

# Demystifying Models: Answers to Ten Common Questions that Ecologists Have about Earth System Models

Rebecca G. Asch<sup>1\*</sup>, Darren Pilcher<sup>2</sup>, Sara Rivero-Calle<sup>3</sup>, Johnna Holding<sup>4</sup>

<sup>1</sup> Princeton University, Program in Atmospheric and Oceanic Sciences, Princeton, NJ 08540.

<sup>2</sup> National Oceanic and Atmospheric Administration, Pacific Marine Environmental Laboratory, Seattle, WA, 98115.

<sup>3</sup> Johns Hopkins University, Department of Earth and Planetary Sciences. Baltimore, MD 21218.

<sup>4</sup> Institut Mediterrani d'Estudis Avançats, Esporles-Mallorca-Illes Balears 07190.

\* Corresponding author: [rasch@princeton.edu](mailto:rasch@princeton.edu)

## Introduction

Every two years the University of Hawaii hosts the Ecological Dissertation in the Aquatic Sciences (EcoDAS) symposium, an ASLO and NSF-funded workshop designed to foster interdisciplinary collaborations among early career aquatic ecologists. During the EcoDAS symposium in 2014, many of the discussions revolved around a growing interest in using global climate models for making projections about how ecosystems and organisms will respond to future changes. However, many EcoDAS participants did not feel they had the required knowledge to employ climate models or participate in a climate modeling study.

Following these discussions, we sent out a questionnaire to the past four cohorts of EcoDAS participants to further understand the general feeling of early career scientists toward climate models. Respondents were most interested in using climate models for determining effects of climate change on ecological processes relevant to specific organisms or habitats and the application of models to ecosystem management and assessment of anthropogenic impacts. However, about one third of respondents reported to have never worked with or were uncomfortable working with climate model output, and only about 10% reported having strong experience working with climate models. Similarly, a survey of marine resource managers

indicated that 90% wanted to use outputs from climate change models, but >40% faced hurdles related to lack of technical expertise, time, or personnel (Halpin et al., 2014).

We concluded that a solution to the issues identified in these surveys could be to provide a brief, practical primer laying out the basics of climate modeling to proficiency in climate model use among ecologists. We hope this will in turn facilitate conversations between modelers and empiricists and serve as a stepping stone for collaboration among aquatic scientists. Here, we explain the basic functioning of climate models, provide advice on how to use climate models in relation to ecological research, outline data availability from models, discuss model limitations, and highlight common errors for beginners. We also suggest some understudied areas which could be a starting point for collaborations between modellers and empiricists.

## **1. What is an Earth System Model (ESM) and how do ESMs work?**

Projections of future climate change are made using two variants of climate models: General Circulation Models (GCMs) and Earth System Models (ESMs). Both types of models represent the global climate system through mechanistic, mathematical equations describing thermodynamics and fluid dynamics. These models divide Earth into a three-dimensional grid representing latitude, longitude and a vertical component (altitude of the atmosphere, ocean depth). At the start of a model run, each grid cell is assigned a value for each of the model's state variables (e.g., atmospheric temperature, ocean salinity). These initial conditions (Table 1) are based on global mean observations. For each time interval of the model simulation, values of the state variables change based on the mathematical equations, which describe either processes occurring within a grid cell (e.g., heating/cooling, evaporation/precipitation) and flows between

cells. The model is first run over an initialization period of hundreds to thousands of years during which the climate system will come into a quasi-equilibrium state. This stage is referred to as “spin-up.” Once this quasi-equilibrium is reached, greenhouse gases, volcanic emissions, aerosols, and land use may be varied reflecting historical trajectories or future scenarios.

