Chapter 10
Predictability

This chapter focuses on prediction of future climate and how models are used to generate predictions. Prediction includes both projections (estimates given scenarios) and forecasts (estimates given the current state of the climate), as discussed in Chap. 9. Prediction can occur over different timescales. One of the key aspects of prediction is that projections and forecasts are often not useful unless they come with an estimate of uncertainty. So we spend a good deal of space in this chapter trying to understand and characterize uncertainty.

We hear predictions about the future all the time: from sporting events to the weather. But generally a prediction is not useful unless we understand how certain it is. Suppose, for example, the forecast high temperature tomorrow is going to be 5 °F (3 °C) above normal (the average long-term mean or climatology for the date). If the uncertainty on that estimate is 10 °F (6 °C), is the forecast really that useful? If you hear of a "chance" of rain, you really need to know what chance (probability): 10 % is a lot different from 90 %. On the other hand, we can be pretty certain the sun will rise tomorrow at the predicted time (whether we can see it or not). So uncertainty becomes critical to the idea of predictability.

This is especially true for climate and climate change. Climate models can provide projections of the future. Scientists have been issuing projections for a generation or more (30+ years) of future climate using climate models (see below). A common prediction is the global average temperature response to a given forcing of the system. This forcing is often a carbon dioxide (CO_2) level of 560 parts per million, or twice the pre-industrial value of 280 parts per million. Climate model predictions of the climate response to this "doubled CO_2" forcing have not changed much in 30 years.[1] But the models have. So what have we learned? We have learned a lot about uncertainty. Incidentally, if you look at the historical temperature record for the past 30 years since many of the forecasts were published, most of the

---

[1] The basic understanding of what will happen to global temperature has not changed much (between 1 and 4 °C global temperature change for doubling CO_2 concentrations in 30 years; see Charney, J. G. (1979). *Carbon Dioxide and Climate: A Scientific Assessment*. Washington, DC: National Academies Press; and Houghton, J. T., Meira Filho, L. G., Callander, B. A., Harris, N., Kattenberg, A., & K. Maskell, eds. (1996). *Climate Change 1995: The Science of Climate Change*. Cambridge, UK: Cambridge University Press.

forecasts have been correct or close to correct.[2] But *correct* and *incorrect* also depend on the degree of uncertainty. A forecast of 5 °F above normal, when the actual temperature is 8 °F above normal, is incorrect if the "uncertainty" in the forecast is 2 °F (a range of 3–7 °F above normal), but the same forecast is correct if the uncertainty is 4 °F (1–9 °F above normal). The small changes over the past 30 years are not enough to sufficiently constrain the future. The models are broadly "correct."

In this chapter, we review key uncertainties in model projections (and forecasts) of climate change. We start by trying to characterize and classify uncertainty, which depends on the physical system and the particular set of problems as well as the time and space scale of climate projection. We then discuss methods for using models to understand uncertainty with multiple simulations and models.

## 10.1 Knowledge and Key Uncertainties

What do we know about climate? Or, perhaps more important, what do we *think* we know, and what do we *not* know about climate? Uncertainty in prediction is tied to several different aspects of the system: (a) the physics of the problem, or how constrained it is; (b) the variability in the system, which is related to the underlying physics; and (c) the sensitivity of the system to changes that might occur. With regard to climate, we discuss each of these aspects in turn.

10.1.1 *Physics of the System*

Uncertainty depends in part on the underlying physics of the problem. Some problems are better constrained than others, and this affects our ability to predict them. In climate simulation, the global average temperature is often used as a metric. This implies an average over time and over space. The global average temperature is a fairly well-constrained number: Global averages of the incoming and outgoing energy at the top of the atmosphere and assumptions about the ocean heat uptake allow a pretty good estimate of the global average temperature. This is because the physics of energy conservation are fairly straightforward. Energy comes in from the sun. Some is reflected and some is absorbed by the surface, and then some is radiated back. There are complex constraints on how much energy is absorbed, is reflected, or remains at the surface, mostly related to cloud processes

---

[2]For an example of a forecast nearly 30 years old, see Hansen, J., Fung, I., Lacis, A., Rind, D., Lebedeff, S., Ruedy, R., et al. (1988). "Global Climate Changes as Forecast by Goddard Institute for Space Studies Three-Dimensional Model." *Journal of Geophysical Research: Atmospheres, 93* (D8): 9341–9364.
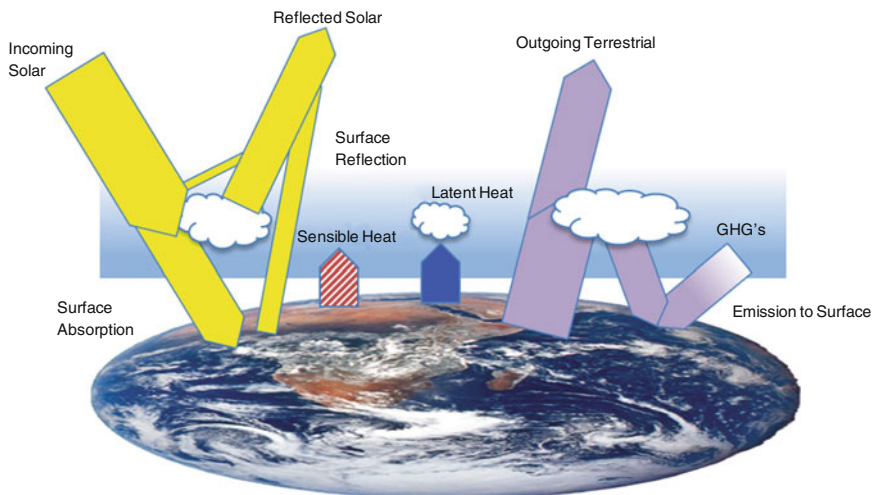
Fig. 10.1 Energy budget. Solar energy, or shortwave radiation (*yellow*) comes in from the sun. Energy is then reflected by the surface or clouds, or it is absorbed by the atmosphere or surface (mostly). The surface exchange includes sensible heat (*red striped*) and latent heat (*blue*, associated with water evaporation and condensation). Terrestrial (infrared, longwave) radiation (*purple*), emitted from the earth's surface, is absorbed by the atmosphere and clouds. Some escapes to space (outgoing terrestrial) and some is reemitted (reflected) back to the surface by clouds and greenhouse gases (*GHGs*)

and the ocean circulation, as illustrated in the energy budget diagram (Fig. 10.1).[3] But the global energy flows have limits, and the physics (conservation of energy) are fairly certain. Some of them offset each other as well. Nonetheless, we cannot predict the year-to-year variations very far in advance (more on this later).

