Projecting Future Climates

In order to consider the biological impacts of climatic change we need to know what these climatic changes will look like. In Chapter 1 we learned how we know about past climate, but how do we know what our future climate will be? Future climates are projected using mathematical models of our climate system, based on atmospheric physics. Given the mixture of greenhouse gases in the atmosphere at any given point in time, these models can deduce what the climate will be like. It is important to understand that these models do not predict the levels of greenhouse gases. Greenhouse gas concentrations are climate model inputs, not model outputs. So where do the greenhouse gas concentrations come from? They come from another type of mathematical model called a ‘scenario model’. This two-step process, using first scenario models and then climate models, is the focus of this chapter. Having some understanding of this process is important if we as biologists are to make intelligent choices of climate projections for use in our studies of climatic impacts.

Throughout the remainder of this book we will, from time to time but particularly when we discuss future climatic change, refer to the ‘IPCC’. This is the Intergovernmental Panel on Climate Change (www.ipcc.ch). The IPCC was established by the United Nations Environment Programme (UNEP) and the World Meteorological Organization (WMO). According to their website:

The IPCC is a scientific body. It reviews and assesses the most recent scientific, technical and socio-economic information produced worldwide relevant to the understanding of climate change. It does not conduct any research nor does it monitor climate related data or parameters. Thousands of scientists from all over the world contribute to the work of the IPCC on a voluntary basis. Review is an essential part of the IPCC process, to ensure an objective and complete assessment of current information. Differing viewpoints existing within the scientific community are reflected in the IPCC reports.

2.1 What are Scenarios?

Emissions scenarios are stories about the future. They are not predictions; they are descriptions of plausible alternative futures. To quote the IPCC’s Special Report on Emissions Scenarios (SRES):

Future levels of global GHG [greenhouse gas] emissions are the products of a very complex, ill-understood dynamic system, driven by forces such as population growth, socio-economic development, and technological progress; thus to predict emissions accurately is virtually impossible.

Predicting future greenhouse gas emissions with any degree of accuracy would require us to know things that we cannot possibly know, like the pace of technological change, the rate of economic development, and the political and cultural will to abate greenhouse gas emissions.

Equally, the scenarios used by the IPCC are also not statements of value. There are no ‘preferred’ or ‘favourite’ scenarios. This position may seem paradoxical. If anthropogenic greenhouse gas emissions are the cause of global warming, and global warming is undesirable, then why do we not prefer emissions scenarios that result in lower future greenhouse gas concentrations? The reason is that growth in greenhouse gas emissions is related to things like the pace of economic development, and economic development has moral justification, because it
meets the needs of a growing population. That is, there is human and moral value in at least some of the activities that lead to higher greenhouse gas emissions. To some extent it is a matter of taste as to whether we prefer improved human development even if it comes with a cost to the environment. In any case, the IPCC is not in the business of telling governments what they should do. The IPCC's mandate involves the science around climatic change, not policy making.

So what are 'scenarios'? Figure 2.1 shows how scenario construction involves making logical connections between known drivers of greenhouse gas emissions and the emissions themselves. Why do we need this? Surely it would be possible to simply run the climate models under a whole range of greenhouse gas emissions and then draw conclusions about the consequences of rising greenhouse gas emissions, without ever considering where those emissions come from. Of course this is possible, but what scenario construction does is to provide tools for policy makers to think about how changes in policy might alter future emissions. By attempting to elaborate on the connection between, say, the mixture of fuels used in the future and greenhouse gas emissions, scenarios provide the policy makers with a tool to think about how their energy policy will impact future climatic change.

Scenarios involve two parts, qualitative storylines and quantitative interpretations of those storylines. They use historical and contemporary information to generate some assumptions about the relationships between driving forces and greenhouse gas emissions. These assumptions are reproducible and internally consistent. Sometimes these assumptions are qualitative, sometimes they are quantitative and sometimes there are explicit mathematical models behind each of them. The qualitative and quantitative assumptions are fed into scenario models that then generate the time course of emissions over the next 100+ years. Figure 2.2 shows the annual emissions for one particular scenario as an example.

Scenario construction starts with something called a 'storyline'. Storylines are narratives that

![Fig. 2.1. Schematic diagram indicating how climate change scenarios are constructed and their role in projecting future climate conditions. An example of the greenhouse gas emissions projections produced from the scenario construction is seen in Fig. 2.2.](image-url)
describe broad pictures of how the future might unfold. The storylines differ in the way that each conceives of how the major geopolitical regions of the world interrelate, how new technologies spread, how different economic regions develop, how environmental protection is conceived and implemented, and how population growth changes globally and regionally. These storylines contain no explicit policies to limit greenhouse gas emissions or to adapt to climatic change.

The IPCC settled on four storylines. These storylines are to be thought of as equally plausible views of the future. There are no ‘business as usual’, ‘best case’ or ‘worst case’ scenarios. In order to avoid any value implications, the IPCC decided to give the storylines neutral names: A1, A2, B1 and B2. The differences between the four storylines are summarized in Table 2.1.

Within each storyline there is a ‘family’ of scenarios. Storylines are qualitative views of the
future, while the scenarios can be thought of as \textit{quantitative} interpretations of the each storyline. The scenarios themselves are developed by interpreting the qualitative storylines through a scenario modelling framework. There are six frameworks used by the IPCC:

- Asian Pacific Integrated Model (AIM) from the National Institute of Environmental Studies in Japan;
- Atmospheric Stabilization Framework Model (ASF) from ICF Consulting in the USA;
- Integrated Model to Assess the Greenhouse Effect (IMAGE) from the National Institute for Public Health and Environmental Hygiene (RIVM), used in connection with the Dutch Bureau for Economic Policy Analysis (CPB) WorldScan model, the Netherlands;
- Multiregional Approach for Resource and Industry Allocation (MARIAS) from the Science University of Tokyo in Japan;
- Model for Energy Supply Strategy Alternatives and their General Environmental Impact (MESSAGE) from the International Institute of Applied Systems Analysis (IIASA) in Austria; and
- Mini Climate Assessment Model (MinCAM) from the Pacific Northwest National Laboratory (PNNL) in the USA.

