Sterman, J. 2015 Learning for Ourselves: Interactive Simulations to Catalyze Science-Based Environmental Activism. In P. E. Stoknes and K. A. Eliassen (eds.) *Science-Based Activism*. Bergen, Fagbokfolaget: 253-279.

Learning for Ourselves: Interactive Simulations to Catalyze Science-Based Environmental Activism

John D. Sterman
MIT Sloan School of Management
jsterman@mit.edu

Humanity has never been as healthy, rich and numerous as it is today, but at the same time, we have never faced greater and more difficult threats to our health, wealth and survival. The developed nations consume disproportionate shares of global resources and seek further economic growth, while billions lack the food, housing, healthcare, electricity, mobility, education, and other resources the affluent take for granted and legitimately seek these basic needs. Yet the global environmental footprint of humanity is already an unsustainable 1.5 Earths (Wackernagel *et al*. 2002, as updated at http://www.footprintnetwork.org). And it is growing: the UN projects world population will exceed 10 billion by 2100, and exponential growth in global gross product is rapidly worsening environmental problems from collapsing fisheries to water shortage to extinction of species to climate change (Meadows, Randers and Meadows 2004, Randers 2012, Rockström *et al*. 2009, Running 2012).

The traditional response to any environmental problem is trust in technological innovation and market responses. Resource depletion or environmental degradation, it is argued, will raise the price of any scarce resources, leading to innovation that solves the problem. Where there are market failures that prevent the price system from functioning, as for climate change and many other environmental problems, it is the role of scientists and other experts to carry out research identifying threats and potential solutions and communicate the results to political leaders, who will then take actions to mitigate the harms through regulation or government-funded technological innovation. That approach, which I call the Manhattan Project model, does not and can not work for the most important environmental problems we face. Why? In 1939, faced with existential threats, Albert Einstein and other scientists directly alerted President Roosevelt to developments in atomic physics promising weapons of unimaginable power. The President listened to the experts. Then, by focusing enough money and genius in the deserts of New Mexico, the US created nuclear weapons in just six years. The appeal of the Manhattan Project model for environmental threats is clear: it is a deus ex machina, a technical fix that offers the prospect of solutions to tough problems without requiring sacrifice, contentious political battles or deep changes in our way of life.

But a Manhattan Project cannot solve the climate problem or other pressing environmental threats. The bomb was developed in by the US in secret, with no role for the public. In contrast, no single nation can solve the climate problem on its own. The climate is a common-pool resource and vulnerable to the Tragedy of the Commons; solutions require collective action at multiple scales, from individual behavior to government actions at local, national and international levels. Cutting Greenhouse Gas (GHG) emissions enough to prevent additional harm from climate change requires everyone to cut their carbon footprints, requiring dramatic changes in the sources and uses of energy—and the political support to enact the policies and legislative changes needed to implement these changes in time. For many years there has been no reasonable doubt that the climate is changing, that most of these changes are caused by human activity, and that continuing on the traditional path creates high risk of serious harm to human welfare. The science is clear. But political leaders have not acted. Emissions grow to record levels every year. Science is no longer the bottleneck to action. The problem is social and political: politicians cannot act because there is not enough public support for the changes needed to secure our future. Changes in people's views and votes create the political support elected leaders require to act on the science. Changes in buying behavior create incentives for businesses to transform their products and operations. The people cannot be ignored.

At the same time, action must remain grounded in and guided by the best available science. How can policymakers and the public learn about complex environmental issues, and do so in ways that catalyze action rather than causing denial or despair? How can learning occur when the size, complexity, time delays, and irreversible consequences of acting in many systems means there is no possibility of learning from experience? What is required not only to generate reliable knowledge about complex problems, grounded in the best available scientific evidence, but also to catalyze the grass roots support needed to implement high-leverage policies and motivate action to address highly politicized issues? Here I describe how systems thinking and interactive simulation modeling, focusing on the field of system dynamics (Forrester 1961, Randers 1980, Sterman 2000), can help, not only to generate knowledge, but to catalyze the adoption of policies to address the pressing challenges we face. As an example I describe interactive simulations that are used around the world, from senior policymakers and business leaders to high school students, to build shared understanding of and action to address one of the most daunting and urgent issues we face: climate change.

1. Policy Resistance

Thoughtful leaders throughout society increasingly suspect that the policies we implement to address these and other difficult challenges have not only failed to solve the persistent problems we face, but are in fact causing them. All too often, well-intentioned programs create unanticipated "side effects." The result is *policy resistance*, the tendency for interventions to be defeated by the system's response to the intervention itself. From the fossil fuels that power our economy but harm human welfare through air pollution and climate change, to the overuse of antibiotics that spread resistant pathogens, to technological innovations that boost catch per boat only to cause faster collapse of fish stocks, to the obesity and diabetes caused by the sedentary lifestyles and cheap calories prosperity affords, our best efforts to solve problems often make them worse (Sterman 2000, 2012).

Policy resistance arises from a failure of systems thinking, from a narrow, reductionist worldview. We have been trained to view our situation as the result of forces outside ourselves, forces largely unpredictable and uncontrollable. Consider the "unanticipated events" and "side effects" so often invoked to explain policy failure. Political leaders blame recession on corporate fraud or terrorism. Managers blame bankruptcy on events outside their organizations and (they want us to believe) outside their control. But there are no side effects—just *effects*. Those we expected or that prove beneficial we call the main effects and claim credit. Those that undercut our policies and cause harm we claim to be side effects, hoping to excuse the failure of our intervention. "Side effects" are not a feature of reality but a sign that the boundaries of our mental models are too narrow, our time horizons too short.

2. Complexity, learning failures and the implementation challenge

Generating reliable evidence through scientific method requires the ability to conduct controlled experiments, discriminate among rival hypotheses, and replicate results. But the more complex the phenomenon, the more difficult are these tasks. Economic, social, medical, and other policies are embedded in intricate networks of physical, biological, ecological, technical, economic, social, political and other relationships. Experiments in complex human systems are often unethical or simply infeasible (we cannot release smallpox to test policies to thwart bioterrorists). Replication is difficult or impossible (we have only one climate and cannot compare a high greenhouse gas future to a low one). Decisions taken in one part of the system ripple out across geographic and disciplinary boundaries. Long time delays mean we never experience the full consequences of our actions. Follow-up studies must be carried out over decades or lifetimes, while changing conditions may render the results irrelevant. The more complex the system and longer the time frame the more potential confounding factors there are, making it harder to find the causal needle in the haystack of spurious correlations. Complexity hinders the *generation* of evidence.

Learning often fails even when reliable evidence is available. More than two and a half centuries passed from the first demonstration that citrus fruits prevent scurvy until citrus use was mandated in the British merchant marine, despite the importance of the problem and unambiguous evidence supplied by controlled experiments (Mosteller 1981). Some argue that today we are smarter and learn faster. Yet, for example, adoption of medical treatments varies widely across regions, socioeconomic strata, and nations, indicating overuse by some or underuse by others—despite access to the same evidence on risks and benefits (Fisher *et al*. 2003, Clancy and Cronin 2005). From airline kitchens to health care, similar firms in the same industry and even different floors of the same hospital exhibit persistent performance differences despite financial incentives, market forces, publications, benchmarking, training and imitation that should lead to broad diffusion of best practices (Gibbons and Henderson 2013, Wennberg 2010). For example, total factor productivity varies by about a factor of two between the 10^{th} and 90^{th} percentile firms in the same 4-digit SIC industries in the US, and by more than a factor of five in China and India (Syverson 2011). Complexity hinders *learning* from evidence.

Many scientists respond to the complexity and learning problems by arguing that policy should be left to the experts. But this Manhattan Project approach, in which experts secretly provide advice to inform decisions made without consulting the public or their elected representatives, fails when success requires behavior change throughout society. Policies to manage complex natural and technical systems should be based on the best available scientific knowledge. However, in democracies (at least), the beliefs of the public, not only those of experts, shape and constrain government policy. When science conflicts with "common sense" people are unlikely to favor or adopt policies consistent with science (Fischhoff 2007, 2009, Morgan *et al*. 2001, Slovic, 2000, Bostrom *et al*. 1994, Read *et al*. 1994). Strong scientific evidence documents the benefits of seat belts, motorcycle helmets, and childhood vaccinations, yet legislation mandating their use took decades. Citizen groups campaign actively against many of these policies, and compliance remains spotty. The connection between actions and outcomes in these cases is far simpler than the connection between GHG emissions, climate and human welfare. People are often suspicious of experts and their evidence, believing—often with just cause—that those with power and authority routinely manipulate the policy process for ideological, political, or pecuniary purposes. Unable to assess the reliability of evidence about complex issues on their own, and frequently excluded from the policy process, citizen noncompliance and active resistance grow, from motorcycle helmet laws to routine immunization. Without effective risk communication, even the best science creates a knowledge vacuum that is then filled by error, disinformation and falsehood—some supplied inadvertently by people without knowledge of the science and some injected deliberately by ideologues and vested interests (Oreskes and Conway 2010). Complexity hinders the *implementation* of policies based on evidence.

