useless arithmetic

Why Environmental Scientists Can't Predict the Future

Orrin H. Pilkey & Linda Pilkey-Jarvis

Columbia University Press New York

Prediction is very difficult, especially if it's about the future.

—*Neils Bohr, Nobel Prize–winning physicist*

chapter two

mathematical models

escaping from reality

War by the Numbers

During World War II, military mathematical modeling, or *operational research*, became a critical tool for analyzing the war experience. One of the more successful applications of mathematical models resulted in a large increase in the sinking of U-boats by the British navy, after studies suggested new tactics and new settings for depth charges. Operational research was also responsible for suggesting that large convoys of merchant ships were safer than small convoys, the opposite of contemporary military thinking on the subject.

However, the low point of the military mathematical model may have come and gone during the Vietnam War, when modeling the battlefield proved difficult and disastrous. Robert McNamara, one of the ten Ford Motor Company whiz kids, instituted a numbers-only mentality in the management of everything from industry to the World Bank. It may have worked quite well at Ford where the whiz kids, hired by Henry Ford II, turned around a foundering company. But this mentality applied to war was another matter.

McNamara, today best known for the fiasco he helped to create in Vietnam, emphasized numbers, costs, and efficiency, while downplaying the role of human intuition. Once when a White House aide said that the war was doomed to failure, McNamara reportedly responded: "Where is your data? Give me something I can put in a computer. Don't give me your poetry." Twenty years after the war was over, Mr. McNamara admitted that his approach to managing the war was "terribly wrong."

Allain Enthoven, now a chaired professor at Stanford University, was the chief whiz kid and systems analyst for McNamara. The McNamara and Enthoven approach to managing war was cold as a fish, quantitative, impersonal, objective, and lacking in intuition and common sense. Events proved that these *rational modelers* had a fatal flaw: they were unable to admit failure.

One infamous part of quantitative warfare in Vietnam was the notoriously inaccurate enemy body count, considered a measure of success in the war. The body count for remote air and artillery strikes was mathematically modeled to determine how many people would be killed by a certain tonnage and type of explosives and the number of napalm canisters, taking into account the terrain, the vegetation, and the density of people on the ground, among other factors. In closer combat involving infantry units, individual combatants tracked the body count. It was a mathematical model vulnerable to manipulation because evaluations, promotions, commendations, and decorations for officers and noncoms were at stake, depending on the results.

As we have already seen, models may be far from objective when human choices and politics play a part in the process. Arriving at high body counts in Vietnam perhaps was easier than going against the grain with more accurate counts, just as going against the grain of an assumed robust cod population on the Grand Banks by reporting more realistic figures proved difficult for fisheries managers. Eventually all dead bodies became enemy dead bodies—"If it's dead and Vietnamese, it's VC" was the gruesome saying of the times.

The body-count modeling problems are obvious, especially in hindsight. They provide important lessons for all quantitative mathematical modeling.

- *Political objectives polluted the models.* The perception that the war was being won was important in order to sustain support back home.

- *The wrong question was asked.* The body count was not a good measure of success of the American army against a highly motivated, disciplined peasant army.
- *No one looked back.* The veracity of the modeling effort should have been confirmed by field checks.

Models, Models, and Models

A *mathematical model* is a description of a process or a prediction about the end result of a process, expressed as an equation or equations. A model is a numerical analogue—a set of equations that describes the relationships between parameters that control a process. In this book we talk mostly about mathematical models that are said to describe or predict with useful accuracy something about large-scale processes on the surface of the earth. This includes both physical and biological processes. All of the model dominions in this book are *applied models,* or societal relevant models used for engineering, policy, financial, or management purposes.

Quantitative mathematical models are predictive models that answer the questions *where, when,* and *how much.* Where will the invasive plant spread next? When will an artificial beach disappear? How much will the sea level rise in the next century?

By contrast, *qualitative mathematical models* are used to predict directions and magnitudes. For example, is a plant likely to be invasive? Will sea level rise or fall? Will the available fish for harvest be large or small? Will the global climate warm or cool? These models also seek the answer to the questions *why, how,* and *what if.* Why is plant species X invasive, while plant species Y is not? How will the nourished beach disappear, and what mechanisms are likely to be responsible for beach loss? What if trawling for a fish species is halted and only long-lining for the species is allowed? What if rainfall increases at Yucca Mountain's proposed nuclear waste storage site?

The distinction between quantitative and qualitative models is a critical one. The principal message in this volume is that quantitative models predicting the outcome of natural processes on the surface of the earth don't work. On the other hand, qualitative models, when applied correctly, can be valuable tools for understanding those processes.

There are a number of other categories of models as well, sometimes rather vaguely defined. *Statistical models* are those based on statistical

studies of past events for the purpose of estimating the probabilistic future behavior of the system. This type of modeling is often used in the social and health sciences. The insurance industry, for example, determines premiums based on statistical models of health data. *Simulations* mimic an event to determine what might transpire. For example, hurricanes, floods, and landslides are often simulated, as are nuclear weapon explosions, battles, and damage to spacecraft in orbit.

Quantitative models may be categorized as either analytical or numerical. *Analytical models* involve simple equations that can be solved rather readily, perhaps using only paper and pencil. *Numerical models* are much more complex, may involve differential equations, and are often solved with complex computer codes.

Determining *model sensitivity* is a method used to resolve the relative importance of the various factors that make a process work. Various components of the equations are changed to see if the outcome of the model changes. Is wave height more important than wave angle or grain size of the sand on a beach in determining sand transport in the surf zone? An important parameter will make a big difference in the final answer and an unimportant factor won't make much difference. *It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature.* If the model is well founded, determining the sensitivity of various parameters is a valid exercise. If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit.