The same storyline is run through more than one model framework. This exercise then results in different quantitative descriptions of each of the qualitative stories, and hence in different quantitative projections of greenhouse gas emissions. Figure 2.3 shows the differences between the model interpretations of each of the storylines for total anthropogenic emissions of CO₂. The error bars show the range of ±1 standard deviation in the distribution of results. So for example, the five model interpretations of the A2 storyline result in total CO₂ emissions in 2100 ranging between 28.19 GtC (MESSAGE) and 34.47 GtC (AIM). A1C and A1G (not shown in Fig. 2.3) are usually combined to form a scenario called A1FI (FI denotes \textit{fuel intensive}) and A1FI is often referred to as the ‘high emissions scenario’, for obvious reasons. A2 is often referred to as the ‘medium high emissions scenario’, B2 as the ‘medium low emissions scenario’ and B1 as the ‘low emissions scenario’.

### 2.2 From Emissions to Climate Projections: General Circulation Models

Formally, future climates are projected using a type of mathematical model called ‘Atmosphere–Ocean General Circulation Models’ or AOGCMs. It is still common to see them referred to simply as ‘General Circulation Models’ or GCMs. GCM is also used to mean ‘Global Circulation Model’ and ‘Global Climate Model’, but for our purposes these terms are all synonymous. Technically, the term GCM harkens back to a period when the ocean and atmosphere portions of climate models were either fully or partially ‘uncoupled’. Today, GCM is used synonymously with AOGCM.

GCMs are computer models that represent the world by a large number of grid cells, typically representing a horizontal scale of between 1° and 5° latitude by longitude. These grid cells cover the
The total anthropogenic CO$_2$ emissions. The ±1 standard deviation error bars denote the variability in how the various scenario modelling frameworks interpret the various storylines.

Fig. 2.3.

The entire surface of land, water and ice. The grid cells are stacked on top of each other so that processes are modelled in layers extending from the top of the atmosphere to the bottom of the oceans. Most models used in the IPCC assessments include from 20 to 30 layers. What happens in one grid cell affects the conditions in surrounding grid cells. The ‘atmosphere’ comprises information about the density, in terms of temperature and pressure, the motion of the air mass up and down and horizontally, and the composition of the air in terms of its concentrations of greenhouse gases (e.g. CO$_2$ and water vapour). There are exchanges of heat and water between the atmosphere and the land and oceans (Houghton, 2004).

GCMs are based on well-established principles of physics, and are able to reproduce recently observed features of our current climate and can reproduce past climate. And as such, ‘There is considerable confidence that Atmosphere-Ocean General Circulation Models (AOGCMs) provide credible quantitative estimates of future climate change, particularly at continental and larger scales’ (Randall et al., 2007: Executive Summary). Such spatial scales are, of course, problematic for biologists who rarely work at such large scales. There is some relief from this limitation in the form of statistical downscaling and regional climate models, considered later in this chapter.

There are currently 18 modelling centres around the world, and 25 GCMs were used in the IPCC’s Fourth Assessment Report (AR4). These modelling centres and model names are shown in Table 2.2.

We might reasonably ask: ‘Why use so many models?’ The fact is that climatologists know more about the mechanisms controlling climate than can be represented in a mathematical model and still have the model run, in a reasonable length of time, on some of the world’s fastest computers. As a consequence, climate modellers must make trade-offs between speed and process representation. Some processes are represented mechanistically, while some are represented phenomenologically. To translate this into terms that biologists can understand, suppose we were modelling plant growth. In general, it will be true that if we add more mineral nitrogen to the soil, the plant will grow more. We could represent this mechanistically by modelling N-uptake by the roots, N-transport, N-use, N-recycling and so on. This representation would take many mathematical equations and demand a considerable amount of
Table 2.2. Modelling centres and model names for GCMs used in the IPCC's Fourth Assessment Report.

