Even being just a preliminary investigation, the example shows how the ENM Components can be combined to produce unique scientific workflows. Additionally, the workflow also shows how to include other tools into the same workflow, such as the currently ubiquitous R, and how to include custom code, which can be transformed into new components whenever necessary. Another possibility for new workflows is to combine components from different areas, such as the phylogenetics components also created during the BioVeL project, or to benefit from other Taverna-related tools, such as the workflow parameter optimization plug-in that can be used with ENM (Holl et al. 2013). There are still many practical uses and research opportunities in ENM that can be explored, and we hope that ENM Components can provide a flexible and powerful alternative for future works in this area.

Acknowledgements – This work is part of the BioVeL project with funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 283359.

References

Ahmed, S.E. et al. 2015. Scientists and software – surveying the species distribution modelling community. – Divers Distrib, 21(3): 258–267.

Altintas, I. et al. 2004. Kepler: an extensible system for design and execution of scientific workflows. – in: Proc 16th Int Conf Scientific and Statistical Database Manage, pp.423–424.

Elith, J. et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. – Ecography, 29(2): 129–151.

Giardine, B. et al. 2005. Galaxy: a platform for interactive large-scale genome analysis. – Genome Res, 15(10): 1451–1455.

Graham, C. H. et al. 2008. The influence of spatial errors in species occurrence data used in distribution models. – J Appl Ecol, 45(1): 239–247.

Guisan, A. et al. 2007. Sensitivity of predictive species distribution models to change in grain size. – Divers Distrib, 13(3): 332–340.

Henfridsson, O. and Bygstad, B. 2013. The generative mechanisms of digital infrastructure evolution. – Mis Quarterly 37(3): 907–931.

Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int J Climatol, 25: 1965–1978.

Holl, S. et al. 2013. On specifying and sharing scientific workflow optimization results using research objects. – in: Proc 8th Workshop on Workflows in Support of Large-Scale Science, New York, ACM Press, pp.28–37.

Leidenberger, S. et al. 2014. Mapping present and future predicted distribution patterns for a meso-grazer guild in the Baltic Sea. – J Biogeogr, advance online publication, doi: 10.1111/jbi.12395

Lobo, J. M. et al. 2010. The uncertain nature of absences and their importance in species distribution modelling. – Ecography, 33:103–114.

Morisette, J. T. et al. 2013. VisTrails SAHM: visualization and workflow management for species habitat modeling. – Ecography 36: 129–135.

Muñoz et al. 2011. openModeller: a generic approach to species' potential distribution modelling. – Geoinformatica, 15: 111–135.

Pahwa, J. S. et al. 2006. Biodiversity World: A Problem-Solving Environment for Analysing Biodiversity Patterns. – in: Proc 6th IEEE Int Symp Cluster Computing and the Grid (CCGRID), Singapore, pp. 201–208.

Pennington, D. D. et al. 2007. Ecological Niche Modelling Using the Kepler Workflow System. – in: Taylor, I.J., Deelman, E., Gannon, D.B., Shields, M. (Eds.), Workflows for e-Science, Springer, Berlin, pp.91–108.

Pérez, F. and Granger, B. E. 2007. IPython: a system for interactive scientific computing. – Comput Sci Eng, 9(3): 21–29.

Peterson, A. T. et al. 2011. Ecological Niches and Geographical Distributions. – Princeton University Press.

Phillips, S. J. et al. 2006. Maximum entropy modelling of species geographic distributions. – Ecol Model, 190: 231–259.

R Development Core Team. 2008. R: A language and environment for statistical computing. – R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Smith, V. S. et al. 2011. Scratchpads 2.0: a Virtual Research Environment supporting scholarly collaboration, communication and data publication in biodiversity science. – Zookeys, 150: 53–70.

