Geophysical Research Abstracts
Vol. 18, EGU2016-9508, 2016
EGU General Assembly 2016

Advances in Geoscience Modeling: Smart Modeling Frameworks, Self-Describing Models and the Role of Standardized Metadata

Scott Peckham
United States (Scott.Peckham@colorado.edu)

Over the last decade, model coupling frameworks like CSDMS (Community Surface Dynamics Modeling System) and ESMF (Earth System Modeling Framework) have developed mechanisms that make it much easier for modelers to connect heterogeneous sets of process models in a plug-and-play manner to create composite "system models". These mechanisms greatly simplify code reuse, but must simultaneously satisfy many different design criteria. They must be able to mediate or compensate for differences between the process models, such as their different programming languages, computational grids, time-stepping schemes, variable names and variable units. However, they must achieve this interoperability in a way that: (1) is noninvasive, requiring only relatively small and isolated changes to the original source code, (2) does not significantly reduce performance, (3) is not time-consuming or confusing for a model developer to implement, (4) can very easily be updated to accommodate new versions of a given process model and (5) does not shift the burden of providing model interoperability to the model developers.

In tackling these design challenges, model framework developers have learned that the best solution is to provide each model with a simple, standardized interface, i.e. a set of standardized functions that make the model: (1) fully-controllable by a caller (e.g. a model framework) and (2) self-describing with standardized metadata. Model control functions are separate functions that allow a caller to initialize the model, advance the model's state variables in time and finalize the model. Model description functions allow a caller to retrieve detailed information on the model's input and output variables, its computational grid and its timestepping scheme. If the caller is a modeling framework, it can use the self description functions to learn about each process model in a collection to be coupled and then automatically call framework service components (e.g. regridders, time interpolators and unit converters) as necessary to mediate the differences between them so they can work together.

This talk will first review two key products of the CSDMS project, namely a standardized model interface called the Basic Model Interface (BMI) and the CSDMS Standard Names. The standard names are used in conjunction with BMI to provide a semantic matching mechanism that allows output variables from one process model or data set to be reliably used as input variables to other process models in a collection. They include not just a standardized naming scheme for model variables, but also a standardized set of terms for describing the attributes and assumptions of a given model. Recent efforts to bring powerful uncertainty analysis and inverse modeling toolkits such as DAKOTA into modeling frameworks will also be described. This talk will conclude with an overview of several related modeling projects that have been funded by NSF's EarthCube initiative, namely the Earth System Bridge, OntoSoft and GeoSemantics projects.