2.1 Dynamic Complexity: Most people define complexity in terms of the number of components or possible states in a system. Optimally scheduling an airline's flights and crews to provide good service at low cost is highly complex, but the complexity lies in finding the best solution out of an astronomical number of possibilities. Such problems have high levels of *combinatorial* complexity. However, most cases of policy resistance arise from *dynamic complexity*—the often counterintuitive behavior of complex systems that arises from the interactions of the agents over time. Where the world is dynamic, evolving, and interconnected, we tend to make decisions using mental models that are static, narrow, and reductionist. Among the elements of dynamic complexity people find most problematic are feedback, time delays, and stocks and flows.

2.2 Feedback: Like organisms, the environment, economy and society are entwined through intricate networks of feedback processes, both self-reinforcing (positive) and self-correcting (negative) loops. However, studies show people recognize few feedbacks; rather, people usually think in short causal chains, tend to assume each effect has a single cause and often cease their search for explanations when the first sufficient cause is found (Dörner 1996, Plous 1993).

For example, the US Corporate Average Fleet Efficiency (CAFE) regulations require the average efficiency of new vehicles to reach 54.5 miles per gallon (mpg) by 2025, a 237% improvement relative to the average US light duty fleet efficiency of approximately 23 mpg in 2012. Such a large improvement would, it is argued, cut US gasoline consumption and GHG emissions substantially, as suggested by Figure 1.

However, in contrast to such open-loop reasoning, the world reacts to our interventions. There is feedback: Our actions alter the environment and therefore the decisions we take tomorrow. Our actions may trigger so-called side effects we didn't anticipate. Other agents, seeking to achieve *their* goals, act to restore the balance we have upset. Yesterday's solutions become today's problems.

Policy resistance arises because we do not understand the full range of feedbacks surrounding—and created by—our decisions. Higher vehicle efficiency is certainly important in reducing greenhouse gas emissions. But the connection between efficiency standards, petroleum consumption and GHG emissions is not as simple as the open loop model in Figure 1 suggests. Figure 2 shows a few of the feedbacks affecting vehicle efficiency and petroleum consumption. As efficiency rises, petroleum demand will drop, causing oil prices to fall, leading people to buy larger, less efficient vehicles and stimulating higher oil consumption in other sectors of the economy (aviation, shipping, heating and industries using oil as feedstocks). Lower oil prices will cut investment in energy efficiency and in solar, wind, and other renewable energy sources we need to reduce petroleum use and GHG emissions. Lower renewable investment and production keeps these sources from achieving the scale and process improvements through learning that lower their costs, further reducing their attractiveness and production in a vicious cycle. As higher efficiency and lower fuel prices lower the cost of driving (per vehicle mile), people will carpool less, use less mass transit, and drive more, undercutting the benefits of the greater efficiency, a process known as the direct rebound effect. As mass transit use and revenue fall, transit systems will be forced to raise fares or cut service, driving still more people into cars and further cutting transit system revenue, another vicious cycle (Sterman 2000 ch. 5). Lower oil prices will erode support for strict efficiency standards that constrain the public and automakers who seek to buy and build larger, more powerful vehicles. The savings from higher efficiency and lower oil prices boost people's real incomes and lead to greater consumption of other goods and services, increasing people's energy use and GHG emissions elsewhere in the economy, a process known as the indirect rebound effect (Herring and Sorrell 2009 and Sorell *et al*. 2009). Indeed, spending our gasoline savings on, say, vacations and the flights to get there may even cause our carbon footprint to rise. The impact of efficiency standards on oil consumption and GHG emissions is determined by a complex network of feedbacks, both balancing and self-reinforcing. Many of these feedbacks offset the intended benefits of efficiency standards.

The existence of these feedbacks does not mean that greater vehicle efficiency is unimportant in the quest for a renewable, carbon-free energy system. To the contrary, creating markets for alternative fuel vehicles that are sustainable both ecologically and economically is essential in limiting GHG emissions and the risks of climate change. But the idea that there is a simple technical fix for any environmental problem is both wrong and dangerous. Effective policies must account for the widest range of feedbacks surrounding the impact of "obvious" solutions, lest we will be blindsided by unintended consequences that undermine their intended benefits and erode the political support needed for them to remain in force long enough to work.

2.3 Time Delays: Time delays in feedback processes are common and particularly troublesome. Most obviously, delays slow the accumulation of evidence. More problematic, the short- and long-run impacts of our policies are often different (smoking gives immediate pleasure while lung cancer develops over decades). Delays also create instability and fluctuations that confound our ability to learn. Driving a car, drinking alcohol, and building a new semiconductor plant all involve time delays between the initiation of a control action (accelerating/braking, deciding to "have another," the decision to build) and its effects on the state of the system. As a result, decision makers often continue to intervene to correct apparent discrepancies between the desired and actual state of the system even after sufficient corrective actions have been taken to restore equilibrium. The result is overshoot and oscillation: stop-and-go traffic, drunkenness, and high-tech boom and bust cycles (Sterman 1989, Randers and Göluke 2007).

People routinely ignore or underestimate time delays (Sterman 1989, 2000). Underestimating time delays leads people to believe, wrongly, that it is prudent to "wait and see" whether a potential environmental risk will actually cause harm. Many citizens, including many who believe climate change poses serious risks, advocate a wait-and-see approach, reasoning that uncertainty about the causes and consequences of climate change means potentially costly actions to address the risks should be deferred. If climate change turns out to be more harmful than expected, policies to mitigate it can then be implemented.

Wait-and-see policies often work well in simple systems, specifically those with short lags between detection of a problem, the implementation of corrective policies, and the impact of those actions. In boiling water for tea, one can wait until the kettle boils before taking action because there is essentially no delay between the boiling of the water and the whistle of the kettle, nor between hearing the whistle and removing the kettle from the flame. To be a prudent response to the risks of climate change, wait-and-see policies require short delays in all the links of a long causal chain, stretching from the detection of adverse climate impacts to the implementation of mitigation policies to the resulting emissions reductions to changes in atmospheric GHG concentrations to radiative forcing to surface warming to changes in ice cover, sea level, weather patterns, agricultural productivity, habitat loss and species distribution, extinction rates, and other impacts. Contrary to the logic of "wait and see" there are long delays in every link of the chain.

More problematic, the short- and long-run impacts of policies are often different (Forrester 1969, Sterman 2000, Repenning and Sterman 2001). Such "Worse Before Better" and "Better Before Worse" behavior is common: credit card debt boosts consumption today but forces austerity when the bills come due; restoring a depleted fishery requires cutting the catch today. The tradeoff between short- and long-run responses is particularly difficult in the context of climate change because the lags are exceptionally long. Standard frameworks for intertemporal tradeoffs such as discounting are problematic because potentially catastrophic events sufficiently far in the future, such as sea level rise from the loss of the Greenland or West Antarctic ice sheets, are given essentially no weight even if discount rates are small. Further, people commonly exhibit inconsistent time preferences (Frederick, Loewenstein and O'Donoghue 2002). For example, people often prefer two candy bars in 101 days over one candy bar in 100 days, but prefer one bar today over two tomorrow, a violation of standard assumptions of rational decision theory. The preference for immediate gratification, which appears to have a neurological basis (McClure *et al*. 2004), often leads people to avoid actions with long-term benefits they themselves judge to be desirable, over and above the usual effects of discounting.

2.4 Stocks and Flows: The process of accumulation—stocks and the flows that alter them—is fundamental to understanding dynamics in general and the climate-economy system in particular. The stock of CO_2 in the atmosphere accumulates the flow of CO_2 emissions less the flow of CO_2 from the atmosphere to biomass and the ocean. The mass of the Greenland ice sheet accumulates snowfall less melting and calving. The stock of coal-fired power plants is increased by construction and reduced by decommissioning. And so on.

People should have good intuitive understanding of accumulation because stocks and flows are pervasive in everyday experience. Yet research shows that people's intuitive understanding of stocks and flows is poor in two ways that cause error in assessing climate dynamics. First, people have difficulty relating the flows into and out of a stock to the level of the stock, even in simple, familiar contexts such as bank accounts and bathtubs. Second, narrow mental model boundaries mean people are often unaware of the networks of stocks and flows in a system.

