International Environmental Modelling and Software Society (iEMSs)
2010 International Congress on Environmental Modelling and Software
Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada
David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.)
http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

End-to-End Workflows for Coupled Climate and Hydrological Modeling

Kathy Saint[a] and Sylvia Murphy[b]
[a]SGI Inc., 17141 Primavera Cir. Cape Coral, FL 33909 (ksaint@sgi.com)
[b]Cooperative Institute for Research in the Environmental Science, University of Colorado
Boulder, Box 216 UCB Boulder, CO 80309-0216 USA (sylvia.murphy@noaa.gov)

Abstract: In order to examine the effects of environmental changes on local watersheds, it has become increasingly important to be able to interoperate between diverse models and document and share the resulting data. To this end, we are prototyping an end-to-end workflow that executes multiple, distributed models, loosely couples the model components, and disseminates the model results and metadata using a data portal. This project is made up of three major components and is implemented in multiple phases. The main components include the interface to the models for configuration and execution, the interface to the portal for data dissemination, and the workflow that provides the glue to make it all work together. The models being coupled are a high performance atmospheric model (currently represented by a stubbed-out ESMF Component) and the hydrological model, SWAT (Soil and Water Assessment Tool). Each model is accessed via standard interfaces (OpenMI, ESMF), set up as web services, and in the final product, the resulting data will be published to the ESG (Earth System Grid) or other data portal. Managing the model execution and coupling is initially handled by the OpenMI Configuration Manager, but may eventually be replaced by a more extensive workflow application, such as Kepler. The resulting product is an end-to-end, self-describing and repeatable workflow that demonstrates one-way interactive systems involving climate and other models can be created to address emerging questions about climate impacts.

Keywords*:* Frameworks; Hydrology; Climate; Modelling

1. INTRODUCTION

Climate change is one of the most pressing scientific questions of the day and climate models have become an important tool in the examination of climate change. While most climate models attempt to resolve the hydrological and water balance cycles, their resolution is usually insufficient to fully resolve climate impacts upon water resources. A separate hydrological model is usually required [Graham *et al.,* 2007]. Studies show, however, that these separate hydrological models perform better when they are coupled in some way to a global or regional climate model [Zeng *et al.,* 2003; Yong *et al.,* 2009]

There are two paradigms for the coupling of climate and hydrology models. The traditional method is to directly force the hydrology model with output from the climate model [Graham *et al.,* 2007]. Attempts are also being made to integrate standalone hydrological models into climate modeling systems so that two-way feedback is possible.

Advances in the direct forcing method are hindered by the technologies in use within the climate and hydrological communities. Climate models run in high performance computing (HPC) environments while many hydrological models run on personal computers. Both communities have, however, a standardized framework, that if utilized, can bridge this technological gap.

Standardized model frameworks have another key advantage in that they remove the barriers to interoperability on a broad scale. Developers are no longer limited to models they have laboriously integrated. Instead, any model that is wrapped in the framework may be used. This greatly expands scientific flexibility.

This paper describes a project that improves the direct forcing of hydrology models by climate models through the integration of the Earth System Modeling Framework (ESMF), and OpenMI. ESMF (http://earthsystemmodeling.org) is a widely used framework for geoscience HPC applications while OpenMI (http://www.openmi.org/) is a hydrological community standard. Both systems, as well as our modeling workflow, will be described in detail in the following sections.

It should be noted that this paper describes the results from the first phase of an ongoing effort to couple a hydrological model with an active atmosphere and to archive and describe the data from both sides of the simulation to a science gateway. The final phase of the project will build upon both of these milestones to achieve two-way communications between the models.

2. SYSTEM DESCRIPTION

This section describes the main pieces of the coupled climate/hydrology system (Figure 1). A high level description of the entire system is provided first. Descriptions of the individual components follow.

At the highest level, this coupled system is comprised of three main components: a hydrological model, an atmospheric model, and a driver application. The atmospheric model is implemented as an ESMF component and is wrapped with an OpenMI interface, which facilitates the communication with the OpenMI-compliant hydrological model. The hydrological model, driver, and the two OpenMI wrappers physically exist on a personal computer running Windows. The ESMF component itself exists on a multi-threaded Linux computer. An ESMF web services interface is used to bridge the two computer platforms.

The data files represented in Figure 1 can exist at any network accessible location. In phase 1 of the project, the files are copied to the PC by the atmospheric model wrapper. In the final phase of the project, this runtime data transfer will occur exclusively by the web services interfaces. At that point the arrow between the PC and the data files in Figure 1 will become uni-directional, representing the archival of model output only.