## **2. What is the difference between GCMs and ESMs?**

The difference between GCMs and ESMs is that the former generally represents physical processes occurring in the atmosphere, ocean, and interactions between these domains. In addition to representing oceanic and atmospheric dynamics, ESMs also include information on biogeochemical cycling in terrestrial and marine ecosystems and allow for these ecosystems to have feedbacks on the circulation. Therefore, all ESMs are GCMs, but not all GCMs are ESMs. By representing the full carbon cycle, including interactions between the atmosphere, ocean, land, and cryosphere, some ESMs do not need to prescribe a given concentration of atmospheric CO<sub>2</sub> associated with emissions of greenhouse gases. Instead, CO<sub>2</sub> can be dynamically partitioned between different reservoirs. Many ESMs also describe the fluxes of oxygen, nitrogen, phosphorus, silica, and iron throughout the earth system. To describe the cycling of these elements, most ESMs include functional groups of phytoplankton and zooplankton, as well as particulate and dissolved organic matter. Due to the inclusion of lower trophic level organisms in ESMs, these models are becoming increasingly useful for examining ecological questions. For example, ESMs have been used to investigate diverse ecological questions, such as how will climate change affect the seasonal timing of phytoplankton blooms (Henson et al., 2009, 2013),

the species range of top marine predators (Hazen et al., 2013), and the boundaries of marine biomes (Polovina et al., 2011).

### **3. What are the advantages and disadvantages of using a model?**

Models provide experimental earth systems that can be used to test an imposed change and reasonably estimate the impacts of that change. Since the impacts of the imposed change are the outcome from numerical equations, models can be effective at elucidating the mechanistic drivers behind changes in oceanic dynamics and ecosystem structure and function. This differs from observational analyses, which rely mainly on correlative significance tests limited in space and time to support a hypothesized process that may lack a mechanistic explanation. Statistical relationships may also break down under future novel climate conditions. Modeling also differs from experimental analysis in that it is less limited in terms of the spatial and time scales of biological responses, and in its ability to manipulate multiple factors while maintaining adequate controls. Unlike most ocean observations, models provide “complete” data coverage, so that model output has no spatial or temporal gaps.

The main disadvantage to using models is that they are dependent on the accuracy of the underlying model structure. While a model provides mechanistic drivers of the underlying processes, how representative these processes are of the natural system is limited by imperfect knowledge. This is why increased collaboration between modelers and empiricists is important. Other disadvantages stem from the spatial resolution and the number of ecological processes resolved by models.

#### 4. How do ESMs incorporate observations?

Observational data may be used at several stages:

- 1) Creating initial conditions from which numerical models “run”. For example, the World Ocean Atlas sets the baseline for the monthly climatology of nutrients, salinity and dissolved oxygen in model grid cells based on field data.
- 2) Parameterizing the model to describe complex processes that cannot be fully represented in a model. For example, a model with a 1° latitude/longitude resolution is too coarse to resolve ocean eddies. However, eddies affect mixing of heat, salinity, and nutrients. To include these effects in a model, parameterization can scale mixing to what one would expect from eddies even though the eddies aren’t explicitly included. See Table 1 for examples of ecological parameterization and parameters.
- 3) Evaluating model performance by assessing its ability to reproduce current or historical conditions. Model performance varies depending on the climatic variable and spatial scale under consideration. Generally, there is greater skill at reproducing global patterns of core climate variables (e.g., sea surface temperature [SST], sea surface salinity [SSS]) compared to skill when evaluating biogeochemistry or regional phenomena. However, it is important to remember that past and present model performance does not always translate into high model skill when evaluating future projections.
- 4) Constraining the model to reality either through data assimilation or tuning (Table 1). Observations are key to tuning or adjusting the model to “reality”. There are several

data assimilation techniques (e.g., interpolation, least squares; Luo et al., 2011). It is important to note that tuning is not a universal feature of models. For example, many modelers are interested in emergent ecological behaviors unrelated to tuning. Most Intergovernmental Panel on Climate Change (IPCC) models assimilate relatively little data, especially for biogeochemical processes (Table 1). An example where data assimilation is used is the NASA Ocean Biogeochemical Model, which can be run freely or with the constraint that the sum of phytoplankton groups has to be consistent with the total chlorophyll *a* from satellite imagery (Nerger and Gregg, 2007).