Poor understanding of accumulation leads to serious errors in reasoning about climate change. Sterman and Booth Sweeney (2007) gave graduate students at MIT a description of the relationships among GHG emissions and atmospheric concentrations excerpted from the SPM in the IPCC's Third Assessment Report. Participants were then asked to sketch the emissions trajectory required to stabilize atmospheric CO_2 by 2100 at various concentrations. To draw attention to the stock-flow structure, participants were first directed to estimate future net removal of CO_2 from the atmosphere (net CO_2 taken up by the oceans and biomass), then draw the emissions path needed to stabilize atmospheric CO_2. The dynamics are easily understood without knowledge of calculus or climate science using a bathtub analogy (Figure 3): the stock of CO_2 in the atmosphere rises when the inflow to the tub (emissions) exceeds the outflow (net removal), is unchanging when inflow equals outflow, and falls when outflow exceeds inflow. Yet, 84% violated these principles of accumulation. Most (63%) erroneously asserted that stabilizing emissions above net removal would stabilize atmospheric CO_2—analogous to arguing a bathtub continuously filled faster than it drains will never overflow. The false belief that stabilizing emissions would quickly stabilize the climate not only violates mass balance, one of the most basic laws of physics, but leads to complacency about the magnitude and urgency of emissions reductions required to mitigate climate change risk (Sterman 2008).

It might be argued that people understand the principles of accumulation but don't understand the carbon cycle or climate context. But the same errors arise in familiar settings such as bathtubs and bank accounts (Booth Sweeney and Sterman 2000, Sterman 2002, Cronin, Gonzalez and Sterman 2009). Moreover, training in science does not prevent these errors. Three-fifths of the participants had degrees in STEM; most others were trained in economics. These individuals are demographically similar to, and many will become, influential leaders in business and government, though with more STEM training than most. Merely providing more information on the carbon cycle will not alter the common but false belief that stabilizing emissions would quickly stabilize the climate.

3. Learning for Ourselves through Interactive Simulations

There is no learning without feedback, without knowledge of the results of our actions. Scientists usually generate that feedback through controlled experimentation, an iterative process through which intuitions are challenged, hypotheses tested, insights generated, new experiments run. When experiments are impossible, as in the climate-economy system, scientists rely on models and simulations, which enable controlled experimentation in virtual worlds (Sterman 1994, Edwards 2010). Learning arises in the process of interacting with the models, hypothesizing how the system might respond to policies, being surprised, forming new hypotheses, testing these with new simulations and data from the real world. Paradoxically, however, scientists, having deepened their understanding through an iterative, interactive learning process, often turn around and tell the results to policymakers and the public through reports and presentations, expecting them to change their beliefs and behavior, then express surprise when these groups—excluded from the process, unable to assess the evidence on their own and presented with claims that conflict with deeply held beliefs—resist the message and challenge the authority of the experts.

When experimentation is impossible, when the consequences of our decisions unfold over decades and centuries, that is, for climate change and many of the important issues we face, simulation becomes the main—perhaps the only—way we can discover *for ourselves* how complex systems work, what the impact of different policies might be, and thus integrate science into decision making. As an example, consider climate change.

In 1992 the nations of the world created the United Nations Framework Convention on Climate Change (UNFCCC) to negotiate binding agreements to address the risks of climate change. Nearly every nation on Earth committed to limiting global greenhouse gas (GHG) emissions to prevent "dangerous anthropogenic interference in the climate system," (http://unfccc.int/essential_background/convention/background/items/1349.php) which is generally accepted to mean limiting the increase in mean global surface temperature to 2°C above preindustrial levels. High hopes were dashed at the 2009 Copenhagen climate conference when face-to-face negotiations among heads of state collapsed. Instead, nations were encouraged to make voluntary pledges to reduce their emissions. Despite some promising developments such as the 2014 US-China agreement, global emissions have risen to record levels since the great recession of 2008. The IPCC's Fifth Assessment Report (2014) concludes:

> "Continued emission of greenhouse gases will cause further warming and long-lasting changes in all components of the climate system, increasing the likelihood of severe, pervasive and irreversible impacts for people and ecosystems. Limiting climate change would require substantial and sustained reductions in greenhouse gas emissions which, together with adaptation, can limit climate change risks." (Synthesis Report, SYR-18, http://www.ipcc.ch/report/ar5/syr/).

The failure of global negotiations can be traced to the gap between the strong scientific consensus on the risks of climate change and widespread confusion, complacency and denial among policymakers, the media and the public (Sterman 2011). Even if policymakers understood the risks and dynamics of climate change—and many do not—in democracies, at least, the ratification of international agreements and passage of legislation to limit GHG emissions requires grass-roots political support.

Historically, information about climate dynamics and risks comes to policymakers, negotiators and the public in the form of reports based on the results of advanced general circulation models such as those used by the IPCC. Such models are essential in developing reliable scientific knowledge of climate change and its impacts. However, these models are opaque and expensive, and neither available to nor understandable by nonspecialists. The cycle time for creating and running scenarios is too long to allow real-time interaction with the models. Consequently, policymakers, educators, business and civic leaders, the media and the general public often rely on their intuition to assess the likely impacts of emissions reduction proposals. However, as shown above, intuition, even among experts, is highly unreliable when applied to understanding how proposals affect likely future GHG concentrations, temperatures, sea level, and other impacts.

Poor understanding of complex systems not only afflicts the public, but the negotiators themselves. In 2008, Christiana Figueres, then lead negotiator for Costa Rica, and named executive secretary of the UNFCCC in 2010, commented

> "Currently, in the UNFCCC negotiation process, the concrete environmental consequences of the various positions are not clear to all of us….There is a dangerous void of understanding of the short and long term impacts of the espoused… unwillingness to act on behalf of the Parties" (personal communication, Sept. 2008).

The C-ROADS (Climate Rapid Overview And Decision Support) model is designed to address these issues and build shared understanding of climate dynamics in a way that is solidly grounded in the best available science and rigorously nonpartisan, yet understandable by and useful to nonspecialists, from policymakers to the public. C-ROADS:

- is based on the best available peer-reviewed science and calibrated to state-of-the-art climate models;
- tracks GHGs including CO_2, CH_4, N_2O, SF_6, halocarbons, aerosols and black carbon;
- distinguishes emissions from fossil fuels and from land use and forestry policies;

- allows users to select different business as usual (BAU) scenarios, e.g., the IPCC Representative Concentration Pathways (RCPs), or to define their own;
- enables users to capture any emissions reduction scenario for each nation portrayed;
- reports the resulting GHG concentrations, global mean temperature change, sea level rise, ocean pH, per capita emissions, and cumulative emissions;
- allows users to assess the impact of uncertainty in key climate processes;
- is easy to use, running on a laptop computer in about one second so users immediately see the impact of the scenarios they test;
- provides an independent, neutral process to ensure that different assumptions and scenarios can be made available to all parties;
- is freely available at climateinteractive.org.

3.1 Model structure and user interface: C-ROADS is a continuous time compartment model with an explicit carbon cycle, atmospheric stocks of other GHGs, radiative forcing, global mean surface temperature, sea level rise and surface ocean pH. Figure 4 shows the overall model architecture, and Figure 5 shows the stocks and flows capturing the carbon cycle. C-ROADS explicitly models CO_2 and other GHGs, including methane (CH_4), nitrous oxide (N_2O), SF_6 and other fluorinated gases (PFCs and HFCs), each with its own emissions fluxes, atmospheric stock and lifetime. Sterman *et al*. 2013 describe the model structure and behavior in detail; complete documentation is available at climateinteractive.org.

C-ROADS includes a variety of climate-carbon cycle feedbacks, including feedbacks from global mean temperature to net primary production and ocean CO_2 uptake. C-ROADS also includes positive feedbacks involving methanogenesis, e.g., CH_4 from melting permafrost, but sets the base-case gains of these feedbacks to zero because they are, at present, poorly constrained by data. Consequently, C-ROADS is likely to underestimate future warming and sea level rise. Users can test any values they wish for these feedbacks. We revise the model as knowledge of climate-carbon cycle feedbacks improves.

C-ROADS simulations begin in 1850. The model is driven by historic CO_2 and GHG emissions and includes the impact of volcanoes, variations in insolation and other forcings. Figure 6 compares C-ROADS to data; the model tracks the data well; Sterman *et al*. 2013 and the full model documentation compare C-ROADS to history for other GHGs and radiative forcing, and to other projections and models.