Another distinction between qualitative and quantitative models is the kind of answer that a model provides. If the answer is a single number, the model is quantitative. For example, a quantitative model might predict that the global atmospheric temperature will rise by 3 degrees Centigrade, plus or minus 1 degree, over the next century, whereas a qualitative model might predict that the temperature will continue to increase over the next century, with a possibility that the rate of temperature rise will accelerate. In another example, quantitative modeling is the prediction that because of sea-level rise, the shoreline will retreat 170 feet, plus or minus 30 feet, over the next century. The qualitative equivalent might be a prediction that the shoreline will continue to retreat and probably the rate of retreat will accelerate over the next century. Whether the path to an answer is analytical or numerical, a quantitative answer comes from a quantitative model. The same goes for qualitative models.

In a qualitative model, because one is determining only the direction of a process or the basic mechanics behind a process, only the most important variables need to be considered. Because of the omission of minor processes, the results of all qualitative models may be imprecise or wrong to some degree, but that does not matter so long as the qualitative question at hand can be reasonably answered. Quantitative models require a great deal more accuracy, and to make an accurate prediction a process must be completely understood. All variables of any importance, including feedbacks, must be accounted for if the model is to answer the question at hand.

The actual model may be expressed in one or several relatively simple equations (see appendix), but the calculations using these equations that apply to a large area of the earth's surface through time may be very complex. The method of calculation required for the application of a model is known as the *computer code.*

A single computer simulation of a natural process over time and space may involve hundreds of lines of equations. Imagine the fifteen-year effort involving a small army of specialists that Microsoft went through to develop the word-processing program used in typing the drafts of this page. Millions of dollars were spent in debugging Microsoft Word, yet as anyone who uses a word processor knows, bugs still exist, albeit mostly very minor ones. Programs behind the models that we discuss in this volume have for the most part not been through a detailed quality assurance program. So the question always exists: does the software or computer code actually model what the authors say it models? Programmers know that inevitably there will be many bugs; the hope is that they will all be minor ones.

In chapter 3 we deal with a complex super model, actually based on hundreds of models, to predict the fate of nuclear waste stored at Yucca Mountain, Nevada. These computer codes must describe hundreds of physical, biological, and chemical events that occur over long periods of time over a wide area of the earth's surface. The potential for computer code error is vast, and it is very difficult to evaluate.

A good modeling approach is to "*open-source*" the codes for any and all who are interested. In a recent controversy concerning the shape of the global warming curve over time, however, the scientists who came up with the curve refused to allow others to inspect their computer code. As a result, a pall of suspicion has fallen over their results.

Equally crucial is providing a list of all important assumptions behind models—but this can be tricky. For example, one might say that the

assumed average wave height in a mathematical model to predict sand transport is six feet. But the story behind that assumption is more complex. To fully understand the average wave height number, one must accept the following sub-assumptions:

- All waves come from the same direction.
- All waves are of the same height.
- Future wave conditions will be the same as those in the past.

Naomi Oreskes, science historian and modeling philosopher of the University of California at San Diego, uses Lord Kelvin to provide an illustration of the hazards created in earlier times by the drive for quantification. William Thomson, otherwise known as Lord Kelvin of Kelvin temperature scale fame (figure 2.1), was one of the leading physicists of the latter half of the nineteenth century. More than 100 years ago, in Lord Kelvin's time, there was much uncertainty about the earth's age. This was before the onset of techniques to determine ages by rates of decay of radioactive elements. Estimates by geologists ranged from 100 million years to hundreds of billions of years, but most geologists, more or less correctly, thought that the age must be on the order of a few billion years. Current thinking is 4.5 billion years. To come to their conclusions, the geologists used a *conceptual model* based on observation, past history, and experience, spiced with a dose of intuition. A conceptual model is a qualitative one in which the description or prediction can be expressed as written or spoken words or by technical drawings or even cartoons. The model provides an explanation for how something works—the rules behind some process.

The conceptual model that provided a qualitative age estimate was based on the *Principle of Uniformitarianism*, which holds that the present is the key to the past. It is assumed that the processes that mold and shape the surface of the earth today must have worked the same way in the past. Judging from the rate at which streams, blowing wind, and glaciers remove and deposit sediment today, and the frequency of volcanic eruptions, the geologists calculated, in extremely rough form, that it must have taken billions of years for the earth to come to its present state.

Lord Kelvin, unconvinced by such a crude approach, obtained an age of 98 million years, on the assumption that the earth had started out as a molten body and had been cooling ever since. This determination could be shown using a simple mathematical model, which could be calculated by hand. Kelvin's method was a quantitative and precise way to get at the

Figure 2.1 Physicist William Thomson, otherwise known as Lord Kelvin, estimated the age of the earth to be 93 million years, on the assumption that the planet began in a molten state. The simple model he used to calculate the age was valid, but the underlying assumption was wrong. Geologists using conceptual models correctly determined that the planet was much older, but Lord Kelvin's age estimate remained credible (until the role of radioactive elements was discovered) for a few years, an early example of the quantitative trumping the qualitative. Photo from answers.com.

earth's age. And since its basis was a principle of physics (cooling rate), the results were widely accepted. Lord Kelvin declared much of geologic thinking about fossils, stratigraphy, and earth history to be invalid. He also cast doubt on Darwin's theory of evolution because, according to his concept, the earth had been at its current temperature for only a short time span, too short for evolution to operate.

Alongside the shaky qualitative conceptual models of the struggling field geologists of his day, Lord Kelvin's number, derived from a valid mathematical model, seemed to be a precise and reproducible thing of beauty. Combined with his forceful personality, Lord Kelvin's declaration plunged geology into a virtual dark age that held back progress in both earth science and evolutionary theory for a few years.

