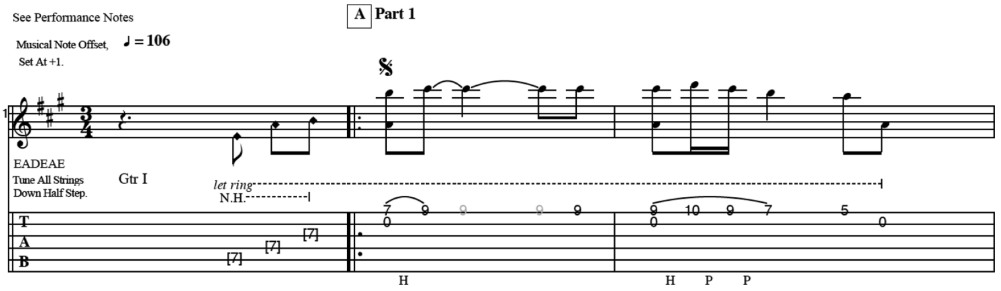
6

The Climate Modelling Hierarchy



One half step at a time.
EADEAE, The Handing Down, Edward Gerhard

Climate phenomena are studied using observations and climate models. Because we cannot investigate most of these phenomena in the laboratory, climate models are a central component of climate research. A wide range of models is in use with on the one hand 'very simple' conceptual climate models and on the other hand 'very complex' state-of-the-art global climate models (GCMs). It would be impossible (and also useless) to try to provide an overview of all the models around. However, general notions on the use and importance of a climate modelling hierarchy can be given and are the main focus of this chapter.

In Section 6.1, the main model 'traits', that is, scales and processes, are described, and we argue that climate models can be roughly distinguished using these two attributes. In Section 6.2, we then address the intricate coupling between model choice and scientific question. Deterministic versus stochastic modelling approaches are described in Section 6.3, and in the last section (Section 6.4), sources of modelling errors are discussed.

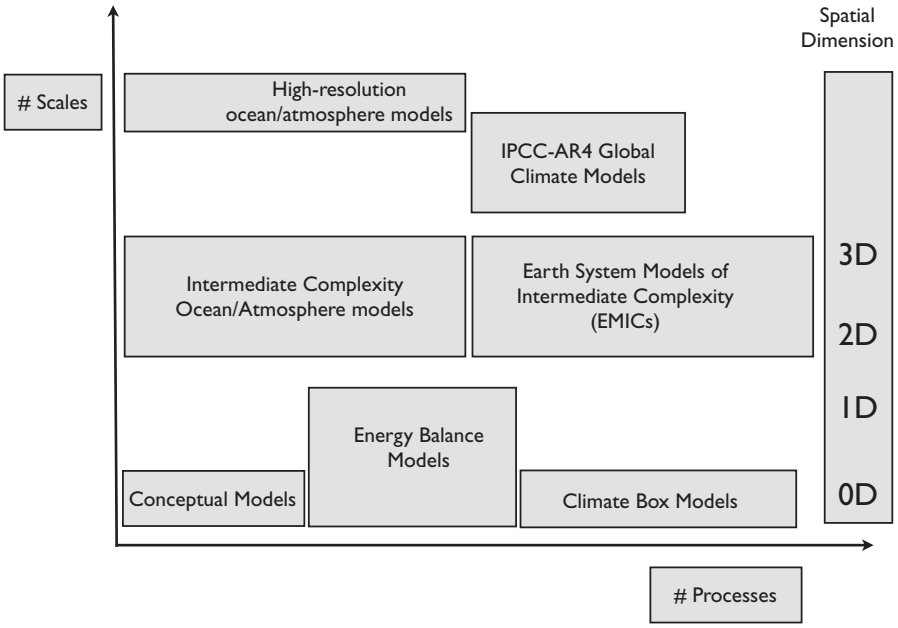


Figure 6.1 'Classification' of climate models according to the two model traits: number of processes and number of scales. There is, of course, overlapping between the different model types, but for simplicity they are sketched here as non-overlapping. Here 0D (zero dimensional), 1D (one dimensional), 2D (two dimensional) and 3D (three dimensional) indicate the spatial dimension of the model.

6.1 Model traits: scales and processes

As we have seen in Chapter 1, climate phenomena occur on different spatial and temporal scales. Climate variations associated with the Pleistocene Ice Ages have typical time scales of 100 kyr and a global spatial pattern. In contrast, thunderstorms affect only a local area and have typical time scales of hours. Based on observations, it is known that distinct processes are often involved in these different phenomena. The processes driving El Niño variability differ substantially from those that are involved in glacial-interglacial cycles.