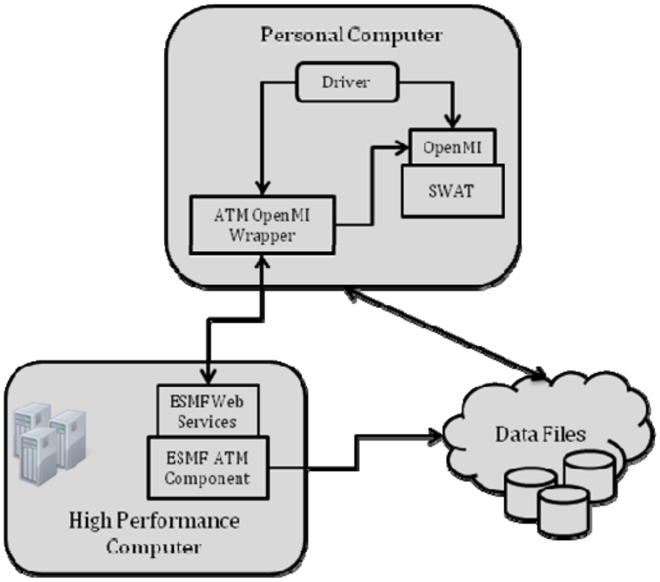

Figure 1. Diagram of the coupled system showing the components on the personal computer and the components on the HPC system as well as their interactions.

2.1 The Frameworks

2.1.1 Earth System Modeling Framework (ESMF)

ESMF is a high-performance, flexible software infrastructure that increases the ease of use, performance portability, interoperability, and reuse of Earth science applications. ESMF provides an architecture for composing complex, coupled modeling systems and includes data structures (e.g. representing physical fields) as well as utilities for developing individual models. The software includes tools to describe and organize models and the ability to interface with model components via web services [Hill *et al.,* 2004].

ESMF components

The terminology "ESMF component" has a particular meaning that is important to define. ESMF components are systems that have been broken into initialize, run, and finalize segments. Each of these segments is called using a specific, and standardized interface.

ESMF data structures

ESMF data structures include Arrays, ArrayBundles, Fields, FieldBundles, and States. ESMF Arrays store multi-dimensional data. They can also store halo points. The ESMF Field represents data defined within a continuous region of space. It includes the physical grid associated with the data and a decomposition that specifies how data associated with points in the physical grid are distributed in computer memory. ArrayBundles and FieldBundles are groupings of Arrays and Fields respectively. States can contain any of the aforementioned data structures and are used to transfer data between components.

ESMF web services

ESMF web services (http://www.earthsystemcurator.org/projects/web_service.shtml) consist of a socket interface library that allows any networked ESMF component (defined in section 2.1.1) to be available as a web service. This is accomplished through a series of Java classes that provide the following: a) access to the networked ESMF component via a java socket client, b) a set of SOAP interfaces that select and execute the component services, and c) a registrar application that the components use to register themselves with the web service.

2.1.2 OpenMI

OpenMI Software Development Kit (SDK) is a software library that provides the hydrological community with a standardized interface that focuses on time dependent data transfer. It is primarily designed to work with systems that run simultaneously, but in a single-threaded environment. Regridding and temporal interpolation are also part of the package [Gregerson *et al.,* 2007].

OpenMI data structures

Components in OpenMI are called LinkableComponents. They perform a variety of functions. The primary data structure in OpenMI is the ExchangeItem, which comes in the form of an InputExchangeItem and an OutputExchangeItem. The former describes what can be accepted at each location while the latter describes what can be provided at each location [Gregerson *et al.,* 2007].

OpenMI GetValues

Data transfer begins in OpenMI when a LinkableComponent requests data of another LinkableComponent via the GetValues method. In a two-way system, the data provider does not run forward in time until it receives this data request. Once it does, the component runs forward in time, stops, and converts its data onto the grid or location of the requesting LinkableComponent [Moore and Tindall, 2005]. Often in the hydrological community, these data requests are for point data.

2.2 System Components

2.2.1 The Hydrological Model

The hydrological model chosen for this project is the Soil Water Assessment Tool (SWAT). It is a river basin scale model developed to quantify the impact of land management practices in large, complex watersheds [Gassman *et al.,* 2007]. It was chosen for this project because it is widely used, is open source, and runs on a Windows platform. The OpenMI version that has been made into a LinkableComponent, was downloaded with permission from the United Nations Educational, Scientific and Cultural Organization-Institute for Water Education (UNESCO-IHE). Inputs to SWAT include maximum and minimum temperatures, daily precipitation, relative humidity, solar radiation data, and wind speed data [Gassman *et al.,* 2007].