But Lord Kelvin was wrong, and it was the discovery of the continuous production of heat by the decay of radioactive elements in the earth's upper layers that finally countered his idea. The present temperature of the earth was not derived from the cooling down of a molten body. Instead, because heat is generated in the crust by radioactive decay of a number of elements, including uranium, the earth has steadily maintained its current temperature for a very long time. Otherwise, we would be looking toward a very cold earth on the not-too-distant horizon. Interestingly, Lord Kelvin's age of the earth is still supported by a number of creationists in their battles with modern earth science.

Lord Kelvin's model was an early example of a quantitative model trumping a qualitative one, a common problem even today. His model of the rate of cooling was perfectly valid—that is, the principles of physics he applied were correct. The cooling of the earth is not a complex process, and a quantitative model can successfully describe it. His mistakes were the underlying assumption of a molten beginning of the earth and the failure to understand the importance of radioactive decay as a source of heat in the earth's crust. The important lesson here is that no model can overcome a series of bad assumptions.

In hindsight, it is hard to see a way that Lord Kelvin could have guessed the truth. His was a *situational bias*, the phenomenon by which our thinking is so obscured by our present state of knowledge and known conditions and observed trends that we are blinded to the future. It is hard to get out of one's own cocoon.

Still another conceptual model of the age of the earth dominated Western thought for more than 1,000 years. It was the biblical chronology model, which began with the publication of *Chronologia* in A.D. 212 by a priest and former Roman soldier named Julius Africanus. The "chronologists" were trying to determine the date of the Second Coming of Christ, in order to understand when the thousand-year reign revealed in the book of Revelation would begin. The widely held assumption at the time was that the Second Coming would occur 6,000 years after the earth was formed. Thus the age of the earth was needed in order to determine the date. The 6,000-year assumption was based on two sources of information. The first was Elijah's prophecy in the Jewish Talmud that the earth would last 6,000 years. The second was the belief that each day in the seven days of creation described in Genesis was in reality 1,000 years and that Christ would return for the seventh day of rest.

Africanus totaled "known" time spans of biblical lives and events, starting with Adam. The Septuagint version of the Hebrew Bible was the source of the data on the early part of the earth's existence, and it revealed that Adam lived for 930 years, Noah for 950, Moses for 120, and Abraham for 175. Adding up all the life spans, Africanus concluded that Christ was born 5,500 years after the formation of the earth and that he would return in the year A.D. 500.

Subsequent chronologists, including Martin Luther, adjusted the date of the Second Coming by a process we would now call *model tweaking*. According to Jack Repcheck's fascinating account of this in *The Man Who Found Time*, "the chronologists [that followed Africanus] were consistent in putting off [the Second Coming] until a couple of hundred years after their own deaths." As is often the case in some modern modeling endeavors, so much uncertainty existed in the original numbers that tweaking was carried out without raising questions of credibility.

The last and perhaps most famous chronologist was James Ussher, the Calvinist archbishop of Armaugh (Ireland), who pronounced in a 2,000-page book published in 1650 that creation of the earth started at noon on Sunday, October 23, 4004 B.C. By his reckoning, Christ should have returned around October 23, 1996.

The age of the earth according to geologists is much greater than Ussher's reckoning. The conceptual model of the chronologists failed for a number of reasons. Like Lord Kelvin's model of a cooling earth, the methodology of the model was reasonable enough, but the underlying assumption was unsound. Counting biblical generations and events is a valid approach (assuming that everything was recorded accurately in the Bible), but the assumption that Adam came along when the earth began has no basis in science.

Faith-based assumptions and conceptual modeling are clearly immiscible. But we will demonstrate that applied mathematical modeling is at times no less biased, skewed, or slanted by political correctness, advocacy, or economic interests than the biblical slant of the chronologists.

Fast-forwarding to the late twentieth century, we confront another celebrated failure of quantitative modeling. It began with the 1972 publication of *Limits to Growth*, a book commissioned by the Club of Rome. The club, a secretive think tank started by a distinguished British research chemist and a successful Italian industrialist in 1966, today consists of around a hundred economists, businesspeople, scientists, and government officials from fifty-two countries on six continents. The club's book famously predicted that within the coming hundred years, there would be widespread natural resource shortages and economic collapses. The authors warned that unless immediate action was taken to control population and pollution, we would not be able to turn the situation around. This doomsday prediction was based on a mathematical model known as the *pessimist model*. Unlike the simple analytical model applied by Lord Kelvin, this was a more complex model called World III and requiring extensive computer calculations. The document argued that population growth and pollution from industrial expansion were leading to total exhaustion of natural resources and massive environmental destruction. It predicted that catastrophes would begin by the year 2000.

There were many problems with the model. It treated the earth's mineral reserves as fixed and unchanging. This decidedly static view of economics and unhistorical understanding of human creativity held that we would run out of oil according to a time schedule calculated from what was then known about reserves and production methods. It ignored the possibility of additional major oil discoveries, advances in petroleum exploration and extraction technology, and the possible contributions of nuclear, solar, or wind energy sources. The model also assumed that food production per unit of land area would remain steady.

Oreskes notes: "In effect [earth] scientists treat the systems they are modeling as though the systems were static. This is not to say that the modelers believe the systems are static—no earth scientist could imagine any system as truly static. Nonetheless scientists often imbed stasis into their models."

University of Manitoba professor Vaclav Smil summed up his view by noting that the *Limits to Growth* report "pretended to capture the intricate [global] interactions of population, economy, natural resources, industrial production and environmental pollution with less than 150 lines of simple equations using dubious assumptions to tie together sweeping categories of meaningless variables."

The problems with the model went beyond the huge technical weaknesses. It was an example of an *advocacy model*. A Club of Rome official stated shortly after the predictions were released that the idea was "to get a message across, and to make people aware of the impending crisis." In other words, the model outcome had been determined before the model was run. Finding the truth according to a preconceived opinion or philosophy is a common flaw in applied mathematical modeling. And it is very similar to finding truth that matches one's religious faith.