## **5. Where can I find ESM outputs and how do I work with these data?**

Data from all IPCC-class models are available for free online through the Coupled Model Intercomparison Project – Phase 5 (CMIP5). Figure 1 provides an overview of how to download and use CMIP5 data. To begin working with data from CMIP5, you will need access to software for visualizing and analyzing your data. MATLAB, R, Python, and IDL are commonly used programs for working with data from climate models. However, all of these require basic scientific computing skills. Ocean Data View (ODV) and ArcGIS are two additional programs that can be used, which have the advantage of not requiring users to have any previous programming knowledge. Working with outputs from ESMs is similar to working with data from satellites or any other type of gridded dataset with information on geographic locations and multiple time periods. The learning curve for people who previously have used such datasets and software programs should not be excessively steep when transitioning to working with data outputted from ESMs.

One potentially confusing aspect of working with ESMs is the proliferation of acronyms describing models. While it is beyond the scope of this paper to provide a comprehensive list to acronyms, we recommend the following resources for further detail: (1) Standardized abbreviations for climate variables used by CMIP5 are described in [http://cmip-pcmdi.llnl.gov/cmip5/docs/standard\\_output.pdf](http://cmip-pcmdi.llnl.gov/cmip5/docs/standard_output.pdf) (last accessed: March 6, 2016); (2) The 2013 IPCC report defines acronyms pertaining to IPCC climate models in [http://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5\\_AnnexIV\\_FINAL.pdf](http://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5_AnnexIV_FINAL.pdf) (last accessed: March 6, 2016).

## **6. How do I choose which model(s) to use?**

There is no single best model for all applications. The strengths of each model will vary depending on the variable of interest, region, temporal or spatial scale, and your scientific question. Since models are built in different ways and have contrasting strengths and weaknesses, it is a good practice to compare outputs from several models. The latest IPCC report used a compilation of over 50 GCMs referred to as CMIP5. Among those, only 16 are ESMs that have an ocean biogeochemistry component that include nutrients (Si, Fe, N, P) and one or more phytoplankton and zooplankton groups. Since different models perform better in different regions, using many models (also known as an ensemble) often improves accuracy on a global basis. Using multiple models also helps to assess uncertainty in predictions due to differences in model structure. The ensemble mean often most closely resembles observations, but it is also important to consider the range of variability between models. Also, the dynamics of an ensemble may not be internally consistent, since it represents a combination of different models.

Currently, there is no agreed upon way of weighting different models in an ensemble. At the regional scale, weighting can be problematic because a model can correctly reproduce small, regional features for the wrong reasons. If this occurs, reproducing such features may be unrelated to the ability to make robust future projections (Stock et al., 2011). The 2013 IPCC report uses a “one model, one vote” approach for calculating ensemble means (Flato et al., 2013). When calculating ensemble means and other statistics, remember that each model is not fully independent, since many models were constructed using similar code and mathematical equations.

Despite advantages of using ensembles, it may be more practical to consider a single model in cases where one wants to compare a large number of projections from different scenarios. This can be useful for conservation or management studies to show the effects of different mitigation strategies.

## **7. What if I am interested in processes occurring at a small spatial scale?**

ESMs are designed to simulate the Earth System at the global scale and reproduce broad regions of physical and biogeochemical variability. ESMs have a coarse resolution ( $\sim 1^\circ \times 1^\circ$ ), although exact resolution varies between models. The atmospheric and oceanic components of a single model may have different resolutions or the resolution of the model grid may vary with latitude (i.e., finer resolution in the tropics to resolve equatorial currents).

Typically ESMs cannot resolve mesoscale processes (e.g., eddies, fronts). Until recently, eddy-resolving models were prohibitively computationally expensive for anything beyond regional, multi-year simulations. Some next generation ESMs, currently under development, will



have a 0.1° to 0.25° resolution that can begin to resolve eddies (Griffies et al., 2015). Overall, smaller scale coastal processes that are of particular interest to aquatic ecologists may not be resolvable using current ESMs.

Furthermore, river inputs in ESMs typically impact the freshwater balance only and do not yet modify ocean biogeochemistry. Thus, for the moment, a regional model will be better at capturing nearshore processes than an ESM. However, current research efforts are attempting to provide a global model of river/watershed export of nutrients (<http://marine.rutgers.edu/globalnews/datasets.htm>; last accessed: March 6, 2016).