<table>
<thead>
<tr>
<th>Originating group(s)</th>
<th>Country</th>
<th>Names of models</th>
<th>Further information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing Climate Center</td>
<td>China</td>
<td>BCC-CM1</td>
<td><a href="http://tinyurl.com/ch3-CSMD">http://tinyurl.com/ch3-CSMD</a></td>
</tr>
<tr>
<td>Bjerknes Centre for Climate Research</td>
<td>Norway</td>
<td>BCCR-BCCM2.0</td>
<td><a href="http://tinyurl.com/ch3-BCCR">http://tinyurl.com/ch3-BCCR</a></td>
</tr>
<tr>
<td>National Center for Atmospheric Research</td>
<td>USA</td>
<td>CCSM3</td>
<td><a href="http://tinyurl.com/ch3-CCSM">http://tinyurl.com/ch3-CCSM</a></td>
</tr>
<tr>
<td>Canadian Centre for Climate Modelling and Analysis</td>
<td>Canada</td>
<td>CGCM3.1(T47), CGCM3.1(T63)</td>
<td><a href="http://tinyurl.com/ch3-CGCM">http://tinyurl.com/ch3-CGCM</a></td>
</tr>
<tr>
<td>Météo-France/Centre National de Recherches Météorologiques</td>
<td>France</td>
<td>CNRM-CM3</td>
<td><a href="http://tinyurl.com/ch3-CNRM">http://tinyurl.com/ch3-CNRM</a></td>
</tr>
<tr>
<td>CSIRO Atmospheric Research</td>
<td>Australia</td>
<td>CSIRO-Mk3.0, CSIRO-Mk3.5</td>
<td><a href="http://tinyurl.com/ch3-CSIRO">http://tinyurl.com/ch3-CSIRO</a></td>
</tr>
<tr>
<td>Max Planck Institute for Meteorology</td>
<td>Germany</td>
<td>ECHAM5/MPI-OM</td>
<td><a href="http://tinyurl.com/ch3-echam5">http://tinyurl.com/ch3-echam5</a></td>
</tr>
<tr>
<td>Geological Institute of the University of Bonn, Geological Institute of KMA, and the Data Group</td>
<td>Korea</td>
<td>ECHO-G</td>
<td><a href="http://tinyurl.com/ch3-echo-G">http://tinyurl.com/ch3-echo-G</a></td>
</tr>
<tr>
<td>LASG/Institute of Atmospheric Physics</td>
<td>China</td>
<td>FGOALS-g1.0</td>
<td><a href="http://tinyurl.com/ch3-fgoals">http://tinyurl.com/ch3-fgoals</a></td>
</tr>
<tr>
<td>US Department of Commerce/National Oceanic and Atmospheric Administration (NOAA)/Geophysical Fluid Dynamics Laboratory</td>
<td>USA</td>
<td>GFDL-CM2.0, GFDL-CM2.1</td>
<td><a href="http://tinyurl.com/ch3-gfdl">http://tinyurl.com/ch3-gfdl</a></td>
</tr>
<tr>
<td>NASA/Goddard Institute for Space Studies</td>
<td>USA</td>
<td>GISS-AOM, GISS-EH, GISS-ER</td>
<td><a href="http://tinyurl.com/ch3-giss">http://tinyurl.com/ch3-giss</a></td>
</tr>
<tr>
<td>(Note: the FORTRAN code for these models can be downloaded from this site)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Istituto Nazionale di Geofisica e Vulcanologia</td>
<td>Italy</td>
<td>INGV-SXG</td>
<td><a href="http://tinyurl.com/ch3-ingv-sxg">http://tinyurl.com/ch3-ingv-sxg</a></td>
</tr>
<tr>
<td>Institute for Numerical Mathematics</td>
<td>Russia</td>
<td>INM-CM3.0</td>
<td><a href="http://tinyurl.com/ch3-inm-cm3">http://tinyurl.com/ch3-inm-cm3</a></td>
</tr>
<tr>
<td>Institut Pierre Simon Laplace</td>
<td>France</td>
<td>IPSL-CM4</td>
<td><a href="http://tinyurl.com/ch3-ipsl-cm4">http://tinyurl.com/ch3-ipsl-cm4</a></td>
</tr>
<tr>
<td>Center for Climate System Research (The University of Tokyo), National Institute for Environmental Studies, and Frontier Research Center for Global Change (JAMSTEC)</td>
<td>Japan</td>
<td>MIROC3.2 (hires), MIROC3.2 (medres)</td>
<td><a href="http://tinyurl.com/ch3-miroc">http://tinyurl.com/ch3-miroc</a></td>
</tr>
<tr>
<td>Meteorological Research Institute</td>
<td>Japan</td>
<td>MRI-CGCM2.3.2</td>
<td><a href="http://tinyurl.com/ch3-mri-cgcm2">http://tinyurl.com/ch3-mri-cgcm2</a></td>
</tr>
<tr>
<td>National Center for Atmospheric Research</td>
<td>USA</td>
<td>PCM</td>
<td><a href="http://tinyurl.com/ch3-NCAR-pcm">http://tinyurl.com/ch3-NCAR-pcm</a></td>
</tr>
<tr>
<td>Hadley Centre for Climate Prediction and Research/Met Office</td>
<td>UK</td>
<td>UKMO-HadCM3, UKMO-HadGEM1</td>
<td><a href="http://tinyurl.com/ch3-HadCM3">http://tinyurl.com/ch3-HadCM3</a></td>
</tr>
</tbody>
</table>

computing power. On the other hand, we could forgo a mechanistic representation and simply capture the 'phenomenon' of higher soil nitrogen concentrations and greater plant growth with a single, in this case statistical, function. Such a phenomenological representation would demand far less computing time but it would sacrifice a mechanistic understanding.

We would have to consider our goals carefully before choosing one or the other approach. Climate modelers make many different decisions about which processes to represent mechanistically and which to represent phenomenologically. In other words, modelers decide which abstractions to make and which not to make.
Another trade-off that climate modellers must consider is computing time and spatial and temporal resolution. Early climate models from the 1970s had very little spatial resolution. They represented the globe with a lattice of grid cells that were 600 km x 600 km, and these grid cells would be stacked six cells high to represent the atmosphere (Pope, 2007). To give a sense of scale, a grid cell this size is about the same size as Germany (357,114 km²).

By comparison, some of the most modern climate models have a spatial resolution of 133 km x 133 km, stacked 38 cells high to represent the atmosphere (see Fig. 2.4 for more details). These models take a minimum of 256 times the computing power required for the early 1970s models (Pope, 2007). Why does the spatial and temporal resolution matter? Bear in mind that space and time are actually continuous, but they must be represented as discrete units in order to be modelled in a computer. The smaller the discrete units, the better the approximation is to the continuous features of the model (space and time).

With that as background, we can now return to the question of 'why so many models?'. Since climate modellers have to make these various abstractions and compromises, there would always be the fear that the resulting projections are highly dependent on the particulars of the model. So the philosophy is to have many groups, all working essentially independently, making different abstractions and different compromises and using different assumptions. We then take the projections from all of these groups and models and compare them. In this comparison we look for 'robust' projections. A projection is robust if it does not depend strongly on the particular climate model. For example, one robust projection is that the global mean annual temperature will increase. All of these models, despite their differences in assumptions, abstractions and compromises, project an increase in mean annual global temperature. None of the models projects cooling. On the other hand, the particular degree of warming is less robust. Figure 2.5 shows that the particular degree of warming is still very much dependent on the assumptions, abstractions and compromises of each model.