There are other aspects of climate and efforts at prediction that are not as certain and do not have strong constraints. Prediction of regional distributions of rainfall, for example, is not as well constrained by total energy budgets. In addition to local evaporation and precipitation, air motions bring moisture sources into a region from various and changing direction. The total available energy in a region for precipitation depends on the global atmospheric circulation. It may also depend on details of cloud processes. So predicting the distribution of rainfall at any location is much less certain than global quantities due to the basic physical constraints.

---

[3]For an overview of the earth's energy budget, see Chap. 2 of Trenberth, K. E., Fasullo, J. T., & Kiehl, J. (2009). "Earth's Global Energy Budget." *Bulletin of the American Meteorological Society, 90*(3): 311–323.

10.1.2   Variability

Another aspect of uncertainty related to the underlying physics is the level of variability. Recall Fig. 9.2, with two distributions having the same mean but different variance (represented by the **standard deviation**). If the variance is high, we may not be very good at prediction, since the quantity is highly uncertain. Global average temperature is an example of a distribution with low variance, whereas regional rainfall is a distribution with high variance. Higher variances are harder to limit and predict, often because our knowledge of the distributions is low. It is hard to tell if a **signal** (change in mean, for example, or change in variance) is significant if there is a large amount of variability. The local temperature example in the introduction to this chapter is an example of this. A temperature deviation from a mean by 4 °F (2 °C) is hard to detect if the variance is large, usually measured as a standard deviation from the mean (see Chap. 9). The probability of the deviation's being "significant" depends on how variable the distribution is, so larger variability (standard deviation) yields less certainty.

The concept of variability is used to help understand the difference between a "signal", and "noise."[4] The signal is a change in the mean: a trend in the climate over time. The noise is the variability that occurs around the mean, such as the variability of a given climate metric from year to year. The larger the variability (noise), the harder it is to see the signal. Higher variability typically occurs on smaller space and time scales. The annual average temperature of a particular location is much more constant than the daily temperature, and the global average temperature varies far less (daily or annually) than does the average temperature in any particular location or region.

10.1.3   Sensitivity to Changes

Finally, and related to the concept of distributions, impacts often depend on the sensitivity to changes of a distribution. Sensitivity has two aspects. First, we may be worried about extreme events, which are low-probability events at the tail of the distribution. Changes to extreme events are hard to predict because they are low-frequency events, and they have high variability. This is another aspect to uncertainty in prediction that stems from the nature of the physics of climate.

Second, we may worry about events that have thresholds. For example, if the temperature is 68 °F (20 °C), and there is a 1–2°(F or C, F is bigger) uncertainty, that might not matter much. But if the temperature is 32 °F (0 °C), being wrong by a few

---

[4]For a good popular treatment of the signal and noise in statistics and in climate modeling, see Silver, N. (2014). *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't*. New York: Penguin. Chapter 12 is all about climate science and climate prediction.
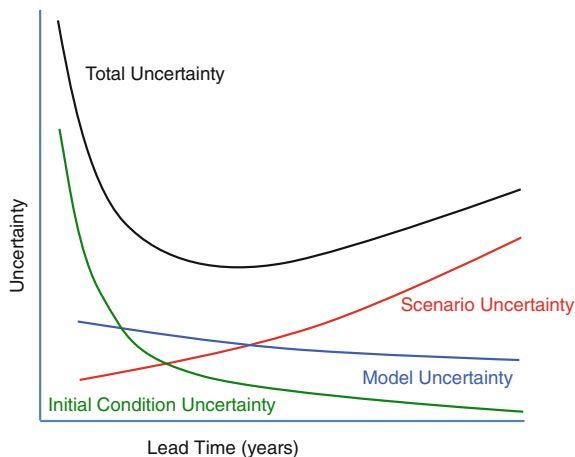
Fig. 10.2 Types of uncertainty. Uncertainty is illustrated here as a function of lead time. Initial condition uncertainty is shown in *green*, model or structural uncertainty in *blue*, scenario uncertainty in *red*, and total uncertainty (the sum of three uncertainties) in *black*. Initial condition uncertainty is large initially, then shrinks, and scenario uncertainty grows over time. *Source* Based on Hawkins and Sutton (2009)

degrees in either direction makes a big difference: If ice and snow are present and remain, then the weather situation is a whole lot different than if a cold rain falls.

10.2   Types of Uncertainty and Timescales

In addition to the physics of a problem contributing uncertainty (which we characterize later in the discussion of model uncertainty), there are uncertainties related to the timescales of prediction that have profound implications for climate modeling. Prediction has different sources of uncertainty on different timescales. These can be divided into uncertainty in predicting the near term (initial condition uncertainty), predicting the next few decades (model uncertainty), and predicting the far future (scenario uncertainty).[5] The categories are the same as introduced in Chap. 1. Understanding uncertainty and what it means is a critical tool for evaluating climate models, that is, for understanding whether they are fit for the purposes of prediction for which the models are used. The different types of uncertainty are illustrated in Fig. 10.2.

---

[5]The definitions follow Hawkins, E., & Sutton, R. (2009). "The Potential to Narrow Uncertainty in Regional Climate Predictions." *Bulletin of the American Meteorological Society, 90*(8): 1095–1107.

10.2.1   Predicting the Near Term: Initial Condition Uncertainty

Short-term climate prediction commonly means predicting the climate from the next season to the next decade or two. In many respects, it is similar to weather forecasting. In the climate context, a decade is "short term." In some sense, the short term is characterized by the period over which the initial state of the system matters: The weather tomorrow depends on the weather today. Other examples include routine seasonal forecasts of expected high and low rainfall one season in advance. Broadly, we are defining "short term" as the period over which the current state of the system matters, and when we can predict the state of some of the internal variability of the climate system.

As described earlier, prediction depends on the problem. Some parts of the system are more predictable than others. In some regions, rainfall in the next season depends on things that are predictable several months in advance, like ocean temperatures. For example, current and predicted temperatures of the tropical Pacific Ocean provide predictive skill for rainfall in western North America.[6] In other places, like Northern Europe, extremes of precipitation or temperature (both high and low) are usually functions of **blocking events**, which are weather patterns that persist but are predictable only a week or so in advance. The extent (persistence) of blocking patterns is difficult to predict at all.

Estimates of the near term (whether a week or a season in advance) made with skill can be valuable in preparation. For example, El Niño commonly brings wetter conditions to Southern California, and when a strong El Niño is developing, precautions are often taken to deal with mudslides, especially in regions recently affected by fire. Forecasts of impending tropical cyclone impacts 24–72 hours in advance enable evacuations and staging of rescue and disaster relief.