Researchers interested in processes operating at small spatial scales can explore dynamical and statistical downscaling, which is reviewed by Harris et al. (2014). Dynamical downscaling uses a global ESM to force the boundary conditions of a regional climate model with finer resolution. In statistical downscaling, empirical relationships between variables are used to increase spatial resolution in cases where fine-scale data are available for some variables.

## **8. Why is process “X” not included in the model?**

Numerical models only include processes closely aligned to the primary goal of model development, which is making climate projections at a global scale. Models also only include processes that can be described with a mechanistic formulation or parameterized. Many empirical studies conclude with an argument that a process needs to be included in future models to improve accuracy (e.g., Davison et al., 2013; Holding et al., 2013, 2015). While there are several important ecological and biogeochemical processes that remain unresolved by ESMs, there is a risk of over-parsing or over-tuning a model by incorporating a process that, although

critical for a particular region or variable, may offset other model processes that were already effectively simulated. Adding new variables can increase computing time, while not improving overall model performance. Consequently, there is a clear trade-off when including new processes. Furthermore, there is a lag time between the discovery of an important process and its incorporation into a model, because models are developed in distinct phases.

## **9. How can I assess model uncertainty?**

With hindcasts, model performance can be assessed through comparisons with observations. However, when working with future projections, there are no observations to which forecasts can be compared, so model uncertainty needs to be assessed in a different way. Climate scientists have developed the following framework to quantify model uncertainty (Hawkins and Sutton, 2009):

- 1) *Scenario uncertainty* stems from uncertainty in future emissions of greenhouse gases and aerosols. In IPCC reports, this uncertainty is quantified by considering multiple emission scenarios.
- 2) *Model uncertainty* exists in terms of what processes are incorporated into ESMs, how those processes are represented mathematically, and what parameters are selected. Model uncertainty is commonly assessed with a multi-model ensemble to quantify differences in projections among models. For a single model, model uncertainty can be *partially* quantified by performing sensitivity tests to determine how parameter choices affect results.

3) *Internal variability* in models reflects natural climate fluctuations, such as El Niño-Southern Oscillation. Internal variability arises from differences in initial conditions at the start of a simulation, which evolve over time mimicking natural climate variability. Internal variability can be quantified by creating an ensemble where a single ESM is run multiple times with slight changes ( $\sim 10^{-14}$ ) in initial conditions (Deser et al., 2012).

## 10. How can empirical ecologists help inform models?

The following are examples of challenges faced by modelers that could be resolved with the aid of observational or experimental ecologists:

- 1) Synthesis of observational data available via open-access portals (e.g., the green ocean project website:  
[http://lgmaweb.env.uea.ac.uk/green\\_ocean/data/index.shtml?d1](http://lgmaweb.env.uea.ac.uk/green_ocean/data/index.shtml?d1); last accessed: March 6, 2016 )
- 2) Focus experiments and field campaigns on the regions most sensitive to climate change and those that are less predictable. Also, coordinate these studies with modelers to assist with selecting sampling locations (e.g., Observing System Simulation Experiments) and to help ensure that, when appropriate, the experimental results will help inform modeling to the maximum extent possible.
- 3) Improved understanding from experiments or observations of the synergistic or antagonistic effects of multiple stressors (e.g., freshening, warming, deoxygenation)

- 4) Improved empirical constraints on ecological parameters used in models. Areas where such improvements are needed include: growth, predation, and sinking rates; physiological response to increasing CO<sub>2</sub>; and improved knowledge of functional group traits (e.g., distinct zooplankton groups [Sailley et al., 2013]).
- 5) Mesocosm experiments to examine community-level changes and experiments examining acclimation, adaptation, and evolution in short-lived organisms.
- 6) Experiments designed to look at full environmental gradients and variability, when possible, and within a range of naturally occurring conditions (e.g., temperature, irradiance, nutrients). It is especially helpful for modelers that experiments include a full range of values, since it is difficult to parameterize a model based on a limited number of experimental treatments that may represent extreme scenarios.