We take up the issue of robust projections in greater detail in Box 2.1. Besides such robust solutions, there are also reasons to believe that averaging over the various differences in the GCMs yields the best model projections. The recent US Climate Change Science Program Synthesis and Assessment Product 3.1 (Bader et al., 2008) concluded:

The CMIP3 (Coupled Model Intercomparison Project 3) 'ensemble-mean' model performs better than any individual model by this metric and by many others. This kind of result has convinced the

Fig. 2.4. Schematic showing the spatial scale of climate models in the 1990s (the IPCC's Second Assessment Report) and present (the IPCC's Fourth Assessment Report). (Figure from the Hadley Centre. © British Crown Copyright 2010, the Met Office, redrawn by the Southwest Climate Change Network, www.southwestclimatechange.org.)
Fig. 2.5. Mean annual change in global surface temperatures predicted from climate models. The bars on the right-hand side show the range of projections for each scenario that are produced by all the different climate models. So we can see that a robust projection is that the mean annual global surface temperature will rise, but that the rise could be anywhere from 1 to 6.5°C depending on the emissions scenario and the particular climate model.

Community of the value of a multimodel approach to climate change projection. Our understanding of climate is still insufficient to justify proclaiming any one model ‘best’ or even showing metrics of model performance that imply skill in predicting the future.

The report goes on to point out that different models perform differently on a host of different metrics (i.e., methods of measuring how ‘good’ they are). No single model does better on all metrics, and without knowing which metric makes a model better or worse at projecting future climate, when forcing variables will be very different from what they are now, it is impossible to say which model is ‘best’.

2.3 Regional Models and the Problem of Downscaling

As we mentioned previously, even today GCMs simulate climate on fairly large spatial scales, so large that their ecological relevance may be questionable for many applications. Even the GCMs with the finest spatial resolution, such as the HadGEM1, use grid cells 135 km on a side. While considerably better than the early models that used grid cells up to 600 km on a side, even these finer-scale models are still quite coarse for our purposes. There are, however, two possible solutions to this problem: (i) the use of regional climate models (RCMs); and (ii) the use of statistical downscaling.
Box 2.1. The issue of consensus.

The media has a lot to say about the issue of ‘consensus’ among climate change ‘experts’, on both sides of the issue. Some argue that there is substantial consensus among experts that the Earth is warming and that we humans are to blame. Others argue that such ‘consensus’ is a fiction, that there really is considerable disagreement among experts, and, besides, science doesn’t work by consensus. Who is right? First, it is worth noting that the most authoritative body speaking about climate change, the IPCC, rarely uses the word ‘consensus’, and when it does it is not about such grand pronouncements but rather about small details. Instead, the IPCC talks about ‘levels of uncertainty’ in our understanding of different aspects of climate change.

There are four big questions when it comes to climate change and GCMs: (i) Has there been detectable warming in the recent past? (ii) Is this warming unusual against the background of historical climate change? (iii) Are humans to blame for this warming? (iv) How much warming will there be in the future if emissions continue to increase? There are many other important questions we might ask, but these are the really big ones. These four questions all require different sorts of evidence but, more importantly, we need to understand that science never provides definitive answers to any questions, let alone these. Science is not about ‘absolute truths’, it is about ‘very likely maybe’s’. All of science, and all applications of science, are the same in this regard — we don’t deal in absolutes. Non-trivial questions in science are ‘weight of evidence’ arguments. For some questions we can give answers for which we have very little uncertainty, and for others there may be a great deal of uncertainty. But notice that there are no answers that we can give with absolute certainty. To feel comfortable with scientific explanations of natural phenomena, one has to feel comfortable with uncertainty. It is a misunderstanding of how science works, by politicians and lay-people, which causes them to seek definitive answers from scientists. There are no definitive answers, there are only ‘very likely maybe’s’.

We expand on these points in Chapter 14.

We mentioned that the IPCC deals with levels of uncertainty, not with consensus and not with absolute definitive answers. This quote from the IPCC AR4 Synthesis Report (IPCC, 2007: 27) explains how it considers uncertainty:

Three different approaches are used to describe uncertainties each with a distinct form of language. Choices among and within these three approaches depend on both the nature of the information available and the authors’ expert judgment of the correctness and completeness of current scientific understanding.

Where uncertainty is assessed qualitatively, it is characterised by providing a relative sense of the amount and quality of evidence (that is, information from theory, observations or models indicating whether a belief or proposition is true or valid) and the degree of agreement (that is, the level of concurrence in the literature on a particular finding). This approach is used by WG III (Working Group III: Mitigation of Climate Change) through a series of self-explanatory terms such as: high agreement, medium evidence; high agreement, medium evidence; medium agreement, medium evidence; etc.

Where uncertainty is assessed more quantitatively using expert judgment of the correctness of underlying data, models or analyses, then the following scales of confidence levels is used to express the assessed chance of a finding being correct: very high confidence (at least 9 out of 10); high confidence (about 8 out of 10); medium confidence (about 5 out of 10); low confidence (about 2 out of 10); and very low confidence (less than 1 out of 10).

Where uncertainty in specific outcomes is assessed using expert judgment and statistical analysis of a body of evidence (e.g. observations or model results), then the following likelihood ranges are used to express the assessed probability of occurrence: virtually certain >99%; extremely likely >95%; very likely >90%; likely >66%; more likely than not >50%; about as likely as not 33% to 66%; unlikely <33%; very unlikely <10%; extremely unlikely <5%; exceptionally unlikely <1%.

WG II (Working Group II: Impacts, Adaptation and Vulnerability) has used a combination of confidence and likelihood assessments and WG I (Working Group I: The Physical Science Basis) has predominantly used likelihood assessments.

The problem with the term ‘consensus’ is that it implies an ‘all or nothing’ judgement — we either have consensus or we do not. But clearly there can be degrees of consensus. Better to just stop using the word at all, and stick to these operational definitions of the degrees of uncertainty.