As scales and processes are such important properties of phenomena, it is important to classify climate models using these two traits (Fig. 6.1). Here the trait 'scales' refers to both spatial and temporal scales, as there exists a relation between both: on smaller spatial scales, usually faster processes take place. 'Processes' refers to physical, chemical or biological processes taking place in the different climate compartments (atmosphere, ocean, cryosphere, biosphere, lithosphere). Both traits affect the dimension of the state vector of the dynamical system in a different way. An increase in scales increases the state vector dimension, with a fixed relation between the state vector variables due to the spatial coupling. For example, when the resolution is doubled of a one-dimensional scalar variable, the dimension increase due to that

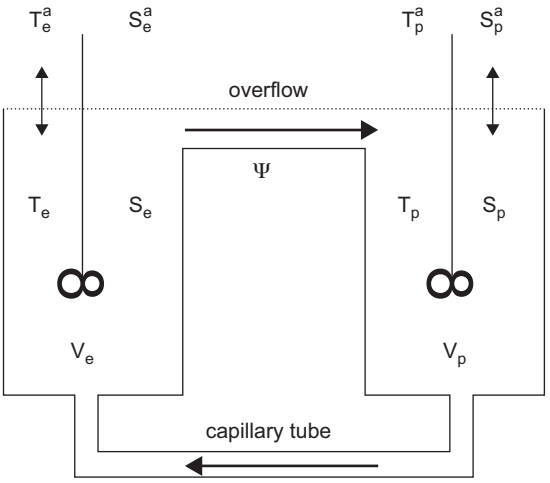


Figure 6.2 Sketch of the two-box model as in Stommel (1961). Two reservoirs contain well-mixed water and are connected through an overflow and a capillary tube. The circulation is driven by density gradients between the boxes, which are set up by the exchange of heat and freshwater at the surface.

variable is doubled. Such a relation is absent when the state vector is increased due to the increase in the number of processes.

Models with a limited number of processes and scales are usually referred to as conceptual climate models. In these models, only very specific interactions in the climate system are described. A typical example is a box model; one of the simplest ones was used by Stommel (1961) to study the stability of the ocean's thermohaline circulation (Fig. 6.2). As we use this model in Chapter 10, the equations are provided in Example 6.1.

Example 6.1 The Stommel two-box model In the model proposed by Stommel (1961), two boxes, having volumes V_p and V_e, contain well-mixed water of temperature and salinity (T_e, S_e) and (T_p, S_p), with the subscripts e and p indicating the equatorial and polar box, respectively. The boxes are connected at the surface by an overflow region and at the bottom by a capillary tube. The flow rate Ψ is directed from high to low pressure and is assumed linearly related to the density difference of the liquid between the boxes, that is,

$$\Psi = \gamma \frac{\rho_p - \rho_e}{\rho_0}, \tag{6.1}$$

where ρ_0 is a reference density and γ a hydraulic constant. Hence the flow rate is taken positive if the liquid is heavier in the polar box. The exchange of properties does not depend on the sign of Ψ, because it only matters that properties from one box are transported to the other box. Because mass is conserved, the pathway (either through the overflow or through the capillary) is unimportant. A linear equation of

state is assumed of the form

$$\rho = \rho_0(1 - \alpha_T(T - T_0) + \alpha_S(S - S_0)), \tag{6.2}$$

where T_0 and S_0 are reference values for temperature and salinity and α_T and α_S are thermal expansion and saline contraction coefficients, respectively.

Exchange of heat and salt in each box due the surface forcing is modelled through a relaxation to a prescribed surface temperature and salinity (T^a, S^a) with relaxation coefficients C^T and C^S. These coefficients differ for each box and for each quantity considered (heat or salt). In this way, the balances of heat and salt in each box are given by

$$V_p \frac{dT_p}{dt} = C_p^T(T_p^a - T_p) + \mid \Psi \mid (T_e - T_p), \tag{6.3a}$$

$$V_e \frac{dT_e}{dt} = C_e^T(T_e^a - T_e) + \mid \Psi \mid (T_p - T_e), \tag{6.3b}$$

$$V_p \frac{dS_p}{dt} = C_p^S(S_p^a - S_p) + \mid \Psi \mid (S_e - S_p), \tag{6.3c}$$

$$V_e \frac{dS_e}{dt} = C_e^S(S_e^a - S_e) + \mid \Psi \mid (S_p - S_e). \tag{6.3d}$$

In the following, we restrict consideration to the case of realistic forcing, for which $T_e^a - T_p^a > 0$ and $S_e^a - S_p^a > 0$. For simplicity, it is assumed that the relaxation times for temperature to the surface forcing in both boxes is proportional to their volume, and hence $C_p^T/V_p = C_e^T/V_e \equiv R_T$ is constant. The same simplification is made for salinity with $R_S = C_p^S/V_p = C_e^S/V_e$.