In conclusion, collaborations between modelers and empiricists are critical for furthering knowledge of how ecosystems will respond to global climate change. For scientists new to scientists new to ESMs, we recommend developing collaborations with climate modelers in order to have a source of support for resolving queries and to get help understanding what features of a model are “real” vs. features that may arise due to idiosyncrasies of model design. Table 2 lists several references that can be consulted for further help with ESMs. Overall, we hope that this paper eases the burden of getting started with using ESMs, as well as spurs a greater dialog between modelers and experimental and field ecologists.

## **Acknowledgements**

281           We acknowledge the support of NSF (grant no. OCE-1356192) and ASLO for funding  
282   our participation in the 2014 EcoDAS XI symposium. We also thank the organizers and host of  
283   EcoDAS XI, Center for Microbial Oceanography: Research and Education (C-MORE) at the  
284   University of Hawaii at Manoa, for making this collaborative work possible. A. Gnanadesikan  
285   and J. Sarmiento reviewed earlier drafts of this manuscript and provided helpful feedback.  
286

## References

- Davison, P.C., D.M. Checkley, J.A. Koslow, and J. Barlow. 2013. Carbon export mediated by mesopelagic fishes in the northeast Pacific Ocean. *Prog. Oceanogr.* 116: 14-30, doi:10.1016/j.pocean.2013.05.013
- Deser, C., A. Phillips, V. Bourdette, and H. Teng. 2012. Uncertainty in climate change projections: the role of internal variability. *Clim. Dyn.* 38: 527-546, doi:10.1007/s00382-010-0977-x
- Flato, G., J. Marotzke, B. Abiodun, and others. 2013. Evaluation of climate models, p. 741-866. In T.F. Stock, D. Qin, G.-K. Plattner, and others [eds.], *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Griffies, S.M., M. Winton, W.G. Anderson, and others. 2015. Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J. Climate* 28: 952-977, doi:10.1175/JCLI-D-14-00353.1
- Halpin, P.N., A.M. Boustany, D.C. Dunn, B. Donnelly, and J. Cleary. 2014. Incorporating climate data into the decision making process for living marine resources. A survey report submitted to the NASA Earth Science Division.
- Hawkins, E. and R. Sutton. 2009. The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.* 90: 1095-1107, doi:10.1175/2009BAMS2607.1
- Hazen, E.L., S. Jorgensen, R.R. Rykaczewski, and others. 2013. Predicted habitat shifts of Pacific top predators in a changing climate. *Nature Clim. Change* 3: 234-238, doi:10.1038/nclimate1686
- Henson, S.A., J.P. Dunne, and J. L. Sarmiento. 2009. Decadal variability in North Atlantic phytoplankton bloom. *J. Geophys. Res.* 114, C04013, doi:10.1029/2008JC005139.
- Henson, S., H. Cole, C. Beaulieu, and A. Yool. 2013. The impact of global warming on seasonality of ocean primary production. *Biogeosci.* 10: 4357-4369, doi:10.5194/bg-10-4357-2013
- Holding, J.M., C.M. Duarte, J.M. Arrieta, R. Vaquer-Suyner, A. Coello-Camba, P. Wassmann, and S. Agustí. Experimentally determined temperature thresholds for Arctic plankton community metabolism. *Biogeosci.* 10: 357-370, doi:10.5194/bg-10-357-2013.

327 Holding, J.M., C.M. Duarte, M. Sanz-Martín, and others. 2015. Temperature dependence of  
 328 CO<sub>2</sub>-enhanced primary production in the European Arctic Ocean. *Nature Clim. Change* 5: 1079-  
 329 1082, doi:10.1038/nclimate2768  
 330

331 Kirtman, B., S.B. Power, J.A. Adedoyin, and others. 2013. Near-term climate change:  
 332 projections and predictability, p. 953-1028. In T.F. Stock, D. Qin, G.-K. Plattner, and others  
 333 [eds.], *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to*  
 334 *the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge  
 335 University Press.  
 336

337 Luo, Y., K. Ogle, C. Tucker, S. Fei, C. Gao, S. LaDeau, J.S. Clark, and D.S. Schimel. 2011.  
 338 Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, 21(5):  
 339 1429–1442, doi:10.1890/09-1275.1  
 340