2.2.3 The Atmospheric Model

In this phase of our project, an ESMF web services prototype containing a stub component representing an atmospheric model was used to send data from ESMF to OpenMI and SWAT. The next phase of the project will replace this system test with an atmospheric model. Using stubs in this first phase allowed us to focus on the OpenMI to ESMF communications and postpone the complexity of dealing with a HPC atmospheric model until those communications were in place.

The ESMF web services code exercises a full component lifecycle (initialize, run, finalize) through a socket interface. In the run portion of the code, the stub component reads in atmospheric data, converts that data into ESMF Arrays, and then outputs the same data, mimicking history output. This output file served as the source file for SWAT climatic inputs.

2.2.2 The System Driver

The system driver controls the application flow and is implemented using OpenMI's Configuration Editor. It loads the models as defined in configuration files (OMI files) and uses a trigger to start the run. The Configuration Editor provides a graphical user interface for setting up and configuring the workflow execution. When a model is loaded into the Configuration Editor, its input and output exchange items are defined. The user can then define how models exchange data by mapping output exchange items in one model to input exchange items in the other model.

The OpenMI Configuration Editor was a convenient tool for the testing of the OpenMI implementations and model interactions used in this project. Since the Configuration Editor is part of the OpenMI Graphical Users Interface, it allowed for the running of the OpenMI wrapped model codes without additional coding. This was a significant developmental timesaver. In the final phase of the project, however, the goal is to include this system within a larger workflow that includes the archival of history files to a science gateway. This final workflow will also contain post-processing steps such as data visualization. The Configuration Editor is unsuited to this task and therefore may eventually be replaced by a more versatile workflow package (e.g. Kepler [https://kepler-project.org/]).

2.3 Logical Workflow

The OpenMI interface standard defines a serial approach to communication between models. Figure 2 provides a detailed view of the method calls for the system. An OpenMI implementation will follow these fundamental steps of execution: initialization and configuration, preparation, execution, and completion. These steps correspond to methods in the LinkableComponent interface: Initialize, Prepare, GetValues, and Finish/Dispose. The Initialize method is where the input and output exchange items are defined and the connection to the ESMF web services is created. In the current system, the Prepare method is where the atmospheric model does most of its work; the model is run and the output data is exported into a file. Once the models are prepared, the driver performs a time step on the SWAT model, which then triggers a request to the atmospheric model (via the GetValues method) to exchange data. The atmospheric model retrieves the data from the data file and returns it to the SWAT model. The time step process repeats until the end time is reached, after which the models are completed using the Finish and Dispose methods.

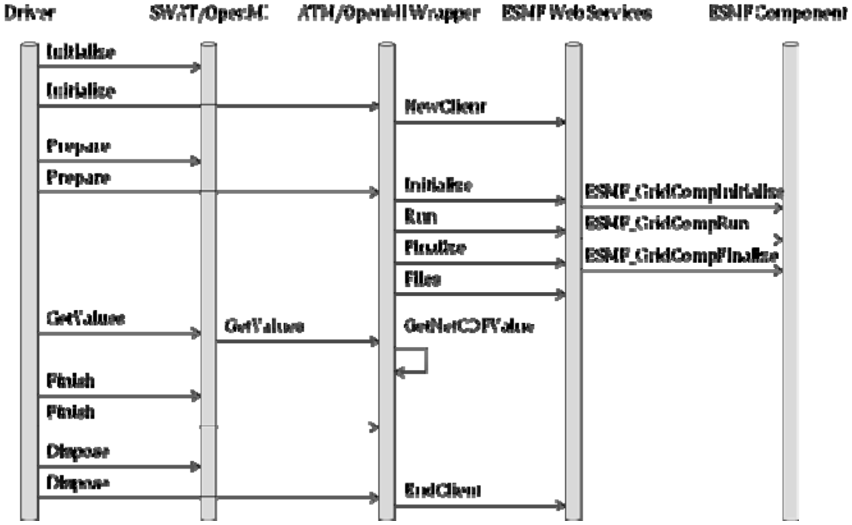

Figure 2. The sequential method calling sequence for the entire system.

2.4 The Data flow

The format of the data in this coupled system changes several times during execution. The data initially starts out as an ESMF Array contained within an ESMF State. That State is exported to an external NetCDF file. Once SWAT requests data via OpenMI's GetValues method, the atmospheric OpenMI wrapper accesses the NetCDF file and converts the data into an OpenMI OutputExchangeItem. This then gets passed into the SWAT OpenMI wrapper where it is converted to an InputExchangeItem. The data is now available for use by SWAT (Figure 3).