When time, temperature, salinity and flow rate are scaled with $1/R_T$, $V_eV_pR_T/(\gamma\alpha_T(V_e + V_p))$, $V_eV_pR_T/(\gamma\alpha_S(V_e + V_p))$ and $V_eV_pR_T/((V_e + V_p)$, respectively, the dimensionless equations become

$$\frac{d\bar{T}}{dt} = \eta_1 - \bar{T}(1 + \mid \bar{T} - S \mid), \tag{6.4a}$$

$$\frac{d\bar{S}}{dt} = \eta_2 - \bar{S}(\eta_3 + \mid \bar{T} - \bar{S} \mid), \tag{6.4b}$$

where $\bar{T} = T_e - T_p$, $\bar{S} = S_e - S_p$ and $\bar{\Psi} = T - S$ is the dimensionless flow rate. Three parameters appear in the equations (6.4), which are given by $\eta_1 = (T_e^a - T_p^a)\ \gamma\alpha_T(V_e + V_p)/(V_eV_pR_T)$, $\eta_2 = (R_S/R_T)(S_e^a - S_p^a)\ \gamma\alpha_S(V_e + V_p)/(V_eV_pR_T)$ and $\eta_3 = R_S/R_T$.

The explicit equations (6.4) show that the dimension of the state vector d is very small ($d = 2$), and, consequently, the model is a very limited representation of scales (box dimension, overturning time scale). In addition, it is clear that the model contains several (here, three) parameters that have to be estimated ad hoc. The ad hoc parameters

arise through simplifying assumptions on the form of the different fluxes (mass, momentum, heat, salt) in the ocean system. ■

The number of processes between the boxes can be extended, for example, by including sea ice, land ice processes and/or biogeochemical processes. In this way, one ends up in the right lower part of the diagram in Fig. 6.1 because the number of scales is still relatively small. Models of this type are abundant in the area of biogeochemical climate modelling, and a prominent example is the model of Gildor et al. (2002), which we use in Chapter 11. An increase in the number of boxes will increase the number of scales represented, but the model will still consist of a system of ordinary differential equations (ODEs). The dimension of the state vector in these models is mainly increased because of an increase in the number of dependent quantities. Another example is the integrated models (Kemfert, 2005), where, apart from physical, chemical and biological processes, societal aspects are also included. A third example is the low-dimensional ODE-type models of glacial-interglacial cycles (Saltzmann, 2001).

Limiting the number of processes, scales can be added by discretising the governing partial differential equations spatially up to three dimensions. A prominent class of early climate models of this type are so-called energy balance models (EBMs) described in North et al. (1981). As we use these models in Chapter 11, the equations of a typical EBM are provided in Example 6.2.

Example 6.2 Energy balance models In the most simple EBMs, the state vector is composed only of the atmospheric surface temperature T_a, and a heat balance gives an equation of the form

$$\rho_a H_a C_{pa} \frac{\partial T}{\partial t} = \nabla \cdot \mathbf{F} + Q_S - Q_L - Q_{oa} - Q_{la} - Q_{ia}, \tag{6.5}$$

where H_a is a typical atmospheric vertical (depth) scale. In addition, ρ_a and C_{pa} are the density and the heat capacity of air, respectively. The quantity Q_S is the incoming shortwave radiation, Q_L is the outgoing longwave radiation and Q_{oa}, Q_{la} and Q_{ia} are the heat fluxes into the ocean, land and ice, respectively. The flux $\mathbf{F}$ is the meridional heat transport, which is mostly accomplished by atmospheric eddy processes.

The various energy balance models differ in the representation of physical processes through the choices of $\mathbf{F}$, the fluxes Q and the spatial dimension of the problem (0D, 1D or 2D). One of the earliest EBMs was the 0D model presented by Budyko (1969). The equation is

$$\rho_a H_a C_{pa} \frac{\partial T}{\partial t} = \frac{\Sigma_0}{4}(1 - \alpha) - \sigma T^4,$$

where σ is the Stefan-Boltzmann constant. In this EBM, only a simplified globally integrated radiation balance is represented.

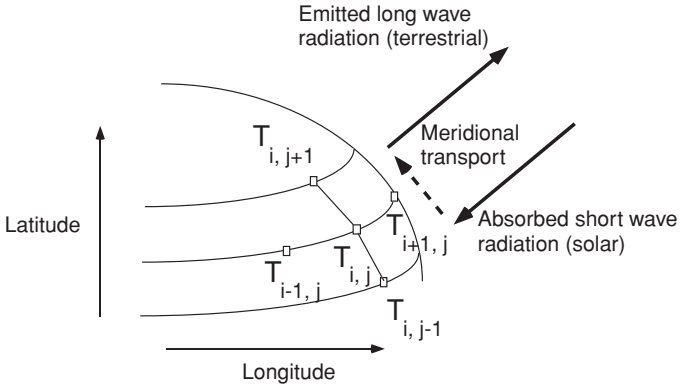


Figure 6.3 Sketch of a two-dimensional EBM, modelling the surface temperature T of the Earth.

