*International Environmental Modelling and Software Society (iEMSs)*
*7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,*
*Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds.)*
*http://www.iemss.org/society/index.php/iemss-2014-proceedings*

# Supporting environmental modelling with Taverna workflows, web services and desktop grid technology

**Ferenc Horváth[1], P. Ittzés[1], D. Ittzés[1], Z. Barcza[12], L. Dobor[2], D. Hidy[3], A. Marosi[4], A. Hardisty[5]**
*[1]Institute of Ecology and Botany, Centre for Ecological Research*
*Hungarian Academy of Sciences, Alkotmány u. 2-4., 2163 Vácrátót, Hungary*
*({horvath.ferenc, ittzes.peter, ittzes.dora}@okologia.mta.hu)*
*[2]Department of Meteorology, Eötvös Loránd University, Pázmány sétány 1/A,*
*1117 Budapest, Hungary (bzoli@elte.hu, doborl@nimbus.elte.hu)*
*[3]Institute of Botany and Ecophysiology, Szent Istvan University,*
*Páter K. u.1., 2100 Gödöllő, Hungary (hidy.dora@mkk.szie.hu)*
*[4]Laboratory of Parallel and Distributed Systems, Institute for Computer Science*
*and Control, Hungarian Academy of Sciences, P.O Box 63,*
*1518 Budapest, Hungary (marosi.attila@sztaki.mta.hu)*
*[5]School of Computer Science & Informatics, Cardiff University, Queen's Buildings,*
*5 The Parade, Roath,Cardiff CF24 3AA, UK (Alex.Hardisty@cs.cardiff.ac.uk)*

**Abstract:** Ecosystem functioning, climate change, and multiple interactions among biogeochemical cycles, climate system, site conditions and land use options are leading-edge topics in recent environmental modelling. Terrestrial ecosystem models are widely used to support carbon sequestration and ecosystem studies under various ecological circumstances. Our team uses the Biome-BGC model (Numerical Terradynamic Simulation Group, University of Montana), and develops an improved model version of it, called Biome-BGC MuSo. Both the original and the improved model estimate the ecosystem scale storage and fluxes of energy, carbon, nitrogen and water, controlled by various physical and biological processes on a daily time-scale. Web services were also developed and integrated with parallel processing desktop grid technology. Taverna workflow management system was used to build up and carry out elaborated workflows like seamless data flow to model simulation, Monte Carlo experiment, model sensitivity analysis, model-data fusion, estimation of ecosystem service indicators or extensive spatial modelling. Straightforward management of complex data analysis tasks, organized into appropriately documented, shared and reusable scientific workflows enables researchers to carry out detailed and scientifically challenging 'in silico' experiments and applications that could open new directions in ecosystem research and in a broader sense it supports progress in environmental modelling. The workflow approach built upon these web services allows even the most complicated computations to be initiated without the need of programming skills and deep understanding of model structure and initialization. The developments enable a wider array of scientists to perform ecosystem scale simulations, and to perform analyses not previously possible due to high complexity and computational demand.

***Keywords**:* Biome-BGC; terrestrial ecosystem; model data fusion framework; ecosystem services; virtual laboratory; biodiversity.

## 1. INTRODUCTION

### 1.1 Science-policy context

As it is stated in the Millennium Ecosystem Assessment (MEA 2005) „Over the past 50 years, humans have changed ecosystems more rapidly and extensively than in any comparable period of time in human history, largely to meet rapidly growing demands for food, fresh water, timber, fiber, and fuel". The extensive exploitation of ecosystems has contributed to substantial net gains in economic

development, but these gains have been achieved at growing costs in the form of biodiversity loss, the degradation of many ecosystems and their human health and well-being related services, and the exacerbation of poverty for considerable groups of people (MEA 2005). From a more global viewpoint, the anthropogenic pressures and growing trends on the Earth's biogeochemical system have reached a scale where safe living of humanity cannot be safeguarded in some regions and even at the global scale (Rockström et al. 2009). Although the future of biodiversity will be determined by society, at the same time science is required to develop and improve effective tools and sound scientific methods for ecosystem classification, valuation, modelling, prediction and ecosystem management technologies, as well as to assess a variety of restoration and trade-off analysis between management options and ecosystem services at even finer spatially explicit resolution (Haines-Young et al. 2012).