The user interface enables rapid experimentation with different policies and parameters. On the main screen users can access instructions, a video tutorial, interactive model structure diagrams and documentation, then select the level of regional aggregation for emissions, including global totals, or 3, 6, or 15 different nations and regional blocs (Table 1). Users interested in examining the impact of emissions from nations not explicitly represented can do so by developing a spreadsheet specifying the emissions projections for these nations; C-ROADS can read such files directly. Users then select a BAU scenario, choosing those of the IPCC or Energy Modeling Forum, or specifying their own. Users can also load prior simulations, carry out Monte-Carlo sensitivity analysis to assess uncertainty and analyze the contribution of any nation's proposals to global outcomes.

Next users define scenarios for anthropogenic CO_2 emissions from fossil fuels and land use and emissions of other GHGs through 2100 for individual countries and regional blocs (Figure 7). Users enter projected emissions for each nation or bloc in one of three modes: numerically, graphically, or from an Excel spreadsheet. Users can specify future emissions relative to a user-selected base year (e.g., emissions in 2020 will be 17% below the 2005 value); relative to the BAU scenario (e.g., emissions in 2020 will be 30% below the BAU value for that year); relative to the carbon intensity of the economy for each nation or bloc (e.g., emissions in 2020 will reflect a 45% reduction in carbon intensity relative to 2005); relative to per capita emissions for that nation or bloc (e.g., emissions in 2050 will reflect 10% growth in emissions per capita over the 2005 level for that nation or bloc); or other options (detailed in the documentation). Input modes, target years and emissions in each target year can differ for each nation and bloc.

Model output updates immediately. Users can select graphs and tables to display, by nation/bloc or globally, population and GDP, emissions of CO_2 and other GHGs, emissions per capita, the emissions intensity of the economy, CO_2 and CO_2e concentrations, CO_2 removal from the atmosphere, global mean surface temperature, sea level rise, ocean pH, and other indicators.

C-ROADS also offers interactive sensitivity analysis. Users can alter the values of key parameters, individually or in combination, and get immediate results.

3.2 Applications: Negotiators, policymakers, scientists, business leaders, and educators are among the many who use C-ROADS. Senior members of the US government including legislators and members of the executive branch have used C-ROADS. The US Department of State Office of the Special Envoy for Climate Change has developed an in-house capability to use C-ROADS and deploy it in the UNFCCC and other bilateral and multilateral negotiations. US Secretary of State, John Kerry, has used the model, and commented at a C-ROADS presentation in 2009 (when he was Chair of the US Senate Foreign Relations Committee)

> "More chilling is the computer modeling [the C-ROADS team] did against the current plans of every single country that is planning to do anything, and it's not that big a group.... They took all of these current projections and ran the computer models against what is currently happening in the science. And in every single case, it showed that we are not just marginally above a catastrophic tipping point level. We are hugely, significantly above it."

Dr. Jonathan Pershing, the former Deputy Special Envoy, commented

> "The results [of C-ROADS] have been very helpful to our team here at the U.S. State Department….The simulator's quick and accurate calculation of atmospheric carbon dioxide levels and temperatures has been a great asset to us. …I have made use of the results in both internal discussions, and in the international negotiations…." (personal communication).

Former staff member Dr. Benjamin Zaitchik elaborates

> "…[P]olicy makers and negotiators need to have a reasonable sense of what a particular action will mean for global climate, when considered in the context of other actions and policies around the world. Previously, we would make these calculations offline. We'd download emissions projections from a reliable modeling source, input them to an excel spreadsheet to adjust for various policy options, and then enter each proposed global emissions path into a model like MAGICC to estimate the climate response. This method…was time consuming and opaque: in the end we had a set of static graphs that we could bring into a meeting, but we couldn't make quick adjustments on the fly. With C-ROADS, we can adjust policy assumptions in real-time, through an intuitive interface. This makes it much easier to assess the environmental integrity of various proposed emissions targets and to discuss how complementary emissions targets might achieve a climate goal…." (personal communication).

Former State Department staffer, Eric Maltzer, commented, "You have been the backbone of our analytic work here" (personal communication, 2013).

C-ROADS is also used in China, where it has been disaggregated to include drivers of CO_2 emissions at the provincial level using assumptions about total energy use and fuel mix, by the United Nations Environment Program (e.g., UNEP 2010, 2011), and by climate policy activists, including 350.org's founder, Bill McKibben (http://www.theguardian.com/environment/cif-green/2009/dec/15/bill-mckibben):

> "…the only people who really understand what's going on may be a small crew of folks from a group of computer jockeys called Climate Interactive. Their software speaks numbers, not spin – and in the end it's the numbers that count."

3.3 Educational Use: A free online version, C-Learn, is widely used in classrooms. C-ROADS and C-Learn are also used in an interactive role-play simulation of the global climate negotiations entitled *World Climate* (Sterman *et al*. 2014). Instructions and all materials needed to run World Climate are freely available at climateinteractive.org.

Participants playing the roles of major nations negotiate proposals to reduce emissions, using C-ROADS to provide immediate feedback on the impacts of their proposals. Participants learn about the dynamics of the climate and impacts of proposed policies in a way that is consistent with the best available peer-reviewed science but that does not prescribe what should be done.

For example, participants commonly believe stabilizing atmospheric CO_2 concentrations and the climate requires only that emissions be stabilized. As discussed above, this erroneous belief arises from people's poor understanding of stocks and flows. C-ROADS enables people to discover the dynamics of accumulation themselves. Initial proposals often stabilize emissions well above current levels. When they simulate these proposals, however, they find that CO_2 concentrations steadily grow, because the "Carbon Bathtub" (Figure 3) is still filled faster than it drains. Through experimentation, they discover that stabilizing atmospheric CO_2 concentrations requires substantial emissions cuts, soon (Figure 8).

World Climate has been used successfully with groups including students, business executives and political leaders. Grass-roots civil society organizations such as the youth-led "COPinMyCity" (http://copinmycity.weebly.com) use C-ROADS and World Climate to educate and inspire. The COPinMyCity program combines *World Climate* with "mobilization, awareness raising and debriefing"—what they call "Simul-action"—to help participants connect the science to action (http://copinmycity.weebly.com/simulating-the-nego.html).

Evaluations show *World Climate* improves participant knowledge of climate science and policy options (Sterman *et al*. 2014). Even more important, the experience can generate hope and catalyze action, as these participant comments, collected in sessions run by Prof. Juliette Rooney-Varga, Univ. of Massachusetts, Lowell, illustrate:

> "This exercise makes me think I will have to tackle climate change more seriously."
>
> "I feel surprisingly excited. All of these problems can be solved.... We have a chance to build a new world."
>
> "I will do what I can to consume less and motivate the people close to me to do the same."
>
> "I would like to reduce my own CO_2 emission[s] and lead a program for this in my future career."

3.4 Limitations and Extensions: C-ROADS enables decision-makers, educators, the media, and the public to quickly assess important climate impacts of particular national, regional or global emissions scenarios and to learn about the dynamics of the climate.

As with any model, C-ROADS is not appropriate for all purposes. To be able to run in about a second on standard laptops, the carbon cycle and climate sectors are globally aggregated. Thus C-ROADS cannot be used to assess climate impacts at regional or smaller scales.

C-ROADS takes future population, economic growth, and GHG emissions as scenario inputs specified by the user and currently omits the costs of policy options and climate change damage. Many users, particularly those involved in negotiations, value the ability to specify pledges and proposals exogenously. But GHG emissions result from complex interactions of energy demand, production, prices, technology, learning and scale economies, regulations and government policies. To address these issues, the Climate Interactive En-ROADS model endogenously generates energy use, fuel mix, and GHG emissions. Stocks of energy producing and consuming capital determine energy production and consumption by fuel type. The model includes construction and planning delays for the development of new energy sources and the possibility of retrofits and early retirement for existing capital stocks. The costs of each energy source are endogenous, including resource depletion and supply constraints that raise costs, and R&D, learning curves, and other feedbacks that can lower costs. Users can test a wide range of policies including carbon prices, regulatory constraints and subsidies for specific technologies. Users can also vary key parameters governing resource availability, technical breakthroughs, cost reductions, construction times and lifetimes for new plant, the potential for efficiency and retrofits, etc. The resulting emissions are then the input to the same carbon cycle and climate structures used in C-ROADS so users immediately see the impact of policies such as carbon prices, efficiency standards, subsidies for renewables, etc., on emissions, GHG concentrations, global average temperatures, sea level rise and other climate impacts. Like C-ROADS, En-ROADS simulates in seconds on an ordinary laptop.