For such a 0D EBM, the number of degrees of freedom $d = 1$. However, when a spatially one-dimensional version is considered, with latitude as spatial dimension, and the latitudinal grid has M points, the number of degrees of freedom will scale with M, that is, the values of T on the grid points. Similarly, discretising the equation (6.5) for the two-dimensional temperature field T (Fig. 6.3) using a finite difference, finite element or spectral method will result in a discrete state vector $\mathbf{x}$. The dimension of $\mathbf{x}$ will increase due to an increase in resolution to present T over the sphere. For example, using a finite difference method on a grid with longitudes $\phi_i = i\Delta\phi$, $i = 1, \ldots, N$ and latitudes $\theta_j = j\Delta\theta$, $j = 1, \ldots, M$ will result in a state vector $\mathbf{x}$ of dimension $N \times M$. ■

A higher spatial resolution and inclusion of more processes will give models located in the right upper part of the diagram. In a GCM, we divide the atmosphere, ocean, ice and land components into grid boxes (Fig. 6.4). Over such a three-dimensional grid box, we consider the budgets of momentum, mass and, for example, heat. Momentum budgets basically follow from Navier-Stokes equations formulated for air and water. The difference on what goes into a box minus what goes out of that box leads to an increase/decrease of a particular quantity, say temperature. Once the distribution of a quantity is known at a certain time, then these budgets provide an evolution equation to determine the quantity some time later.

The advantage of more boxes is that we resolve the temperature better (more points in a certain area). With an increasing number of grid boxes, however, the time development of an increasing number of quantities (at each grid box) has to be calculated. The same holds for the number of processes included in a GCM: more processes simply mean more calculations. Also, the longer time period over which we want to compute the development of each quantity, the longer it takes to do the calculation on a computer. An overview of the IPCC-AR4 models is provided in

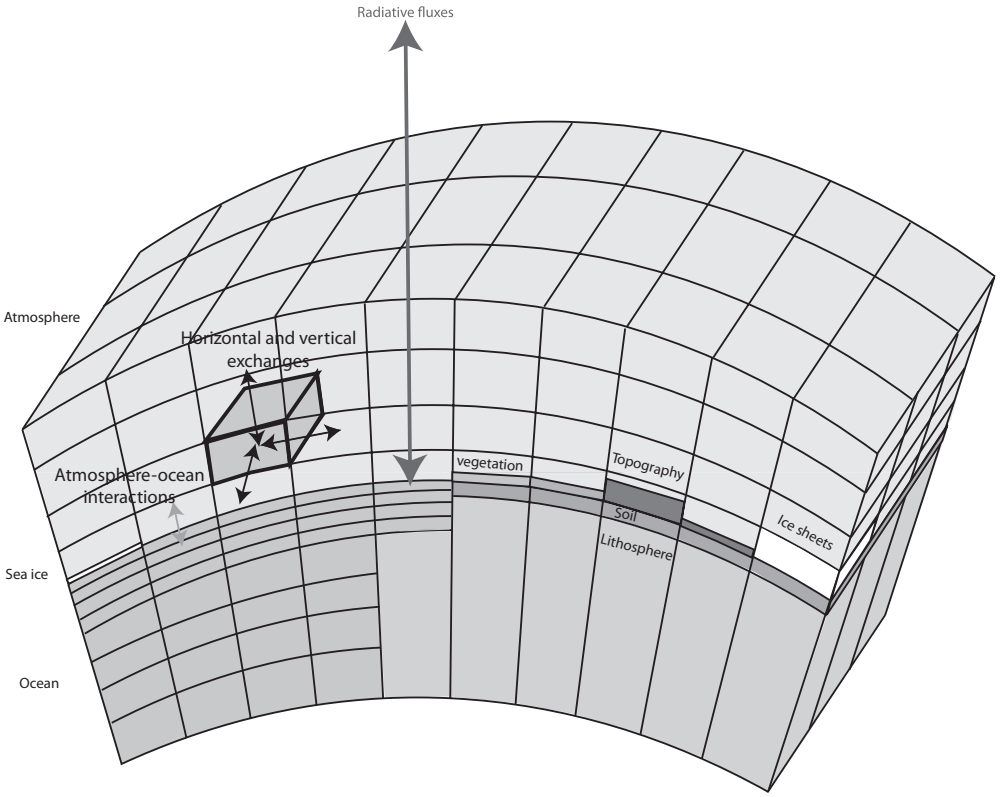


Figure 6.4 A typical structure of a GCM; the number of grid boxes in each of the components determines the spatial resolution of the model (figure from http://stratus.astr.ucl.ac.be/textbook/ and courtesy of Hugues Goosse). (See Colour Plate.)

Chapter 8 (pp. 597–600) of WGI (Meehl et al., 2007). A data archive of the results of these model simulations is at PCMDI (http://www-pcmdi.llnl.gov/).