But high variability that makes predictions uncertain means that they are less useful. As a counterexample, El Niño also alters the frequency of occurrence of tropical cyclones in the Atlantic (hurricanes), because the upper-level winds in the Atlantic blow the opposite way from usual (and in the opposite direction to surface winds). However, while storms are less frequent, even a single storm in an El Niño year can be damaging. In 1992 after a large El Niño, there were only four hurricanes that year in the Atlantic and one major storm (average is six hurricanes and three major ones), but that single storm was Hurricane Andrew, the second most devastating (in financial terms) storm in recent U.S. history. Because the variability of tropical cyclones is extremely large, and the small-probability events are devastating, the improvement in understanding and forecasting may not be that valuable for prediction.

Near-term prediction is very much focused on the current state of the system: What El Niño will do in 6 months depends strongly on the state of the system now.

[6]The tropical Pacific temperatures are due to El Niño; see Chap. 8.

It is very similar to trying to predict the weather a few hours or day ahead. Most of the uncertainty is in the initial state of the system, called **initial condition uncertainty**. As an example, the El Niño forecast depends strongly on the current state of the atmosphere and ocean; it does not depend on the CO_2 concentration in the atmosphere in 6 months.

Initial condition uncertainty is a common problem between weather and climate prediction: The climate system is usually sampled sparsely with uncertain observations, and we do not know key things to help us "constrain" the system to forecast the weather more than a few days in advance. If we had more and better information (observations of temperature and wind) in the right places, we would be able to reduce initial condition uncertainty.

For climate prediction (e.g., What will the distribution of weather states be next year, in the year 2020, in 2080, or 2200?), initial condition uncertainty dominates over other uncertainty out to decadal scales. This is because there are long-term patterns in the climate system, mostly from the transfer of heat to and from the deep ocean. At scales beyond a few months, it is the evolution of the ocean that governs how the system responds. Figure 10.2 illustrates that initial condition uncertainty is very large and dominates the uncertainty initially but averages out over long periods of time. The weather tomorrow is dependent on today, and El Niño next year is dependent on this year. But the weather and El Niño's state in 2080 is not dependent on today.

10.2.2   Predicting the Next 30–50 Years: Scenario Uncertainty

As the example of the weather or El Niño suggests, there are certain scales for which we know the state of how the system is "forced": Tomorrow's weather does not depend much on the greenhouse gas concentration. Over the course of a few years, the earth's energy budget does depend on the greenhouse gas concentration. But this concentration is fairly well known. As we described in Part II, there are sources and sinks of carbon, and these balance to leave a reservoir of carbon in the atmosphere (see Fig. 7.6). The CO_2 concentration (and that of methane) changes slowly, with variable growth rates, but in the near term—next year, or in the next 10 years—we have a pretty good idea of what will happen. The total use of fossil fuels also changes slowly over time. How much energy did you use in your house this year? How much gas did you use either in your own car, or as part of sharing a ride or using public transportation? Absent major life changes (new house, new job, new car), you can probably project that your use next year will be similar. So it goes for societies in general. Figure 10.3 illustrates the global emissions of carbon. It is a line that trends up a little bit from year to year. You can make a good estimate of next year's emissions from this year's, and projecting forward with a line is not too hard.
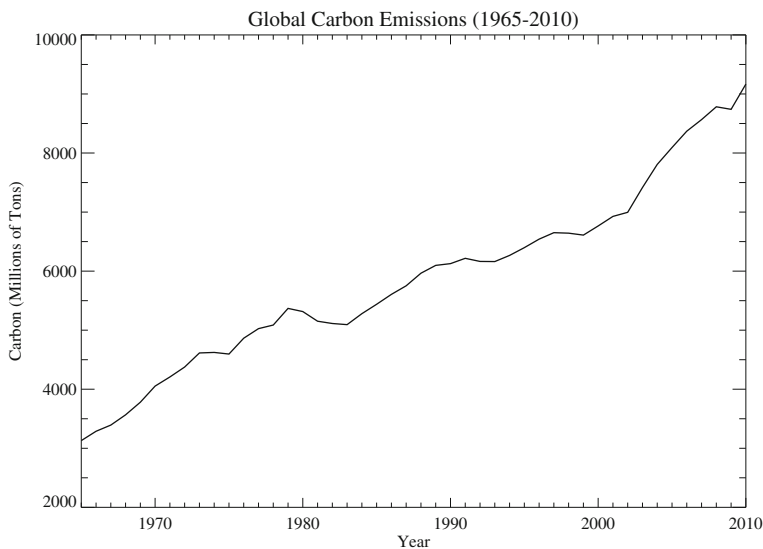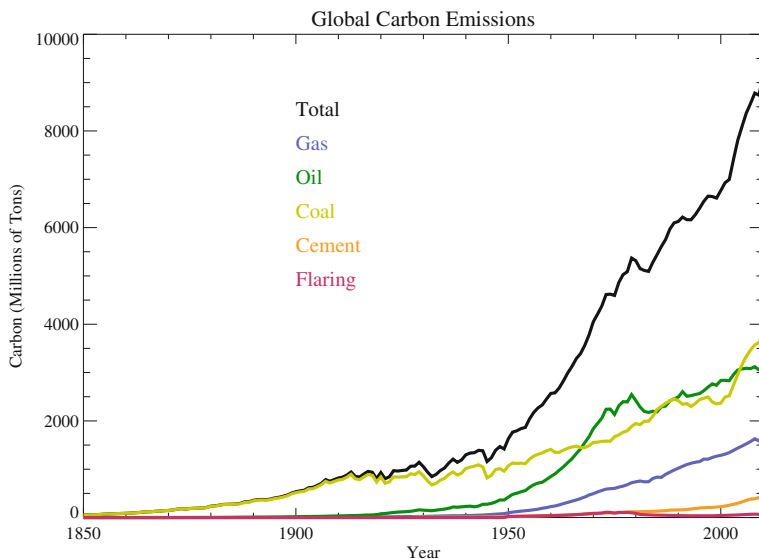
Fig. 10.3 Recent CO_2 emissions. Total global carbon emissions over time from 1965 to 2011 in millions of tons. Data from the NOAA Carbon Dioxide Information Analysis Center (NOAA)

But over longer periods of time, things change. What will your personal energy use be in 10 years? That depends on where you live, your job, your house, if you have a family. It also depends on how you will use energy. For example, next year you might put some more efficient lights in your house. But in 10 years you might have an electric car, or you might telecommute, or you might have solar panels, and it might be the same house or a different house. So the uncertainties on your energy use increase. And the uncertainties on your carbon emissions get larger. For instance, you will probably still have a refrigerator. We can probably guess it will use a similar amount of energy as today (maybe you will have the same refrigerator), but maybe you (or your utility) will get the energy for that appliance from solar or wind or gas rather than coal. Those changes mean more variability in the future. Figure 10.4 shows the same data as Fig. 10.3, but now from 1850 to 2011. Now the use is harder to project into the future. If the plot stopped in 1900 or 1950, the "projection" might be very different from what you would assume today. The mix of fuels would be different: a projection of petroleum use in 1970 would be very different from one made today. And you can guess in the future that it might also be very different, for any number of reasons.