Climate change is undoubtedly one of the greatest challenges of the 21st century. Huge amounts of carbon dioxide (and other greenhouse gases) have been released to the atmosphere via burning of fossil fuels, cement manufacturing, land use conversion and other processes, which can unbalance the energy budget of the Earth-atmosphere system and can trigger global warming (IPCC, 2013). At present terrestrial biosphere (and also the oceans) mitigate climate change as biosphere and oceans together sequester about half of the $CO_2$ that is emitted to the atmosphere annually by human activity. Climate change has already modulated the carbon balance of terrestrial biosphere (e.g. Ciais et al. 2008; Churkina et al. 2010), and it is reasonable to assume that it will further interfere with the carbon cycle (Friedlingstein et al. 2010). Lack of appropriate knowledge about the future evolution of terrestrial carbon sequestration is a major problem which causes the main uncertainty in the climate projections. In order to reduce this uncertainty efforts are clearly needed in the field of carbon cycle research. This involves the improvement of existing biogeochemical models, and creation of information technology solutions to foster ‘in silico’ experiments that are computationally intensive.

Biodiversity informatics (or ecoinformatics in the sense of Dengler et al. 2011, Michener et al. 2012) plays a central role in the research community's efforts to address the urgent questions around land-use, environmental change, sustainability, food security and ecosystem services that are facing Governments worldwide. It can provide ultimate IT tools, but it is clear that some barriers to progress are also sociological, basically persuading people to use the technological solutions that are already available. This is best addressed by developing more effective coordinated coupled modelling systems that deliver immediate benefit to the user, hiding the majority of the technology behind simple user interfaces relating to our use of the natural environment (Hardisty et al. 2013).

### 1.2 The model-data fusion framework

Due to the well-recognized uncertainties related to the future evolution of the global carbon cycle improvements are needed in biogeochemical models that are essential components of the state-of-the-art Earth System Models (IPCC 2013). Model development and validation would not be possible without measurement data. Due to considerable progress in the availability of experimental data in the past few decades (e.g. FLUXNET data, biomass measurements, plant trait databases, remote sensing, etc.) new possibilities opened in biogeochemical model development.

The concept of model-data fusion (MDF; Williams et al. 2009) involves all important steps in biogeochemical model development and application. As it is stated by Williams et al. (2009) biogeochemical model calibration, evaluation of model performance, model testing and model improvement are all key elements of the MDF concept. Most importantly, experimental data plays a central role in model development, optimization, and validation. Biogeochemical models are highly complex, nonlinear mathematical tools that usually contain many simplifications and empirical equations. These simplifications may hamper the reliable estimation of the response of ecosystems to changing environmental conditions. In this sense MDF inevitably means construction of models with ever increasing complexity (and an increasing number of adjustable model parameters). Increasing model complexity usually means higher computational demand and implementation of complex mathematical tools for model application. These requirements raise the need to find IT solutions for the different components of the MDF procedure. Within the framework of the BioVeL project our emphasized aim is to address the main components of MDF, jointly with state-of-the-art computational solutions.

### 1.3 Biome-BGC model developments

Biome-BGC – developed by the Numerical Terradynamic Simulation Group, University of Montana – is a widely used, biogeochemical model that simulates the storage and flux of water, carbon, and nitrogen between the ecosystem and the atmosphere, and within the components of the terrestrial ecosystems (Thornton, 2000). Several researchers used and modified the original Biome-BGC model in the past (Churkina et al. 2003; Vetter et al. 2008; Trusilova et al. 2009). Most recently our research group developed Biome-BGC to improve the ability of the model to simulate carbon and water cycle in managed herbaceous ecosystems (Hidy et al. 2012). The modifications included structural improvements of the model (e.g., the simple, outdated, one-layer soil module was replaced by a multilayer soil module; drought related plant senescence was implemented; model phenology was improved). Since that publication additional modules were developed to simulate annually varying cropland management (i.e. ploughing, planting, harvesting, application of fertilizers). Forest thinning was also implemented and included as a possible human intervention, and dynamic whole plant mortality was implemented in the model to enable more realistic simulation of forest stand development. In the most recent model version separate pools have been defined for fruit following the method of Ma et al. (2011) to support cropland related simulations. Our improved model is called Biome-BGC MuSo v2.2 (where the abbreviation refers to Multilayer Soil module).