The *optimist model* emerged in a 1976 book titled *The Next 200 Years*, by Herman Kahn and others. This volume presented a view of the future that could be briefly stated as "necessity is the mother of all invention." Kahn basically argued that when the need for more food arises, better technology will save the day. When the price of oil soars out of sight, other sources of energy will come to the fore. This model is a qualitative conceptual model, based simply on a number of scenarios devised by the authors.

Both the pessimist model and the optimist model were derived from the same database. The difference is in the assumptions made and in the personal views of the modelers. *Personal view models* are those that are slanted to prove the belief of the modeler.

Ideally, comparing model results with a real-world situation, a process known as *calibration or validation*, tests a model. That is, an attempt is made to "predict" an event that has already occurred using the model in question. For example one could *hindcast* the cod failure on the Grand Banks.

However, one successful calibration or one successful prediction does not mean the next attempt at calibration will also pass muster. As Naomi Oreskes argues, successful reproduction of an event in a complex natural system is no guarantee that the model will accurately predict or describe the next such event. In fact, she argues that most likely it won't make a successful subsequent prediction. Calibration may show that a model fails to reproduce a situation, but the converse is not always true. Leonard Konikow and John Bredehoeft, geologists with the U.S. Geological Survey, made the same point in a famous 1992 paper titled "Ground-Water Models Cannot Be Validated." The Konikow and Bredehoeft paper received the Meinzer award from the Geological Society of America, but their paper and the views of Oreskes seem to have had minimal impact in the modeling community. Model calibration and validation are alive and well.

In some types of modeling, a second calibration, known as verification, is carried out. The model is first calibrated with one set of events and then verified with a second set. It could work like this: The model is tweaked so it successfully predicts shoreline erosion along a stretch of coast that occurred between 1950 and 1970. The tweaked model is then verified by application to the known erosion rate between 1970 and 1990. If it successfully predicts the erosion rate between 1970 and 1990, the model is said to be verified and can be used to predict the future.

Perhaps the single most important reason that quantitative predictive mathematical models of natural processes on the earth don't work and can't work has to do with *ordering complexity*. Interactions among the numerous components of a complex system occur in unpredictable and unexpected sequences. In a complex natural process, the various parameters that run it may kick in at various times, intensities, and directions, or they may operate for various time spans. Chapters 5 and 6 provide examples of this phenomenon, with lists of dozens of parameters that may

affect the natural processes of shoreline erosion and longshore transport of beach sand, respectively. Parameter after parameter kicks in and out—who knows when, where, and for how long. Complicating things even more are positive and negative feedbacks.

Complexity is a big thing in today's modeling world. There are circulating newsletters, books, technical journals, societies, scientific meetings, and branches of funding agencies that are concerned almost exclusively with complexity. The term is formally defined in a number of (complex) ways, but we will stick with the (relatively) simple explanatory description in the previous paragraph.

William Sherden is a marketing consultant, a Stanford University professor, and the author of *The Fortune Sellers*, a book that provides a skeptical view of stock market forecasting. He notes that complex systems are so highly interconnected with numerous positive and negative feedback loops that they often have counterintuitive cause-and-effect results, as when the "addition of a new highway to alleviate a traffic jam causes the traffic jam to become worse," a *negative feedback*. The rich getting richer and the poor getting poorer are both examples of *positive feedbacks*.

Global warming could lead to melting of the Arctic Ocean ice cover, leading to increased evaporation of ocean water, leading to more precipitation in the Arctic region. Increased snowfall leads to increased accumulation of ice leads to a new ice age. Thus, global warming leads to global cooling, a negative feedback of global proportions.

One reason why earth systems are complex has to do with the relationships between the variables that make a system work. A *linear relationship* is one in which variables increase or decrease at a uniform rate—a straight line on a graph. Most relationships between parameters in a complex earth surface process, however, are *nonlinear relationships*. As one variable changes, another may change exponentially. What complicates the relationships between the numerous parameters that control any natural process even more is the fact that a number of them may change simultaneously as a natural process unfolds. A relationship that may be linear in isolation may be nonlinear in the context of simultaneous changes in other parameters. The reality of any natural process on the earth's surface is a convoluted bird's nest of interrelationships. Complexity reigns, and that is the beauty of the natural world. An example is Yucca Mountain (chapter 3), where as the downward rate of water flow increases through the rocks, the volume of water transported increases disproportionately. This is a positive feedback and a nonlinear relationship.

In classic physics, by contrast, the systems being dealt with are usually not complex, in the sense used here. Modeling in physics is labeled *determinism*, in that prediction of events is possible. Thus we are successful in prediction of the future positions of the planets, the times and dates of eclipses, the rate at which radioactive elements decay, and the time it will take a ball to roll down an inclined plane.

The *New York Times* on June 7, 2004, noted: "In New York City sunrise will be at 5:25 a.m. Eastern time on Tuesday, and Venus is to begin leaving the solar disc at 7:06 a.m., when the sun is 17 degrees above the horizon. The planet's final contact with the sun's edge should occur about 7:26 a.m. when the sun is 20 degrees high. There will be another transit on June 6, 2012. After that, the next ones will occur in 2117 and 2125." What a contrast to prediction of events in complex systems like beaches, global climate, fisheries, rivers, the stock market, and invasive plants!

The same predictive success is possible in the engineering design of bridges and elevated water tanks. The laws of physics apply well to steel and concrete. Plus, humanity has accrued a great deal of experience with these materials to sharpen predictions.