Figure 10.5 shows an example of a plot that does *not* continue to increase. Figure 10.5 is of another greenhouse gas emission: CCl_2F_2, dichlorodifluoromethane, also called CFC-12,[7] a refrigerant used in home and industrial coolers from refrigerators in homes to air-conditioning units in cars or buildings. It also

[7]CFC-12 data from McCulloch, A., Midgley, P. M., & Ashford, P. (2003). "Releases of Refrigerant Gases (CFC-12, HCFC-22 and HFC-134a) to the Atmosphere." *Atmospheric Environment, 37*(7): 889–902.

Fig. 10.4 CO$_2$ emissions by sector. Global carbon emissions by sector from 1850 to 2011 in millions of tons. Total emissions are the sum of the other sectors. Gas (methane, CH$_4$), Oil and Coal sectors are burning of these fossil fuels. Flaring is burning methane at oil wells, and CO$_2$ from cement production occurs due to the chemical process and the fossil fuels used for energy to drive the process. Data from the NOAA Carbon Dioxide Information Analysis Center (NOAA)

happens to contain chlorine (Cl), and when it decomposes in the upper atmosphere this chlorine destroys stratospheric ozone, resulting in the Antarctic ozone hole. As a result of this effect, CFC-12 use has effectively has been banned. Again, projecting the emissions of this substance in 1960 or 1980 into the future would be very different from what actually happened. In the case of CFC-12, it was regulated under the 1986 Montreal Protocol, and eventually production had stopped over most of the planet.

You can see where this is going. Over short periods of time, the basic inputs to a climate model—concentrations of greenhouse gases—are fairly predictable. But over longer periods of time, due to external factors (political oil shocks, new technologies) or regulation (limits or bans on productions or emissions), the inputs can change dramatically. For the historical period, whether for carbon or CFCs, we have a pretty good handle on emissions. However, what does this mean for the future? Will oil use rise, as in Fig. 10.4? Will the Chinese have as many cars per person as in the United States? Or is oil use going to look more like the curve in Fig. 10.5, because either everyone is going to have solar panels on their roofs or we run out of economical supplies of oil.

These are not questions for climate models to answer. They are beginning to be questions for integrated assessment models (see Chap. 8). But the answers (how much oil will be used, and how much CO$_2$ be will emitted by humans in the year 2080) are highly uncertain. They depend on myriad socioeconomic choices. They are not so much predictions as projections and they represent scenarios. Oil use in the future is dependent on a lot of interrelated things: population, population density,
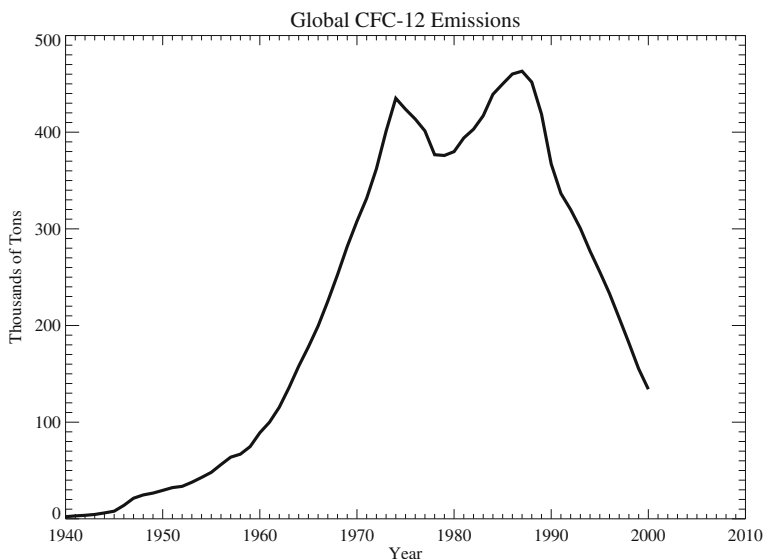
Fig. 10.5 Global CFC-12 emissions from 1930 to 2000 in thousands of tons. Emission controls began in 1986. Data from McCulloch et al. (2003)

available economical supplies of oil, new technologies, and how all these factors interact. One of the problems of scenarios is predicting the application and spread of **disruptive technologies**. Integrated assessment models are better at predicting the improving efficiency of refrigerators, or power plants, or solar panels. They are less able to predict the impact of smart phones or creative financing for solar panels.

It should be clear that **scenario uncertainty** (the red line in Fig. 10.2) grows over time and the inputs to a climate model (emissions of gases that affect climate) become highly uncertain. This occurs for timescales longer than a human genera- tion (20–30 years), which also corresponds to the depreciation lifetime of many fixed investments (like a power plant). The major drivers of these uncertainties are slow-accumulating things, like global population, or commonly used energy sys- tems. They are usually subject to the results of different societal decisions: popu- lation growth in China, gas extraction technologies in the United States, economic growth in developing countries (Brazil, India, China), and disruptive technologies (internet, global manufacturing, and so on).