4. Conclusion

Policies to address pressing challenges often fail or worsen the problems they are intended to solve. Evidence-based learning should prevent such policy resistance, but learning in complex systems is often weak and slow. Complexity hinders our ability to discover the delayed and distal impacts of interventions, generating unintended "side effects." Yet learning often fails even when strong evidence is available: common mental models lead to erroneous but self-confirming inferences, allowing harmful beliefs and behaviors to persist and undermining implementation of beneficial policies. When evidence cannot be generated through experiments in the real world, virtual worlds and simulation become the only reliable way to test hypotheses and evaluate the likely effects of policies.

Most important, when experimentation in real systems is infeasible, simulation is often the only way we can discover for ourselves how complex systems work. Without the rigorous testing enabled by simulation, it becomes all too easy for policy to be driven by ideology, superstition, or unconscious bias. The alternative is rote learning based on the authority of an expert, a method that dulls creativity and stunts the development of the skills needed to catalyze effective change in complex systems.

Interactive simulations such as the models developed by Climate Interactive, and by MIT Sloan (see https://mitsloan.mit.edu/LearningEdge/simulations) enable people to develop their systems thinking capabilities and understand the science and politics of complex issues, then use these capabilities to design effective policies and build the collective commitment needed to implement them.

References

Booth Sweeney, L. and Sterman, J. (2000) Bathtub Dynamics: Initial Results of a Systems Thinking Inventory. *System Dynamics Review,* 16(4), 249-294.

Bostrom, A., Morgan, M. G., Fischhoff, B., and Read, D. (1994) 'What Do People Know About Global Climate Change? Part 1: Mental models,' *Risk Analysis*, 14(6), 959-970.

Clancy CM, Cronin K. (2005) Evidence-Based Decision Making: Global Evidence, Local Decisions *Health Affairs*. 24(1):151-162.

Cronin, M., Gonzalez, C., Sterman, J. (2009) Why Don't Well-Educated Adults Understand Accumulation? A Challenge to Researchers, Educators, and Citizens. *Organizational Behavior and Human Decision Processes,* 108(1), 116-130.

Dörner, D. (1996) *The Logic of Failure*. New York: Metropolitan Books/Henry Holt.

Edwards, P. (2010) *A Vast Machine*. Cambridge: MIT Press.

Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL (2003) The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Annals of Internal Medicine* 138(4):273-287.

Fischhoff, B. (2009) Risk Perception and Communication. In R. Detels, R. Beaglehole, M. Lansang, and M. Gulliford (Eds), *Oxford Textbook of Public Health*, 5th Ed., pp. 940-952. Oxford: Oxford University Press.

Fischhoff, B. (2007) Non-Persuasive Communication about Matters of Greatest Urgency: Climate Change. *Environmental Science & Technology*, 41, 7204-7208.

Forrester, J. W. *Industrial Dynamics*. Cambridge, MA: MIT Press; 1961.

Forrester, J. W. (1969) Urban Dynamics. Cambridge MA: MIT Press.

Frederick, S., Loewenstein, G., O'Donoghue, T. (2002) Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*. 40(2), 351-401.

Gibbons, R. and Henderson, R. (2013) What do managers do? In Gibbons, R. and Roberts, J. (eds). *Handbook of Organizational Economics*. Princeton: Princeton University Press: 680-731.

Herring, H. and Sorrell, S. (2009) Energy efficiency and sustainable consumption: the rebound effect. Palgrave Macmillan.

McClure, S., Laibson, D., Loewenstein, G., Cohen, J. (2004) Separate Neural Systems Value Immediate and Delayed Monetary Rewards. *Science*. 306 (15 October) 503-507.

Meadows D, Randers J, Meadows D (2004) The limits to growth: the thirty year update. Chelsea Green, White River Junction, VT

Morgan, G., Fischhoff, B., Bostrom, A., Atman, C. (2001) *Risk Communication: A Mental Models Approach*. Cambridge, UK: Cambridge Univ. Press.

Mosteller F. (1981) Innovation and evaluation, *Science*. 211:881-886.

Oreskes, N., Conway, E. (2010) *Merchants of Doubt*. Bloomsbury Press.

Plous, S. (1993) *The Psychology of Judgment and Decision Making*. New York: McGraw Hill.

Randers, J. (ed) (1980) Elements of the System Dynamics Method. Cambridge, MA: MIT Press.

Randers, J. (2012) 2052: A Global Forecast for the Next 40 Years. Chelsea Green, White River Junction, VT.

Randers, J. and Göluke, U. (2007) Forecasting turning points in shipping freight rates: lessons from 30 years of practical effort. *System Dynamics Review* 23(2/3): 253-284).

Read, D., Bostrom, A., Morgan, M. G., Fischhoff, B., and Smuts, D.: 1994, 'What Do People Know About Global Climate Change? Part 2: Survey studies of educated laypeople,' *Risk Analysis*, 14(6), 971-982.

Repenning, N. and Sterman, J. (2001) Nobody Ever Gets Credit for Fixing Problems that Never Happened: Creating and Sustaining Process Improvement. *California Management Review* 43(4), 64-88.

Rockström J *et al*. (2009) A safe operating space for humanity. *Nature* 461:472–475.

Running, S. (2012) A Measurable Planetary Boundary for the Biosphere. *Science*. 337: 1458-1459.

Slovic, P. (ed.) (2000) *The Perception of Risk*. London: Earthscan.

Sorrell, S., Dimitropoulos, J., & Sommerville, M. (2009). Empirical estimates of the direct rebound effect: A review. Energy Policy, 37, 1356–1371.

Sterman, J. (1989) Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment. *Management Science,* 35(3): 321-339.

Sterman, J. (1994) Learning In and About Complex Systems. *System Dynamics Review* 10(2-3): 291-330.

Sterman, J. (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill.

Sterman, J. (2002) All Models are Wrong: Reflections on Becoming a Systems Scientist. *System Dynamics Review,* 18(4), 501-531.

Sterman, J. (2008) Risk Communication on Climate: Mental Models and Mass Balance. *Science,* 322, 532-533.

Sterman, J. (2011) Communicating Climate Change Risks in a Skeptical World. *Climatic Change*, 108: 811–826

Sterman, J. (2012). Sustaining Sustainability: Creating a Systems Science in a Fragmented Academy and Polarized World. In M. Weinstein and R. E. Turner (eds) *Sustainability Science: The Emerging Paradigm and the Urban Environment*, Springer: 21-58.

Sterman, J. and L. Booth Sweeney (2007) Understanding Public Complacency About Climate Change: Adults' Mental Models of Climate Change Violate Conservation of Matter. *Climatic Change,* 80(3-4): 213-238.

Sterman, J., Fiddaman, T., Franck, T. Jones, A. McCauley, S. Rice, P. Sawin, E. Siegel, L. (2013). Management Flight Simulators to Support Climate Negotiations. *Environmental Modelling and Software*. 44: 122-135.

Sterman, J., *et al*. (2014). World Climate: A Role-Play Simulation of Global Climate Negotiations. *Simulation and Gaming*. DOI: 10.1177/1046878113514935.

Syverson, C. (2011) What Determines Productivity? *Journal of Economic Literature* 49: 326–365.

UNEP (2011) Bridging the Emissions Gap. United Nations Environment Programme (UNEP). Available at www.unep.org/publications/ebooks/bridgingemissionsgap.

Wackernagel M., Schulz N., Deumling D., *et al*. (2002) Tracking the ecological overshoot of the human economy. *PNAS* 99:9266–9271.

Wennberg, J. (2010) Tracking Medicine: A Researcher's Quest to Understand Health Care. Oxford University Press, New York.

Figure 1. Open-loop mental model: Tightening vehicle efficiency standards should raise average fleet efficiency, lowering oil consumption and GHG emissions. No feedbacks or so-called "side effects" of higher efficiency are recognized. Arrows indicate causal influence; arrow polarity, e.g., $x \longrightarrow^{+} y$ indicates an increase in x raises y above what it would have been otherwise; $x \longrightarrow^{-} y$ indicates an increase in x lowers y below what it would have been otherwise (Sterman 2000, ch. 5 provides formal definitions and examples).