## 2. THE KEY INFORMATICS TECHNOLOGIES

### 2.1 Taverna scientific workflow management system and the BioVeL Portal

Workflow technology plays a key role as a means for managing complex distributed computational experiments in virtual laboratories (Gil et al. 2007; Littauer et al. 2012). Reusing and re-purposing scientific workflows allow scientists to do new analysis faster, since the workflows capture useful expertise from others. As workflow libraries grow (i.e. myExperiment for collections of workflows written in Taverna or other systems), scientists face the challenge of finding workflows appropriate for their task, understanding what the workflow does, and reusing relevant portions of a given workflow (Garijo et al. 2013).

An open source, domain independent, fully featured, extensible and scalable workflow management system: the Taverna workflow tool suite (includes the Workbench – desktop client application; the Command Line Tool for a quick execution of workflows from a terminal; the Server for remote execution of workflows; the Player – web interface plugin for submitting workflows for remote execution; integrated with myExperiment – social networking and workflow sharing environment for scientist of different domains; and integrated with a curated catalogue of web services) was selected. It is designed to combine distributed Web Services and local tools into complex analysis pipelines (Wolstencroft et al. 2013). Taverna can invoke generic WSDL-style, soap or rest style web services, local Java services (Beanshell scripts), local Java API (API Consumer), R scripts on an R server (Rshell scripts). It can also access others, such as BioMoby, BioMart and SoapLab services, and import data from a csv or Excel spreadsheet and interact with users in a browser. These pipelines can be executed on desktop machines or through larger infrastructure, using the Taverna Server. The main advantage of using workflows and distributed services, however, is that the majority of computational processing in the workflow occurs remotely with the Web Service providers. There is no requirement to install tools and data sources locally, which reduces local infrastructure and maintenance costs and enables rapid workflow development and testing.

The BioVeL portal (powered by Taverna Server) running at http://portal.biovel.eu manages and runs the workflows developed by the Biodiversity Virtual e-Laboratory project (http://www.biovel.eu/). These include workflows for taxonomy, phylogenetics, metagenomics, population modelling, ecological niche modelling and ecosystem modelling. Workflows can set to be private (only visible to the owner), shared with project or institution members or public. The portal provides (or reaches) all the software tools and IT services required to manage and run workflows, so the users doesn't need to install any software components or to deal with constraints of resources (e.g. memory, storage, computation capacity). All runs and results can be managed on the portal and shared with other users or downloaded.

### 2.2 Web service technology

The workflow and distributed web service concept is historically used in many different science domains (i.e. particle physics, bioinformatics, cosmology, mathematics). Although it has great potential to support ecosystem modelling, we have found only a few projects related to this technology. The ARIES and DOPA projects (Bagstad et al. 2011; Dubois et al. 2011) can be mentioned here as examples. There were no existing re-usable web services of terrestrial ecosystem modelling, so it was an open and opportunistic 'niche' to develop some according to our user's demands. The existing functions of a simple or a complex executable model, such as Biome-BGC can be wrapped into a single web service according to the 'model-as-a-service' (MaaS) concept of Roman et al. (2009), but a much more complex processing chain or a hierarchically embedded set of functions can be developed as a separate web service as well.