Engineering design and prediction always have a large *safety factor* to allow for human error and to assure that structural safety predictions will be right. Designs are intended to last a certain length of time, to withstand a certain wind velocity or an earthquake of a specified magnitude.

Modeling in any system that results in a single answer (right or wrong) without any indication of the possible range of error in the answer is a *deterministic approach*. In most applied quantitative modeling of earth processes, the results should be *probabilistic* to express the uncertainties involved. That is, the answers should have an error bar or a plus or minus expression of the possible range within which the correct answer must lie.

It is a catch-22 situation. Modelers view error bars as a valid response to critics of quantitative mathematical models, but you can't determine accurate error bars for a prediction without having the same level of model accuracy that is needed to get accurate deterministic answers. Furthermore, an invalid model doesn't provide a valid answer, whether you use error bars or not.

The *error envelope*, or *cone of uncertainty*, on a predicted hurricane path (figure 2.2) is an example of a very useful error bar for quantitative mathematical models. The National Hurricane Center and the Weather Channel produce maps showing an ever-widening funnel of possible storm impact areas in the direction of storm movement. The funnel is centered on the most likely predicted line of storm movement for as

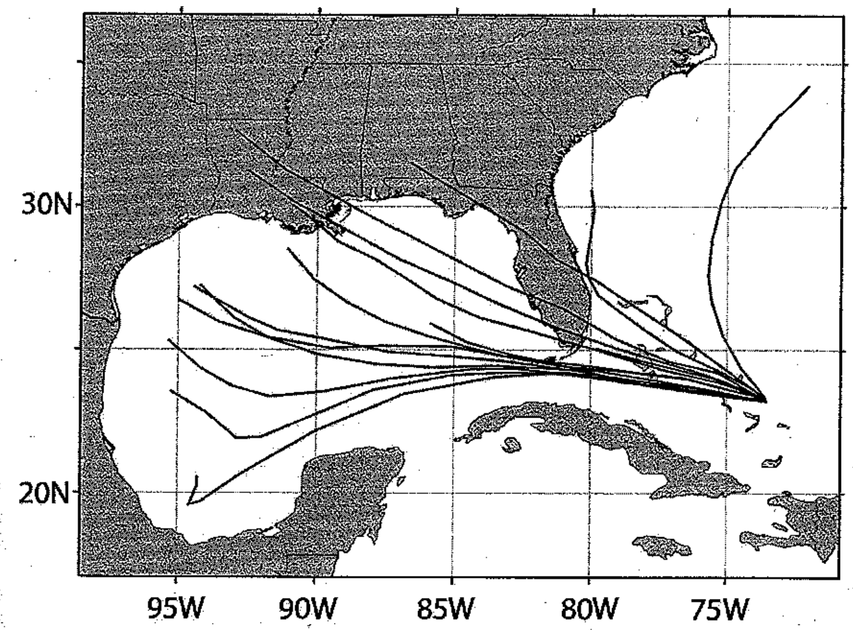

Figure 2.2 Modeled hurricane tracks for Hurricane Rita in 2005. The tracks form a cone of uncertainty, which, while frustrating to coastal dwellers, is a straightforward way to represent the uncertainties of hurricane track predictions. Diagram from Colorado State University Department of Atmospheric Sciences.

much as three days in advance. Most people in hurricane-prone areas probably have an intuitive feeling about the accuracy of hurricane model predictions because in the past many have stocked up with extra food supplies and/or evacuated their homes only to find that the sun is shining and a gentle breeze is blowing on the predicted day of storm arrival.

Meteorologists are up front about the uncertainties of their hurricane path predictions, which are high, even though the models have an excellent statistical or experience base. There must have been much gnashing of teeth at the Hurricane Center when Hurricane Dennis (1999), located off Cape Hatteras, halted and then reversed its path and began moving south, a most unusual path. Teeth gnashing of an even higher order must have occurred when Hurricane Ivan (2004) passed across the Gulf of Mexico shoreline, through the state of Virginia, and then made a wide southerly arc out into the atmosphere over the Atlantic Ocean. Eventually the remnants of Ivan returned to the Gulf of Mexico and crossed the Gulf shoreline for a second time!

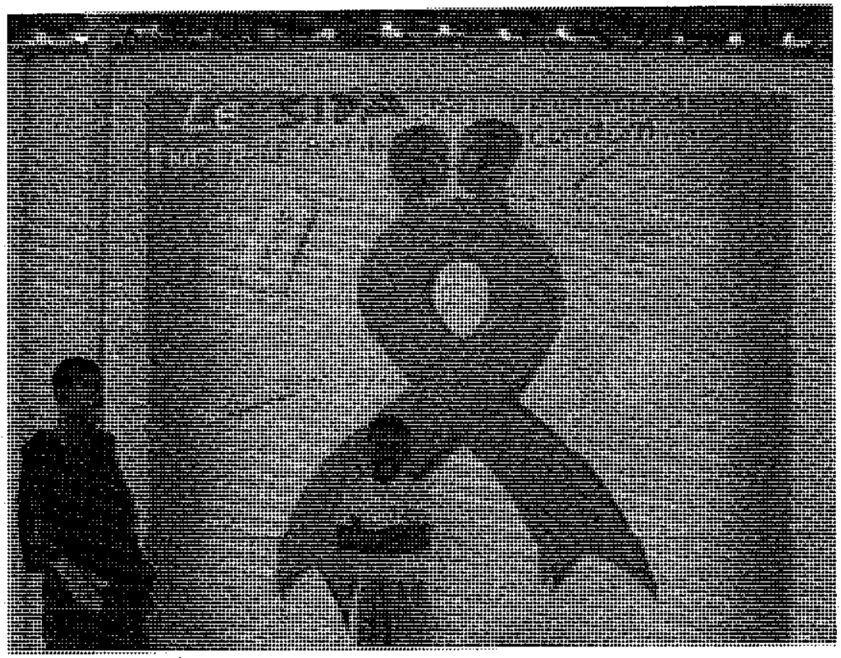

Figure 2.3 AIDS education poster in Mali, Africa. Countries in Africa with extensive education campaigns are having some success in holding down HIV numbers. Mathematical modeling by the United Nations of the extent of this societal catastrophe in Africa illustrates the problem of introduction of a sympathy bias into the numbers. Photo courtesy of the United Nations.