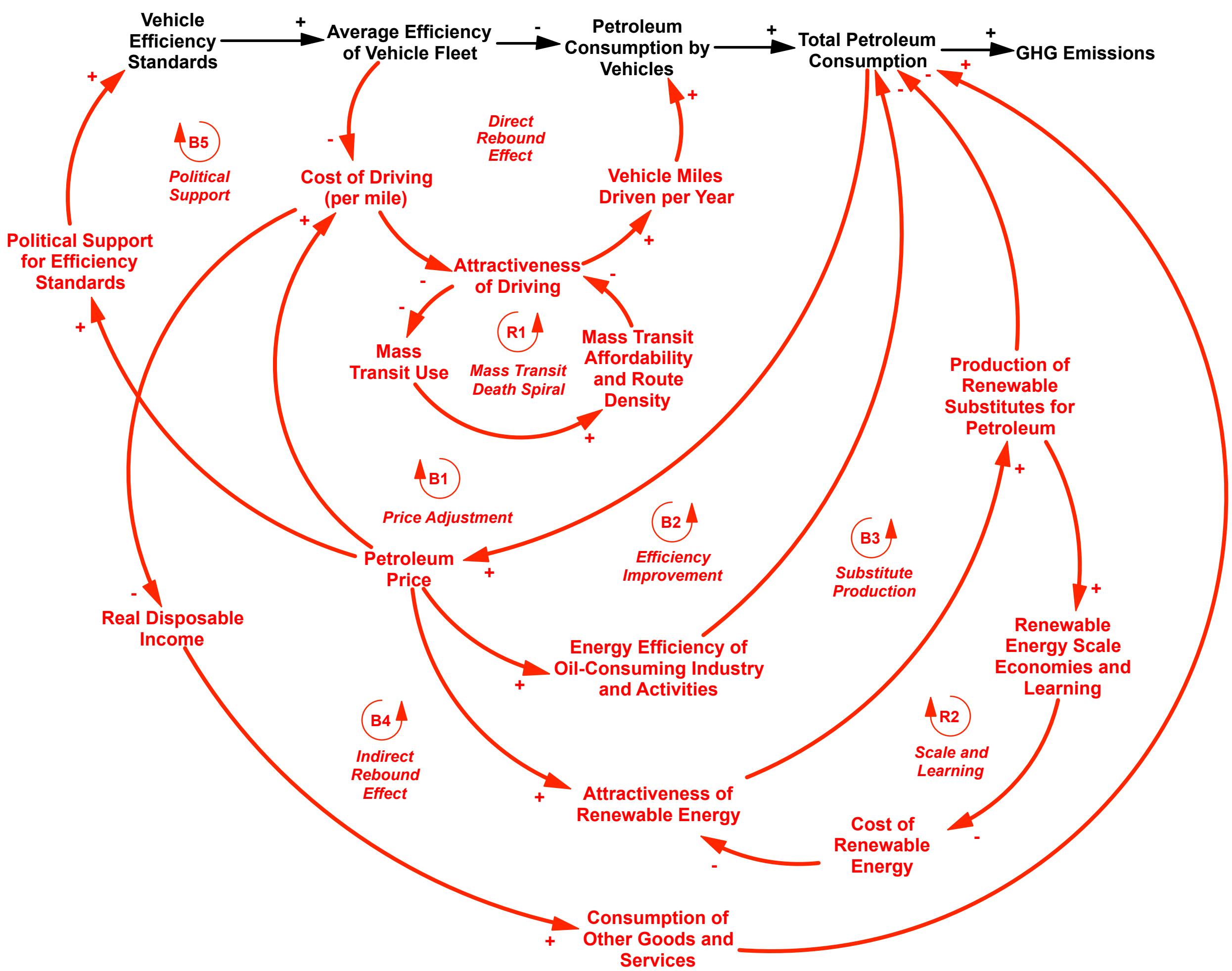

Figure 2. A few of the feedbacks affecting efficiency, oil consumption and GHG emissions. Feedback loops are shown by loop identifiers and names. Balancing (negative) feedbacks are shown by "B"; reinforcing (positive) feedbacks are shown by "R". (see Sterman 2000 for formal definitions and examples). Delays are not shown.

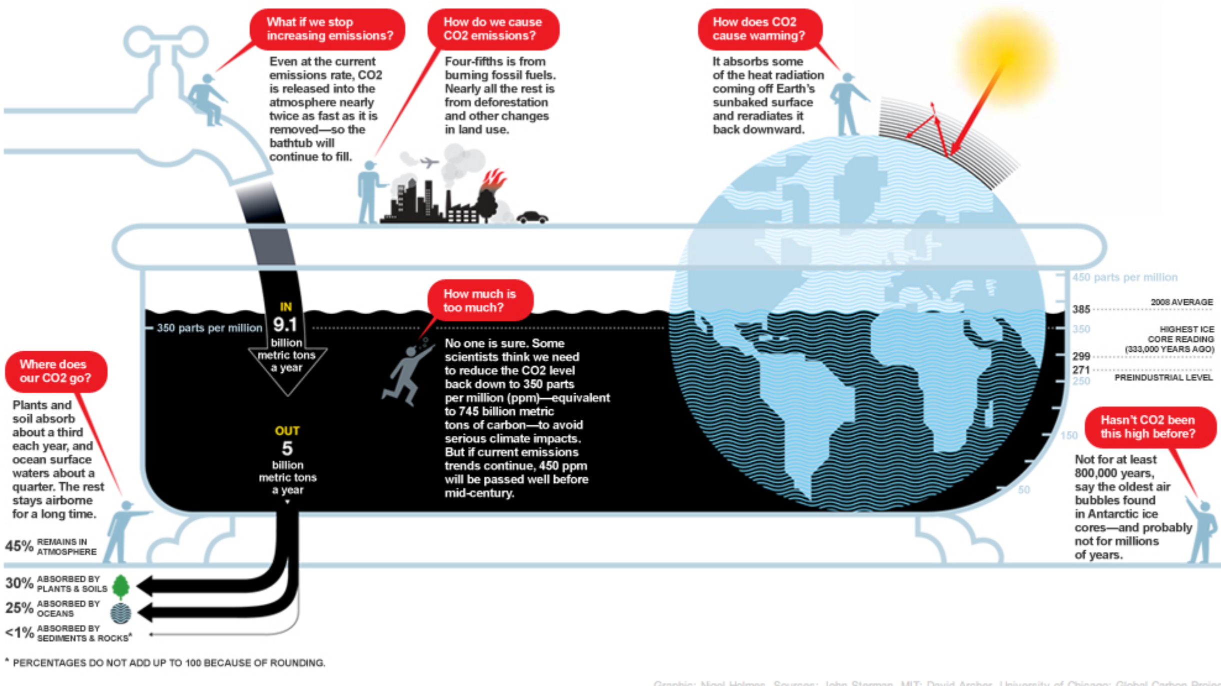

Figure 3. Portraying stocks and flows: The "Carbon Bathtub" (Source: *National Geographic*, December 2009; available at ngm.nationalgeographic.com/big-idea/05/carbon-bath).

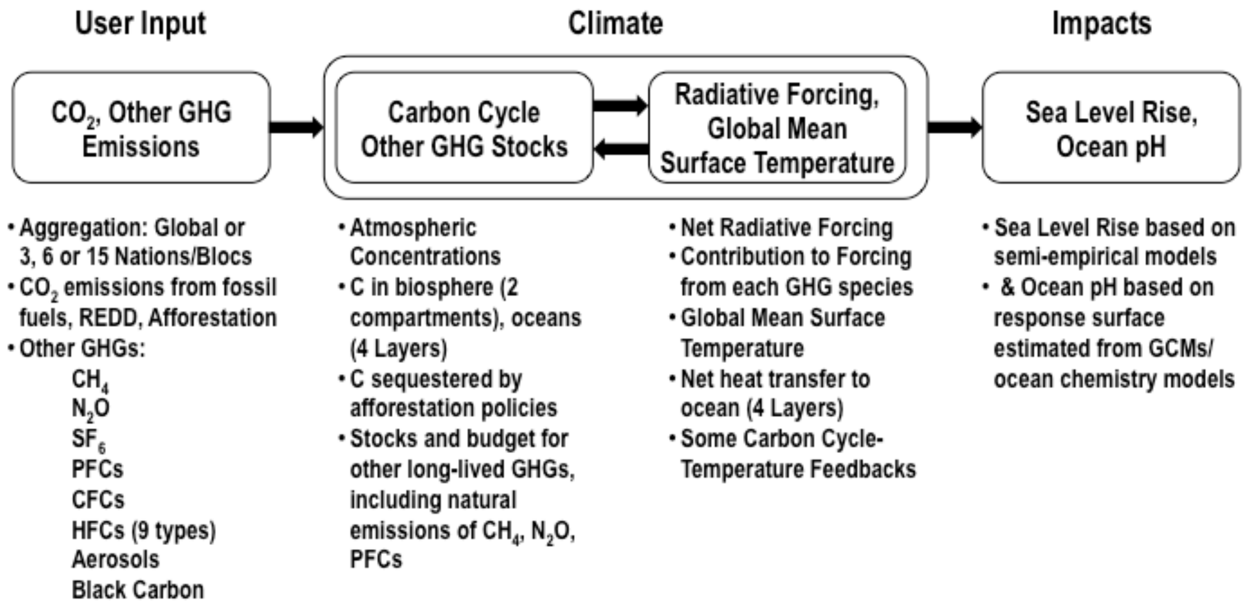

Figure 4. C-ROADS Overview. User-specified scenarios for GHG emissions affect atmospheric concentrations and the climate, which in turn drive impacts including sea level and ocean pH. The model includes climate-carbon cycle feedbacks.