## 3. RESULTS

### 3.1 The Biome-BGC specific web services in relation to the 'Model–Data–Fusion' framework and the 'Biodiversity Catalogue'

As a user's group of the model, we come up to the decision of following the 'Model–Data–Fusion' framework at planning and developing the Biome-BGC model based web services. The *rest*-style web service applications are written in Python and based on the Flask micro framework running on the ECOS – Ecosystem functioning & valuation linux server. *Table 1* gives an overview of the developed service set. All new services must be well-documented and discoverable. Hence, a fully curated, well-founded catalogue or registry of web services for biodiversity science has been established by the BioVeL project: the 'Biodiversity Catalogue' (http://www.biodiversitycatalogue.org/). This catalogue provides four basic functions: i) discovering and finding the right web services; ii) registering new web services; iii) annotating and documenting; and iv) monitoring, curating of web services.

**Table 1.** Overview of the Biome-BGC model based web services

| Service Name | Description |
|---|---|
| Get local meteorological datasets | It provides a local daily meteorological dataset for modelling at given locations from a database, using e.g. the FORESEE database (under construction). |
| Get coarse spatial, gridded meteorological datasets | It provides spatial/gridded meteorological datasets or scenarios for modelling at landscape level requested from a database, such as FORESEE (under construction). |
| Generate fine spatial, gridded meteorological datasets | The service runs MTCLIM model to produce a fine spatial meteorology dataset, which consider terrain interaction and measured weather data. |
| Biome-BGC CARBON | The service runs a single Biome-BGC model version instance. Different model versions, parameter sets and input datasets can be selected. |
| Biome-BGC Ecosystem Service Indicators | The service runs a single Biome-BGC model instance and provides a set of Ecosystem Service Indicators. |
| Biome-BGC Monte Carlo Experiment (on grid) | It runs on the ECOS server and interact with the EDGeS@Home desktop grid to manage executions of a high number of instances with randomly sampled input parameter sets. |

*Table 1 continued*

| Service Name | Description |
|---|---|
| Biome-BGC Sensitivity Analysis | It runs on the ECOS server, gets the results of a Monte Carlo Experiment executed previously, than calculates parameter sensitivity of selected output variables. |
| Biome-BGC Parameter Estimation | It provides a 'calibration' of model settings with field measurements, based on comparison and evaluation of the results of a Monte Carlo Experiment and measured data sets. |
| Biome-BGC Spatial Extension (on grid) | It runs a bunch of Biome-BGC model instances on the EDGeS@Home desktop grid. Different model versions, parameter sets and input datasets can be selected (under construction). |

For the latest and current information about deployed and published services, developers are advised to consult the Biodiversity Catalogue (http://www.biodiversitycatalogue.org/) directly.

### 3.2 The Biome-BGC Projects database (BBGCDB)

Originally the model is an executable, that requires input and various parameter settings in a given directory/file/plain text coding format way (Thornton 2000) difficult to control and causing long learning process for specialized users. A Drupal (content management platform) database with user friendly web interface was developed as a kind of middleware between Biome-BGC, web services and the workflows using them. This Biome-BGC Projects database (http://ecos.okologia.mta.hu/bbgcdb/) provides also interaction pages for Taverna workflows and helps to manage data files associated with the workflows. The main aims of BBGCDB are: i) to support developments of ecosystem modelling web services and workflows; ii) to help manage and launch single run or multiple run Biome-BGC investigation projects embedded in Taverna workflow environment; iii) to shorten learning curve of the preparation phase of model simulations and provide quality assurance on initial datasets, input files and related settings; iv) to help integration and interaction with desktop grid and/or GIS technology; and v) to support management of users and scientific investigation projects.

BBGCDB is presented in Figure 1 in the centre of the graph, because it is interrelated with all of the other components due to its central data and investigations management functions.

Two versions of the model were implemented and deployed: Biome-BGC v4.1.1 Max Planck Institute (Trusilova et al. 2009) and MuSo v2.2 (Hidy and Barcza 2013). Different types of Biome-BGC investigations are managed in the database. There are single run investigations of the model (CARBON and Ecosystem Service Indicators) and there are investigations where multiple Biome-BGC runs are handled together. Monte Carlo Experiment (MCE) and Spatially Explicit (GIS) investigations are those types that consist of several Biome-BGC runs with different parameterisation.