The state-of-the-art GCMs are located above the Earth System Models of Intermediate Complexity (EMICs) because they represent a larger number of scales (Claussen et al., 2002). Compared with GCMs, the ocean and atmosphere models in EMICs are strongly reduced in the number of scales. For example, the atmospheric model may consist of a quasi-geostrophic or shallow-water model, and the ocean component may be a zonally averaged model. The advantage of EMICs is therefore that they are computationally less demanding than GCMs, and hence many more long-time scale processes, such as land-ice and carbon cycle processes, can be included. Each of the individual component models of EMICs may also be used to study the interaction of a limited number of processes. Such models are usually referred to as Intermediate Complexity Models (ICMs). A prominent example is the Zebiak-Cane model of the El Niño/Southern Oscillation phenomenon (Zebiak and Cane, 1987). In time, the GCMs

of today will be the EMICs of the future, and the state-of-the-art GCMs will shift towards the upper right corner in Fig. 6.1.

6.2 Scientific questions and model choice

There is no 'unified' climate model with which one is able to tackle any scientific question on the climate system, and there will never be. Model choice is therefore tightly coupled to (i) the specific question one wants to answer and (ii) practical limitations, such as the computational platform that is available.

For example, regarding the El Niño/Southern Oscillation (ENSO) phenomenon, important scientific questions include the following. Which processes are responsible for the dominant ENSO time scale (about 4 years)? Can we predict the development of an ENSO over a certain time period, say a few months? How does ENSO affect the weather over the globe? Will the dominant period and amplitude of ENSO change under an increase of the global mean surface temperature?

How do climate scientists, motivated by one or more of these questions, make a model choice? For example, a conceptual model (implemented on a PC) may be very successfully used to understand particular aspects of ENSO (i.e., the dominant time scale), but it may be useless for an adequate quantitative prediction. However, a GCM (implemented on a supercomputer) may be fully capable of simulating ENSO under changing external conditions, but the analysis of the output may not lead to a detailed understanding of the dominant time scale of variability.

Let us restrict the discussion here to questions related to the understanding of processes (physical mechanisms) giving rise to a particular phenomenon. Such questions are usually about consistent relations between dependent quantities that eventually lead to a theory of the phenomenon, for example, a theory of ENSO. By choosing a certain model out of the hierarchy presented in the previous subsection, one is limited to the range of questions that can be 'usefully' posed. Or the other way around: with a specific question as target, not all models will be suited to answer it. Of course, one is tempted to answer questions with results of models that are not suited for their task, but this usually leads to confusion rather than understanding.

Obviously, observations are crucial for theory development, as otherwise there is no way to falsify the proposed theories. However, theories based only on observations do not have to be consistent with the underlying conservation laws that must be satisfied in the system. Models are therefore always needed to interpret the observations, and the whole model hierarchy (from conceptual models to GCMs) contributes to this interpretation. With the conceptual models, where the causal chains are precisely known, specific relations between quantities can be discovered, whereas with the GCMs, the degree of realism (the quality of model-observation agreement) can be improved.

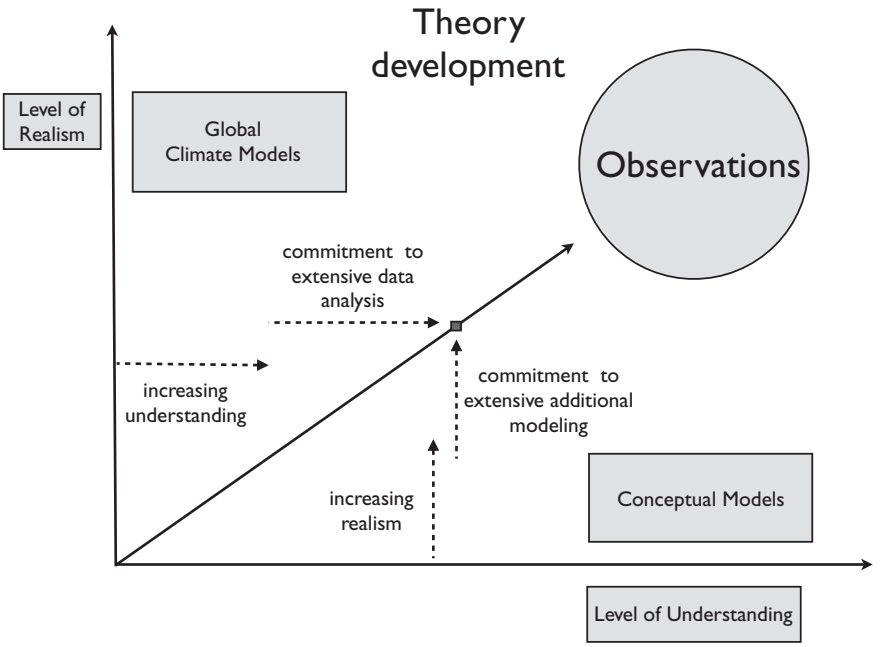


Figure 6.5 Model usage for theory development on phenomena in the climate system.