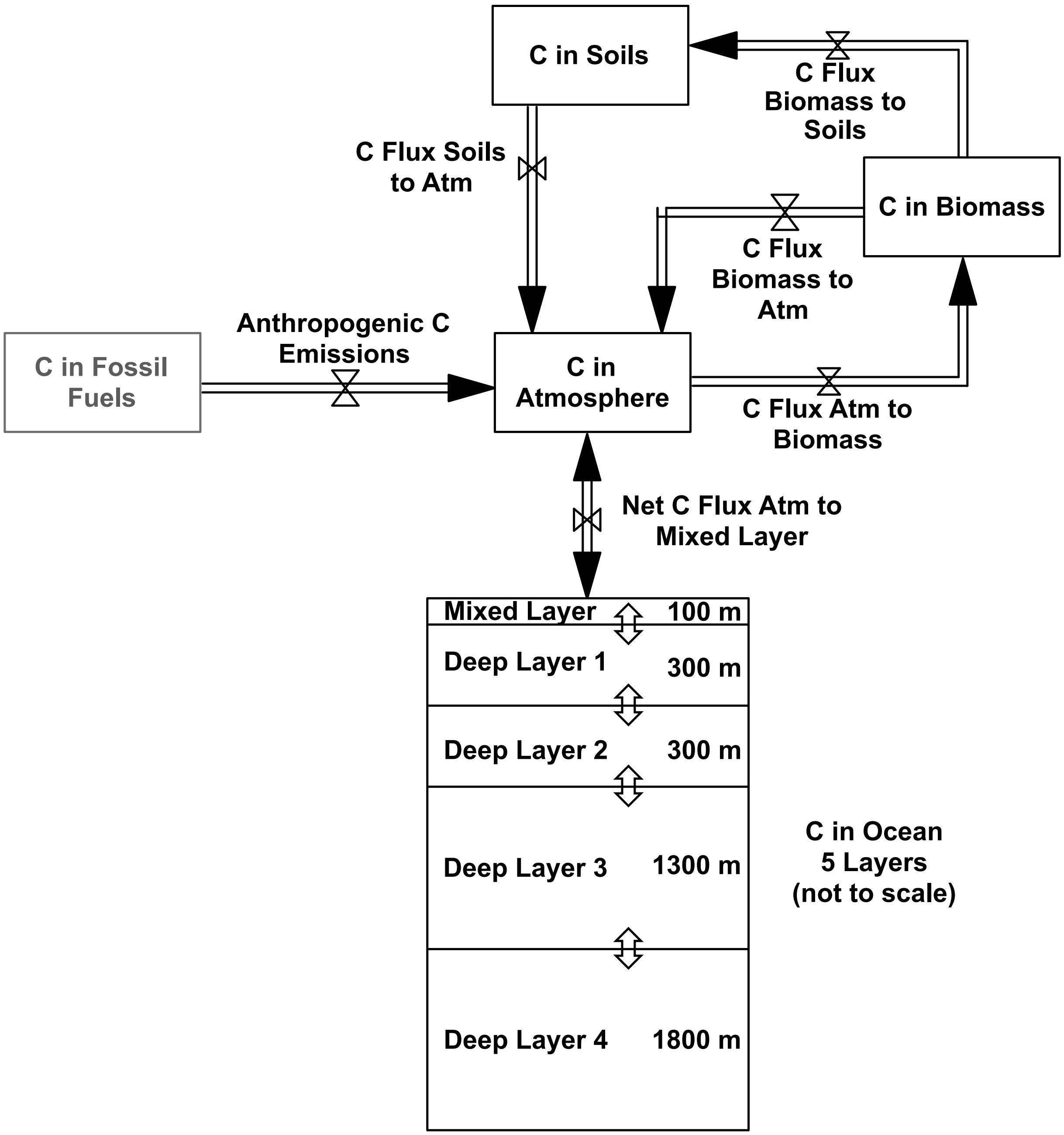

Figure 5. C-ROADS carbon cycle. CH_4 fluxes and atmospheric stock and C fluxes and stocks due to deforestation/afforestation are represented explicitly but are aggregated in this simplified view.

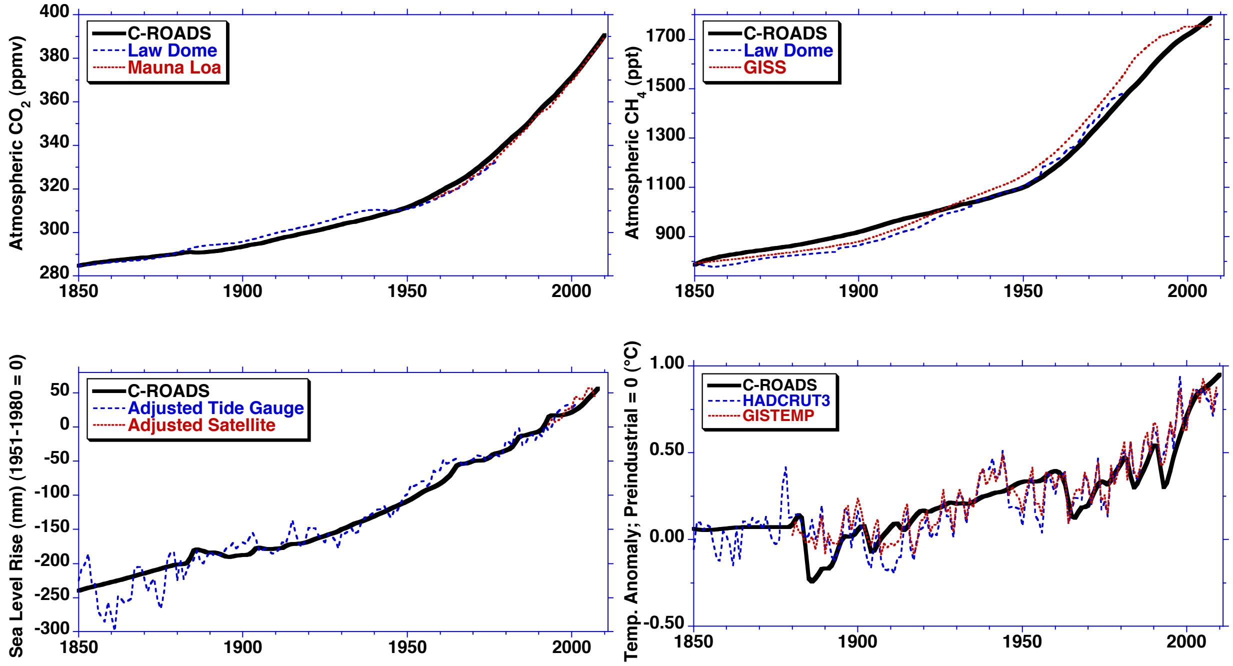

Figure 6. C-ROADS fit to historical data. Clockwise from top left: Atmospheric CO_2, CH_4, temperature anomaly, sea level.

3 Regions[a]	6 Regions	15 Regions
Developed All developed nations Developing A Rapidly developing nations (Brazil, China, India, Indonesia, Mexico, South Africa and other large developing Asian nations) Developing B Rest of world: least developed nations in Africa, Asia, Latin America, Middle East, Oceania	China European Union India United States Other Developed Nations Australia, Canada, Japan, New Zealand, Russia/FSU/ Eastern Europe, South Korea Other Developing Nations Brazil, Indonesia, Mexico, South Africa; Other Africa, Asia, Latin America, Middle East, Oceania	Australia Brazil Canada China European Union India Indonesia Japan Mexico Russia South Africa South Korea United States Developed non MEF[b] nations Other Eastern Europe & FSU, New Zealand Developing non MEF nations Other Africa, Asia, Latin America, Middle East, Oceania

[a] The three region level of aggregation is available in C-Learn, the online version of C-ROADS.

[b] Major Economies Forum on Energy and Climate; www.majoreconomiesforum.org.

Table 1. In addition to the global level, C-ROADS users may choose 3, 6 or 15 nation/region levels of aggregation.