### 3.3 Integration with desktop grid technology

Monte Carlo Experiments and Spatial Extension investigations require large numbers of executions of Biome-BGC model (e.g. several hundreds of thousand). Such high computational demands can be supported by desktop grid services. Both versions of the model have been ported to the EDGeS@Home desktop grid system, which is built on BOINC volunteer desktop grid technology. The model is running with the help of the GenWrapper environment, that is a generic BOINC wrapper for legacy applications utilizing GitBox: use POSIX like shell scripting and built-in commands (Balaton et al. 2007, Marosi et al. 2009), and the jobs are submitted through the web service interface of a 3G Bridge component to the BOINC server (Kacsuk et al. 2009). All these components are integrated into Monte Carlo Experiment and Spatial Extension services.

The porting was performed by MTA Centre for Ecological Research with the support of MTA SZTAKI. It was an iterative effort, because of consequential need for fixing the problems revealed by the

'porting – testing – improving' process. The Biome-BGC applications were internally tested first on a small scale private desktop grid before being deployed on a test infrastructure: "EDGI Demo" Desktop Grid provided by MTA SZTAKI. Then the 'matured' applications were deployed on the hosting EDGeS@Home Desktop Grid infrastructure (http://home.edges-grid.eu/home/). The on grid services speed up the Biome-BGC execution by 20-50 times according to the first benchmark testing, however it highly depends on output settings (writing daily outputs generate big files to transfer).

### 3.4 Biome-BGC modelling based workflows

All workflows utilize combinations of the web services listed in *Table 1*, and served by Biome-BGC Projects database due to the interaction page mechanism of Taverna workflow. Biome-BGC CARBON and Ecosystem Service Indicators workflows are based on single run instances executed 'immediately' by the ECOS server, while workflows exploiting grid-based and/or time consuming services (i.e. Monte Carlo Experiment, Sensitivity Analysis, Parameter Estimation, Spatial Extension) launch the service, than leave it to run in the background. The BBGCDB manages these services and provides information about the status of work launched. Sensitivity Analysis and Parameter Estimation workflows require the raw (preliminary) results of a Monte Carlo Experiment, so these workflows are consecutively linked.

For the latest and current information about workflows, the authors are advised to consult the Biodiversity Virtual e-Laboratory group on myExperiment directly (http://www.myexperiment.org/groups/643.html).

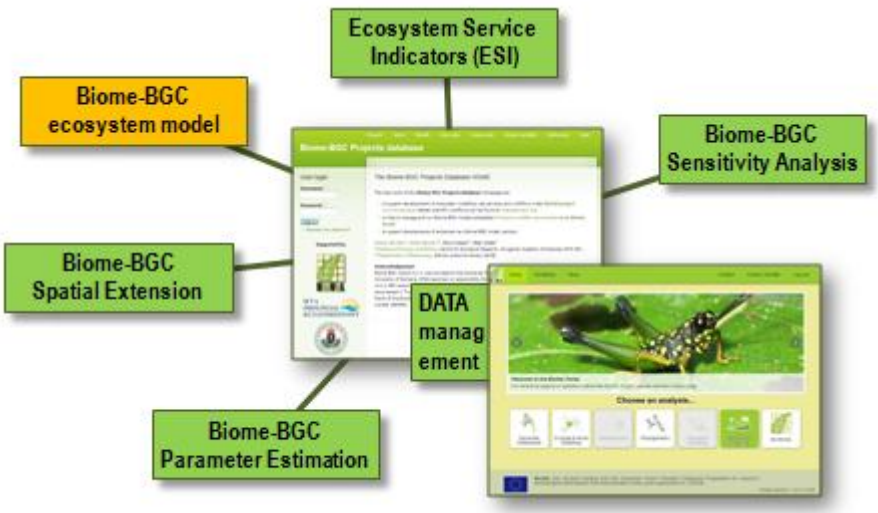


**Figure 1.** Biome-BGC workflow and web service functionalities developed in relation to the MDF framework according to Williams et al. (2009)

## 4. CONCLUSIONS

By supporting service discovery, workflow design, reuse and execution, the Taverna tool suite enables the exploitation of distributed environmental data and analysis methods. The Biome-BGC based services are interlinked with each other, so they can be considered as a specific service network providing the most important and flexible functionalities through the 'Model–Data–Fusion' framework. Although the web services are embedded and deployed on rather different computing infrastructures, workflows can easily orchestrate the web services to fulfil the ultimate scientific goals: effective, flexible and fully reproducible investigations.