One way to look at this is to look at past projections of the present. Figure 10.6 is a reprint from a 2014 energy policy report from the Office of the President of the United States. The report is political, but the data presented in the figure are simply the consumption of gasoline for vehicles in the United States. In 2006, it was projected that in 2013 the United States would use 10 million barrels (550 million gallons, or about 2 billion liters) of gasoline per day. The projection for 2030 was 12 million barrels per day. But the actual value for 2013 was closer to 8 million barrels per day, and the 2030 forecast estimated in 2014 is only 6 million barrels per day. These are big numbers, with big implications. If we wanted to be 40 % below
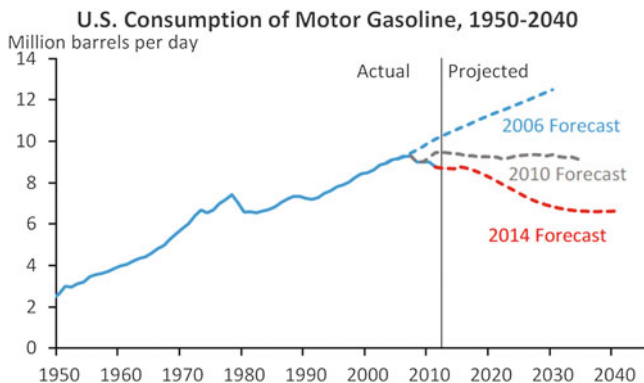
Fig. 10.6 Changing forecasts. Historical and projected U.S. consumption of gasoline in millions of barrels per day. Different forecasts started from different years are shown with the *dashed lines*. *Source* Executive Office of the President, 2014, "The All of the Above Energy Strategy as a Path to Sustainable Economic Growth," Fig. 2.2b. Data from Energy Information Agency

2006 consumption in 2030, that would take extreme measures for the 2006 forecast, but that would be the expected scenario in 2014 (without doing anything). So scenarios are forecasts as well, and they are often wrong.

Projections of economic and social indicators span a range of predictability. It is probably pretty easy to estimate population now and for the next few years. But economic growth or prices of commodities? Or use of gasoline? Economic growth estimates are rarely even known until well after the fact (the U.S. government revises all growth figures a few months after they are issued). These errors all compound to become part of the uncertainty in the input scenarios used to drive climate models.

So how do we deal with these scenario uncertainties in climate projections? These are the uncertainties in the forcing of climate. In the absence of making a prediction about what society might do, generally social scientists and economists estimate what is possible, and a range of possibilities (scenarios) are derived. The inputs to climate models are typically outputs of integrated assessment models, driven more by a story translated into quantitative inputs. These scenarios have evolved in complexity over time, but the values and spread have not changed that much. But they are not forecasts, because they depend on human policy decisions and choices (like the CFC-12 curve in Fig. 10.5).

Most climate models now use commonly developed scenarios.[8] Figure 10.7 illustrates the latest set of scenarios (**Representative Concentration Pathways**, or RCPs) used to force most climate models.[9] Figure 10.7 illustrates the scenarios in

---

[8]For a background on scenario development, consult Nakicenovic, N., & Swart, R., eds. (2000). *Special Report on Emissions Scenarios*. Cambridge, UK: Cambridge University Press. See also https://www.ipcc.ch/pdf/special-reports/spm/sres-en.pdf.

[9]A full description of the RCP scenarios is at Van Vuuren, D. P., et al. (2011). "The Representative Concentration Pathways: An Overview." *Climatic Change, 109*: 5–31.
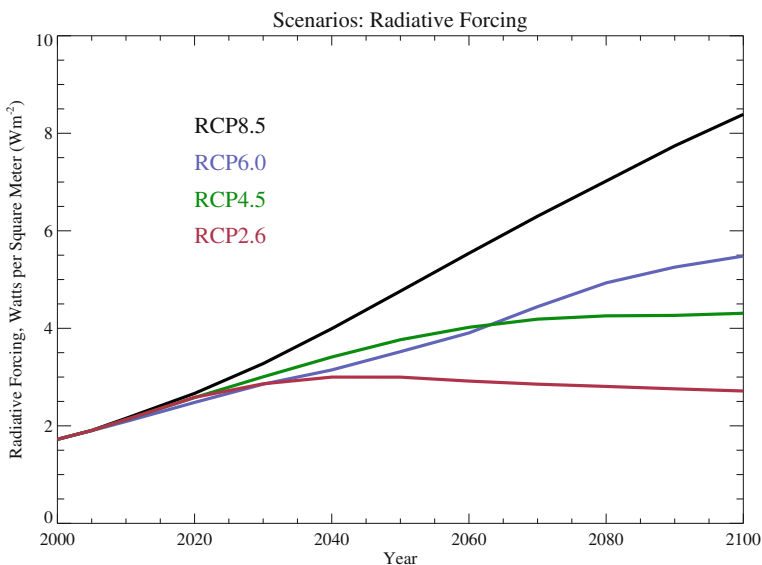
Fig. 10.7 Radiative forcing (in Watts per square meter) over time from 4 representative concentration pathways (*RCPs*) illustrating the most common scenarios used to force climate models. RCP8.5 (*black*), RCP6.0 (*blue*), RCP4.5 (*green*), RCP2.6 (*red*)

terms of the radiative forcing of climate in watts per square meter (Wm^{-2}) of the earth's surface. Recall that a watt is the same as the watt in lightbulbs: 60 Wm^{-2} is the energy of a typical incandescent bulb, and the solar input is about 160 Wm^{-2}.[10]

Each RCP scenario was designed to reach a specific radiative forcing target from greenhouse gases by a particular date, based on different predictions of what society might do. The levels of CO_2 and other greenhouse gases in the atmosphere that create the forcing are derived for each RCP (Fig. 10.8). CO_2 concentrations are in parts per million (ppm). One part per million means one molecule of CO_2 for every million molecules of air. The current concentration of CO_2 is about 400 ppm, and the level in 1850 before most industry developed was 280 ppm CO_2. That is a 40 % increase. Note that these scenarios are possible futures, just like the different forecasts in Fig. 10.6. We have been on the black curve (high emissions), but Fig. 10.6, and others like it, suggest that we might be transitioning to one of the more moderate curves.

Using common scenarios implicitly says that the scenarios are a major driver of uncertainty, and we need to compare models with common scenarios. Otherwise, the uncertainty due to the scenarios would be hard to assess.

As we shall see, a central theme of future climate prediction is to realize that most of the uncertainty lies not in the models themselves, but in the scenario for emissions (i.e., in the scenario uncertainty). The climate of the future is dominated