So what does the IPCC actually say about any ‘consensus’ projection of climate change? Here is its statement on questions (i) and (ii), about whether there has been detectable warming in the recent past and whether this warming is unusual against a background of historical climate change:

Average Northern Hemisphere temperatures during the second half of the 20th century were very likely higher than during any other 50-year period in the last 500 years and likely the highest in at least the past 1300 years. (WG I.6.6, SPM) (IPCC, 2007: 30)

Its answer to question (iii), on whether humans are to blame, is:

There is very high confidence that the global average net effect of human activities since 1750 has been one of warming, with a radiative forcing of +1.6 ± 0.6 to +2.4 W m⁻². (WG I.2.3, 6.5, 2.9, SPM) (IPCC, 2007: 37)

And one way that it answers question (iv), regarding how much climate change will there be in the future, is:

Climate sensitivity is likely to be in the range of 2 to 4.5 °C with a best estimate of about 3°C, and is very unlikely to be less than 1.5°C. Values substantially higher than 4.5°C cannot be excluded, but agreement of models with observations is not as good for those values. (WG I.8.6, 9.6, Box 10.2, SPM) (IPCC, 2007: 38)

Climate sensitivity is the equilibrium change in surface temperature that results from a doubling of atmospheric CO₂ concentrations. More details on this measure are provided elsewhere in this chapter.
RCMs are dynamic models in the same way that GCMs are, but they cover only a specific region (often of land surface). They use either observed climate or that generated by a relevant GCM to provide the RCM with so-called ‘boundary conditions’. RCMs have considerably finer spatial scales, on the order of 50–100 km on a side, with some of the best having grid cells of just 20 km on a side. This is still large by ecological standards, but considerably better resolution than GCMs can provide. Like GCMs, RCMs are based on physical principles of the climate system and they are capable of reproducing contemporary climate for many regions of the globe. This ability gives us some confidence in their ability to downscale future climates, too.

The chief disadvantage of RCMs is that despite the fact that they cover a much smaller portion of the globe than GCMs, RCMs can be as computationally demanding or even more so than GCMs. Because of the finer spatial scale, a given region will have many more grid cells than the same region in a GCM. More grid cells mean more calculations; but also because of the finer spatial scale, a shorter time step is required for the integration in order to achieve numerical stability (Bader et al., 2008). Partly because of the computational costs, these models are rarely ever run for as long as GCMs. While GCMs are used to project climate for the next 100–200 years, RCMs are rarely used for periods longer than a few tens of years.

Statistical downscaling is much simpler. With statistical downscaling, we take contemporary relationships between observed large-scale climate features and observed smaller-scale climate features and we use these relationships to translate large-scale projections into smaller-scale projections. The principal disadvantage with statistical downscaling is that it assumes the relationships between large- and small-scale climate features remain constant through time. The principal advantage to statistical downscaling is that it is computationally cheap. Where comparisons have been made between statistical downscaling and RCMs, neither has been found to be consistently superior. Both have strengths and weaknesses and there is currently a place for both approaches (Bader et al., 2008).

A word of caution is probably useful at this point. There is a temptation to think of downscaled or regional climate projections, particularly those generated by RCMs, as being in some sense ‘more accurate’ than projections from GCMs. This temptation probably stems from the higher spatial resolution that tends to produce a ‘smoother’ picture of projected climate. The finer spatial scale allows more accurate representation of topography for example, and topography can influence local climate. Sometimes this confidence is justified. For example, there is reason to believe that RCMs do a better job of capturing the spatial variability in precipitation, something that often varies on a spatial scale much finer than GCMs operate at. Notwithstanding such examples, improved fit does not always happen and in some cases the fit can actually be worse (see Bader et al., 2008: 32 for further discussion). RCM projections are also totally dependent on the source of their boundary conditions. These conditions usually come from GCMs and, as we see below, GCMs vary a good deal in their abilities, strengths and weaknesses.

### 2.4 Hindcasts and Model Validation

Model validation is the process of determining if a model is, in some sense of the word, ‘good’. Model validation differs a bit between different kinds of models and models built for different purposes. In general, model validation would involve using the model to make a series of predictions and then comparing those predictions with what is actually observed. However, because what we are actually trying to predict with GCMs is the future climate, model validation must necessarily utilize surrogate data.

There are three general methods for validating GCMs: (i) the model is run backward through time to generate recent past climate and then these simulations are compared with historical climate data collected from weather stations all over the world; (ii) the models are run for a period in the Earth's history when climate forcings and other key variables were substantially different, and we see if the model can adequately reproduce the climate we think was present at these more distant times (see Chapter 1); and (iii) the models are used to predict the effects of large perturbations on the formation of major circulation events, such as El Niño. We concentrate only on the first method in this section. The others are interesting, but perhaps more esoteric than necessary for our purposes.

In any GCM, the equations are integrated to see how conditions change in time. To predict future climate we integrate forward in time but, as mentioned above, we have nothing to compare these results with, so they will not do for model validation. Mathematically, there is no reason that the integration needs to be carried out forward in time; these equations can be
integrated backward from the present time, and this produces ‘hindcasts’ (the opposite of forecasts) that can be compared with past climate (recall our discussion in Chapter 1 of how we know about past climates).

Let us start by looking at the ‘anomalies’ in global mean temperature for the past century. When talking about climate change (past or future) it is common to express the measure in terms of an anomaly, i.e. a difference between the measure's value at one point in time and its value during a particular reference period of time. Figure 2.6 shows the results for 58 simulations of 14 GCMs (shown in yellow). The mean of all 58 runs is shown in red. The observed temperature anomalies are shown in black. Vertical lines denote major volcanic eruptions (Randall et al., 2007).