The developments enable a wider array of scientists to perform ecosystem scale simulations and to perform analyses not previously possible due to high complexity and computational demand. Detailed

documentation, sample datasets and well prepared learning materials and training events can help to overcome knowledge barriers.

The desktop grid technology can provide a sustainable solution for computationally demanding tasks, even in an environment of continuously changing/evolving IT.

The authors hope to open new perspectives for biodiversity and environmental analysis. However several gaps remained unbridged and we have to count some constraints also: i) further shortage in relevant dataset services; ii) difficulties in finding appropriate data and services; iii) facing the inter-disciplinary gap; iv) data inconsistency; and finally v) model construction constraints - according to the well-known bon mot of George E. P. Box *„Essentially, all models are wrong, but some are useful”*.

## 5. ACKNOWLEDGMENTS

The research was supported by the BioVel (Biodiversity Virtual e-Laboratory Project, FP7-INFRASTRUCTURES-2011-2, project number 283359), the Hungarian Scientific Research Fund (OTKA K104816) and the EU FP7/2007-2013 under grant agreement no 312297 (IDGF-SP).

## 6. REFERENCES

Bagstad, K.J., Villa, F., Johnson, G.W., Voigt, B., 2011. ARIES - Artificial Intelligence for Ecosystem Services: A guide to models and data, version 1.0.

Balaton, Z., Gombás, G., Kacsuk, P., Kornafeld, Á., Kovács, J., Marosi, A.C. et al., 2007. SZTAKI Desktop Grid : a Modular and Scalable Way of Building Large Computing Grids, in: Proceedings of the 21st International Parallel and Distributed Processing Symposium (IPDPS 2007), 26-30 March 2007, Long Beach, California, USA. pp. 1–8.

Bhagat, J., Tanoh, F., Nzuobontane, E., Laurent, T., Orlowski, J., Roos, M., Wolstencroft, K., Aleksejevs, S., Stevens, R., Pettifer, S., Lopez, R., Goble, C. 2010. BioCatalogue: a universal catalogue of web services for the life sciences. Nucleic Acids Research 38: 689-694.

Churkina, G., Tenhunen, J., Thornton, P., Falge, E.M., Elbers, J.A., et al. 2003. Analyzing the Ecosystem Carbon Dynamics of Four European Coniferous Forests Using a Biogeochemistry Model. Ecosystems 6: 168-184.

Churkina, G., Zaehle, S., Hughes, J., Viovy, N., Chen, Y., Jung, M. et al., 2010. Interactions between nitrogen deposition, land cover conversion, and climate change determine the contemporary carbon balance of Europe. Biogeosciences 7: 2749-2764.

Ciais, P., Schelhaas, M.J., Zaehle, S., Piao, S.L., Cescatti, A., Liski, J., et al., 2008. Carbon accumulation in European forests. Nature Geoscience 1: 425-429.

Dengler, J., Ewald, J., Kühn, I., Peet, R.K., 2011. Ecoinformatics and global change - an overdue liaison. J. Veg. Sci. 22, 577–581.

Dubois, G., Jesus, J. De, Peedell, S., Hartley, A., Nativi, S., Santoro, M., Geller, G., 2011. eHabitat: a Contribution to the Model Web for Habitat Assessments and Ecological Forecasting, in: 34th International Symposium on Remote Sensing of Environment, Sydney, Australia, 2011. pp. 1–4.

Fielding, R.T. and Taylor, R.N., 2002. Principled Design of the Modern Web Architecture. ACM Transactions on Internet Technology 2:115–150.