Mathematical models can be used to boost causes, both bad and good. A troublesome example of *good cause modeling* is the prediction and monitoring of the spread of HIV/AIDS around the world, especially in Africa, where the disease is taking its worst toll (figure 2.3). UNAIDS, a sub-agency of the United Nations World Health Organization, takes the responsibility for tracking the disease, which it does in large part through the use of mathematical models. UNAIDS now claims that 30 million Africans suffer from the disease. Rian Malan, a descendant of the Malans who instituted apartheid in South Africa, author, anti-apartheid activist, and now an investigative South African reporter, argued in a startling article in the December 14, 2003, *Sunday Telegraph* that the UN models may have distorted the extent of the AIDS epidemic in Africa.

Quantitative mathematical models are universally used to keep track of and to predict the future courses of diseases. But, of course, models require extensive *ground-truthing*, or *field-checking*. In most of southern Africa, record keeping is poor to nonexistent, and with the exception

of South Africa there simply is no dependable real-world information to run checks on model results. UNAIDS predicted (in hindsight) that 250,000 South Africans died of AIDS in 1999. This figure was determined by the use of the *Epimodal Model*, the same model that was used to predict AIDS deaths all over Africa. Although the number who died of AIDS is unknown, according to Malan it is accurately established that 375,000 South Africans died of all causes in 1999. The number of AIDS victims predicted by Epimodal is far too large a proportion (two-thirds) of the total deaths. Other public health scientists, using the *ASSA 600 model* (Actuarial Association of South Africa), predicted (again in hindsight) that 143,000 South Africans died of AIDS in 1999. In 2001 the "much advanced" *ASSA 2000* model concluded that there must have been 92,000 AIDS deaths. A run of the new *MRC model* (Medical Research Council) came up with 153,000 deaths in 2001–2002 from AIDS in South Africa.

This is not to negate the importance of AIDS modeling, for the disease is a dreadful global plague that kills before middle age and has left orphans by the many thousands. Indeed, there are real difficulties in determining AIDS death rates because the weakened immune system can result in death from a number of other causes. In addition, doctors may not report the involvement of HIV in order to spare stigmatization of relatives or to prevent invalidation of insurance policies.

Robert Guest, in his book *Shackled Continent* (2004), argues that two kinds of orphans are produced by AIDS—the young and the old. In most African societies the elderly expect their children to care for them in their final years. Instead, the old ones are caring for their dying children and then inheriting their grandchildren. The tragic AIDS-related 1995 death of Nelson Mandela's son brought home to South Africans that no one is safe from the disease. In Durban, South Africa, where the disease is particularly widespread, there are now 600 funerals per weekend compared to 120 five years ago, and graves are being "recycled." And the worst may be yet to come, as the disease appears still to be on the increase in Africa.

But the experience in South Africa suggests that the AIDS disaster may not be as far advanced as previously assumed by the United Nations. Certainly this is a point worth considering, because research on other, more ravaging diseases in Africa, such as malaria, is said to be underfunded because of the anticipated AIDS calamity.

Malaria experts say that 900,000 deaths from malaria occur every year in sub-Saharan Africa. Seventy percent of the dead are children under five years of age. Where did the numbers come from? The estimates of both of these dreadful diseases in Southern Africa suffer from the same lack of local health records. The models have a poor database.

The possibility that a true global human disaster is just around the corner unfortunately provides an unparalleled opportunity for modeling that jacks up the numbers to draw attention and funding. Failure to make a simple reality check allowed the results to become accepted "facts." The apparent sophistication of the models dampened criticism, as did the huge outpouring of sympathy for the afflicted.

In this case the models are probably perfectly good, but the answers they come up with are problematic because:

- the database was poor;
- the models were polluted by a huge "sympathy" bias;
- no one looked back.

Reporter Malan noted: "They told me that AIDS had claimed 250,000 lives in South Africa in 1999 and I kept saying this can't possibly be true. What followed was very ugly—ruined dinner parties, broken friendships, ridicule from those who knew better, bitter fights with my wife. After a year or so she put her foot down. 'Choose,' she said. 'AIDS or me.'" He dropped the subject for more than a year but couldn't resist returning to the question, presumably with his wife's reluctant approval. Malan discovered the going-with-the-grain truth about models, that modeling results are easier to live with if they follow preconceived or politically correct notions.

Sometimes the results from mathematical models are pushed aside by public opinion for reasons bearing no relationship to the veracity of the models. *Epidemiology models* provide standard and widely applied methods to evaluate the causes of human illnesses. *Relative risk* is determined by comparing one population that is affected by something (cigarette smoking, coffee drinking, polluted water consumption, and so on) with the general population. The models must take into account a complex array of *confounders* that could affect the results, such as age, sex, economic status, race, location, allergies, and nationality, among others. Still, the results of these data-rich *statistical models* are widely accepted. Statistical models are based on the assumption that past behavior of a system is a guide to future behavior.

The secondhand smoke (SHS) problem is an example of the public's refusal to accept model results because the results are unpopular.

This is buttressed, of course, by the fact that health problems related to smoking are indisputable and widely recognized and that public tolerance for SHS has rapidly diminished in recent years. Imagine the response of a crowd in an elevator when the big guy in the opposite corner lights up a cigarette!