The validation shown in Fig. 2.6 is nice, but non-spatial. Do the models capture the spatial variation with any degree of accuracy? Figure 2.7 shows the multi-model mean simulated contemporary temperatures minus the observed temperatures. The IPCC's Working Group I report (Randall et al., 2007: 608) summarizes the correspondence between simulated and observed climate as follows:

With few exceptions, the absolute error (outside polar regions and other data-poor regions) is less than 2°C. Individual models typically have larger errors, but in most cases still less than 3°C, except at high latitudes. Some of the larger errors occur in regions of sharp elevation changes and may result simply from mismatches between the model topography (typically smoothed) and the actual topography. There is also a tendency for a slight, but general, cold bias. Outside the polar regions, relatively large errors are evident in the eastern parts of the tropical ocean basins, a likely symptom of problems in the simulation of low clouds. The extent to which these systematic model errors affect a model's response to external perturbations is unknown, but may be significant. In spite of the discrepancies discussed here, the fact is that models account for a very large fraction of the global temperature pattern: the correlation coefficient between the simulated and observed spatial patterns of annual mean temperature is typically about 0.98 for individual models. This supports the view that major processes

Global mean near-surface temperatures over the 20th century from observations (black) and as obtained from 58 simulations produced by 14 different climate models driven by both natural and human-caused factors that influence climate (yellow). The mean of all these runs is also shown (thick red line). Temperature anomalies are shown relative to the 1901 to 1950 mean. Vertical grey lines indicate the timing of major volcanic eruptions. Figure adapted from IPCC WG I Chapter 9, Figure 9.5. Refer to corresponding caption for further details: http://tinyurl.com/fig2-6-suppl. (From: Climate Change 2007: The Physical Science Basis. Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Figure 8.1, Cambridge University Press.)

**Fig. 2.6.** The match between GCM-simulated global temperature anomalies (yellow lines, red line shows the mean) and observed anomalies (black line).
(a) Observed climatological annual mean SST and, over land, surface air temperature (labelled contours) and the multi-model mean error in these temperatures, simulated minus observed (colour-shaded contours). (b) Size of the typical model error, as gauged by the root-mean-square error in this temperature, computed over all AOGCM simulations available in the MMD at PCMDI. The Hadley Centre Sea Ice and Sea Surface Temperature (HadISST; Rayner et al., 2003) climatology of SST for 1980 to 1999 and the Climatic Research Unit (CRU; Jones et al., 1999) climatology of surface air temperature over land for 1961 to 1990 are shown here. The model results are for the same period in the 20th-century simulations. In the presence of sea ice, the SST is assumed to be at the approximate freezing point of seawater (~1.8°C). Results for individual models can be seen in the Supplementary Material, Figure S8.1. (From: Climate Change 2007: The Physical Science Basis. Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Figure 8.2. Cambridge University Press.)

Fig. 2.7. (a) The sea-surface temperature or the surface air temperature (contour lines) and (b) the difference between the multi-model mean predictions and the observed values (shaded contours).
governing surface temperature climatology are represented with a reasonable degree of fidelity by the models.

These models then seem to capture global mean temperature, at least for the recent past, with reasonably good accuracy. What about other aspects of climate? Let us consider precipitation. Figure 2.8 shows the observed pattern of annual precipitation (Fig. 2.8a) and the multi-model mean hindcast (Fig. 2.8b). The IPCC (Randall et al., 2007: 611–612) summarizes the correspondence between simulated and observed precipitation as follows:

Models also simulate some of the major regional characteristics of the precipitation field, including the major convergence zones and the maxima over tropical rain forests, although there is a tendency to underestimate rainfall over the Amazon. When considered in more detail, however, there are deficiencies in the multi-model mean precipitation field. There is a distinct tendency for models to orient the South Pacific convergence zone parallel to latitudes and to extend it too far eastward. In the tropical Atlantic, the precipitation maximum is too weak in most models with too much rain south of the equator. There are also systematic east–west positional errors in the precipitation distribution over the Indo-Pacific Warm Pool in most models, with an excess of precipitation over the western Indian Ocean and over the Maritime Continent. These lead to systematic biases in the location of the major rising branches of the Walker Circulation and can compromise major teleconnection pathways, in particular those associated with El Niño (e.g., Turner et al., 2005). Systematic dry biases over the Bay of Bengal are related to errors in the monsoon simulations. Despite the apparent skill suggested by the multi-model mean, many models individually display substantial precipitation biases, especially in the tropics, which often approach the magnitude of the mean observed climatology.

So these models would seem to produce pretty accurate reconstructions of recent past climate, when we know with reasonable accuracy the relevant climate forcings, at least for the global mean temperature. In and of itself, this is not too surprising because the models actually go through a parameter ‘tuning’ phase to produce better matches in the energy balance parts of the models. It is also possible for a model to perform well on average, even if it performs poorly in detail. A better test of accuracy is how the models capture the spatial and temporal variability around the mean (Figs 2.7 and 2.8). But even here the models perform quite well, with overall correlations between hindcast and observed temperature being ≥95%. There still exist some local errors and these errors can be quite large, but generally the models do a good job at reconstructing temperatures. On the other hand, hindcasts of monthly mean precipitation are not as well correlated with observed pattern; overall correlations are typically 50–60%, and the fit tends to be worse in the tropics than at higher latitudes. Nevertheless, GCMs do a pretty good job at matching the observed large-scale patterns and inter-annual variability in precipitation (Bader et al., 2008).

### 2.5 Model Results and Projections

Figure 2.9 shows the so-called ‘time-evolving’ change in global mean surface temperature and percentage change in global mean precipitation for the various GCMs under a few different scenarios. The first thing that is evident from Fig. 2.9 is that there is considerable inter-annual variability in the projection for any given model. This is something little understood by the media, who made such a hoopla over the observed downturn in global average temperatures in 2008. The same is true for the mean model projections, although much of the inter-annual variability is cancelled out when such a mean is calculated. This becomes even more evident when we look at Fig. 2.10, in which the global mean temperature anomaly is summarized. A robust conclusion from these models is that global mean temperatures will increase by between 1.5 and 3.5°C by the end of this century (barring the increasingly unlikely commitment scenario being realized). Also note that there is little difference between the scenarios until around the middle of the present century.