When one starts with a model high up in the hierarchy, it may be difficult to connect the description of causal chains to well-understood 'building blocks' as present in conceptual models. In this case, there must be a commitment to in-depth analysis of the model results. When starting with a model low in the hierarchy, the connection to observations will be difficult. Hence there must be a commitment to demonstrate that the processes identified in the conceptual model are also dominant in models higher up in the hierarchy. Theories then develop in stages and towards the interpretation of the observations, using successively better models of which the behavior is better understood (Fig. 6.5).

6.3 Deterministic versus stochastic models

In principle, we know the equations governing the climate system's development of the physical and chemical processes. In other words, no new physical 'laws' have to be discovered. However, because of the large-scale differences between processes in the climate system, we are unfortunately not able to use these equations. Consider, for example, the flow of salt water in an ocean basin such as the Atlantic. The flow is governed by the Navier-Stokes equations. The horizontal spatial scale is typically $L = 1{,}000$ km and the vertical dimension $D = 1$ km, whereas the smallest (Kolmogorov) scale is typically $l = 0.1$ m. The ratio of scales $L/l = 10^7$ in only one

horizontal dimension, 10^{14} in the horizontal plane and the ratio of scales in the vertical $D/l = 10^4$. Hence one needs at least 10^{18} grid points to resolve the smallest scale processes in the ocean. At the moment there is no computer that can even store one state vector of a solution to these equations.

This effectively means that the governing equations of the physical processes are not known and have to be developed for each problem. An additional complication is that when including biological processes in the marine and/or terrestrial biosphere in climate models, we lack a first-principles description guideline (such as we have for the physical processes), and many of the governing equations are semiempirical. Again, the questions asked and problems targeted strongly determine which model is useful. In addition, should we opt for a stochastic description or for a deterministic approach?

6.3.1 Deterministic modelling approach

When we are restricted to questions about understanding, a deterministic approach will be appropriate if the (parameterized) small-scale processes are not crucially involved in the phenomenon. In many cases later in this book, small-case processes are, for example, represented as friction or diffusion, which are needed for consistency (closing balances), but these are not the crucial processes controlling the particular climate phenomenon.

One way to understand the processes driving a climate phenomenon using this approach is to start low in the hierarchy of deterministic models, say with a conceptual model. One identifies a basic mechanism consisting of the interaction of certain processes giving rise to the phenomenon. The next step is then to identify characteristics of this mechanism, so-called mechanistic indicators, which can be used to monitor whether this mechanism is present in models up in the model hierarchy. It is then crucial that the studies up to a model in the top of the hierarchy are really carried out and the properties of the mechanistic indicators investigated. When indeed the mechanism proposed from the conceptual model can be traced by its characteristics in a GCM and the latter provides an adequate simulation of observed phenomena, then we proceed in the right direction towards a theory of the phenomenon (Fig. 6.5). In the following chapters of the book, examples of this approach are given.

6.3.2 Stochastic modelling approach

When a stochastic approach is followed, it is assumed that the small-scale processes (at least the statistics of these processes) are crucial for the understanding of the phenomenon. In this case, a statistical model of the small-scale processes has to be explicitly provided.

The approach followed by Hasselmann (1976) starts in abstract form from the following set of equations:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{y}), \tag{6.6a}$$

$$\frac{d\mathbf{y}}{dt} = \mathbf{g}(\mathbf{x}, \mathbf{y}), \tag{6.6b}$$

where $\mathbf{f}$ and $\mathbf{g}$ are given vector functions of the 'slow' (climate) variable $\mathbf{y}$ and the 'fast' (weather) variable $\mathbf{x}$. For example, the variable $\mathbf{x}$ can be associated with baroclinic instability in the atmosphere with a characteristic response time scale τ_x of less than a few days. In Example 6.3, these would be processes affecting T_a. The slow variable $\mathbf{y}$ is that associated with the long-term development of climate, say on a typical time scale τ_y; in Example 6.3, this is the sea-surface temperature T.

We want to know the statistics of the variable $\mathbf{y}$ under the condition of a time scale separation, that is, $\tau_x \ll \tau_y$. Let $\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{y}_0$, where $\mathbf{y}_0$ is the initial condition for $\mathbf{y}$. If $t \ll \tau_y$, then $\mathbf{y}$ can be assumed constant in the right-hand side of (6.6b) and hence

$$\frac{d\tilde{\mathbf{y}}}{dt} \approx \mathbf{g}(\mathbf{x}). \tag{6.7}$$

Let us now consider an ensemble of realizations over the short time scale τ_x, then by taking the ensemble average $\langle\ \rangle$ of (6.7), we find

$$\langle\tilde{\mathbf{y}}\rangle \approx \langle\mathbf{g}(\mathbf{x})\rangle t, \tag{6.8}$$

and hence for the fluctuations $\mathbf{y}' = \tilde{\mathbf{y}} - \langle\tilde{\mathbf{y}}\rangle$, it follows that

$$\frac{d\mathbf{y}'}{dt} \approx \mathbf{g}(\mathbf{x}) - \langle\mathbf{g}(\mathbf{x})\rangle = \mathbf{g}'(\mathbf{x}), \tag{6.9}$$

with $\langle\mathbf{g}'(\mathbf{x})\rangle = 0$. Hence on the short time scales $\mathbf{y}'$ will fluctuate rapidly compared to $\mathbf{y}$ and can be represented by a stochastic process.