341 Nerger, L. and W.W. Gregg. 2007. Assimilation of SeaWiFS data into a global ocean-  
 342 biogeochemical model using a local SEIK filter. *J. Mar. Syst.* 68(1-2): 237-254,  
 343 doi:10.1016/j.jmarsys.2006.11.009.  
 344

345 Polovina, J.J., J.P. Dunne, P.A. Woodworth, and E.A. Howell. 2011. Projected expansion of the  
 346 subtropical biome and contraction of the temperature and equatorial upwelling biomes in the  
 347 North Pacific under global warming. *ICES J. Mar. Sci.* 68(6): 986-995,  
 348 doi:10.1093/icesjms/fsq198  
 349

350 Sailley, S.F., M. Vogt, S.C. Doney, and others. 2013. Comparing food web structures and  
 351 dynamics across a suite of global marine ecosystem models. *Ecol. Model.* 261-262: 43-57,  
 352 doi:10.1016/j.ecolmodel.2013.04.006  
 353

354 Stock, C.A., M.A. Alexander, N.A. Bond, and others. 2011. On the use of IPCC-class models to  
 355 assess the impact of climate on Living Marine Resources. *Prog. Oceanogr.* 88: 1-27,  
 356 doi:10.1016/j.pocean.2010.09.001  
 357

358 Taylor, K.E., R.J. Stouffer, and G.A. Meehl. 2011. An overview of CMIP5 and the experiment  
 359 design. *Bull. Amer. Meteor. Soc.* 93: 485-498, doi:10.1175/BAMS-D-11-00094.1

360 **Table 1.** A climate modeler's vocabulary toolbox  
361

Vocabulary	Definition
Climate model	A general term that applies to both GCMs and ESMs, but usually implies a model that includes ocean, atmosphere, and ice components.
Coupled/uncoupled	The term “coupled” is generally used to describe different model components (e.g., atmosphere, land, ocean, cryosphere) that fully interact, rather than one component forcing the other. For example, in a coupled model, atmosphere temperatures will influence those of the surface ocean, and vice versa. In contrast, in an uncoupled model, atmospheric conditions may determine or force surface ocean conditions, but not vice versa.
Data assimilation	Historical observations of oceanic and/or atmospheric conditions are used to adjust model output at each time step of a model run. When data assimilation is performed, climate variability in a model's output will be in phase with the observed variability in oceanic and atmospheric conditions (e.g., El Niño events occur during the same years as in observations). This contrasts with normal ESM output where climate variability will be out of phase with observed variability if there is no data assimilation.
Earth System Model (ESM)	ESMs expand on the GCM by including the biosphere and chemosphere, thus capturing interactions and feedback between biological, chemical and physical processes. ESMs principally include the cycling of key elements and nutrients (i.e., O <sub>2</sub> , CO <sub>2</sub> , N, P). This allows climate scientists to see how much CO <sub>2</sub> stays in the atmosphere for a given amount of greenhouse gas emissions rather than having to force the model with a prescribed concentration of atmospheric CO <sub>2</sub> . It is helpful to remember that all ESMs are GCMs, but not all GCMs are ESMs.
Emissions scenarios	Scenarios describing how greenhouse gas and aerosol emissions may change in the future. Currently, there are two sets of emissions scenarios that are widely used: Representative Concentration Pathway (RCP) vs. Special Report on Emission Scenarios (SRES). RCP scenarios were used for IPCC Fifth Assessment Report and are based on projected changes in radiative forcing (units: W/m <sup>2</sup> ) by the year 2100. SRES scenarios were used in previous IPCC reports and are based on several economic growth and climate change mitigation scenarios. Furthermore, SRES scenarios prescribed certain atmospheric concentrations of CO <sub>2</sub> during a given year, whereas RCP scenarios can allow CO <sub>2</sub> concentrations to evolve due to exchange of gases between the atmosphere, land, and ocean.
General Circulation Model (GCM)	Three-dimensional model that numerically describes the physical processes in the atmosphere, cryosphere, ocean, and land surface and their interactions across a latitudinal/longitudinal grid that covers the globe. GCMs are a mathematical representation of fluid motion on a rotating sphere, integrated through time using the thermodynamic and Navier-Stokes equations.