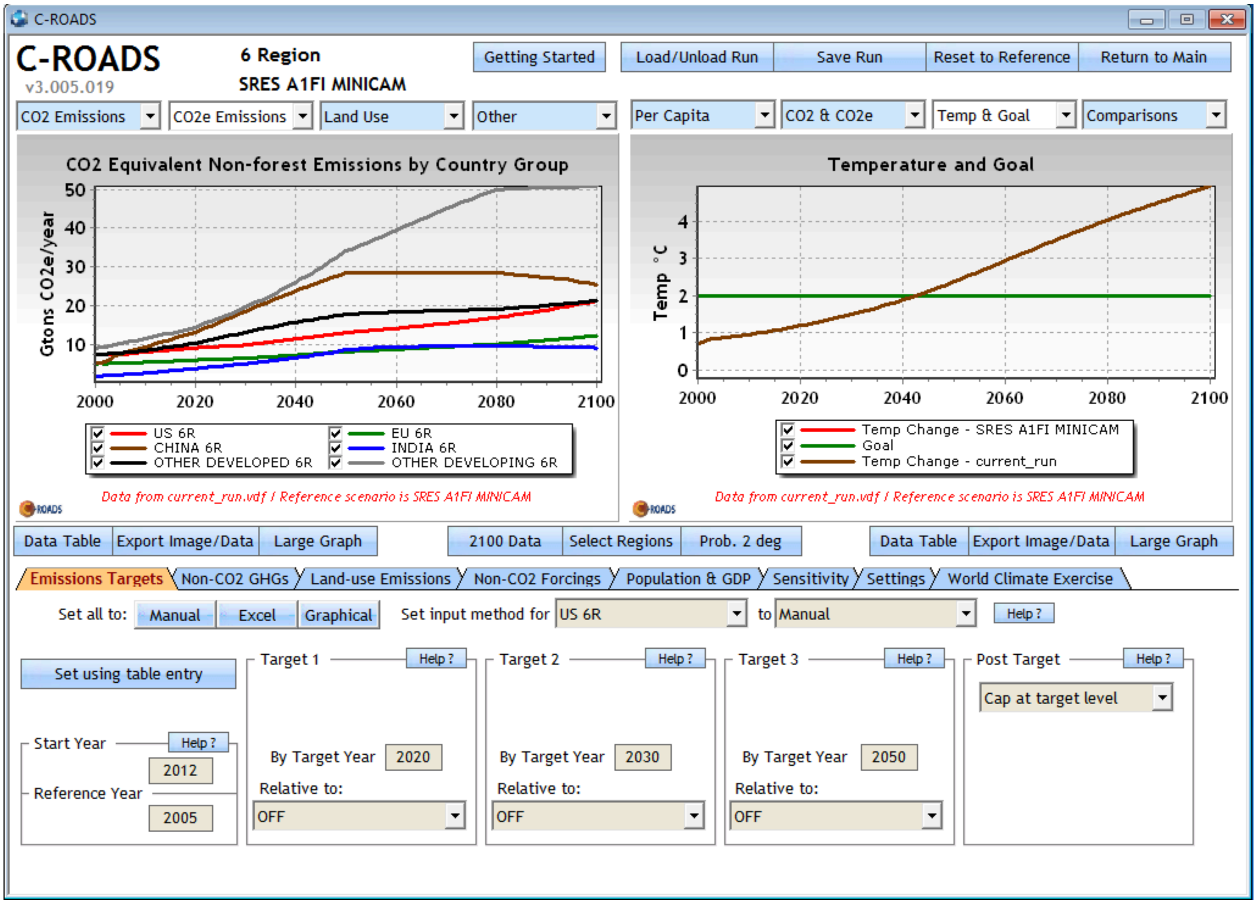

Figure 7. C-ROADS interface, showing emissions (left graph; 6 region level of aggregation) and global average surface temperature (right graph, °C above preindustrial levels). Tabs provide users with different ways to enter emissions pathways and proposals for each nation and bloc, pathways for GHGs other than CO_2, the ability to select different scenarios for population and economic output per capita, sensitivity analysis, and a variety of other graphs, tables, and options so they can carry out the experiments and use the assumptions they want to test.

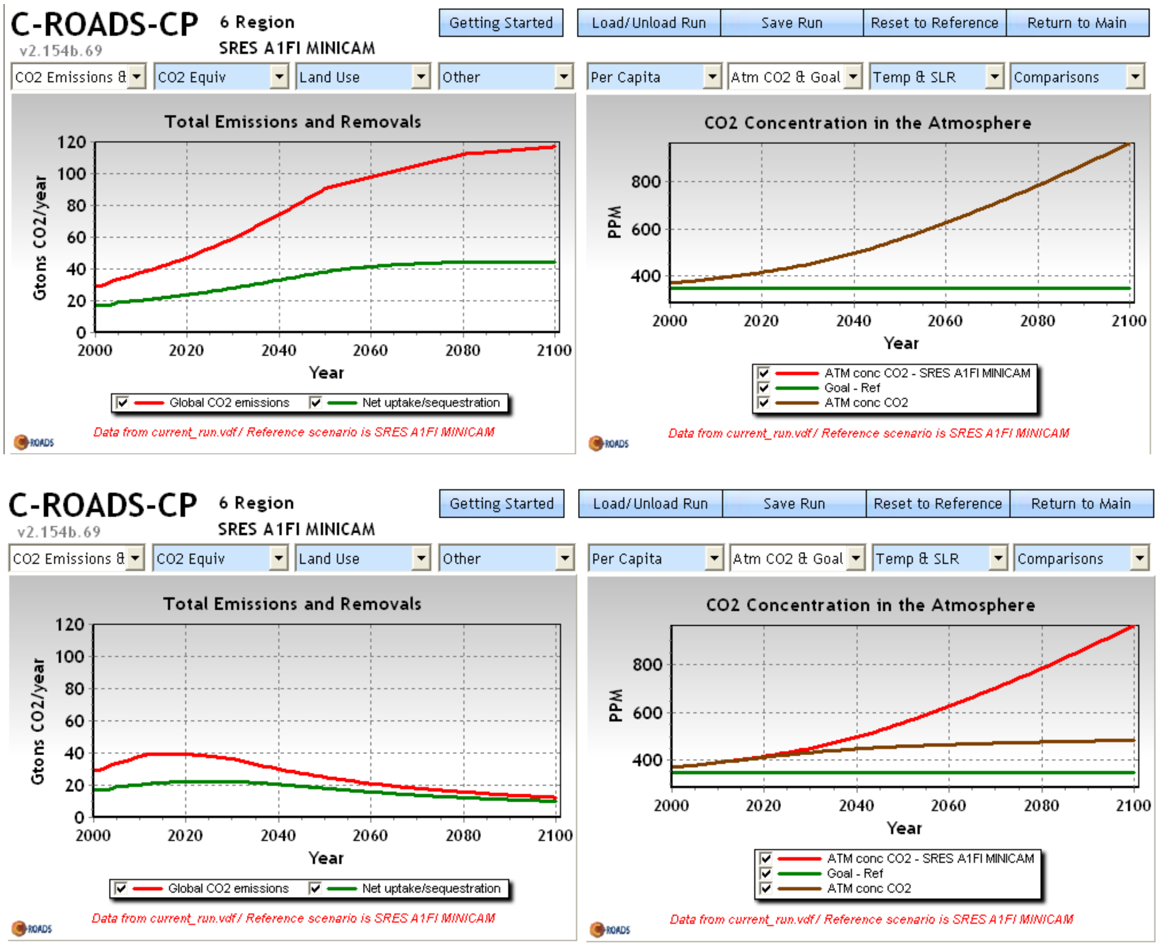

Figure 8. Carbon mass balance or "bathtub dynamics" illustrated by C-ROADS. Top: The graph on the left shows the inflow to the stock of atmospheric CO_2 (global CO_2 emissions; red line) and the outflow of CO_2 from the stock in the atmosphere (net CO_2 removal as it is taken up by biomass and dissolves in the ocean; green line). The inflow always exceeds the outflow, so the level of CO_2 in the atmosphere rises continuously. The gap between inflow and outflow increases over time, so concentrations rise at an increasing rate, reaching 965 ppm by 2100. To stabilize atmospheric CO_2 concentrations, emissions must fall to net removal. Bottom: A scenario in which global emissions peak around 2020 and fall to roughly a third of the 2005 flux by 2100, by which time emissions and net removal are nearly in balance, so that CO_2 concentrations nearly stabilize (in this scenario, at about 485 ppm). Note that net CO_2 removal from the atmosphere falls in the stabilization scenario: lower atmospheric CO_2 concentrations reduce net uptake of CO_2 by biomass and the oceans.