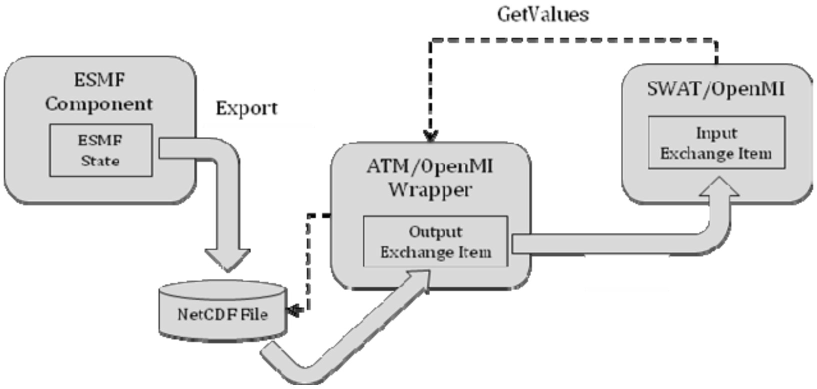

Figure 3. The flow of data through the system.

3. IMPLEMENTATION

To reiterate, this paper represents a snapshot of an ongoing effort to create a two-way coupling between SWAT and an atmospheric model and to archive and describe the data from this simulation on a science gateway. The following sections describe the implementation details of the current phase and then the plans for migrating to a more complete implementation

3.1 Current Implementation Details

This project is being implemented using an incremental prototype approach, and the current phase focuses on the interfaces between each of the system components. Some of the major differences between the current implementation and the final system design include the following: a stubbed-out ESMF component is used to represent the atmospheric data model; data is transferred between this component and the OpenMI wrapper using NetCDF files instead of streamed across the interface; and output data is not currently published to a data portal. Detailed instructions to reproduce this work can be found on our open source repository at http://esmfcontrib.cvs.sourceforge.net/viewvc/esmfcontrib/HydroInterop/.

3.2 Future Directions

The remaining tasks involved in this project will be implemented incrementally, and will involve several releases of the prototype system. By using an incremental approach to implementing this project, we can also learn from the early implementations, and use these lessons learned to modify and improve the final system architecture.

The major tasks to be implemented in the near future include replacing the stubbed out ESMF component with the Community Atmospheric Model (CAM) that is part of the

Community Climate System Model (CCSM), or another atmospheric model. In addition, the NetCDF data transfer will be replaced by an online data transfer through the ESMF web services. Finally, the Earth System Grid (http://earthsystemgrid.org) or another science gateway will be brought into the system to host and describe the resultant coupled simulation data.

4. CONCLUSION

The first phase of this project has successfully demonstrated a coupled ESMF/OpenMI modeling system. Both SWAT and the ESMF atmospheric stub component were initialized and run, and data was transmitted, on request, to SWAT via an external climate data file. The code for this system and instruction to reproduce our results is publically available at http://esmfcontrib.cvs.sourceforge.net/viewvc/esmfcontrib/HydroInterop/.

Enabling OpenMI and ESMF to interface is an important milestone. These two infrastructures are both widely used within their respective communities and their integration has the potential for enabling diverse geoscience and hydrological modeling systems to interact in a more flexible way.

REFERENCES

Gassman, P.W., Reyes, M.R., Green, C.H., Arnold, J.G., The Soil and Water Assessment Tool: Historical development, applications, and future research directions. *American Society of Agricultural and Biological Engineers*, 50(4), 1211-1250, 2007.

Graham, L.P., Hagemann, S., Jaun, S., Beniston, M., On interpreting hydrological change from regional climate models. *Climate Change*, 81 Suppl. 1, 97-122, 2007.

Gregersen, J.B., Gijsbers, P.J.A., and Westen, S.J.P., OpenMI: Open Modelling Interface. *Journal of Hydroinformatics*, 9(3), 175-191, 2007.

Hill, C., C. DeLuca, V. Balaji, M. Suarez, and A. da Silva. Architecture of the Earth System Modeling Framework. *Computing in Science and Engineering*, 6(1), 18-28, 2004.

Moore, R.V and Tindall, I., An Overview of the Open Modelling Interface and Environment (the OpenMI). *Environmental Science & Policy,* 8(2005), 279-286, 2005.

Yong, B., Ren, L., Xiong, L., Yang, X., Zhang, W., Chen, X., and Jiang, S., A study coupling a large-scale hydrological model with a regional climate model. New Approaches to Hydrological Predication in Data-sparse Regions (Proc. Symposium HS.2 a the Joint IAHS and IAH Convention, Hyderabad, India). IAHS Publ. 333, 203-210, 2009.

Zeng, X., Zhoa, M., Su, B., Tang J., Zheng, Y., Gui, Q, and Zhou, Z., Simulations of a Hydrological Model as Coupled to a Regional Climate Model. *Advances in Atmospheric Sciences*, 20(2), 227-236, 2003.