---

[10]Trenberth et al., "Earth's Global Energy Budget.".
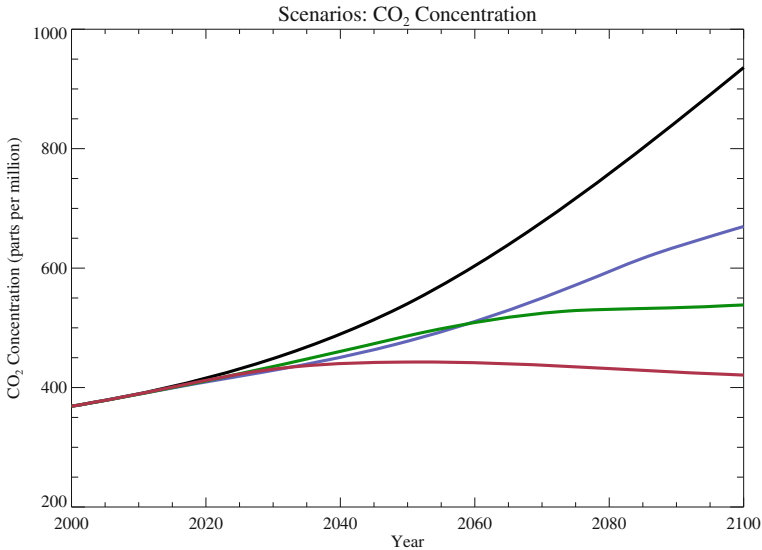
Fig. 10.8 CO_2 concentration scenarios. Projections of atmospheric CO_2 concentration (parts per million) over the 21st century from different representative concentration pathways (*RCPs*) used for climate model scenarios. RCP8.5 (*black*), RCP6.0 (*blue*), RCP4.5 (*green*), RCP2.6 (*red*)

by scenario uncertainty (the red line in Fig. 10.2), and that uncertainty is from human systems. In the end, climate is our choice, even if it is simply to make no choice about our emissions and continue present policies.

10.2.3   Predicting the Long Term: Model Uncertainty Versus Scenario Uncertainty

Thus far we have discussed sources of error in climate projections for the far future (beyond several decades) and in the near term (within a decade or two). Throughout this entire period there is a constant source of uncertainty, and this is what we typically think of as the uncertainty in a climate model: the model uncertainty, or the structural uncertainty in how we represent climate (the blue line in Fig. 10.2).

Model uncertainty is our inability to represent perfectly the coupled climate system. It is present in representing each of the components, and each of the processes within the system. The uncertainty has many different forms, and different values. Some uncertainties are large and some are small, mostly related to the physics of the processes. We have discussed a lot of these physical and process uncertainties in detail. Many uncertainties stem from the scales that models must resolve: Small-scale processes that are highly variable below the grid scale of a model are difficult to resolve or approximate. In the atmosphere, clear skies are easier to understand and represent than cloudy skies. In the land surface, variations in soil moisture and vegetation at the small scale interact with moisture and energy fluxes to make things complex.

Complexity arises due to nonlinear effects. In a nonlinear process, the average of results is not the same as if the inputs are averaged. As a simple example, let's say that the rate of evaporation (R) is the square of the soil moisture ($R = S^2$), where the moisture (S) varies from $0 \rightarrow 10$. If there are two equal-sized areas of a grid box, one with soil moisture of $S = 0$ and one of $S = 10$, then the average soil moisture $S = 5$. But calculating the evaporation from the average yields $R = 5^2 = 25$, whereas calculating the rate from each part and averaging (0 and $10^2 = 100$, divided by 2) yields 50: double the rate.

Thus there are many sources of model uncertainty. In the present context of trying to understand the sources of uncertainty, note that model uncertainty is constant over time: It does not increase. Because in the future there are better large-scale constraints than small-scale constraints, uncertainty in the future might even decrease.

Model uncertainty can be broken into parametric and structural uncertainty.[11] **Parametric uncertainties** are those that arise because many processes important for weather and climate modeling cannot be completely represented at the grid scales of the models. Therefore, these processes are parameterized, and there are uncertainties in these representations. An iconic parameterized process is that associated with convective clouds that are responsible for thunderstorms. Thunderstorms occur on much smaller spatial scales than can be resolved by global climate models. Therefore, the net effect of thunderstorms on the smallest spatial scales that the climate model can represent is related to a set of parameters based on physical principles and statistical relationships. Recall that the spatial scale that is resolved by a climate model is several times larger than the spatial size of a single grid point. The conflation of errors of different spatial scales, dependent on parameterizations, is difficult to quantify and disentangle.

Structural uncertainties are those based on decisions of the model builder about how to couple model parameterizations together and how they interact. One class of these errors is the always-present numerical errors of diffusion and dispersion in numerical transport schemes; that is, the errors of discrete mathematics. But a more important structural uncertainty perhaps is how the different parameterizations are put together. For example, let's say an atmosphere model has convection, cloud microphysics, surface flux, and radiation parameterizations. Typically there is a choice in how to couple parameterizations. Either all the parameterizations are calculated in parallel from the same initial state, called **process splitting**. Or each one could be calculated one after the other in sequence, called **time splitting**. Typically in a model with time splitting, the most "important" processes are estimated last. For weather models, this is usually cloud microphysics to get precipitation correct. For climate models, this is usually radiation. In either case, the last process is the most consistent with the sequential state. Either method, however, may create a structural uncertainty: If clouds and radiation are calculated

---

[11]Tebaldi, C., & Knutti, R. (2007). "The Use of the Multi-Model Ensemble in Probabilistic Climate Projections." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365*(1857): 2053–2075.

separately in a process split model, then the radiation may not reflect the clouds, and the temperature change from heating or cooling may not add up correctly. These are structural uncertainties from coupling processes. The same uncertainties come from coupling the physical processes to the dynamics in the atmosphere or ocean, and from coupling different components.

The result is an uncertainty graph that looks like Fig. 10.2. Internal variability, which is defined from the initial conditions (green line), is the largest contribution to short-term variability (the next tropical cyclone or blocking event or ENSO does not care much about the level of greenhouse gases), with a significant component of model variability (blue line). Over time (a few cycles of the different modes of variability), this fades, and the uncertainty is dominated by structural uncertainty in the model (blue line). But over long time periods, where the input forcings are themselves uncertain, the scenario variability (red) dominates. The latter break point is approximately where the level of uncertainty in the forcing (i.e., uncertainties in emissions translated into concentrations of greenhouse gases, translated into radiative forcing in Wm^{-2}) is larger than the uncertainty in the processes (expressed in radiative terms, Wm^{-2}). For climate models, the largest uncertainty is in cloud feedbacks. The uncertainty in cloud feedbacks is about 0.5 Wm^{-2} per degree of surface temperature change. So far the planet has warmed about 1–2 °C so the uncertainty due to cloud forcing is about 1 Wm^{-2}. This uncertainty would indicate that scenario uncertainty will dominate model uncertainty when the uncertainty in future radiative forcing between scenarios reached about 1 Wm^{-2}. In Fig. 10.7, this would be about 2040–2050 and beyond. The practical significance of this situation will be demonstrated when climate model results are discussed in Chap. 11. But first a few words about using models together in ensembles.