Example 6.3 Red noise spectrum of SST A prominent example of stochastic modelling is the explanation of the red-noise background in the spectrum of sea-surface temperature (SST) anomalies in observations (Hasselmann, 1976). To explain the spectrum of SST, the response of the oceanic mixed layer temperature T to an atmospheric heat flux Q_{oa} is considered. If we assume that the mixed layer depth h is fixed, the governing equation for T is

$$\rho C_p h \frac{dT}{dt} = Q_{oa}, \tag{6.10}$$

where ρ and C_p are the constant density and heat capacity of the ocean water, respectively (Fig. 6.6).

The heat flux can be approximated as

$$Q_{oa} = \mu(T_a - T), \tag{6.11}$$

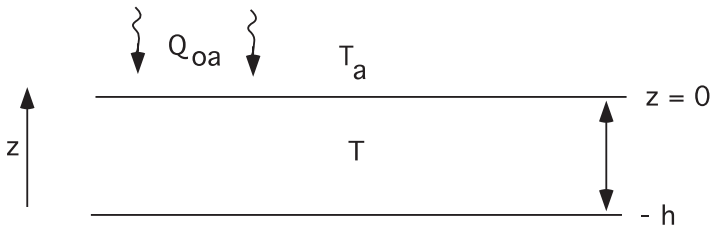


Figure 6.6 Sketch of an upper ocean mixed layer forced by a heat flux Q_{oa}. The mixed-layer depth is indicated by h.

where μ ($\mathrm{Wm^{-2}K^{-1}}$) is a heat exchange coefficient and T_a the atmospheric temperature just above the ocean surface. Combining (6.10) and (6.11), we obtain

$$\frac{dT}{dt} = -\gamma T + \gamma T_a \; ; \; \gamma = \frac{\mu}{\rho C_p h}. \tag{6.12}$$

In general, the forcing part of this equation (here represented by T_a) will be very irregular (have energy in a wide range of frequencies), which can be represented by white additive noise.

This implies that we can write the temperature equation as the Itô SDE of the form

$$dT_t = -\gamma T \; dt + \sigma \; dW_t, \tag{6.13}$$

where σ represents the amplitude of the noise.

The spectrum of T follows directly from the solution of (6.13), which was given in (3.49) as (with initial condition T_0)

$$T_t = e^{-\gamma t}\left[T_0 + \sigma \int_0^t e^{\gamma x} \; dW_x\right], \tag{6.14}$$

with variance given by (3.57) as

$$Var(T_t) = e^{-2\gamma t}\left(Var(T_0) + \sigma^2 \frac{e^{2\gamma t} - 1}{2\gamma}\right) \tag{6.15}$$

and a spectrum

$$\hat{f}(\omega) = \frac{\sigma^2}{\gamma^2 + \omega^2}. \tag{6.16}$$

For high frequencies, $\omega \to \infty$, this spectrum decays as ω^{-2}, but for low frequencies $\omega \to 0$, it approaches a constant. These features of the spectrum are also seen in observations of SST anomalies (Dommenget and Latif, 2002). Hence when the 'noise' is represented as being temporally uncorrelated ('white'), the response shows higher energy in the low frequencies than in the high frequencies (hence it is 'red').

In a way, the slow dynamical (ocean) component is integrating the high-frequency (atmospheric) noise. ■

From Example 6.3, it is clear that it is important to have an appropriate statistical model of the small-scale processes available. Whether to use an Itô or Stratonovich formulation is less crucial. As the representations of small-scale processes come from physics for which classical rules of calculus should be valid, a Stratonovich formulation is a priori preferred. However, as shown in Chapter 3, an equivalent Itô formulation can always be found from this Stratonovich formulation.

6.4 Model error

Apart from many conceptual models, a deterministic climate model (ICM, EMIC, GCM) consists of a set of partial differential equations (representing the basic balances of momentum, energy, etc.) that can always be formulated into the operator form

$$\mathcal{M}\frac{\partial \mathbf{u}}{\partial t} + \mathcal{L}\mathbf{u} + \mathcal{N}(\mathbf{u}) = \mathcal{F}, \tag{6.17}$$

where $\mathcal{L}, \mathcal{M}$ are linear operators, $\mathcal{N}$ is a nonlinear operator, $\mathbf{u}$ is the vector of dependent quantities and $\mathcal{F}$ contains the forcing of the system. To get a well-posed problem, appropriate boundary and initial conditions have to be added to this set of equations.