Global mean values obscure seasonal and spatial variability. Compare Fig. 2.9 with Fig. 2.11, where we put back in the spatial (but not seasonal) variability. We see that the global mean increase in temperature is not distributed uniformly in space. In Fig. 2.12 we see the seasonal variation for the middle right-hand column result from Fig. 2.11. But even Fig. 2.12 hides variability on even shorter time scales, particularly for precipitation. One topic that is not well developed in AR4 but is potentially really important for ecologists predicting the likely impacts of climatic change is the distribution of extreme weather events. In AR4 (Meehl et al., 2007: 778) the authors make reference to two studies and conclude that there will be: ‘a decrease in temperature variability during the cold season in the extratropical NH
Annual mean precipitation (cm), observed (a) and simulated (b), based on the multi-model mean. The Climate Prediction Center Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1997) observation-based climatology for 1980 to 1999 is shown, and the model results are for the same period in the 20th-century simulations in the MMD at PCMDI. In (a), observations were not available for the grey regions. Results for individual models can be seen in Supplementary Material, Figure S8.9. (From: Climate Change 2007: The Physical Science Basis. Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Figure 8.5. Cambridge University Press.)

Fig. 2.8. The observed pattern of annual precipitation (a) and the hindcast pattern of annual precipitation (b).
[Northern Hemisphere] and a slight increase in temperature variability in low latitudes and in warm season northern mid-latitudes. While for precipitation they conclude that there will be an increase in monthly mean precipitation variability in most areas, both in absolute value (standard deviation) and in relative value (coefficient of variation). Figure 2.13 shows the changes in extreme precipitation events based on the multi-model mean anomalies. 'Precipitation intensity' is defined as the annual total precipitation divided by the number of wet days. We can see from Fig. 2.13a that variability in precipitation intensity increases quite substantially over the course of the century, regardless of the scenario. We also see from Fig. 2.13b that this variability is not uniformly distributed across the globe. 'Dry days' is
defined as the annual maximum number of consecutive dry days. We see from Fig. 2.13c that the variability in the distribution of dry days increases across the century in a way that is positively correlated with the differences in emission scenarios. Finally, again, we see from Fig. 2.13d that this increased variability in the distribution of dry days is not expected to be uniformly distributed across the globe.

Box 2.2 provides an example of how scenarios and GCMs are used in studies of the biological impacts of climate change. Notice that any attempt to make statements about the impacts of climate change ought to be made in the context of the scenario and GCM to which they may refer, rather than being sweeping statements about the impacts of climate change per se.

2.6 Conclusions

Our job as biologists is to say something meaningful about the likely biological impacts of climatic change. As will become clear throughout the rest
of this book, those impacts are unlikely to be simple, but will depend on factors like how much warming occurs, when during the year the warming occurs, how, when and where precipitation will change, and so on. That is, our conclusions about the biological impacts of climatic change will be context specific. We won’t be able to say things like: ‘In a warmer world, species X will definitely decrease in abundance’. We won’t be able to say things like this because it will turn out that the answer depends on how much warmer, when that warming occurs and, in many cases, how it interacts with changes in precipitation and other climate variables, and other ecological contingencies. So to say something specific and meaningful (as an impact of climatic change) we need to be able to consider specific changes in climate – not just global annual mean changes, which, let’s face it, are ecologically meaningless since no organism ever lives its whole life at the global mean value for any climate variable. So specifics matter. That is why the understanding we tried to convey in this chapter matters so much. If details matter, we need to understand that all climate scenarios, models and projections are not equal and we need to be able to give an accounting of how sensitive our predictions about climate change impacts are to our choice of emissions scenario and general circulation model. A really thorough study would consider multiple combinations of scenarios and GCMs (or RCMs). At a minimum, when reading the results of an impacts study, one has to beware of what ‘climate change’ the study is considering.
Multi-model mean changes in surface air temperature (°C, left), precipitation (mm day⁻¹, middle) and sea level pressure (hPa, right) for boreal winter (DJF, top) and summer (JJA, bottom). Changes are given for the SRES A1B scenario, for the period 2080 to 2099 relative to 1980 to 1999. Stippling denotes areas where the magnitude of the multi-model ensemble mean exceeds the inter-model standard deviation. Results for individual models can be seen in the Supplementary Material for this chapter. (From: Climate Change 2007: The Physical Science Basis. Working Group I. Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Figure 10.9, Cambridge University Press.)

Fig. 2.12. The multi-model mean projections for mean temperature anomaly, mean precipitation anomaly and mean sea level pressure anomaly (relative to 1980–1999) and at each point in space for the A1B scenario at the end of this century. DJF denotes the average over the months of December, January and February, while JJA denotes the average over the months of June, July and August. The stippling denotes places where the magnitude of the multi-model mean is greater than the inter-model standard deviation. This is a reasonable indication that the value of the anomaly at that point is likely to be different from zero.
Changes in extremes based on multi-model simulations from nine global coupled climate models, adapted from Tebaldi et al. (2006). (a) Globally averaged changes in precipitation intensity (defined as the annual total precipitation divided by the number of wet days) for a low (SRES B1), middle (SRES A1B) and high (SRES A2) scenario. (b) Changes in spatial patterns of simulated precipitation intensity between two 20-year means (2060–2099 minus 1980–1999) for the A1B scenario. (c) Globally averaged changes in dry days (defined as the annual maximum number of consecutive dry days). (d) Changes in spatial patterns of simulated dry days between two 20-year means (2060–2099 minus 1980–1999) for the A1B scenario. Solid lines in (a) and (c) are the 10-year smoothed multi-model ensemble means; the envelope indicates the ensemble mean standard deviation. Stippling in (b) and (d) denotes areas where at least five of the nine models concur in determining that the change is statistically significant. Extreme indices are calculated only over land following Frich et al. (2002). Each model's time series was centred on its 1980 to 1999 average and normalised (rescaled) by its standard deviation computed (after de-trending) over the period 1960 to 2099. The models were then aggregated into an ensemble average, both at the global and at the grid-box level. Thus, changes are given in units of standard deviations. (From: Climate Change 2007: The Physical Science Basis, Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Figure 10.18, Cambridge University Press.)