10.3   Ensembles: Multiple Models and Simulations

Different strategies can be used to better quantify the different types of uncertainty in climate model projections. One of the best, which can be used for quantifying all three types of uncertainty, is to run **ensembles**, or multiple simulations. These simulations are typically done to gauge one kind of uncertainty, but they can be appropriately mixed and matched to understand all three types of uncertainty.

The use of ensembles is now also common in weather forecasting. Weather forecasting suffers from model uncertainty and especially initial condition uncertainty, but usually not scenario uncertainty as we have described it. However, sometimes external events such as fire smoke can significantly affect the weather. In weather forecasting, ensembles of a forecast model are typically run in parallel: multiple model runs at the same time. Then the spread of results of these parallel runs is analyzed. The simulations usually differ from each other by perturbing the initial conditions (slight variations in temperature, for example) and/or by altering the parameters in the model. Sometimes forecasters consult different models, creating another sort of ensemble. Altering initial conditions tests the initial condition

uncertainty, whereas altering the model or the model parameters within a model tests model uncertainty. Ensembles are one way to assign probabilities for forecasts: An 80 % chance of rain, for example, comes directly from an ensemble with 10 members, equally likely, where 8 of the 10 predict rain. Ensemble forecasting provides an estimate of uncertainty: If you run a forecast model once and get an answer (sunny or rainy), how do you know how good that answer is (what is the uncertainty)? If you change the inputs or the model slightly and you get the same answer every time, it is probably pretty robust. If the answer changes in half the ensemble "members," then maybe it is not so robust.

The same methods can be applied to climate forecasting,[12] to estimate model uncertainty and initial condition uncertainty. In addition, different scenarios can be used to test the scenario uncertainty.

To start, one way of testing initial condition uncertainty in a climate model projection is to perform several different simulations of the same model, with the same forcing (scenario) but different initial conditions. This is often done to assess the internal variability of the model. For example, start up a model in 1850, use observations of how the atmosphere, sun, and earth's surface have changed, and see what the results are for the present day. Because of lots of different modes of variability in the system, some lasting decades, it is important to get a complete sample of the possible states of these modes.

Model uncertainty can then be added either by perturbing the model or by using multiple models. Technically, each perturbation to the model parameters is a different model, albeit one that is quite close to the original. But using different models developed quasi-independently in different places by different people provides a pretty good spread of answers. This can show the range of different possible states given a set of input parameters and a particular scenario. One must be careful because many models are not independent: They share common components and common biases.

Finally, models (either the same or different, probably both) can be run with different forcing scenarios, and the results compared. This is commonly done for the future, but it is also effective to do for the past. A nice illustration of this technique (Fig. 10.9) comes from the 2007 Intergovernmental Panel on Climate Change (IPCC) report[13] (see Chap. 11), where a set of models was run for two historical scenarios, one including human emissions of greenhouse gases and particles, and one using only natural sources. The results of the two different ensembles of models are different, and only the one with human emissions matches temperature observations over the 20th century.

In Chap. 11, we show combinations of these ensembles. For example, several different simulations of a single model or multiple models with the same scenario

---

[12]Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., & Meehl, G. A. (2010). "Challenges in Combining Projections From Multiple Climate Models." *Journal of Climate, 23*(10): 2739–2758.

[13]Solomon, S., ed. (2007). *Climate Change 2007: The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC*, Vol. 4. Cambridge, UK: Cambridge University Press.
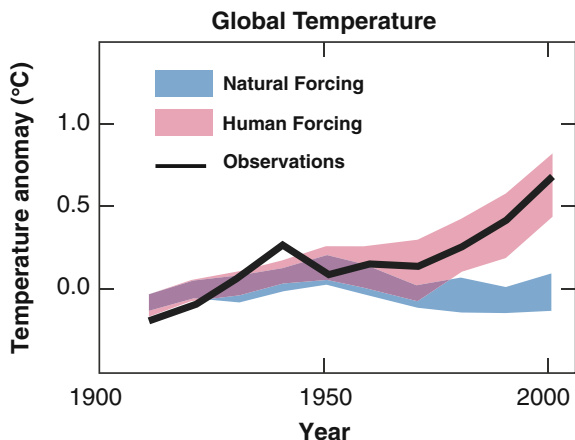
Fig. 10.9 Global average temperature anomalies from 1900 to 2000 based on climate model simulations. Only the models with natural and human emissions (*pink*) match the observations (*black*). Models with only natural emissions (*blue*) do not match observations. Adapted from Solomon et al. (2007), Summary for Policy Makers, Fig. 4. *Note* Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, eds. (2007). *Climate Change 2007—The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC (Climate Change 2007)*. Cambridge, UK: Cambridge University Press

can illustrate the different possible climate outcomes (i.e., the different simulated changes in global average surface temperature) with a single scenario. Ensembles can be built the same way with other scenarios and compared, so that the different contributions to initial condition uncertainty (variations within a single scenario and model but with different initial conditions), model uncertainty (variations within a single scenario using different models), and scenario uncertainty (either of the first two single-scenario ensembles with single or multiple models, but now for multiple scenarios) can all be assessed, and to some extent quantified.

There is much discussion about how to evaluate and weight models in a group: If a model resembles observations of the present, is it better and should it be given more weight[14]? If a model does not meet certain tests or standards, or has known problems, should it not be analyzed? In general, one problem with weighting models is that present performance on one metric in the past is not necessarily a good indicator of skill for the future. It is the same as the statement at the bottom of reports on the history of a financial investment: "past performance may not be an indicator of future returns." Models with fundamental flaws are sometimes removed from analyses or given less weight. However, many times these "flaws" are designed to constrain the models in some way, and are often justifiable. There is still much debate in the scientific community over how exactly to construct weighted averages across ensembles with multiple models.

---

[14]Knutti, Reto. (2010). "The End of Model Democracy?." *Climatic Change, 102*(34): 395–404.

But broadly, there is the interesting property that the wisdom of crowds applies. If the models have random mistakes and their biases are uncorrelated, then the average of the models (the multi-model mean) tends to be better than most of the models. This is a big "if," and many models have dependencies, but in practice (and in synthetic statistical tests), because the climate system is complicated, and most models meet basic measures of representing the system (conservation of energy and mass), the multi-model mean is usually a pretty good statistic. This makes ensembles quite valuable.

One thing to remember is that there is lots of chaotic internal variability in the climate system, and the observed record is only one possible realization of that internal variability. There may or may not be an El Niño event this year, but the possibility is that there could have been. So with the real climate system, we have only one ensemble. We expect models to be different from observations in any given year, or even decade, if they are fully coupled and internally consistent. This also confounds prediction and evaluation. It is one reason why the present and the past do not fully constrain the future (see box).