Under any numerical method, the discretised equations can be written as a nonlinear system of ODEs (in many cases with algebraic constraints), which has the form

$$\mathcal{M}_N\frac{\partial \mathbf{x}}{\partial t} + \mathcal{L}_N\mathbf{x} + \mathcal{N}_N(\mathbf{x}) = \mathcal{F}_N, \tag{6.18}$$

where $\mathbf{x}$ indicates the total N-dimensional vector of unknowns. The operators depend on parameters, and their subscript N indicates that they are discrete equivalents of the continuous operators.

To systematically discuss the errors of such a numerical climate model, it is instructive to start with the situation that the discretized equations (6.18) model a laboratory fluid experiment. The properties of the fluid such a viscosity and thermal diffusivity are accurately known, and we can use Navier-Stokes and/or Boussinesq equations to model the flow. The discretised equations describing the evolution of the state vector $\mathbf{x}$ are then known. In this case, the model errors will consist of two components:

(i) *Discretisation error*. Suppose we use a uniform grid size Δx and a time step Δt in a finite-difference method using direct numerical simulation. When the method is, for example, second order in time and space (such as a central difference Adams-Bashforth method), then there is a discretisation error $\mathcal{O}(\Delta x^2, \Delta t^2)$, as we are basically solving slightly different equations than the continuous equations. We can choose Δx and Δt to achieve a desired accuracy of the solutions.

(ii) *Context error*. There are inaccuracies in representing the physical situation such as the dimensions of the problem, the forcing functions, the liquid properties, the influence of background laboratory conditions and so forth. These context errors enter the equations through uncertainties in the parameters occurring in the operators and the forcing $\mathcal{F}_N$.

When all the scales of the flow can be captured by the numerical method (i.e., Δx and Δt can be chosen sufficiently small), very accurate solutions can be obtained, and hence a close correspondence with observed phenomena can be achieved. In this case, one can develop benchmark problems where different numerical approaches can be compared and their convergence to the 'true' solution can be tested. In many fluid laboratory experiments, however, the flow is turbulent, and not all relevant scales can be captured. There are then two options: the scales are resolved by choosing a certain Δx (e.g., the direct numerical simulation [DNS] method) or a separate model of the behavior of the small scales is used (e.g., by using a large eddy simulation [LES] method). In either case, an additional error appears, which may be called a

(iii) *Scale representation error*. Part of the turbulent cascading process in fluids is not represented, which may be crucial for certain phenomena. We may also misrepresent specific instability phenomena as the particular scales are not resolved. This scale representation essentially differs from the discretisation error, as the latter error is only connected to the numerical representation of the actual equations used.

With a scale-representation error, not all equations for the complete state vector $\mathbf{x}$ in (6.18) are known. Hence $\mathbf{x}$ is split into $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_u$, where the subscripts r and u indicate resolved and unresolved, respectively. In DNS the equations for $\mathbf{x}_u$ are simply omitted, whereas in LES, a separate model (a parameterization) is used, which results in $\mathbf{x}_u = P(\mathbf{x}_r)$.

In climate models the three error types (i)–(iii) above are all present. The context error is, for example, associated with uncertainties in the forcing of the model, which may arise from uncertain observations. The scale representation error appears in many ocean and atmosphere models due to a lack of sufficient horizontal and vertical resolution. In many GCMs and EMICs, for example, the effects of ocean eddies on the transport of heat and salt are represented by the Gent-McWilliams parameterization (Gent and McWilliams, 1990).

However, in climate models there is an additional error related to the representation of processes, that is,

(iv) *Process representation error*. Relevant processes for the particular phenomenon are neglected or have an inadequate representation; for example, there is no representation based on first principles or not even based on semiempirical studies.

In this case either the state vector **x** may be too limited or the operators contain errors. An example is the omission of processes controlling the carbon cycle in GCMs used in IPCC-AR4, such as those in the marine and terrestrial biosphere.

Suppose we have decided on our model type, with appropriate parameterizations, fixed parameters and chosen forcing functions. Even when realising that such a model contains severe idealisations and has all types [(i)–(iv)] of errors, there are still important a priori checks that can be done before using such a model.

(1) *Consistency*. When the continuous formulation is chosen (with all its idealizations and errors) and the discretisation method is chosen, it is very important to check that the grid and time step used provide sufficiently accurate results to these equations. Hence for fixed parameters, it is always necessary to perform calculations with different grid sizes and time steps to investigate whether one solves the 'correct' equations (the equations one intends to solve).
(2) *Integral balances*. For every model, one can integrate the continuous equations used over the domain and obtain usually much more simplified integral equations, representing overall conservation of mass/energy and heat/salt components. For a solution of the discrete model, the discrete integral balances should be satisfied up to a certain accuracy, say a relative error (with respect to the largest term) of a few percentages. Numerical solutions where mass, energy and so forth have large artificial sources or sinks in the system are meaningless.

With this overview of the hierarchy of models and their error sources, we are now ready to tackle several important problems in climate physics using dynamical systems theory as will be done in the next chapters.