Fig. 2.13. Changes in the distribution of precipitation intensity and in the distribution of dry days. (a) and (c) show the time series for the change in the variability of the distribution of these two metrics, while (b) and (d) show the spatial distribution for just the A1B scenario in the time period 2080–2099. Stippling denotes areas in which five of the nine models used in this analysis agree that the change is statistically significant. The solid lines in (a) and (c) denote the multi-model mean and the shaded areas denote the standard deviation.
Box 2.2. Potential carbon sequestration in cultivated soils.

Lugato and Berti (2008) conducted a modelling study of the interactions between the choice of SRES scenario, GCM and different management practices on the potential soil carbon sequestration (see Chapters 7, 9 and 11) in cultivated soils in north-east Italy. We chose this study because it was unusually thorough in that it not only considered the impacts of the choice of scenario, but also the impacts of the choice of four GCMs.

Lugato and Berti considered the impacts on soil carbon sequestration of changing farming practice from ‘business as usual’, to either ‘reduced tillage’ or ‘farmyard manure’. They considered a crop rotation of maize–wheat–maize–soybean. ‘Business as usual’ and ‘reduced tillage’ used 240 kg of mineral nitrogen fertilization per hectare, while ‘farmyard manure’ used mineral-N at 140 kg/ha plus an additional organic-N at 100 kg/ha applied as farmyard manure.

To model the carbon sequestration Lugato and Berti used the Century Model (Parton et al., 1988). The Century Model simulates C, N, P and S dynamics on monthly time intervals. It can be used for either natural or cultivated lands, and it has been used extensively in studies of climate change impacts (see Chapter 3).

Table 2.3 shows the change in soil organic carbon (SOC) that results from switching farming practices. For comparison they also looked at the carbon sequestration that could be obtained by converting that land to a grassland reserve (GR). They compared the output from this analysis for two different time periods: the first commitment period from the Kyoto Protocol (2008–2012) and in the more distant future (2080).

Table 2.3 shows that switching ‘business as usual’ to ‘farmyard manure’ is better for SOC in the short

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Management</th>
<th>Had3</th>
<th>GCM2</th>
<th>CSIRO2</th>
<th>PCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009–2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1FI</td>
<td>RT</td>
<td>113</td>
<td>109</td>
<td>107</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>180</td>
<td>168</td>
<td>167</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>231</td>
<td>260</td>
<td>247</td>
<td>255</td>
</tr>
<tr>
<td>A2</td>
<td>RT</td>
<td>113</td>
<td>109</td>
<td>111</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>178</td>
<td>168</td>
<td>174</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>231</td>
<td>260</td>
<td>239</td>
<td>253</td>
</tr>
<tr>
<td>B1</td>
<td>RT</td>
<td>115</td>
<td>105</td>
<td>107</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>161</td>
<td>165</td>
<td>171</td>
<td>163</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>227</td>
<td>254</td>
<td>227</td>
<td>258</td>
</tr>
<tr>
<td>B2</td>
<td>RT</td>
<td>112</td>
<td>106</td>
<td>108</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>179</td>
<td>166</td>
<td>171</td>
<td>164</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>230</td>
<td>255</td>
<td>229</td>
<td>259</td>
</tr>
<tr>
<td>2008–2080</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1FI</td>
<td>RT</td>
<td>323</td>
<td>484</td>
<td>614</td>
<td>545</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>197</td>
<td>391</td>
<td>529</td>
<td>440</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>1589</td>
<td>1346</td>
<td>1287</td>
<td>1350</td>
</tr>
<tr>
<td>A2</td>
<td>RT</td>
<td>391</td>
<td>509</td>
<td>597</td>
<td>542</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>266</td>
<td>432</td>
<td>516</td>
<td>451</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>1493</td>
<td>1384</td>
<td>1278</td>
<td>1373</td>
</tr>
<tr>
<td>B1</td>
<td>RT</td>
<td>420</td>
<td>502</td>
<td>502</td>
<td>475</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>285</td>
<td>461</td>
<td>461</td>
<td>416</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>1553</td>
<td>1553</td>
<td>1308</td>
<td>1399</td>
</tr>
<tr>
<td>B2</td>
<td>RT</td>
<td>442</td>
<td>511</td>
<td>546</td>
<td>485</td>
</tr>
<tr>
<td></td>
<td>FM</td>
<td>318</td>
<td>459</td>
<td>469</td>
<td>418</td>
</tr>
<tr>
<td></td>
<td>GR</td>
<td>1455</td>
<td>1296</td>
<td>1262</td>
<td>1404</td>
</tr>
</tbody>
</table>

RT, reduced tillage; FM, farmyard manure; GR, grassland restoration.
A1FI is a combination of the A1C and A1G scenarios; the ‘FI’ stands for ‘fuel intensive’.

Bold figures denote the management practice that maximizes carbon sequestration for each scenario–GCM combination. Bold italic figures denote the GCM–management option that results in the maximum carbon sequestration.
term, but in the long term the result is opposite and 'reduced tillage' is ultimately the best. Furthermore, there is an interaction between management practices and the climate change scenario. The biggest difference between 'reduced tillage' and 'farmyard manure' occurs for the A1FI scenario (a combination of the A1C and A1G scenarios, where the 'FI' refers to 'fuel intensive'). Notice too that the conclusion one can reach from this study can differ by more than twofold depending on one's choice of GCM. Certainly, Lugato and Berti's results argue for the use of both multiple GCM projections for the same scenario, and multiple scenarios for the same GCM, in order to have any idea of how robust a study's conclusions are to these choices (see also Chapter 14, 'Which climate change?', p. 239).