Why the Present and Past DO NOT Constrain the Future

It is often assumed that a model must be able to represent the present (or past climate) correctly to represent the future. This is true. However, while necessary, it is not a sufficient condition to constrain the future. Suppose that a model represents the present for the wrong reasons. Perhaps there is a large compensation of errors: Maybe an error in the radiation code letting in too much energy is compensated for by an error in the sea-ice model that reflects too much energy. If the planet warms up, the sea ice will go away, but the error in the radiation code will remain. It is also assumed that the more accurate a model is at representing the present (or the past), the better it will be at representing the future. This is also true, but it requires that "better" representations be for those areas that matter for future climate, and this is not necessarily the case.

What does that mean in practice? Let's say you have two measures of model performance. One is based on clouds, and one is based on temperature. If model A scores 80 % on both temperature and clouds relative to observations, and model B scores 100 % on clouds but 20 % on temperature, then a simple average score says that model A scores 80 % and model B scores 60 %. But if present clouds are more important for future climate than temperature, model B may be better. In the previous example, the present-day model may score well on sea ice and lower on the radiation code, but the radiation code is more important for the future. Also, knowing the mean state today does not imply that you can predict the change in that state accurately in the future. It is usually a good indicator, but not guaranteed.

That covers why the present does not necessarily constrain the future. Because with present climate we do not know what the response to a forcing (i.e., feedbacks) will be. But what about using the past: If we know

observations of the past state, can't we use them, and run models with them, to help us understand the future? Yes, we can, and we do. A good test of a climate model is to run it for situations representing paleoclimates like the last glacial maximum (peak of the last ice age), or the change from the last glacial period to today. This does give some hint as to the potential changes in the system, and whether a model can represent them.

But it does not tell the whole story, and there is currently quite a bit of confusion in the community over the value of the past. Many times, observations and models are examined to look at changes to climate and to estimate the response. If the forcing is known, then this also tells us something about the feedbacks. That is, if we know CO_2 changed a certain amount between the last ice age and today (180–280 ppm), and we know what the temperature change is −15 °F (−8 °C) in the ice core record; see Fig. 3.2), can't we just back out a relationship for temperature change for the next 100 ppm of CO_2? Or more appropriately, can we calculate the radiative effect of that extra CO_2 and we then get the sensitivity of the climate system (°C per Wm^{-2})?

The problem is twofold. First, the causes and effects are different. CO_2 changes were likely a response to changes in the earth's orbit that ended the ice age. The forcing started with changes to the total absorbed radiation from the sun, and CO_2 responded to changes in the ocean circulation. Second, the sensitivity (and the feedbacks) are dependent on the state of the climate system. There was a lot more ice and snow back then and less water vapor in a colder atmosphere, so the ice albedo feedback and the water vapor feedback were likely much different than today. The water vapor feedback was likely smaller, but the albedo feedback larger. It is also worth noting that the processes and pathways are different: Cloud feedbacks, for example, may respond differently to a change in solar radiation (which primarily heats the surface) than to a change in infrared (longwave or terrestrial) radiation, which heats the atmosphere as well as the surface.

10.4   Applications: Developing and Using Scenarios

The development of scenarios is a good example of the interactions between uncertainty and the importance of understanding how models are constructed for appropriately using them. For application purposes, Representative Concentration Pathways (RCPs) were developed by using integrated assessment models. RCPs were produced by making assumptions about the desired radiative forcing at the end of the 21st century, and different integrated assessment models were used with different input assumptions. These assumptions include population and economic growth, as well as technology improvement. The implication is that each RCP with

different emissions implies a different future forecast for society. RCPs with high forcing, like RCP8.5, have large emissions of greenhouse gases, but also assume a large and wealthy global population that would produce these emissions. This means that if you are using an RCP, it should be consistent with the assumptions for the application about the future of the anthroposphere. Many critics argue that the RCP8.5 is impossible to reach: The emissions are too high given what we expect about population growth.

With RCPs, the information about the societal pathway consistent with emissions is not very clear. In recognition, new scenarios are being produced using a different methodology, called **Shared Socioeconomic Pathways** (SSPs).[15] SSPs were developed to have a consistent story. They start with assumptions about the anthroposphere: population, economic growth, technology, and efficiency improvement. These assumptions are more widely available and published, and they are used in several different integrated assessment models to generate the emissions consistent with those pathways.

What this means for applications is that the appropriate SSP (or RCP) output from a set of climate models to use depends on your assumptions about the future state of humanity, and that for specific applications, not all the scenarios should be treated equally, and the application needs to be consistent with the scenario. For example, if the aim is to look at impacts on forests, then an RCP with high emissions, which might assume that many forests have been cut down, is not appropriate. The climate that results from the RCP (with fewer forests specified) may not be appropriate for looking at climate impacts on forestry. Since climate changes in currently forested regions may assume that they become another ecosystem type (like cropland), and assuming in the future that these locations still represent forests is wrong. The same caution exists for the SSPs, but these newer scenarios try to make explicit the assumptions about the assumed evolution of the anthroposphere.

10.5 Summary

Projections in climate models need to have uncertainty attached to them. Uncertainty can be usefully broken up into three categories: initial condition uncertainty (also called internal variability), model uncertainty (also called structural uncertainty), and scenario uncertainty. On short timescales, from days to decades, the initial conditions or internal state of the system matter. Aspects of the climate system are predictable on different timescales, out to a decade (when the deep ocean circulation is involved). Model uncertainty affects all timescales and is not that dependent on timescale. Model uncertainty can be divided into uncertainties in parameters

[15]O'Neill, B. C., et al. (2014). "A New Scenario Framework for Climate Change Research: The Concept of Shared Socioeconomic Pathways." *Climatic Change, 122*(3): 387–400.

(parametric uncertainty) and larger questions of how models are constructed and what fields they contain (structural uncertainty). On long timescales, longer than the internal modes of variability in the climate system (one to two decades), scenario uncertainty dominates climate projections. Scenarios are an exercise in predicting and making our own human future. Scenarios are often produced with models that have their own uncertainty, and scenarios have changed over time as human society changes. The history of the climate system is one way to look at a scenario, but this may not be a complete representation of the climate system.

One way of teasing out the different uncertainties using climate models is to run multiple simulations called ensembles using combinations of different initial conditions, different models, and different scenarios. We explore these results from the latest generation of models in Chap. 11.

Key Points

- Models have different sources of uncertainty, which are important at different timescales and for different problems.
- Scenarios are highly uncertain, and the future of human systems can change rapidly.
- Multiple models and multiple simulations, called ensembles, can be used to sample uncertainty.