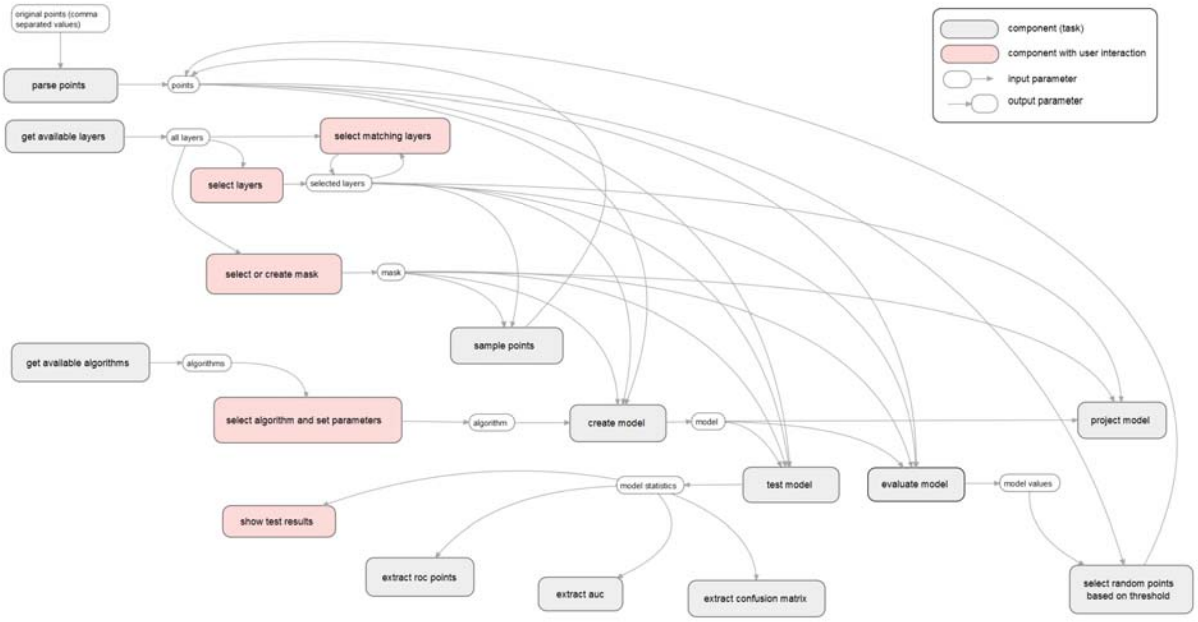
Soberón, J. and Peterson A. T. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. – Biodiversity Informatics 2: 1–10.

Taylor, I. et al. 2003. Triana Applications within Grid Computing and Peer to Peer Environments. – J Grid Comput 1(2): 199–217.