Hindcast	A climate model simulation that examines historical conditions. Hindcasts are often used to evaluate how closely the model can reproduce historical observations. According to the IPCC definition, a true hindcast will only assimilate data from the period prior to a particular event that is being predicted by the hindcast (Kirtman et al., 2013). In this way, hindcasts differ slightly from climate reanalyses.
Initial conditions	Initial values of variables set prior to model spin-up. Typically based on mean, observed values or a climatology.
Model spin-up	Model initialization phase conducted prior to the beginning of a simulation by using repeated, stable forcing. Spin-up is required to achieve a baseline state of consistent internal model dynamics from the initial conditions.
Model validation/ verification	Comparing hindcast model performance against observed data to assess overall skill or accuracy.
Parameters	Numerically defined properties of a system, such as in mathematical equations where parameters define the relationships between variables. Some examples from biological systems include: (a) phytoplankton growth rate where growth rate is parameterized as a function of light, nutrients or temperature; (b) Michaelis-Menton enzyme kinetics where the half saturation constant ( $1/2 V_{\max}$ ) is a parameter that can be included in a model. The way a modeler formulates these relationships is referred to parameterization.
Parameterization	Simplification of an observed process or mechanism which is necessary due to computational constraints, lack of a full mathematical description, or substantial process knowledge uncertainty.
Prediction	The most likely way in which climate will change in the future. Predictions are often made over shorter periods of time (i.e., months-to-years) than projections so that it is not necessary to rely on future scenarios regarding greenhouse gas emissions or changes in other climate forcing agents. Predictions are usually made in a probabilistic manner.
Projection	A quantitative description of changes in future climate based on a climate model simulation and a scenario where greenhouse gases and/or other climate forcing agents change in a prescribed manner. A projection does not include any probabilistic information on the likelihood that a given scenario or the projected changes will occur.
Reanalysis	A model simulation of past conditions that uses data assimilation through the entire period of the simulation. Reanalysis products are typically used to more fully reconstruct time series in cases where there are gaps in the spatial and temporal coverage of observations.
State variables	The principal variables used to describe the climate system, such as atmospheric temperature, oceanic temperature, humidity, wind speed, pressure,

	and salinity. Climate models focus on looking at the changes in these variables spatially and temporally. In a model, a set of mathematical equations are formulated around the state variables.
Tracers	An indicator that is used to examine changes over time in water mass properties due to advection, diffusion, mixing of water masses, and biological and chemical interactions.
Tuning	Adjusting the value of unknown parameters in a model post-hoc so that model outputs match observations to the greatest degree possible. This process is sometimes referred to as model calibration.

363 **Table 2.** Some additional resources for learning about and beginning to use Earth System  
364 Models  
365

Resource	Description	Reference
Taylor et al. (2011)	A description of the model experiments performed in conjunction with the 2013 IPCC report	Taylor, K.E., R.J. Stouffer, and G.A. Meehl. 2011. An overview of CMIP5 and the experiment design. Bull. Amer. Meteor. Soc. 93: 485-498.
CMIP5 ensemble	Complete collection of model output used in the 2013 IPCC report	<a href="http://cmip-pcmdi.llnl.gov/cmip5/">http://cmip-pcmdi.llnl.gov/cmip5/</a> (last accessed: March 6, 2016)
Community Earth System Model (CESM)	Output from an ESM developed by the National Center for Atmospheric Research (NCAR). A "large ensemble" experiment was performed where this ESM was run 40 times in order to quantify the model's internal variability (Deser et al. 2012)	<a href="https://www.earthsystemgrid.org/home.htm">https://www.earthsystemgrid.org/home.htm</a> (last accessed: March 6, 2016)
MARine Ecosystem Model Intercomparison Project (MAREMIP)	Intercomparison between all ESMs that have a marine biogeochemistry and ecosystem sub-model	<a href="http://pft.ees.hokudai.ac.jp/mare mip/index.shtml">http://pft.ees.hokudai.ac.jp/mare mip/index.shtml</a> (last accessed: March 6, 2016)

366

## Figure captions

Figure 1. Schematic illustrating the process of downloading, reading, and pre-processing data files from IPCC-class models.