In 1992 the EPA released its famous report classifying SHS as a class A carcinogen. In this example of *politically correct modeling*, the EPA announced that secondhand smoke was responsible for approximately 3,000 lung cancer deaths each year in nonsmoking adults. On the other hand, in 1998, on the basis of a study involving people from six European countries, the World Health Organization reported no significant cancer risk from SHS.

The EPA report has come under intense criticism. It was based on a *meta-analysis*, a summary of 33 previous investigations, mostly epidemiological mathematical model studies, by others, mostly non-EPA scientists. The number of studies actually used was reduced to 11, yielding a risk factor of 1.19. A risk factor of 3 or 4 is usually required before the EPA considers something a risk to humans. The EPA had announced the 3,000 annual American cancer deaths figure before the study was completed, and when the study did not back up the numbers, it doubled the statistical margin of error to come up with something close to 3,000 in order to save the day. The EPA increased the size of the error bars, thus "enclosing" the number 3,000 between the plus and minus extremes of the prediction.

In 1998 federal judge William Osteen declared the EPA study to be null and void. He noted there was evidence in the record that the EPA had cherry-picked its data (chose only the most favorable studies) and the agency was "publicly committed to a conclusion [3,000 lung cancer deaths from SHS] before the research had begun." In a 2003 speech, Michael Crichton, author of *Jurassic Park* and the highly controversial anti-global-warming novel *State of Fear*, called the EPA study "openly fraudulent science." EPA administrator Carol Browner responded to the judge's 92-page scolding of the agency that "the American people certainly recognize that exposure to SHS brings a whole host of health problems" (probably because of the EPA campaign against SHS). "Consensus trumps science," says Crichton.

SHS opponents routinely claim that the models prove a strong cancer health risk. Not true. Such dishonesty is accepted by our society because the cause (prohibition of SHS) has become a moral issue, not a scientific one. In addition, there are real and significant health problems associated with SHS, including heart disease, pneumonia, and bronchitis, especially among children and asthma sufferers.

Useless Arithmetic on Wall Street

Howard Kurtz and William Sherden, along with many others who have written about stock market prophecy, give innumerable examples of erroneous stock market predictions' being presented with great confidence, the aftermath of which produces no loss of prestige to the failed analyst. Phillip Tetlock in his recent book, *Political Judgment*, shows with statistical analyses that experts in general, and experts on the stock market in particular, are no better than educated non-experts at predicting the future.

Hope springs eternal, however, and market prediction is a field that is becoming ever more quantitative. In 1997 the Nobel Prize for economics was awarded to Myron Scholes and Robert Merton, who, collaborating with economist Fischer Black (who died in 1995), developed a mathematical model for stock market derivatives. Black and Scholes derived the original equation, and Merton is said to have improved it in such a way as to make the model applicable in the real world of Wall Street. The equation involved four variables: duration of an option, prices, interest rates, and market volatility.

The Nobel Prize that year was controversial from the start, although few doubted the genius of the equation. The controversy arose over the question of whether helping rich people get richer was elevating mankind in the sense of Alfred Nobel's original intentions. The next year, possibly in atonement for such insensitivity, the Nobel Prize Committee voted for Professor Amartya Sen, known best for his work on the causes of famine, poverty, and social inequality. The Nobel Prize Committee said Sen "has restored an ethical dimension to the discussion of vital economic problems."

The Black-Scholes equation was widely adopted to calculate the value of options in complex derivative dealings. Derivatives are financial instruments that have absolutely no value on their own, but instead "derive" their value from other assets. The use of derivatives reduces risk and uncertainty in profit, for example, the risk of unexpected price fluctuations. There are derivatives that are contracts or obligations for future delivery, called futures, and there are derivatives that give an opportunity (but not an obligation) to buy or sell at an established price, called options.

The Nobel laureates Scholes and Merton were founding partners in the now infamous Long-Term Capital Management (LTCM) hedge fund that helped fuel the explosion of derivatives trading on Wall Street. At its height, LTCM was the darling of Wall Street, a monetary fund comprising a dream team of Nobel Prize–winning founders and complex financial models who seemingly had developed a clean, highly rational way to earn high returns with little risk, using models and supercomputers to identify investments. Some book titles give clues to the fate of LTCM. The story is told in *Too Big to Fail*, by Kevin Dowd, and in *When Genius Failed*, by Roger Lowenstein. In 1998 the fund that was "too big to fail" suffered catastrophic losses that threatened the stability of money markets worldwide. Lowenstein said the cause was "the disease of perfect belief."

According to Lawrence Summers, former secretary of the U.S. Treasury, "The efficient market hypothesis is the most remarkable error in the history of economic theory." Yet two underlying assumptions behind LCTM's market models were

- that markets are always liquid (e.g., you can always sell an asset at a reasonable price); and
- that markets are efficient and they tend toward equilibrium.

For four years, starting in 1994, LTCM showed incredible returns of about 40 percent per year. Stephen Rhodes (a pseudonym) notes that with about 100 employees, LCTM made more money than McDonald's global hamburger business. All this money and not a useful product in sight.

In the global economy today, international markets are closely linked. A trend in one nation's market can quickly spread to the next. The demise of the LTCM hedge fund began on August 17, 1998. Russia defaulted on its debt, and the worldwide financial markets lost their logical order. Investors fled to more-secure investments, and the firm lost about $3.6 billion in five weeks. On one single day, the firm lost $550 million. The collapse of the hedge fund brought little sympathy from the American media. "We're So Rich, We Can Be Dumb," headlined the *San Francisco Chronicle*.

