

WHY IS CIM-EARTH AN OPEN MODEL

An April 4, 2009, EPA analysis of the Waxman-Markey Discussion Draft of the American Clean Energy and Security Act presents a range of findings about the impacts of the proposed legislation on the U.S. economy. For example, “average annual household consumption is estimated to decline in a range of \$98 to \$140 dollars per year, relative to reference scenario.”

These findings are obtained via the use of a suite of models, including: multiple economic models, an integrated assessment model, an agricultural model, and an electricity sector model.

Given the importance of these findings and the complexity of the economic system, many might want to study the EPA’s methodology and results to verify whether they are correct in their assumptions and implementations, or to compare and contrast with results obtained under different assumptions. For example, the EPA study assumes a 5% discount rate, no international carbon leakage, and monotonic increases in energy efficiency. What happens if we change these assumptions? Indeed, how do we even know that the EPA’s models have implemented them correctly? Broad study and debate of these questions would, in our view, be healthy for science, policymaking, and democracy.

Unfortunately, such study and debate is essentially impossible, for the simple reason that the models used by the EPA and most other organizations that study climate change are *closed*, meaning that they are not freely available to others. The models are, to be sure, documented—in some cases, in sufficient detail that an interested party could, albeit with substantial effort, create a similar model. However, no model documentation is sufficiently detailed to permit complete replication. The overall situation is that no outside party can study, validate, run, or modify these models, or meaningfully compare and contrast the results from one model with those of other models.

Why have other integrated assessment models been closed?

The current preference for closed models appears to be driven by a variety of motivations, including the following.

Fears of misuse: Some model creators think that unqualified people may use the model incorrectly and obtain meaningless results that discredit developers and hinder scientific progress. However, experience in other fields, such as climate modeling, suggests that such problems are rare in practice.

Risk of embarrassment: Model developers may be concerned that openness will lead to criticisms from others for their choice of numerical methods, their programming style, or perhaps even errors. Such concerns are surely not a valid reason for withholding access to code.

Rent seeking: Model developers may be concerned that if models are open, anyone can use them without compensation. Thus, the developers lose the ability to fund their continued work, and perhaps also the ability to get credit for work based on the model. These concerns are real, but need to be balanced against the benefits of open models, and also against the public interest when model development is supported by public funds.

What does it mean for models to be open?

The term open model, like the related term “open source software,” can be used to mean a variety of things. In the most expansive sense of the term, an open model is one whose implementation is freely available to others, who may:

1. Read the model: study the source code, for example to learn about the methods used to construct the implementation, or to check the implementation for errors.
2. Modify and run the model: modifying the model implementation, its input data, and/or its configuration parameters, to study, for example, the effect of different assumptions.
3. Redistribute the model: make modified code available to others for their use.

A model can be made available under terms that allow for some but not all of these usages: for example, we may be able read and modify a model, but not redistribute it. In general, we will reserve the term open for models that satisfy all three conditions.

Why are open models important?

Open models are important, because they:

- **Permit reproducibility:** At the core of the scientific method is the idea that a scientist who publishes a new result is expected to also provide sufficient details concerning the method used to obtain that result that other investigators can repeat the experiment. This discipline not only helps to detect errors, it also helps one investigator to build on the results of another.
- **Allow for examination by third parties:** The methodology, software, and data used by one investigator can be reviewed by others. They can examine the data and software, test the model with different data, compare with other models, and extend the model in new directions.
- **Encourage collaboration among researchers:** A popular open model can form the basis for a community of contributors and users, who develop different components and apply the resulting model to different problems.

Where have open models been successful?

Open models are certainly not universal in science, but they are becoming far more common. One particularly successful open model is the Community Climate System Model (CCSM), developed by a teams led by the National Center for Atmospheric Research. CCSM is downloadable by anyone and is used by many investigators worldwide. It is the foundation for a vibrant community of both users and developers, who experiment with alternative formulations of different components and test the model in different scenarios. Clearly the existence of this community is not due solely to CCSM’s openness—it is a product also of substantial funding—but it is hard to imagine such a community forming and prospering if CCSM was closed.