Friedlingstein, P., Prentice, I., 2010. Carbon-climate feedbacks: a review of model and observation based estimates. Current Opinion in Environmental Sustain-ability 2: 251-257.

Garijo, D., Alper, P., Belhajjame, K., Corcho, O., Gil, Y., Goble, C., 2013. Common motifs in scientific workflows: An empirical analysis. Futur. Gener. Comput. Syst.

Gil, Y., Deelman, E., Ellisman, M., Fahringer, T., Fox, G., et al., 2007. Examining the Challenges of Scientific Workflows. IEEE Comput. 40, 24–32.

Haines-Young, R., Potschin, M., Kienast, F., 2012. Indicators of ecosystem service potential at European scales: Mapping marginal changes and trade-offs. Ecological Indicators 21: 39-53.

Hardisty, A., Roberts, D. et al. 2013. A decadal view of biodiversity informatics: challenges and priorities. BMC Ecol. 13, 16.

Hidy, D., Barcza, Z., Haszpra, L., Churkina, G., Pintér, K., Nagy, Z., 2012. Development of the Biome-BGC model for simulation of managed herbaceous ecosystems. Ecological Modelling 226: 99-119.

Hidy, D., and Barcza, Z. 2013. User's Guide for Biome-BGC MuSo v2.2, Manuscript - revision: 10 December 2013

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change In: Stocker, T.F. et al. (eds.) Cambridge University Press, Cambridge, UK & New York, NY, USA, 1535 pp.

Kacsuk, P.K.. Kovacs, J., Farkas, Z., Marosi, A.Cs., Gombas, G., Balaton, Z., 2009. SZTAKI desktop grid (SZDG): a flexible and scalable desktop grid system. Journal of Grid Computing, 7 (4). pp. 439-461.

Littauer, R., Ram, K., Ludäscher, B., Michener, W. and Koskela, R., 2012. Trends in Use of Scientific Workflows: Insights from a Public Repository and Recommendations for Best Practices. Int. Journal of Digital Curation 7(2) 92-100.

Ma, S., Churkina, G., Wieland, R., Gessler, A., 2011. Optimization and evaluation of the ANTHRO-BGC model for winter crops in Europe. Ecological Modelling 222: 3662-3679.

Marosi, A.Cs., Balaton, Z., Kacsuk, P., 2009. GenWrapper: A Generic Wrapper for Running Legacy Applications on Desktop Grids. Proceedings of the 23rd IEEE International Parallel and Distributed Processing Symposium, pp. 1-6.

Millenium Ecosystem Assessment, 2005. Ecosystems and Human Well-being: Synthesis. Island Press, Washington, DC. p. 137.

Michener, W.K., Jones, M.B., 2012. Ecoinformatics: supporting ecology as a data-intensive science. Trends Ecol. Evol. 27, 85–93.

Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin, F.S.I. et al., 2009. Planetary Boundaries: Exploring the Safe Operating Space for Humanity. Ecology and Society 14(2): 32

Roman, D., Schade, S., Berre, A.J., Bodsberg, N.R., Langlois, J., 2009. Model as a service (MaaS). In Maué, P., Kiehle, C. (eds.) AGILE Workshop: Grid Technologies for Geospatial Applications.

Thornton, P.E., 2000. User's Guide for Biome-BGC, Version 4.1.1.

Trusilova, K., Trembath, J., Churkina, G. 2009. Parameter estimation and validation of the terrestrial ecosystem model Biome-BGC using eddy-covariance flux measurements. Max Planck Institute für Biogeochemie, Technical Reports 16: 1-60.

Vetter, M., Churkina, G., Jung, M., Reichstein, M., Zaehle, S., et al., 2008. Analyzing the causes and spatial pattern of the European 2003 carbon flux anomaly using seven models. Biogeosciences 5: 561-583.

Williams, M., Richardson, A.D., Reichstein, M., Stoy, P.C. et al., 2009. Improving land surface models with FLUXNET data. Biogeosciences 6, 1341–1359.

Wolstencroft, K., Haines, R., Fellows, D., Williams, A. 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. 1–5.