The collapse of Long-Term Capital Management threatened to create a panic on Wall Street, since many major banks had lent it and other such funds huge sums of money. Almost 50 percent of the world's top banks were involved in rescuing the hedge fund. The consortium gave LTCM $3.6 billion in exchange for 90 percent of the firm. Shareholders retained a 10 percent holding, valued at $400 million, and the dream team kept their jobs. Unfortunately, as some in the media have noted, by sparing shareholders, creditors, and fund managers some of the pain of the loss, we seem likely to see a repeat of the behavior that produced the crisis in the first place.

Economic models applied to the stock market do not work because human emotion and action are unpredictable. Is it not obvious that the stock market is not predictable? It shouldn't surprise us that panic, overconfidence, underconfidence, fraud, ignorance, success, and all kinds of other aspects of human nature control the market.

The lesson to be learned from the Nobel Prize–winning equation and its application by LCTM was forcefully expressed by financial guru and founder of Numa Financial Systems, Ltd., Stephen Eckett, who said, "I regard the Black-Scholes model as one of the most dangerous inventions of the twentieth century. This is not to blame Black and Scholes obviously: the danger is always in the application. But what happened was that one single equation—and mathematically the model is simple—seemed to offer the possibility of quickly understanding and controlling derivatives risk. This encouraged thousands of banks to employ bright mathematicians who had little knowledge of the financial markets but nonetheless started furiously programming their spreadsheets on which billions of dollars were gambled."

Cathy Minehan, president of the Boston Fed, is quoted as saying, while introducing a behavioral economist, "All our models and forecasts say we will have a better second half. But we said that last year. Now don't get me wrong. Mathematical models are wonderful tools. Standard economics would argue that people are better off with more options. But behavioral economics argues that people behave less like mathematical models than like—well, people."

The scandalous bankruptcy at the Enron Corporation holds modeling lessons as well. Economist Keith Cooley describes one of the ways that Enron was able to jack up its apparent profits: "At the heart of the so-called innovative trading at Enron was an accounting rule. When Enron agreed to supply power to a company or municipality at a fixed-price contract, it made projections on the level of future prices and the likely profit over the lifetime of the contract. Under the accounting rule in question, it was then able to report that profit as soon as the contract was signed rather than booking the gains over time."

The problems with Enron's prediction of profits under newly acquired contracts were

- the models were " undisclosed";
- the predictions were always highly optimistic;
- credence was provided through approval by an "independent" accounting firm (the firm, Arthur Anderson, lent the model results an air of credibility, but it had to sell itself off in pieces when the scandal unfolded);
- no one looked back.

A Look Back

Modeling equations are sometimes modified and altered (*tweaked* or *tuned*) until the model correctly "predicts" an already known natural event. Frequently in the modeling literature, however, this is considered a model application or prediction. Although the model may have reproduced something in nature, it is a jimmied equation, one that was adjusted bit by bit to fit a single event or to arrive in the approved range. According to Peter Haff, Duke University model critic, model philosopher, and physicist turned geologist, this approach is better termed *model development.* Haff notes that modeling when the outcome is already known is not the same as a true model prediction before an event occurs. It is in no sense a model prediction.

A good tweaked model example is the modeling of the artificial floods that took place in the Grand Canyon in 1996. Water was purposely released from the Glen Canyon Dam in Colorado to imitate the floods that occurred before the dam smoothed out the peaks in flow volumes (figure 2.4). The problem was that the dam had smoothed out floods and had also trapped almost all the sand coming down the river. In addition, the river was expected to become more difficult and dangerous for rafters because mounds of sediment ranging up to boulders in size, brought to the river by flooding tributaries, were staying in one place, piling up and creating dangerous rapids. Normally, the floods flattened out these rock piles.

The hope for this experiment was that the floods would leave behind new sandbars, just like the old floods once did. Sand would be derived from the stream channel, and the new sandbars would provide much-needed new campsites for river rafters and stream habitats for native fish species that are fast disappearing.

The 1996 experimental flood, probably the first of many to come (a second release was carried out in 2004), didn't completely succeed (politicians declared it a success, but scientists had a different view). Few new sandbars were formed. After the fact, however, geologists were

Figure 2.4 An artificial flood in progress. Water is being released at a high rate from the Glen Canyon Dam in an attempt to provide additional sandbars for river boaters in Grand Canyon. Predictive models of the sandbar configuration that resulted from the water release were unsuccessful. Photo courtesy of the U.S. Geological Survey.

able to tweak the model and come up with the same sandbar configuration on paper that actually remained on the canyon floor after the flood. The modelers confidently suggested that the model would now be useful to predict what will happen in planned future water releases from the dam. Tempting as it might be to believe that the model was now valid, agreement between the model and a single event is not an indication of model validity in a complex system, as noted by Naomi Oreskes. The modelers had developed a new model by tweaking the old one, but it probably won't predict sandbar formation in the next flood. It probably won't be even close.

Today's scientists have substituted mathematics for experiments, and they wander off through equation after equation and eventually build a structure which has no relation to reality.

—*Nikola Tesla, inventor and electrical engineer extraordinaire, 1934*

chapter three
yucca mountain
a million years of certainty

Waste Disposal: A Troubled History

The development and use of nuclear technology began in the early 1940s. Americans grandly entered into a nuclear age that ended a world war and promised permanent supplies of cheap energy. Many cultural images from this time linger with us today: dancing the atomic boogie, drinking atomic cocktails, building backyard bomb shelters, and practicing "duck and cover" drills in schools. Even today the mushroom-shaped cloud remains the high school symbol for the "Bombers" of Richland, Washington, located near the Hanford nuclear plant.

Over time, our perception of the bright promise dimmed as the hazards and by-products of nuclear use became apparent. Some nuclear waste products produce radiation that persists for long periods of time; other waste products can be used to make nuclear weapons. The antinuclear movement has been a vocal, visible presence in the United States for decades. Government failures are largely responsible for the current high level of skepticism that the American public holds for our regulation