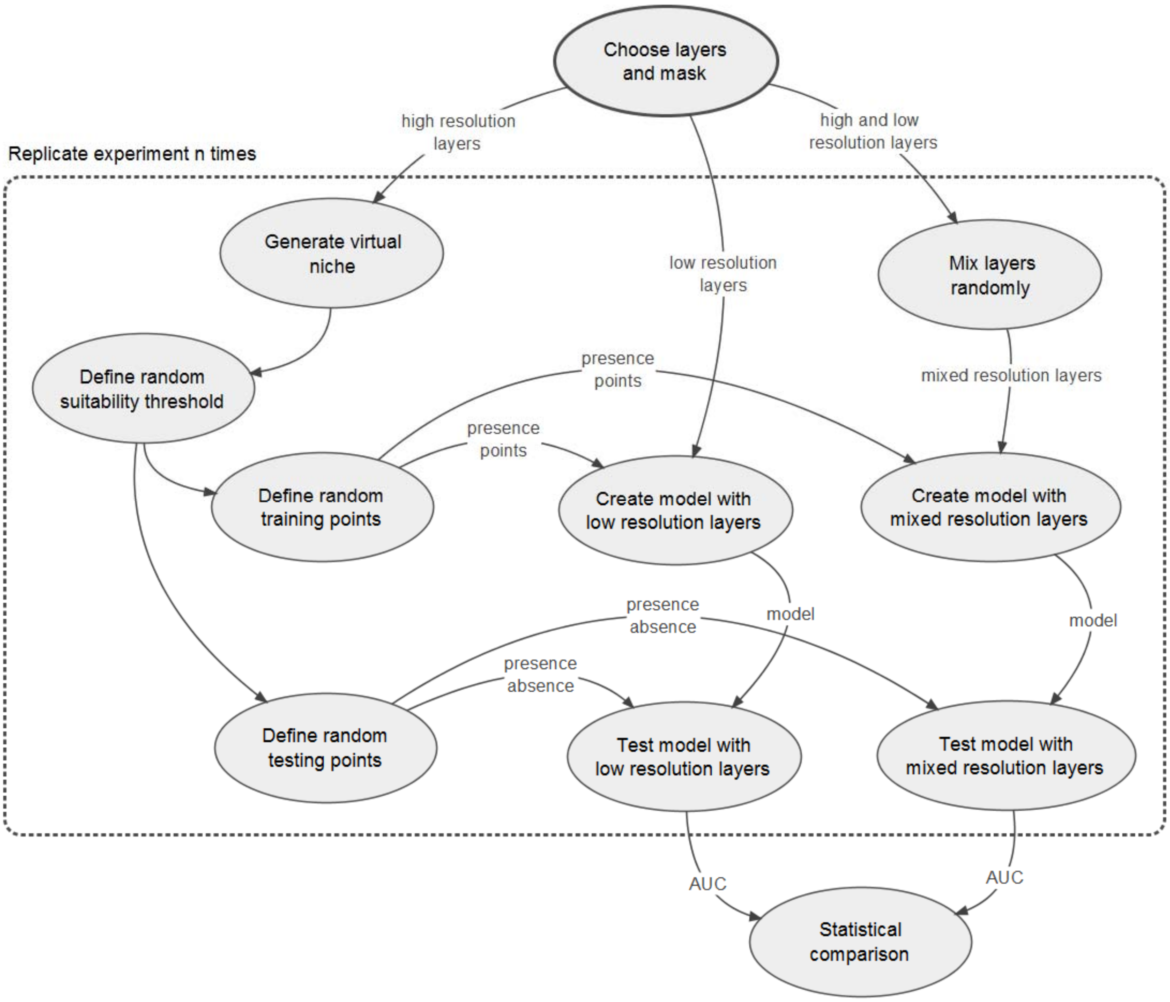
Wolstencroft, K. et al. 2013. The Taverna workflow suite: designing and executing workflows of Web Services on the desktop, web or in the cloud. – Nucleic Acids Res 41(1): 557–561.

Figures

Figure 1:



390 Figure 2:



391

392

393

Figure 3:

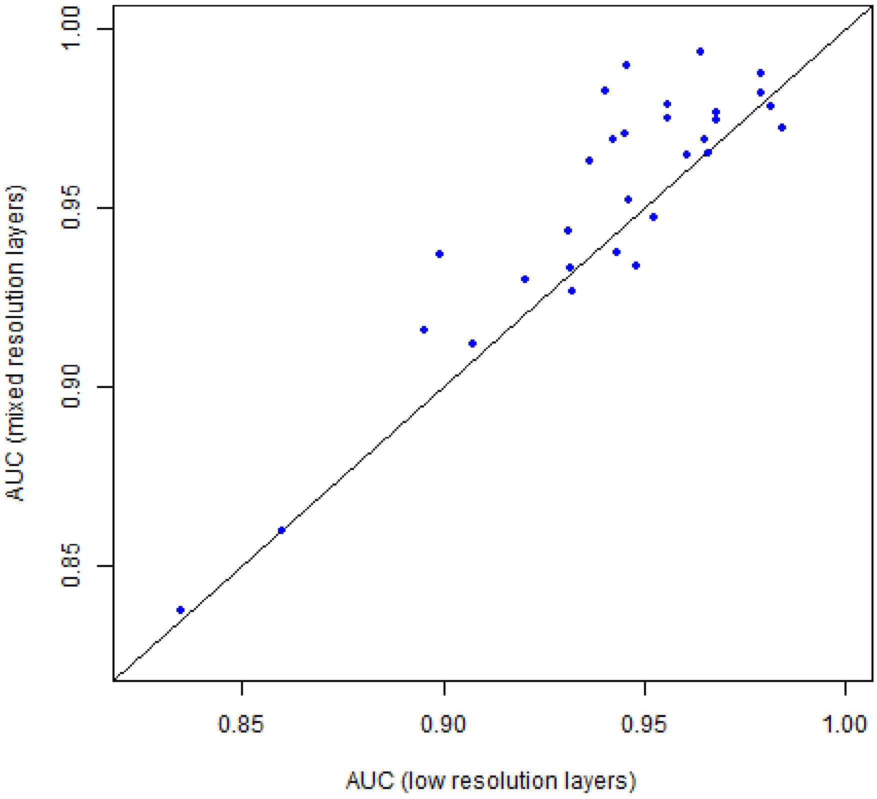

AUC (mixed resolution layers)
0.85
0.90
0.95
1.00
AUC (low resolution layers)

Figure 4:

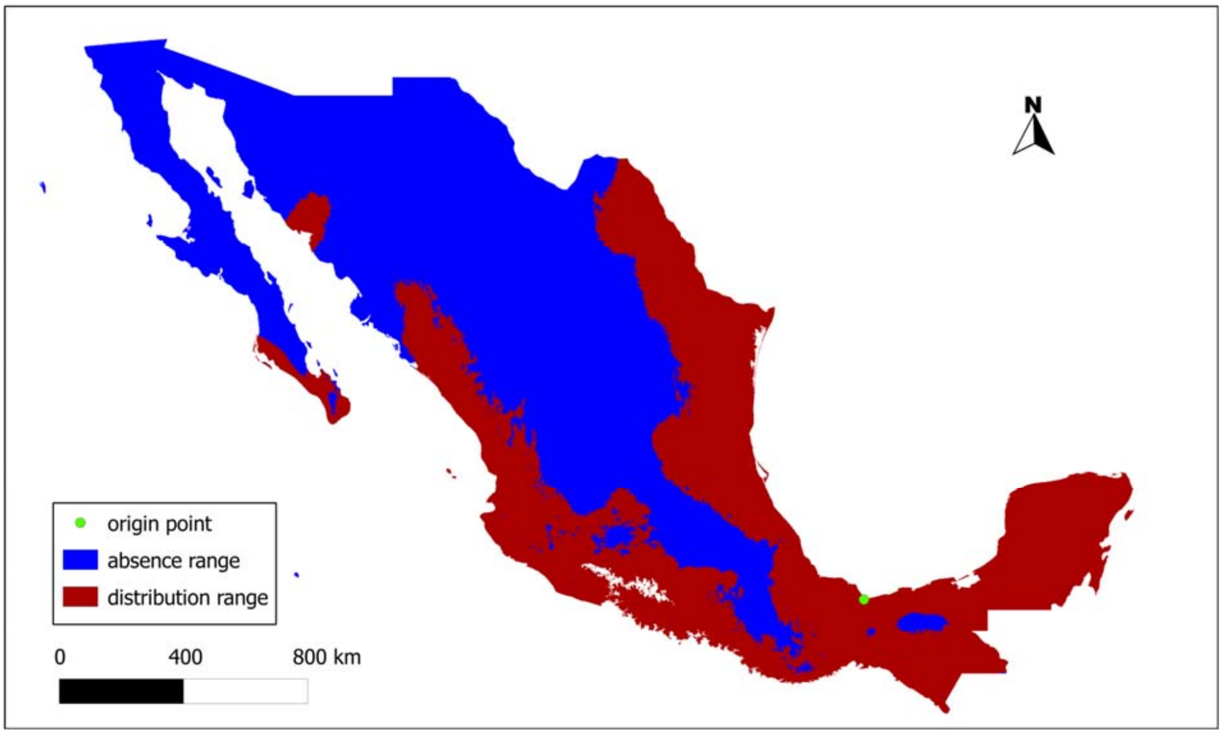

N
origin point
absence range
distribution range
0
400
800 km

Figure 5: