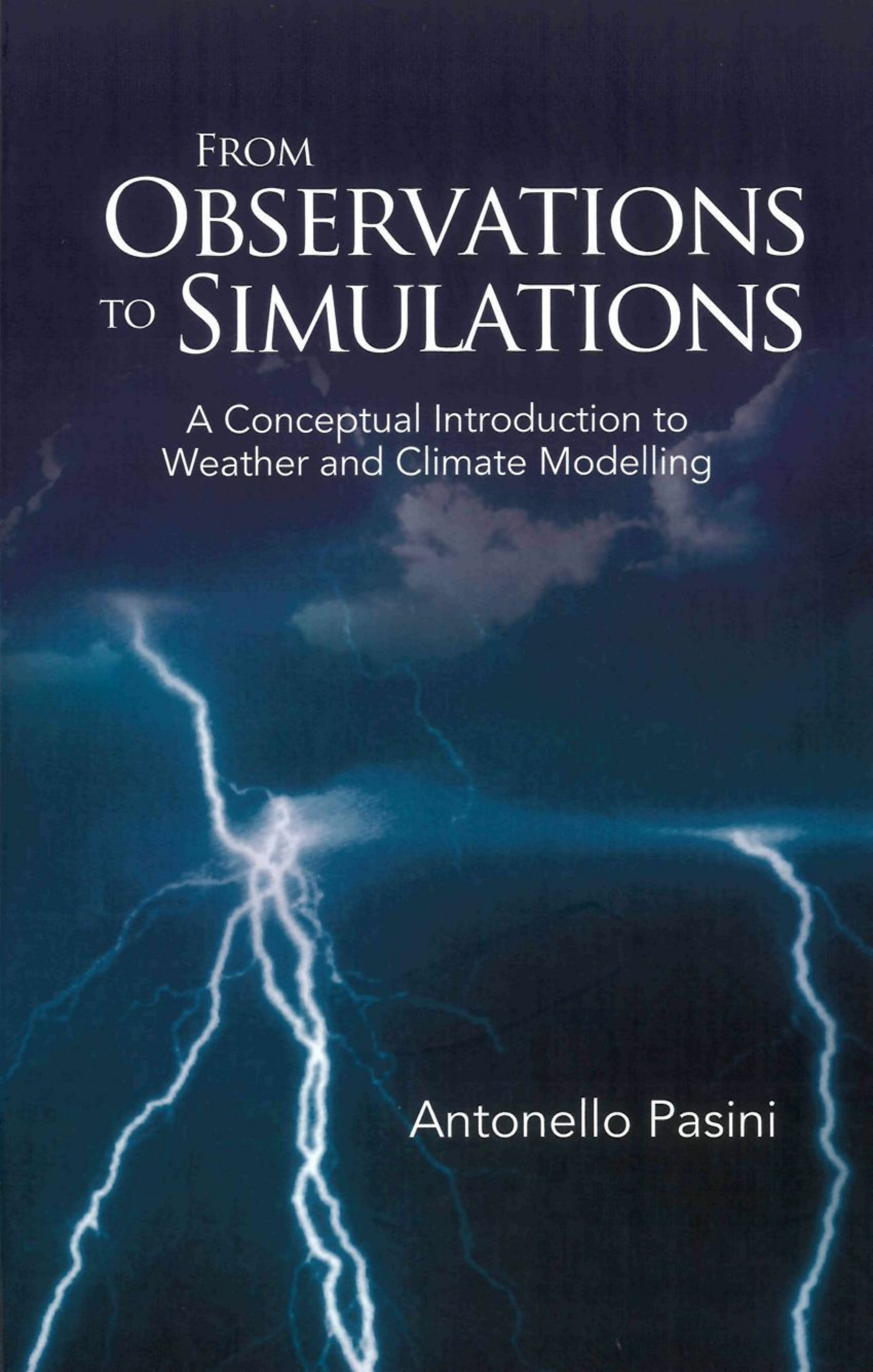


FROM OBSERVATIONS TO SIMULATIONS

A Conceptual Introduction to Weather and Climate Modelling

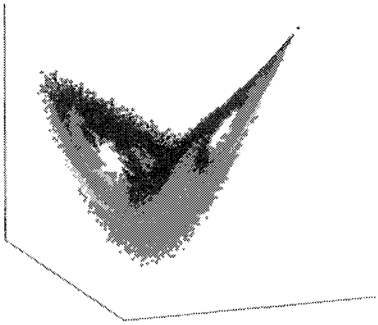
Antonello Pasini

FROM OBSERVATIONS TO SIMULATIONS

This page is intentionally left blank

FROM OBSERVATIONS TO SIMULATIONS

A Conceptual Introduction to Weather and Climate Modelling



Antonello Pasini
CNR - Institute of Atmospheric Pollution, Rome, Italy

Translated by
Francesca Sofri

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Original Italian edition:
I CAMBIAMENTI CLIMATICI — METEOROLOGIA E CLIMA SIMULATO

FROM OBSERVATIONS TO SIMULATIONS
A Conceptual Introduction to Weather and Climate Modelling

ISBN 981-256-475-6

SEPS
SEGRETARIATO EUROPEO PER LE PUBBLICAZIONI SCIENTIFICHE

The translation of this book has been funded by SEPS: Segretariato Europeo per le Pubblicazioni Scientifiche, Via Val d'Aposa 7, 40123 Bologna, Italy, seps@alma.unibo.it, www.seps.it

Printed in Singapore by Mainland Press

In memory of my father Elio,
who first taught me to play with the physical world

This page is intentionally left blank

Acknowledgements

Many friends and colleagues helped me with their advice during the writing of this book: it is impossible to mention all of them. For the enormous amount of time they dedicated to me, I would like to particularly thank Pier Francesco Coppola and Fausto D'Aprile.

Special thanks are due to Eugenia Kalnay and Malaquias Peña, who allowed me to use a code of theirs for the creation of the figure of the Lorenz attractor that appears on the cover of the book.

Preface

When my children were little, I found myself several times watching them with wonder and admiration while they were playing with mechanical toys. In fact, if you give a child a toy, and the child is curious enough and has enough time, he or she will eventually open it up in order to see how it works, then will try to reconstruct it in order to play with it again. This childish attitude is usually lost with age; but modern scientists behave exactly like this with the systems they are studying: it is not by chance that they are often called "grown-up babies".

In a certain sense, this book means to be precisely a journey into the ideas that have led the scientists who study the weather and climate to recover this childish outlook.

In the history of science, after the period of Greek philosophers and their medieval Epigones (during which people confined themselves to observing reality, looking for regularities that might explain its behaviour), with Galileo Galilei scientists began to control and manipulate reality in the laboratory, in order to induce nature to give specific answers to specific questions. This led to great cognitive progresses in the domain of the so-called "hard sciences", such as physics and chemistry.

Obviously, this childish tendency to open up a toy in order to look inside it — to disassemble it, then to reassemble its parts — is pursued nowadays in all the areas of science, including the study of the atmosphere and climate. As a rule, the activity of decomposing a system in order to study its individual elements and their basic interactions does not pose any particular problem: in the laboratory, for instance, we can easily study the absorption of infrared radiation by carbon dioxide molecules (which contributes to the so-called "greenhouse effect"); or, regarding air as a fluid or as a mixture of gases and water, we can analyse the movements of portions of air in simplified cases or study the main thermodynamic processes that take place in the atmosphere. But when we try to reconstruct the whole "toy" in the laboratory, though this is usually possible for mechanical systems, we find that it is extremely difficult for the atmosphere and for the Earth system: we will see this in the course of our journey.

So, up to a few decades ago, meteorology and climatology were still purely observational disciplines, characterised by a lot of difficulties in achieving theoretical syntheses. Then the fruitfulness of the Galilean experimental method (though transferred to a different set-up) was retrieved in these fields too, and now computers and simulation models may be regarded as "virtual laboratories" where the weather and climate are studied. In a model, formed of equations that represent our theoretical knowledge (and can be solved numerically) and of variables that refer to the real data, it is possible to reconstruct the complexity of reality, though in a simplified manner. In particular, we can simulate the evolution of the climate system on the basis of scenarios observed in the past or surmised for the future; and all this can be done in very little time (tens of hours for decades of real evolution) and with the possibility of carrying out "numerical experiments".

In this book we will deal precisely with this "methodological revolution", which underlies our present understanding of the behaviour of many complex systems, including the climatic one.

In our journey from observations to simulations, we will follow the typical route of a scientific investigation, and will encourage the reader to become qualitatively aware of the characteristics of the atmosphere and of the Earth system, gradually finding different explanatory schemes. We will also re-examine some classical, well-known concepts such as "causality" and "prediction", in the light of the models and of new concepts pertaining to the theory of dynamical systems.

To sum up: this book presents a research into complex systems that has a huge range of practical applications, and is also contributing to a substantial change in our outlook on nature.

A. Pasini

Contents

Chapter 1

Introduction

All of us who are a part of the present-day society of global information are inundated with a constant flow of news of all sorts, often relevant to events that have taken place a few hours or minutes before, in some remote corner of the world. Among these reports, sometimes we receive some news-flashes about natural events of a meteorological or climatic type, in many cases extreme events (hurricanes, floods, droughts), that, either directly or as an immediate consequence, have spread destruction and death.

Sometimes the news is not so immediately tragic, but brings anxiety about a more remote future. I refer, for instance, to the news of the separation from the Antarctic continent of an enormous iceberg (called Larsen B, larger than the State of Rhode Island), in March 2002. Subsequently other, smaller icebergs broke away, and many people thought it was reasonable to link these events to the world-wide warming of the oceans and lower layers of the atmosphere during the last century (a subject we will discuss further on). Should this trend go on, other icebergs might be expected to break away; and should the ice that melts or is released into the sea come from the cap that covers the Antarctic land, these “plunges” into the ocean would concur in raising the mean global level of the sea and in causing the interface between the land and the sea, i.e. the coasts, to recede.

There are also some pieces of news that rarely reach the western society, because they are not immediately catastrophic and concern extremely remote places. I refer, for instance, to the fact that an entire people, the approximately 11,000 inhabitants of the archipelago of Tuvalu, a small south-Pacific island-state, is negotiating with the governments of other states (in particular Australia and New Zealand), in order to be allowed to take sanctuary in their territory. This application was made necessary by the increasingly frequent flooding of their atolls and by the rising level of the sea, which suggests the possible need for an evacuation within the next few years.

After the immediate, facile stir caused by certain pieces of news, there begins (but not for all of them) a stage of deeper elaboration. In particular, a traditional deterministic analysis leads us to wonder which causes have produced (or are producing) a certain flood, the detaching of a certain group of icebergs or the appearance of "environmental refugees". The rationale behind these questions springs from the common-sense observation that, once the cause of an undesirable effect is removed, that effect disappears as well. At this point scientists come into play.

In the history of science, causal relationships have always been studied carefully. Within the sphere of Greek natural philosophy, whose *summa* is represented by Aristotle's work, there coexist basically two types of physical causes: efficient cause and final cause. The former is the current meaning of the word: it is what comes before a certain phenomenon and determines it; the latter is the goal towards which the caused thing tends, in the future. In modern science, finalistic thinking — according to which a certain phenomenon or process takes place because it tends towards a final situation, its goal — was abandoned[1]. The present-day causal outlook establishes a "time's arrow": all causes come before the effect they produce (they are situated within the light cone of the past, to put it in relativistic terms). Moreover, within the sphere of classical (not quantum) physics, the theoretical approach to evolutionary phenomena is based on differential equations (ordinary or with partial derivatives), more or less implicitly assuming that the future state of the system under consideration is univocally determined starting from a known past state, by means of evolution equations.

---

[1] In actual fact, in the mathematical physics of the eighteenth century there survives a finalistic explicative concept, with the minimum-action principle or Maupertuis' principle, within the sphere of variational calculus in dynamics: for a critical analysis and a modern outlook that connects it to a causalistic approach, see, e.g., Yourgrau and Mandelstam (1979).

We will analyse this paradigm further on, after having applied it to the atmosphere system, or, more generally, to the Earth system. At present it will be sufficient to point out that, as a rule, in this pattern the scientific explanation of a phenomenon or of the evolution of a process with time involves first of all the identification of its causes, then the analysis of their way of combining to give rise to the phenomenon or process under consideration.

The area that acts as a prototype for these causality analyses is classical mechanics, in particular laboratory-controlled experimental situations[2]. If we recall the simple dynamics problems we had to solve in secondary school and the easy experiments we had to carry out, we notice that in those cases the few forces that act on a material body are easy to identify, and the total effect they produce on the body (e.g. its acceleration) are nothing but the vectorial sum of the effects (accelerations) that each force would produce if it were applied individually. This property of a physical system is called linearity. If the system under examination is correctly described by an equation such as $\vec{a} = \vec{F}/m$, the solution of this equation, in the case of the composition of several forces, is the sum of the solutions of the individual cases in which, each time, only the action of a single force is considered. So once a certain number of concurrent causes has been identified, their composition is quite simple and the problem under examination is easy to solve.

In some systemic ecology studies begun in the nineteen-seventies[3], for the cases that have just been described, linear causality is mentioned, as opposed to a so-called circular causality that is considered typical of living systems, where it becomes essential to consider the intricate relationships and interconnections between the various elements of a system. Though we shall not go so far as to examine the dynamics of living systems, we must point out that also in the atmosphere, and, more generally, when dealing with the overall Earth system, the relationships between causes and effects can no longer be interpreted in terms of the

---

[2] These considerations will be resumed and extended in Chapter 5, when we will discuss the Galilean experimental methods.

[3] E.g., see Bateson (1980).

simple linear causality of classical mechanics. In actual fact, what undermines the simple linear pattern is the presence of feedback, i.e. chains of circular two- or multi-element interactions, in which an effect acts in turn on the cause that has generated it, increasing its effect (positive feedback) or decreasing it (negative feedback). In Chapter 4, when we will analyse the current theoretical knowledge of the Earth system, we will consider some concrete examples of these complex cause-effect relationships (which are called non-linear).

From what we have just explained, it is evident that there does not exist an answer, in terms of a linear composition of concurring causes, to the questions we had previously posed about the floods, icebergs and "environmental refugees". If to this we add the fact that our theoretical knowledge of atmospheric phenomena is linked to a description in terms of systems of equations that can be solved analytically only in very particular, simplified cases, we can understand why, up to a few years ago, giving sensible answers to these questions was quite unthinkable. To make the situation even more complicated, in many cases these phenomena should be regarded as "extreme events", that is statistically improbable events, and this makes it difficult also to approach their description and prediction in terms that are not dynamic but statistical.

Within the picture we have just given of the situation, whose elaboration lends itself to be carried out from various angles, we must now delimit the goals that this book has set out to achieve.

When discussing phenomena relevant to meteorology, the climate and its changes during the last few decades or centuries, the subject can be tackled from the viewpoint of our scientific knowledge in this area (essential for any other discussion of the problem), from the viewpoint of the impact of these phenomena on nature and mankind (including studies on the vulnerability of the latter and on possible adaptation strategies), and — if there seems to be the possibility of acting in a concrete manner to reduce the causes of the most negative phenomena — from the viewpoint of mitigation studies relevant to what is usually called sustainable development (a possible example is the development of energy production methods that have a lower impact on the environment). The third viewpoint is the one where the decisions to be taken are the most delicate, because they affect the world-wide social, economic and political sphere[4].

The viewpoint adopted in this book belongs to the first of these three areas. In brief, we will begin by analysing the current scientific knowledge of meteorological and climatic phenomenology, both from the angle of observations (Chapters 2 and 3) and from that of the theoretical description of the Earth system and its atmosphere subsystem (Chapter 4). In doing this, we will pay attention particularly to the conceptually relevant aspects that reveal the complexity of the system under examination and make it so different from the physical systems that our school education has made familiar to us. This first part of the book is somehow preparatory to what follows, because it supplies the grounding required to understand the change of paradigm in the researches in this field that is described in the second part of the book.

Chapter 5, which at first sight may seem a digression, discusses the Galilean experimental method. The motivation of this brief excursus into the physics of the seventeenth century is that the application of this method was precisely what allowed physics and other so-called "hard" scientific disciplines to achieve extremely important results and a very evident progress in the understanding of nature. In the meantime, meteorology and climatology, as described in the first three chapters, remained observational disciplines, and had much trouble in attempting a theoretical description, because of the complexity of the system that was being studied. At this point it is natural to think that, if we could recover a Galilean way of carrying out scientific researches, other disciplines that up to now were purely observational might achieve important progresses in understanding the phenomena within their province.

Chapter 6, in which simulation models are introduced, analyses precisely the entrance of some observational disciplines into the category

---

[4] It is not fortuitous that this schematisation follows, rather faithfully, the one proposed by the IPCC (Intergovernmental Panel on Climate Change) in the three reports it has recently published on the state of climatic research in the world. The IPCC was established in 1988 by the World Meteorological Organisation (WMO) and the United Nations Environment Programme (UNEP) with the main purpose of periodically drawing up technical reports (and résumés for policy-makers) about the state of the art of scientific, technical, social and economic knowledge of climate changes and their consequences. Further on we will refer to one of these reports.

of Galilean-type sciences. Here a new paradigm in the way of performing a scientific research, simulation, is discussed, particularly as regards the possibility of constructing a "virtual laboratory" for the study of complex systems.

In Chapter 7, we will describe the structure and operation of the models for weather forecasts, highlighting their strong points and weaknesses. In particular we will discuss the appearance of what is called deterministic chaos, which leads to a revision of the concept of deterministic prediction for complex systems, such as the atmosphere.

In Chapter 8, the system that is being studied in its dynamic evolution is extended to fully include the oceans and some phenomena that are neglected or not dealt with dynamically, for instance the so-called carbon cycle (photosynthesis, respiration, storage). The purpose here is not to predict the weather during the next few days, but to correctly reconstruct the climates of the past and make it possible to analyse future scenarios in relation to important climatic variables such as the mean temperature and the precipitations in a certain area of the world. The positive results and current limits of the models are evaluated.

The last chapter contains a general discussion of the importance of the simulation-based approach to the study of the weather and climate. This approach is evaluated from a conceptual and epistemological[5] point of view, and the prospects of future development within this paradigm, and out of it, are presented.

The great importance that meteorology has in everyday life and the enormous publicity that is given to the debate about climate change, essentially based on the results of predictive models, are accompanied by a general lack of information on the scientific practice (of which simulation-based modelling is an integral part) that characterises these disciplines. I hope, on the one hand, that my book will be able to bridge this gap, on the other hand that the conceptual and epistemological

---

[5] In the literature relevant to this area, the adjective "epistemological" is used with meanings that sometimes differ from each other, depending on the exact meaning ascribed to the noun "epistemology" (e.g., see Greco (1998)). Here "epistemology" is understood as the critical and philosophical study of nature and of the procedures of scientific activity.

reflections by which it is accompanied will explain the intellectual appeal of a change of paradigm that has recently allowed these disciplines to be included in the category of Galilean-type sciences. Furthermore, there is a lot of talk about complexity and chaos, and one is led to believe that the phenomena relevant to them are confined in some obscure sector of physics. On the contrary, the study of the atmosphere as a subsystem of the Earth system is an ideal and very concrete case study of complex system, and makes it possible to evaluate the conceptual and practical scope of a model-based approach to these themes.

Chapter 2

Meteorological and Climatic Observations

The beginning of meteorology is conventionally traced back to the year 1643, when Evangelista Torricelli constructed the first mercury barometer that made it possible to demonstrate the existence of atmospheric pressure. This was also the period of the invention of the thermometer and of the improvement of the hygrometer (for measuring humidity), anemometer (for measuring the intensity of the wind and the direction from which it comes), and pluviometer (for measuring the amount of rain that has fallen during a certain period of time). In substance, meteorology, understood as a discipline that studies the Earth's atmosphere and the phenomena that take place in it, by means of instrumental measurements, began in Italy at the court of the Medici in Florence, and was developed within the sphere of the Accademia del Cimento, the first example of a scientific society established in Europe, which included a great number of disciples of Galileo Galilei.

In actual fact, it was not possible to speak of meteorology in the modern sense of the term or to have some hope of also being able to predict the future conditions of the weather until individual observations were integrated in spatially extensive networks that made it possible to have an overall (we might call it synoptical) three-dimensional view of the state of the atmosphere at a certain instant in time. The forerunner of this modern vision appeared, once again, in seventeenth-century Florence: the grand duke Ferdinand II adopted this method in 1654 and established the first network of meteorological observations (obviously surface-based), with data coming even from transalpine territories. Upper air instrumental measurements began much later with the development of flight, to which the recent great upsurge in the study of meteorology is chiefly due. It is interesting to notice, in any case, that the first sector that was benefited by meteorology and began to work in synergy with it was agriculture, once again in the Grand Duchy of Tuscany. In the eighteenth century, with the Accademia dei Georgofili and the reformist project of the grand duke Peter Leopold for the development of Tuscan farming, modern agrometeorology began.

In spite of this brief outline of the beginning of instrumental meteorology, we do not mean, here, to retrace the historical development of this discipline[1]. It will be sufficient for us to give an accurate description of what is available now, in terms of observations, for determining the weather and climate in a certain region.

2.1 The "State" of the Weather

In physics, a system is regarded as defined at a certain instant if its "state" at that instant is known. As a rule, this state is an entity that is not observable and this fact does not make it possible to extract all the information about the system by means of measurements. The unobservability of the state depends on the indetermination inherent in the measuring process and on the complexity of the system. The state of a simple dynamic system of interacting particles (without any internal structure) is known at a certain instant in time if the position and speed of all the particles at that instant are known; in a simple thermodynamic system, it is necessary to be acquainted with the quantities of pressure, volume and temperature. In the former case it is a matter of a microscopic description in terms of basic constituents, in the latter of a macroscopic description in terms of mean quantities[2]. It is well known that statistical mechanics is a bridge between these two descriptions and makes it possible to interpret the macroscopic variables on the basis of a microscopic description: for instance, the absolute temperature of a portion of gas, which in the statistical mechanical definition is directly

---

[1] Some good texts of history of meteorology are available. For its beginning as an independent discipline, the recent essay-novel by Hamblyn (2001) can be read.

[2] For further details about dynamic and thermodynamic systems, it is advisable to consult a basic physics text, such as that by Fermi (1937). A particularly interesting case is that of the concept of state in a quantum system: see, for instance, Ghirardi (2003).

proportional to the mean kinetic energy of the molecules, is interpreted as a macroscopic measurement of this energy.

In the atmosphere, where an enormous number of molecules are present even in a small portion of air, obviously we cannot apply a dynamic description in microscopic terms. Observational meteorology therefore produces a thermodynamic description (inevitably an approximate one) of the physical state of the atmosphere above certain observation sites, by means of measurements (in terms of instantaneous values or of mean values over a period of time) of physical quantities such as temperature, humidity, pressure and wind. This description is completed by information relevant to the possible presence of meteorological phenomena such as rain, fog or mist. If several sampling points can be used on the territory at the same time, these measurements give an idea of the state of the physical atmosphere system in the domain represented by that region.

The concept of state will become crucial when we will start dealing with the evolution of the physical atmosphere system. Then we will take up that concept again and extend the discussion of its importance. At present it is sufficient to state that, through our measurements and discrete sampling, we are only able to achieve an approximate determination of the state of the system under examination.

2.2 A Definition of Climate

What we have discussed up to now refers to the determination of the weather. What about the climate? By climate we mean the set of physical and meteorological conditions that characterise, on the average, a certain area of the world over a certain period of time: at least 30 years, as specified by the World Meteorological Organisation. To put it more accurately, in order to achieve a description of a climate it is necessary to know the mean values and variability of relevant quantities such as temperature and precipitations. In this sense, meteorological observations may be used for determining the climate; the only provision that is required is their specific post-processing in order to highlight not only the mean statistical value of the individual variables over various decades, but also some elements that determine their rate of variability, such as the scattering of their values around the mean (estimated, for instance, by means of the standard deviation) and the frequency of the occurrence of extreme events.

Climatologists, in actual fact, demand more than that, because they also require information about periods in which collected and coded instrumental measurements did not exist yet. As we will explain, this is possible if we rely heavily on our theoretical knowledge of some phenomena, in order to retrace long historical series of important variables, such as temperature.

2.3 An Overview of Meteorological and Climatic Observations

The overall situation of the currently available meteorological and climatic observations is shown diagrammatically in Figure 1. The observations we have called direct consist of the subjective ones reported in historical-period chronicles and documents, and of the instrumental ones that began (as we have already stated) in the middle of the seventeenth century, first in a sporadic, isolated way, and then became increasingly integrated in observational networks. The former should be regarded as impaired by a considerable degree of uncertainty, if nothing else because of the subjective manner in which the weather is perceived (the fact, reported in the chronicles of several periods, that people have always been complaining that spring and autumn have disappeared, should alert us). The latter, obviously, are more accurate and objective, and become also more reliable as the observational networks spread all over the world with more homogeneous instruments.

The observations that have been indicated as indirect in the figure have been obtained by means of the reconstruction of climatic data (called "proxy data") from long local historical series of some quantities that can be interpreted by means of our theoretical knowledge of physical or biophysical principles. This way the values of meteorological and climatic variables such as temperature, even in the remote past, can be inferred indirectly. Some examples of these historical series are those relative to the annual rings of centuries-old trees, the characteristics of corals and the geologic core sampling of Antarctic ice. Further on we will briefly return to this topic, in order to elucidate in a more detailed manner what can be actually obtained from these data and with which rate of uncertainty.

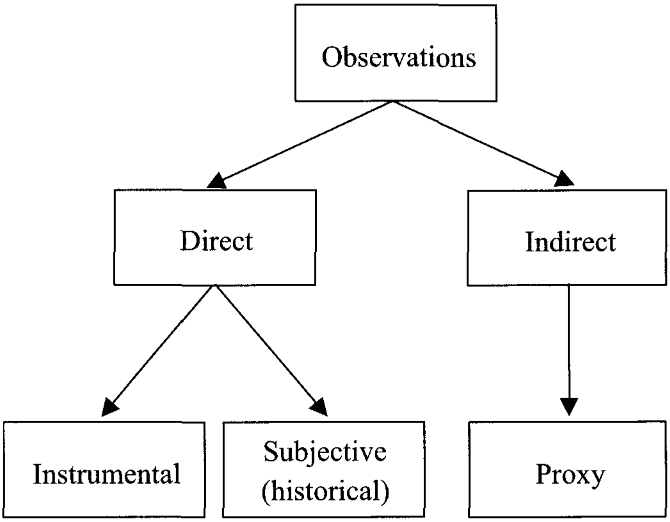


Fig. 1 Diagram of meteorological and climatic observations.

For the time being, it is interesting to focus on the most typically meteorological data that nowadays are routinely collected worldwide: our purpose now is to give the reader an overall view of the constant monitoring to which the atmosphere is subjected. Various sources of information are available to the meteorologists, ranging from automatic or manned meteorological ground stations, to radiosondes that rise into the atmosphere in order to probe the properties of one of its vertical columns, and to satellites that orbit at various heights in order to observe the Earth and its atmosphere by means of many types of different instruments.

We can easily understand the need to have meteorological data coming from an observational network that is extended over the territory or even includes the whole world, if we consider that the weather in a certain region (for instance one of the United States) is often determined by the thermodynamic characteristics of the air masses that a few days before were over another region, thousands of kilometres away (for instance the Pacific Ocean), and have been carried to the region under consideration by upper air currents. If there are no data about the thermodynamic characteristics of the air mass over its place of origin, and perhaps also of the state of the ground over which it is conveyed, it becomes very difficult to produce a weather forecast even for a period of only 24 or 48 hours. From a climatological point of view, the problems due to a lack of data about some regions of the world are equally evident: in case of conspicuous gaps in the observational network, climatic changes can be mentioned only with a lower degree of statistical confidence.

For about 50 years, the international-level efforts to harmonise a global observational network have pertained chiefly to the World Meteorological Organisation (WMO), an agency of the United Nations whose headquarters are in Geneva and that is now co-ordinating a global observational system, both for meteorological observations and for more strictly climatic measurements, including chemical and atmospheric-composition instrumental surveys.

As is always the case when discrete measurements are performed on continuous processes, the spatial distribution of the observational network and the timing of the sampling must be calibrated considering the scale of the phenomena under examination. Since at a meteorological level the prime goal is to obtain reliable so-called medium-range forecasts (1 to 7 days), the WMO recommends a spatial resolution of 50 and 300 km respectively for ground level and upper air observations, and a sampling interval not exceeding 3 and 12 hours, respectively. For climatic observations, the required space-time resolution is obviously much lower (looser network and less frequent sampling), because the processes to be monitored are slower and more homogeneous in space.

An aspect whose importance should be stressed is that the observations are carried out all over the world in a synchronous manner, that is at the same hour: in all the planet, any meteorological observation always refers to the solar hour of Greenwich (GMT: Greenwich Mean Time). This way, for instance, every day at 12 GMT the thermodynamic characteristics of the atmospheric fluid are sampled in order to obtain approximate information about its physical state.

2.4 Conventional Observations

We will now briefly consider the meteorological parameters that are covered by this monitoring. The meteorological stations on the ground and the ships employed for this service on the seas and oceans issue observation bulletins at least every three hours (they are called SYNOPs for the ground stations and SHIPs for the ships). These bulletins contain the encoded information about the pressure, temperature and humidity of the air, the direction and speed of the wind, the clouds (in terms of extension of coverage, height of base and type), the visibility, the quantity of rain that has fallen during a certain period of time, the temperature of the surface (of the soil or water), the thickness of the layer of snow, if any, etc. The bulletins are transmitted to a world-wide telecommunication network; on the average, there are approximately 15,000 of them at a main hour such as midnight GMT. Plate 1 shows an example of the global distribution of these observations: it has some gaps, even considerable ones, on the oceans and on the African continent. These gaps in the ground level observational network are only partly bridged by data that come from automatic buoys installed in the oceans (not shown in the plate).

The upper air observational network is obviously less close, partly as a consequence of the greater homogeneity of the atmospheric fluid far from the ground, which allows a lower-resolution monitoring, and partly because of the high cost of the installation and use of the radiosonde stations, which in many cases is unaffordable for developing countries. In these stations, a sounding balloon full of helium gas is raised into the atmosphere; it carries a box containing electronic instruments that measure the pressure, temperature and humidity at various vertical levels. On the basis of triangulation methods such as LORAN or GPS, which make it possible to locate the ascending system with a 1-metre precision, every 10 seconds of ascent of the balloon the distance covered by it during that period is calculated. Supposing that the horizontal movement of the ascending system follows that of the high-altitude wind, an estimate of the average wind relevant to the layer that is crossed (approximately 40 metres thick) is obtained; in actual fact, the speed of the wind is slightly underestimated. At present, the average number of soundings performed in the world at a main hour is approximately 700.

The radiosonde network is supplemented by the so-called PILOTs, in which a small sounding balloon is used only for estimating the speed of the wind at several altitudes, and by more modern instruments called "wind profilers", which perform the same measurement by means of an apparatus on the ground that emits electromagnetic pulses into the atmosphere and analyses the return echoes and the relevant Doppler shift. Moreover, observations and measurements at a single vertical level are usually performed by intercontinental scheduled flights, particularly on the North Atlantic routes.

2.5 Satellite Observations

All the measurements mentioned up to now (except the ones coming from the wind profilers, which cannot be specifically discussed in this book) are conventional measurements, i.e. they are performed with the classical instruments that exist since the beginning of meteorology (barometers, thermometers, hygrometers, anemometers and pluviometers), though in some cases the sensors that are used currently are different from the original ones. Now, however, there are also other types of measurements that are called non-conventional: they are chiefly those performed by the instruments aboard meteorological or Earth observation satellites. We will briefly explain the additional help offered by these measurements towards a more accurate monitoring of the atmosphere.

In this book we cannot fully discuss the physical principles of the so-called satellite "remote sensing"[3]. It will be sufficient to state that this

---

[3] For a discussion that is more complete, but still fairly comprehensible to an uninitiated reader, consult, for instance, Pease (1994).

kind of survey analyses the behaviour of the electromagnetic radiation that is emitted, absorbed and scattered through the atmospheric medium. Aboard these satellites there are both active and passive instruments: the former are real RADARs that emit an electromagnetic radiation and analyse its return spectrum and Doppler shift, if any; the latter may be regarded as simple digital cameras that are sensitive to various wavelengths of the ingoing radiation, basically from ultraviolet to microwaves, including the visible and infrared range. As a rule, the radiation that reaches the satellite depends on the thermodynamic state of the surface of the Earth and of the various layers of atmosphere that are crossed (and on their composition as well). The problem of how to use the analysis of the radiation received by the instruments aboard the satellite in order to obtain values of parameters that define the state of the ground or atmosphere should be tackled in each case on the basis of what one wants to obtain. Here we will only briefly mention the fact that, as we shall explain concisely further on, it is possible to estimate some quantities such as the temperature of the various layers of the vertical column under examination. The limit of these estimates consists in the fact that they allow us to determine only mean temperatures of very thick atmospheric layers; this limit is only partly technological, because it also involves some limitations inherent in the physics of the problem of remote sensing.

During the last few years, there has been a boom of initiatives for the programming, designing, construction and launching of satellites for meteorology and the observation of the Earth. Without entering into details, we will now briefly discuss only those observations that currently allow an extensive monitoring of the state of the atmosphere.

As regards their orbit around the Earth, satellites can be divided into two great categories: polar ones and geostationary ones. Geostationary satellites are put into orbit above the equator, at a height (approximately 36,000 km) that allows them to orbit the Earth along the equator at its same angular speed; this way they always observe the same portion of the planet. Polar satellites are brought to a lower height (approximately 800–900 km), and their orbit passes near the two poles. While their orbital plane remains constant with respect to the fixed stars, the Earth revolves under them[4]: this way they observe ever-different portions of the planet. Their orbital period is usually 100 minutes or slightly more, and during this period the Earth revolves by 25° or slightly more.

From the point of view of observation, geostationary satellites ensure the constant monitoring of a certain area of the globe, but with a low resolution (because of the considerable height at which the instruments are). On the contrary, polar satellites achieve a monitoring that is more discontinuous but has a higher resolution. As a result of these characteristics, geostationary and polar satellites complement each other in the global observational network. At present, for exclusively meteorological observations there exist 5 geostationary satellites (such as the European METEOSAT, known also to the general public) e 3 polar satellites. A further difference between the observations performed by the two different types of satellites consists in the fact that, whereas geostationary satellites can collect data at predetermined hours (e.g. 12 GMT) simultaneously over a vast area (basically most of the hemisphere over which they are), polar satellites are bound to the limited area they can observe at a certain instant. In order to be able to use the data of the polar satellites for an estimate of the state of the atmosphere at a certain instant, it is necessary to find the way to obtain the values relevant to that instant also on areas over which the satellite has passed a short time before or afterwards[5].

Besides supplying images that may be quite fascinating but are often in themselves unusable for a quantitative analysis, geostationary and polar satellites also issue routine data for meteorology: for this purpose there even exist encoded observation messages that can be identified in the global meteorological telecommunication system by the acronyms SATOB and SATEM.

In the so-called "thermal infrared channels" the passive satellite sensors are sensitive to the temperature of bodies (land, seas and oceans, clouds); the SATOB, observation message from geostationary satellites, supplies information about the temperature of the top of the clouds and

---

[4] The same thing happens, for instance, in the experiment of Foucault's pendulum: in both cases the reference system of the Earth reveals its character of non-inertial system.

[5] This problem will be examined in Chapter 7, when the four-dimensional analysis of the meteorological data will be discussed.

the direction and speed of the wind, calculated through the movements of the clouds, which are taken as a valid tracer of the flow of the atmospheric fluid[6]. If the data of a radiosounding performed near the region to which the SATOB refers are available, it is easy to obtain the height of the clouds by means of the double-entry chart represented by the vertical temperature curve supplied by the sounding: this way the direction and speed of the wind at a certain height can be estimated[7].

The SATEM is a message that comes from the polar satellites of the NOAA series. These satellites carry the TOVS (Tiros Operational Vertical Sounder), whose instruments measure the so-called radiance (practically the intensity of the radiation emitted along a vertical path and detected at the top of the atmosphere) in several ranges of the electromagnetic spectrum. From the viewpoint of the physics of radiative transmission in the atmosphere, when the thermal state of the vertical air column is known, it is easy to obtain the radiance that falls on the satellite. The inverse problem is trickier, but can be solved successfully in many cases. When the radiance at single frequencies is known, it is possible to determine the mean temperatures of several vertical layers of air. It is necessary to point out that these layers are very thick. So the TOVS actually achieves a vertical thermal sounding of an air column, but this sounding is much more averaged than the one achieved by classical radiosoundings: it has severe limitations when thermal structures that have a small vertical scale must be measured. These data, therefore, are very useful when there are no data coming from conventional radiosoundings, but their vertical resolution may be insufficient and they may actually be useless when more precise data are available[8].

The real great advantage of satellite observations consists in their global coverage (shown in Plates 2 and 3), which obviously is not affected by hostile logistic conditions (oceans, remote environments) or

---

[6] In this case too, as in the estimate of wind obtained on the basis of the movement of the radiosondes, the speed is usually somewhat underestimated.

[7] Besides this information, the SATOB also supplies the surface temperature, the percentage of cloudiness of a certain region, and data relevant to the humidity and to the ingoing and outgoing radiation.

[8] The SATEM also supplies data relevant to the precipitable water content of the clouds.

financial problems (management of the ground network by developing countries). From this point of view, meteorological satellites make it possible to eliminate the previously mentioned gaps in the conventional observational networks at ground level and, even more, in the upper air ones. In particular, the surveys performed by geostationary satellites cover, globally and with synchronous observations, all the area approximately from latitude 60° south to 60° north (above and below these latitudes, the data are not usable because of what is called a parallax error, that is because those areas are observed too much "on the skew" for the data to be valid). The polar satellites, moreover, cover these high-latitude areas with an excellent frequency and supply more details in other areas as well; there remains, however, the problem of the non-synchronicity of their readings. Lastly, the number of observations both of geostationary satellites and of polar ones is really enormous, and incommensurably greater than that of conventional observations.

The satellites we have mentioned up to now were designed exclusively for meteorological purposes. During the last few years, however, some satellites with more extensive purposes of observation of the Earth have been launched: some of the instruments they carry also supply meteorological information. An example of data that are now used routinely for meteorology is that of the surface wind fields on the seas and oceans, obtained by means of instruments called scatterometers, installed aboard the American/Japanese polar satellite ADEOS–II and on the European ones ERS–2 and ENVISAT.

Without entering into details, we may briefly state that the scatterometer is a radar that emits electromagnetic pulses and picks up the return signal reflected by the surface of the sea. The energy of this signal depends on the state of the sea: a rough or very rough sea reflects more energy than a sea that is almost calm or only slightly choppy. A series of devices, which we cannot describe in detail here, make it possible for this instrument also to obtain the direction and speed of the wind on the surface of a certain area. This information is crucially important, both because it cannot be obtained so extensively with conventional observations, and because the winds are measured near the interface between the air and the sea, where the influence of the ocean on the atmosphere appears.

2.6 Meteorological or Climatic Observations?

As we have previously stated, the countless meteorological observations that have been discussed in these pages also have an immediate climatic value if they are subjected to a post-processing in order to highlight the mean values, their variability and any possible trend over a period of at least a few decades. For this purpose, for instance, meteorological observation stations on the ground and radiosonde stations emit, with a respectively daily and monthly frequency, specifically climatic bulletins that summarise and highlight some quantities that are important for the reconstruction of the climate of the site under examination. However, as we shall see when we examine our theoretical knowledge of the Earth system and of the factors that affect the changes in the parameters and meteorological phenomena, it is necessary to consider the monitoring of other elements, such as the radiative exchange between the Earth and the outer space, the concentration of the constituents of the atmosphere, the characteristics of the oceans (from a physical, chemical and also biological point of view), and the characteristics of the Earth's ecosystems (including lakes and rivers, ice, flora and fauna, and the presence of human activities).

A more strictly climatic atmospheric monitoring is carried out with the help of observers usually situated in remote areas and by means of satellite observations: this monitoring obtains information relevant to the stratospheric ozone and ultraviolet radiation, ozone at the surface, solar radiation, concentration of certain gases called greenhouse gases, aerosol and dust in suspension, and acid rain. In particular the satellite surveys, combined with ground level observations, considerably help us to keep the overall condition of the entire planet under control, also from the viewpoint of the monitoring of the oceans and of the Earth's ecosystems.

The examples of satellite observations that have a climatic value are countless. We should mention the altimetry activity, which began as early as 1973 with the SKYLAB, and goes on, with increasingly sophisticated instruments, up to the latest altimeter installed on the ENVISAT: this has led to a world-wide monitoring of the mean level of the sea that is characterised by an excellent degree of accuracy. Another field in which the "eyes" of satellites are very useful is the evaluation of the extension of the snow and sea ice cover. As regards ice, we should mention the intensive monitoring undergone by Antarctica during the last few years: this has led, among other things, to an immediate alert when the previously-mentioned detaching of the Larsen B iceberg occurred, and (obviously even more important) to the evaluation of the disgregation process in that part of the Antarctic pack (see Plate 4, where this process is shown as it has evolved from 1992 to 2002).

Inter alia, precisely the polar pack is the main object of the investigation performed by the recent ICESAT of the NASA. On ground areas that are not covered by water or ice, it is extremely important to carry out an analysis of the vegetation cover, with particular reference to the monitoring of the phenomena of drought and desertification, and of the anthropogenic changes in the use of soil: all this contributes, in particular, to the estimate of the so-called "albedo", i.e. the ratio of the energy reflected in the space by the Earth, clouds and atmosphere to the total incident energy coming from the sun.

Without going into the details of more strictly meteorological satellite observations, which have also been mentioned previously, we should point out that it is possible to obtain an estimate of the precipitations on areas not covered by ground level pluviometers (particularly on the oceans). The solar activity is also monitored with a greater accuracy than that offered by ground level observations, since the latter are troubled by the interposed atmosphere: it has been possible to accurately measure the total solar irradiation, which was previously called the solar "constant" and on the contrary was found to have evident fluctuations within an 11-year cycle. Moreover, the various instruments on the most recent satellites (once again we should mention the European ENVISAT) make it possible to monitor the ozone (both the stratospheric one and the one closer to the ground), to estimate the emission spectra and concentration of climatically relevant gases, and to perform a colourimetric analysis of the seas and oceans. The latter is important for determining the thermal state of the upper part of the seas and oceans and for obtaining information about the oceanic part of the so-called carbon cycle, which will be briefly discussed further on.

As the reader has probably understood, satellites make it possible to perform observations of the Earth that are undoubtedly more global, and sometimes also more accurate than those carried out by ground level stations or instruments in the atmosphere. The limit in the use of these data from the angle of climatic researches consists in the fact that they are quite recent: at best, as in the case of altimetric observations, we have historical series that are about 30 years long, but in most cases the surveys do not go further back than one or two decades. Sometimes we have data relevant to a very limited number of years.

2.7 Proxy Data

We have explained that nowadays the global observational network allows an accurate monitoring of the meteorological and climatic health of our planet; but which instruments do we have for understanding how the climate was when this network was not so extensive, or even when there did not exist any instrumental measurements? Climatologists, we have said, are particularly exacting on this point. We cannot blame them! How can we judge the changes that are taking place during the last few decades if we cannot compare them with those of other periods (when, moreover, human activity was not able to disrupt the balance of nature)? An attempt to remedy this lack of information has been made by analysing long local historical series of quantities from which climatic data can be reconstructed (called "proxy data"). We will now give a few short examples of these reconstructions.

The data that are potentially most interesting and make it possible to go far back in time (so far back that sometimes the term "paleoclimatology" is used), are those that come from the deep oceanic sediments and those from vertical soundings (called core samplings) in the very thick layers of ice mostly present in Antarctica and Greenland. In the former case, an examination is carried out on the remains of the shells of small animals, such as foraminifers, that have accumulated on the bottom (these shells are chiefly composed of calcium carbonate). In the latter case, a direct examination is performed on the ice and on what has been trapped in its interstices. In both cases a particular attention is given to the analysis of the oxygen atoms present respectively in the calcium carbonate and in the ice.

The reason for this is that in nature oxygen appears essentially in the form of two stable isotopes[9]: ^{18}O and ^{16}O. On the average, the ratio of the concentration of ^{18}O to that of ^{16}O is approximately 1/500, but it changes slightly with variations in environmental factors such as temperature[10]. This leads to the conclusion that a careful analysis of the oxygen in the calcium carbonate stored in the plankton shells and extracted from the surrounding water may supply information about the temperature of that water during the lifetime of the little animal under examination. Likewise the analysis of ice should reveal the temperature of the snow that had formed in the air above it during the period under consideration. Obviously these analyses must be combined with a reliable dating method.

In actual fact the situation is more complicated! Out of the controlled conditions of a closed laboratory system, it is possible for the ratio between the two oxygen isotopes to change simply because a certain quantity of one of the two is removed from the system. In the seas and oceans that supply oxygen for the storage of calcium carbonate, many water molecules are removed because they evaporate, while others return with precipitation. In particular, during the evaporation stage, on the average more "light" oxygen atoms (^{16}O) than "heavy" ones (^{18}O) are removed. Considering the matter exclusively at a global level, if all the evaporated water returns to the sea with the precipitation (which therefore contain more ^{16}O), the cycle is closed and the ratio between the isotopes remains constant. But what happens during particularly cold periods, when the precipitation (snow) thickens the layers of the polar and mountain glaciers or forms new layers? In this case, the concentration of ^{16}O in the sea decreases and the ratio between the isotopes changes.

---

[9] The reader is reminded that a chemical element is determined by the number of protons contained in its nucleus, and that, if this number remains equal, but the number of neutrons changes, various isotopes of the same element are obtained. In the case under consideration, ^{18}O has two neutrons more than ^{16}O, i.e. it is heavier.

[10] The theoretical change in this ratio is 0.2 parts per million with a 1-degree variation in temperature: though it is very small, it can be revealed by the currently available technology.

In brief, the effects on the variations in the ratio between the isotopes under examination are due to changes in temperature, to evaporation and precipitation, and to the world-wide quantity of accumulated ice: the only measure that is available to us, the ratio between the isotopes, is therefore a function of these variables, which, *inter alia*, are not mutually independent (for instance, a decrease in temperature usually corresponds to an increase in the formation of ice). The usual approach — which consists of unravelling the skein by reversing the problem in order to infer water surface temperature values from isotope measurements performed in this highly interacting system — requires a high degree of theoretical knowledge of the system and of the interactions between the various phenomena. Considerations of this sort are valid also for core sampling.

By examining the samples of air trapped in the interstices of the ice, core sampling also supplies a quantitative evaluation of the presence of some gases in the atmosphere, in particular carbon dioxide (CO_2) and methane (CH_4), during the period under consideration. Coral skeletons, which are composed of carbonate, supply information about the temperature of tropical seas, when the same method that has been adopted for deep-ocean sediments is used. The same technique is also used to analyse organic sediments in lakes: it is worth pointing out that this has recently made it possible to link changes in climate to historical problems, such as the study of the disappearance of ancient civilisations[11].

A comparative analysis of the climates of the past can be performed also by examining the so-called "fossil pollen" in the sediments mentioned above. This makes it possible to find out which types of plants were present during a certain period: from their northward or southward shift in various geological sites (or from the presence of pollen belonging to different types of plants in the same site), we can approximately infer the type of climate of the various eras.

Lastly, plants contribute to the estimate of the climates of the past by means of the analysis of the annual rings of the trunks. Obviously this method can go back only a few centuries. However, since trees produce a

---

[11] About the disappearance of the Mayas, consult Hodell *et al.* (1995).

ring every year, at least in the temperate regions where there is a definite growing season, these historical series have a high temporal resolution. The growth of trees depends on the temperature and precipitation, so when there are warm, rainy years the rings are wide, and when there are cold, dry years they are narrow. If trees are analysed in areas with distinct characteristics, it is possible to obtain information only about the temperature or only about the rainfall: for instance, in typically warm and dry low-latitude regions the critical parameter for the growth of trees is the rainfall (so wide rings mean rainy years and narrow rings mean dry years); in typically cold and humid higher-latitude areas, the critical parameter is the temperature (so wide rings mean warmer years and narrow rings mean colder years).

Since, as we have explained, all these reconstructions lead to more or less accurate estimates of climatically important parameters (often independent of each other and based on different sites), during the last few years "multi-proxy" syntheses have been considered for the estimate of the global or hemispheric surface temperature, achieving various high-resolution reconstructions. This way, it has been possible not only to obtain estimated values for the temperature over the last 1,000 years, but also to define the error bars to be attributed to these values, which are essential for evaluating the reliability of the estimates. For the northern hemisphere during the last 1,000 years, for instance, the error can be estimated at ± 0.5°C up to the year 1600 and at ± 0.2°C from 1600 to the end of the nineteenth century: during the last century the error further decreases, gradually. The estimates obtained by means of proxy data that go further back in time are obviously impaired by a greater uncertainty, and usually also by a lower temporal resolution.

Reconstructions obtained by means of proxy data, in short, are indispensable for obtaining information about the remote past. Within the limit of the error connected with an estimate — which is obviously greater than that connected with direct instrumental measurements — proxy-data reconstructions in any case can detect marked world-wide changes in climate. From a conceptual point of view, it is important to point out that, in order to obtain these reconstructions, evidence based on empirical observation is not sufficient: it is necessary to rely on a theoretical knowledge of the mechanisms of the interactions between the various processes and phenomena that take place in the system under examination. We have seen an example of this in the analysis of marine sediments, where it was found to be necessary to know the theoretical balance between ^{16}O and ^{18}O with changes in temperature and the dynamics of the water cycle (evaporation, precipitation, ice formation).

2.8 Is There Any Evidence that the Climate is Changing?

Once the entire database we have just outlined is available, it becomes interesting to investigate whether these data, in themselves, give an indication of the position of the climatic changes of the last decades in relation to the variability of the more or less recent past. An analysis of this type is performed with statistical methods that have been well-known for a long time and have been applied extensively to climatic data during the last few years. Since here we cannot go into a detailed analysis of the statistical methods that are used, we will translate the results that are obtained into an everyday language, basing our explanation on the concept of probability and supposing that the reader is acquainted with it. Moreover, while referring the reader, for a more detailed analysis, to Chapter 2 of a recent report of the IPCC[12], we will only consider the results that are believed to be verified and are currently regarded as a common property of the international scientific community.

By post-processing the observation data for climatic investigation purposes, we can obtain information about the mean values of some parameters, their variability with time and the occurrence of extreme events. On the basis of these parameters, it is possible also to reconstruct the estimated mean states and trends of the atmospheric and oceanic circulation. As for the variability of the climate, it can be studied on various time scales: there exist an *interannual* variability, a ten-year or hundred-year variability, and a variability on a paleoclimatic scale. The first one is characterised by evident oscillation phenomena in the oceanic circulation interacting with the atmospheric component (such as ENSO, *El Niño Southern Oscillation*), or, more directly, in the atmospheric

---

[12] Houghton *et al.* (2001).

circulation (such as the NAO, *North Atlantic Oscillation*); the third variability is relevant to the transition from glacial eras to interglacial periods and vice versa, during the last few million years. If we wish to attempt to place the climatic changes of the last few decades inside or outside a natural climatic variability, and to understand whether during this period the influence of human activities has been decisive for these changes, we may in actual fact temporarily disregard these two scales of variability[13] (for which the reader is referred to the existing literature[14]) and concentrate on the ten-year or hundred-year variability.

While referring the reader, once again, to the previously-mentioned IPCC report for a more complete overview, we should mention some of the safest observational results: the statements that follow are regarded as very probable (there is a less-than-10% possibility of their being disproved) or practically certain (the possibility of error is less than 1%). These highly reliable climatic indicators are:

- an increase in the temperature of the surface of the sea by 0.4 to 0.8°C since the end of the nineteenth century;
- an increase in the temperature of the air over the surface of the sea by 0.4 to 0.7°C since the end of the nineteenth century;
- an increase in the temperature of the air over the land surface by 0.4 to 0.8°C since the end of the nineteenth century;
- a massive shrinking of mountain glaciers during the twentieth century;
- approximately two weeks less of ice formation in high-latitude lakes and rivers since the end of the nineteenth century;
- a 10% decrease in the spring snow cover in the northern hemisphere since 1987, in comparison with the mean values of the period 1966–1986;

---

[13] As a matter of fact, while obviously in the paleoclimatological sphere the influence of man is completely negligible, a debate is currently under way about the degree to which this influence can affect the cyclic course of the ENSO and NAO and perhaps also intensify them.

[14] About interannual variability, and in particular about ENSO, see Philander (2004). About variability on a paleoclimatic scale and the glacial eras, information can be found in the excellent book by Alley (2002).

- a 5 to 10% increase in the precipitation at high and medium latitudes since 1900, in many cases due to very intense events;
- no significant global change in the frequency and intensity of tropical cyclones.

Without presuming to give a comprehensive picture of the present situation in relation to the last few decades or centuries, it is interesting to consider the sole parameter of the temperature and to examine its variations during these periods. In order to achieve an integrated evaluation of the contributions to the global warming or cooling, we can reconstruct a combined historical series of the temperature of the air near the Earth's crust and of that of the surface of the sea, obtained by means of direct instrumental measurements[15]. This way we can obtain graphic representations like the one in Plate 5. It shows the so-called annual thermal anomalies of the period 1860–2000 with reference to the mean of the period 1961–1990 (taken as zero on the ordinate axis). The downward red columns indicate colder years in comparison with the mean of the period 1961–1990; the ones that move upwards towards positive values indicate warmer years in comparison with the same mean. In any case, the values of the departures from the mean of 1961–1990 can be inferred by the length of the columns and by the relevant values that can be evaluated by interpolating them on the ordinate scale. Each column in the histogram is associated with an error bar that is affected by the reliability of the measurements, particularly in relation to the degree of global coverage of the observational network: as a rule, therefore, these intervals gradually decrease from 1860 to now[16]. Lastly,

---

[15] The reason for this choice will become fully intelligible in Chapter 4, when the temperature variations on the surface of the land and sea will be discussed. At present it will be sufficient to state that this historical series is an average of more evident yearly fluctuations (on the land surface) and less marked variations (on the sea). In any case, the sign of these fluctuations is the same in almost all the years that have been considered (except for a small number of events, in which the absolute value of the variations is extremely small).

[16] We must point out, however, that in the error bars, together with the measurement sensitivity and global coverage, an increasing uncertainty is considered: it is due to the phenomenon of urbanisation and to the so-called "urban heat islands", which may affect the measurements obtained in constantly-developing cities. Here the temperature may increase year after year for local reasons, connected with the fact that the cities, made out of materials that "entrap" heat (basically asphalt and cement), are becoming more

the black curve shown in Plate 5 basically represents an average trend of the temperature that makes it possible to "filter" its yearly fluctuations, since the latter, per se, might be connected only with the intrinsic variability of the climatic system and therefore not be very valuable for the purpose of identifying a trend in the global temperature.

The general trend revealed by Plate 5 is undoubtedly that of an increase. More in detail, there are two *plateaux* of almost-constant values from 1860 to 1910 and from 1945 to 1976. Two periods of increase in temperature by approximately 0.15°C per decade appear from 1910 to 1945 and from 1976 to 2000: these increases are statistically significant, because they amply exceed the indetermination limits allowed by the error bars associated with the measurements. To this we may add (to satisfy the reader's curiosity) that the data relevant to the years 2001 to 2004, not included in the plate, would appear in this histogram as the warmest years after 1998. Obviously it is not possible to regard the data relevant to individual years as statistically significant, but, for instance, the fact that 2001 was a particularly warm year all over the world is considered important, because, contrary to 1998, when ENSO was going through the El Niño stage, with the tropical part of the Pacific Ocean very warm[17], in 2001 this part of the ocean was considerably colder.

Is the evidence supplied by the observations summarised in Plate 5 sufficient to allow us to state that, at least during the last 30 years, we have been experiencing a period of global warming that is not due to the natural variability of the climate? Though it is not the purpose of this book to intervene polemically in the current debate about global warming and its causes, it is worthwhile to give the reader a further contribution of observations to be considered.

As we have previously explained, the analysis of proxy data, though it relies heavily on our theoretical knowledge of the system under examination, supplies information about the global temperature during the last few centuries. Considering only the northern hemisphere (whose sites are the sources of approximately 95% of these data), Plate 6 shows

---

extensive: this may hide the effects of a global warming or cooling. In Plate 5, all this leads to the fact that, starting from the years of the Second World War, though the observation network is constantly growing, the error bars tend to increase slightly.

[17] For a more accurate analysis of this episode, see Philander (2004).

a reconstruction of the temperature anomalies of the last millennium, with reference to the usual thirty-year period 1961–1990, in association with error bars. Notice that the trend of the temperature of the first 9 centuries is almost constant (or slightly decreasing) and that there is a tendency to a marked increase during the twentieth century. Moreover, all the data of the last few years exceed any previous error bar: this means that these years have been the warmest in the millennium (in a statistically significant manner) and that the values of the present warming exceed the climatic fluctuations typical of the ten- or hundred-year variability of the climate.

All this obviously does not lead to a univocal conclusion, but at least it narrows down the range of possibilities to two hypotheses (which we are not yet able to explore on the basis of the information contained in this chapter): the present warming may be the result of a broader-scale natural variability (that we have called paleoclimatic); or it may be the evidence of a perturbation in the natural variability due to human activities. In order to be able to lean to one of these two hypotheses, we need a knowledge that greatly exceeds the limits of what we have endeavoured to describe in this chapter; in particular, we will have to break away from a paradigm that is purely based on observation, to acquire a stock of sound theoretical knowledge about the system under examination, and also to adopt a new approach to scientific research.

Before tackling this subject, like good detectives who must solve a difficult case, in the next chapter we will look for further observation-based clues that may lead us to a track for discovering the motive, i.e. the causes that may have led to this global warming. It is obviously premature to try to find out whether the serial killer may kill again in the future, i.e. whether the global warming may go on in the course of this century.

Chapter 3

Naive Meteorology, Coincidences and Correlations

The previous chapter helped the reader to understand which instruments and methods of observation are required for achieving a (necessarily approximate) definition of the thermodynamic state of the atmosphere system. In doing this, it identified some significant variables in what we perceive every day as the weather. These variables, analysed on a broader time scale, make it possible to define the climate of the place where we live. Finally, the determination of some time series of these variables gives an idea of how the state of the system has evolved in the more or less recent past.

Obviously all this information about the system under examination is extremely important, indeed essential, because it forms the grounding needed for any further analysis of the system. However, like all empirical data pertaining to systems slightly more complex than those we studied at school (which were based on the simple Galilean mechanics), these data do not allow us to easily find, among the variables, evidence of the existence of specific relationships that may be regarded as causal and may therefore lead to an "explanation" of the "operation" of the system under examination[1].

---

[1] The reader is referred to the concise analysis of the concept of causality in the Introduction, and reminded that, from a common-sense point of view, a phenomenon is explained if its causes and their way of combining to produce it are known. Moreover, the operation of a system is understood if the relationships between the variables that define its state (both at that time and in their evolution in time) are known. In any case, the concept of the understanding or intelligibility of a system will be discussed more thoroughly in Chapter 4.

It is clear that the problem can be tackled from the angle of basic sciences such as physics and chemistry. In this context, for instance, the air that forms the atmosphere is nothing but a fluid, a mixture of several gases and water[2], plus other constituents such as aerosols (suspended particles), to which it is possible to apply the theoretical knowledge relevant to disciplines such as fluid dynamics and thermodynamics. Obviously this is done, and this will be the subject matter of the next chapter (together with the recognition that the system to be considered in order to understand the dynamics of the weather and climate exceeds the limits of the atmospheric fluid). For the time being, it will perhaps be more enlightening to perform a different exercise.

3.1 Approaching an Analysis of the Data and of Common Experience

Probably all of us have noticed that present-day society is often defined as based on information and knowledge. As we have already remarked, the flow of news and data that sweeps over us every day is enormous, and the possibility of finding information intentionally and independently is almost unlimited, thanks, above all, to Internet and the global network. However, the use and interpretation of all this information depend on our knowledge of the individual subjects.

In this context, it is reasonable to believe that having the know-how required for a certain subject is more important than learning by rote and "accumulating" uninterpreted data. Possessing know-how about a certain system, as a matter of fact, means knowing the rules that determine its behaviour and therefore mastering a paradigm for the explanation and interpretation of the phenomena that may take place within it. The so-called "experts", who are often interviewed on television or in the newspapers, are people who possess a thorough knowledge of a single area of interest; their domain of competence, because of the extreme specialisation currently present in science, is usually quite limited.

In actual fact an interesting phenomenon is taking place at present, and is extremely evident, for instance, in the labour market: rather than a

---

[2] Water in its three aggregation states: solid, liquid and vapour.

person's specific know-how in a certain area, what is considered important is his or her mind-set, moulded by the experience of school and work. A methodological approach to data analysis and problem solution, and flexibility in applying this approach to unknown systems, are regarded as more important than the acquisition of competence in a specific area. This is particularly to the benefit of science graduates (above all physicists), who sometimes are forced to change area of interest, but manage to find jobs just the same, because of their general qualities as problem solvers. When this occurs, they have to tackle systems about which there generally is a good empirical knowledge but a limited theoretical understanding[3].

What is so special about the mentality of a young person who "studies to become a scientist", though perhaps he/she will never eventually work as a scientist? Undoubtedly it is this person's ability to analyse a certain system, even if it was previously unknown to him/her, in order to find a paradigm that explains it and may make it possible to act upon it. There is a saying that scientists are big babies. It is true that the behaviour of a scientist when faced with an unknown problem is very similar to that of a small child when interacting with the external environment during the course of his logic (inductive and deductive) development. He looks for regularities and causal relationships, performs tests if the system allows it, and works out his subsequent actions on the basis of his previous experience. In particular he develops his own naive physics, which is sometimes erroneous and, unless it is investigated more thoroughly, may lead to typical common-sense errors[4].

At this point of our discussion, we are really in the condition of a child who is facing something unknown. Let us therefore attempt to

---

[3] A concrete example is that of some physicists who have started dealing with economy or finance, applying to this field some methods and techniques that pertain to theoretical physics: nowadays this is called *econophysics*. Conferences are held and books are written about this topic. At a more practical level, banks and companies that act in the stockmarket are obviously interested in this type of analysis.

[4] In international literature, naive physics is now covered by a great number of studies (see, e.g., Hood (2004)). As regards the difficult relationship between modern physics (particularly relativity and quantum mechanics) and common sense, chiefly due to the different physical domains, it is advisable to consult classical writings like Einstein and Infeld (1967), and the previously mentioned Ghirardi (2003).

analyse the system we are considering, about which we possess a considerable quantity of empirical evidence, in order to catch some relationships between the previously-mentioned variables, if possible in a shrewd, discerning manner. The "big baby" who will accompany you in this trip has already covered a similar ground, when, after having taken a degree in physics with a thesis about the unified theories of gravity and other interactions, he proceeded, initially only for occupational reasons, to take an interest in meteorology and climate.

The exercise we are about to perform is not pointless. This is demonstrated by the fact that some people who only have a naive knowledge of physics — for instance old farmers and fishermen — manage to forecast the short-term weather in their territory in an accurate and sometimes surprising way, on the basis of observations they have accumulated during their lives and of interpretations they have worked out over the years. In comparison with these experts of local meteorology, we are at a disadvantage, because we are no longer used to examining the sky to find clues and forewarning signs. On the other hand, we possess a quantity of meteorological and climatic data that are unknown to them and a better knowledge of basic physics, at least as it is acquired nowadays during the compulsory-schooling years[5].

In what follows, we will consider only an extremely limited number of observations: they come both from assessments that everybody can make about phenomena connected to the weather, and from instrumental data about the weather and climate now available to us. This will lead us to point out some particular coincidences, which we will try, as far as possible, also to interpret on the basis of common experience or of the knowledge of basic physics that we can take for granted. Since, as we have already explained, by definition the climate is the sum of the meteorological events (in terms of averages of the individual variables, scattering of values around the averages and number of extreme events),

---

[5] Obviously, even if we are analysing a system that is "unknown" *a priori*, we are doing it with the conceptual instruments that are available to everybody: it would be nonsensical to pretend to be ignorant of certain basic physical properties that by now are common knowledge. As we shall show further on, this does not avoid the risk of giving incorrect interpretations of what takes place in the atmosphere.

we will start by analysing some meteorological observations, then we will examine some data more specifically relevant to the global climate.

3.2 A Naive Interpretation and Its Problems

It is common belief, and also an established scientific result, that the Sun (together with water and the atmosphere) is what makes the existence of life possible on the Earth. The Sun, in particular, supplies the radiant energy required to preserve a temperature that is suitable for life. The feeling of the sunbeams penetrating our body and warming it up is undoubtedly one of the most appreciated pleasures of life: this is demonstrated, for instance, by the presence of crowds of Italian and foreign tourists on the Italian beaches.

This common experience of "body warming" by "absorption" of the beams coming from the Sun obviously leads us to believe that the same thing happens to the air masses. This opinion is confirmed by several pieces of evidence coming from observation. For instance, if we consider the same period of the year and temporarily disregard the direction from which the wind comes, we notice that when the sky is clear and the weather is sunny the air is usually warmer than on cloudy days. In particular, during the morning the air becomes gradually warmer as the Sun rises above the horizon.

The latter fact becomes comprehensible when we consider that, if the quantity of energy emitted by the Sun during a certain lapse of time (e.g. an hour) is regarded as constant, before it reaches the ground (or our body), a certain part of this energy is probably absorbed by the atmosphere through which it passes. At this point, when the Sun is low on the horizon, the path of its beams in the atmosphere is much longer than when it is near its zenith. This way, in the morning and evening the amount of energy (heat) per time unit that reaches the air near the ground is presumably much smaller than the amount that reaches the ground in the central hours of the day, when the Sun is higher on the horizon.

The effect we have just described also has obvious seasonal consequences: it is well known that, at the same hour of the day, during the summer the Sun is higher on the horizon than it is in the intermediate seasons and, even more, in the winter. We must also allow for the fact that the duration of the day is longer in the summer. The result is that the average temperatures in the summer are higher than in the spring and autumn, and, even more, in the winter. So these facts, too, seem to confirm our vision.

A further consequence of what we have said is that we can predict that the average warming of the air (near the ground) will be greater in the tropical areas than in the air at the intermediate latitudes and at the poles (at the same altitude). And once again this prediction is fulfilled. If, at this point, we consider a phenomenon we have neglected up to now, the horizontal conveying of air from a region of the globe to another, in particular between regions at different latitudes, we notice that, at medium latitudes and in the northern hemisphere, the winds that come from the south are warm and those that come from the north are cold. This is consistent with what has been discussed just now about the warming undergone by the air at several latitudes. Obviously, therefore, the temperature of the air at ground level in a certain site is determined by the combined effect of solar irradiation on the place under examination and of the changes in temperature due to the arrival of air whose thermal characteristics are different from those of the air previously present in the site.

This way, on the basis of an analogy with the common experience of the warming of our body by the Sun and of observations that can certainly be shared by everybody, we have acquired a certain qualitative, specific knowledge about the phenomenon of the warming of the air; we have practised naive meteorology and now possess an explanatory scheme of this phenomenon.

Within the sphere of physical sciences, it is commonly repeated that an explanatory scheme (which, if formalised, is more pompously called "a theory") is regarded as valid until an empirical piece of evidence that cannot be explained by it is found[6]. So let us now widen the scope of our

---

[6] In actual fact there may be other reasons for which a theory is dropped even if it is not in contradiction with empirical data: an example is the excessive complication of its theoretical system in comparison with that of a more simple, elegant theory. For many years there has been a controversy about the influence (or absence of influence) of the cultural, social and economical climate on the development of science and in particular of

observation and find out whether other pieces of empirical evidence are compatible with our idea of the air-warming mechanism.

Up to now we have always considered daytime situations in which the Sun shines on the atmosphere. What happens during the night, when the Sun is below the horizon? It is reasonable to suppose that when the energy that comes from the Sun and causes an increase in temperature is absent, the air dissipates the accumulated heat towards the outer space and the ground. If you ask an inhabitant of a continental region (i.e. one that is distant from the sea) how the winter nights are there[7], you will find that the temperature of the air near the ground is very low when the sky is clear and starry, whereas it is higher if the sky is overcast. Disregarding, to begin with, the influence of the stars on the temperature values (a study we will gladly leave to the astrologers), it will be necessary, in any case, to consider a new element in our explanatory scheme. Clouds, which during the daytime screen out a part of the solar radiation, thus justifying the decrease in daytime warming when the sky is cloudy, during the night seem to produce the opposite effect. It is arguable that they act as a screen against the loss of heat (in the form of radiant energy that comes out of the air molecules) towards outer space, but we are not in possession of information that allows us to understand their interaction with this energy (maybe they reflect it back?).

Thus the study of night-time phenomenology shows how important it is to know something more about the interactions between radiation and matter, particularly clouds. We must remember, in any case, that the interaction between the solar radiation and the air has not been studied in this scheme. This does not mean *a priori* that our explanatory scheme is incorrect; but it does confirm that the scheme needs to be re-examined in the light of the new knowledge, in order also to explain the phenomenon of the more or less marked night-time cooling.

---

explanatory theories. We do not mean to take part in this controversy here, and only wish to state that in any case a conflict with a piece of empirical evidence that is not included in the explanatory scheme leads to the rejection of that scheme, at least in its original form (the scheme may sometimes survive with some changes, provided they are not made *ad hoc*).

[7] As a matter of fact, this phenomenon is not limited to the continental regions, but occurs wherever there is a solid surface; a continental region is only a place where this phenomenon is more evident, for reasons that will be explained in the next chapter.

Up to this point we have considered only observations and experiences that are shared by everybody, including farmers and fishermen. Though obviously the latter are more attentive observers of atmospheric phenomena than we are, we have the advantage of being able to exploit more "exotic" experiences and the currently available instrumental data. For instance, people who frequently travel by plane have certainly noticed that the information usually supplied by the pilot of the plane includes the temperature of the external air at the flying altitude: it is normally about –60°C. And the data coming from the radiosondes, which have been discussed in the previous chapter, reveal that the temperature of the air decreases with the increase in altitude[8]. How can we fit these facts into our explanatory scheme?

From our interpretation of air warming by solar radiation there should follow immediately, at least during the daytime, that there is a practically uniform warming on a vertical column, or that the warming is actually greater in the upper layers of the air, where a part of the radiation that falls first is presumably absorbed, and where clouds (which are at a lower level) cannot screen out the beams coming from the Sun. Perhaps, as we did previously for clouds, we should include some new element in the scheme under consideration: for instance, what is the influence of the fact that, in the atmosphere, pressure decreases as altitude increases? Can this piece of evidence (which appears clearly, once again, in the data coming from the radiosoundings) explain the lower temperature in conditions of lower pressure at high altitudes? In a closed room, too, the air is layered and the pressure decreases with height, yet the warmer air is near the ceiling and the colder one near the floor.

If these data — some of which are not available to fishermen — make trouble in our explanatory scheme of atmospheric warming, it is natural to suppose that the solution may come through our greater knowledge of basic physics, in comparison with the fisherman's knowledge. Since the time in which Isaac Newton "unified" the terrestrial and celestial motions, for instance reducing the explanation of ballistic motions to a particular case of his laws of universal gravitation, physicists have never

---

[8] At least up to an area, called tropopause, where this tendency is inverted.

stopped believing in the universality of the laws of nature[9]. This means, for instance, that we believe that masses attract each other in the same way on the Earth and in space. In particular, in a different ambit, if I consider a certain fluid, for instance air, I believe that its behaviour is described by the same laws both in the free atmosphere and in my house. What may change, obviously, are the initial conditions of the system and the so-called "boundary conditions" (for instance, the atmosphere does not have a ceiling, while my room does).

We can believe, therefore, that my room is a good place for studying the properties of air — the same air that forms the atmosphere. In particular, the vertical thermal layering is due to an effect known as Archimedes' principle, whereby a "bubble" of fluid undergoes an upward thrust if its density is lower than that of the surrounding fluid (by applying the law of state of gases, it is possible to demonstrate that for air this is equivalent to the fact that the temperature of the "bubble" is higher than that of the surrounding fluid). An example of the validity of this principle is obtained by observing the ascent of the warm air coming from a hot radiator. An accurate measurement would also reveal that the warm air that rises from the radiator actually cools down slightly before it reaches the ceiling; this effect cannot be explained simply by molecular diffusion, i.e. by the fact that the air on the surface of the "bubble" gets mixed with the surrounding air. With an extrapolation, we might infer from this that if there were no ceiling, the ascending warm air would go on cooling, thus creating a vertical thermal profile in which the temperature decreases with height. This, however, can occur only if there is a source of heat in the lower layers.

The above "domestic" example is a prototype of an explanatory scheme that might enlighten us on the causes of the actual vertical thermal structure of the atmosphere and on the dynamics of its formation. But it requires a source that supplies heat from below, and this is not consistent with the previously-developed paradigm of solar warming, which explained "most" of the empirical evidence. We have thus reached a stalemate. Once again it seems to be important not only to perform a

---

[9] About the concept of law of nature, consult the excellent Barrow (1988).

more careful analysis of the motions of air fluid, but also to reconsider the interaction between radiation and matter in the atmosphere.

For the time being, it is worthwhile to point out that the decrease in temperature with the increase in height is not the only phenomenon that is not explained by our scheme. I would like to mention the fact that (contrary to the previous statement that at the medium latitudes of the northern hemisphere the winds coming from the north are cold) in the regions immediately south of the central and western Alps and Prealps there occasionally is a northern or north-western wind that is warm (the so-called *föhn*). This is particularly surprising if we consider that this wind comes directly from the Alpine range, that is from altitudes where the temperature is decidedly lower than in the Po Valley.

3.3 Coincidences and Correlations in Available Data

At this point, since we are not able to supply an explanatory scheme that is self-consistent and reproduces all the above-mentioned phenomena and properties of the atmosphere, we are at a cross-roads: either we delve more deeply into a general study (if possible an instrumental one) of the interactions between radiation and matter, or we consider other empirical pieces of evidence that help us define the problem better and give us some clues on the variables that are important in this study. In actual fact, in the course of the history of science there have repeatedly been alternations of periods of great theoretical syntheses and periods in which there prevailed data that came from observation and had not been completely understood from a theoretical viewpoint. In the latter case, particularly in observation-based disciplines such as meteorology, the search for evidence and correlations between data relevant to different variables helps us to establish correct relationships between them and to understand their interdependence, though it does not often lead to the establishment of causal relationships.

In particular, the concept of "correlation" (or "cross-correlation" in the case of the analysis of the so-called forewarning signs) is precisely the one that farmers and fishermen use unconsciously when they point out relationships between observations of different kinds, for instance between the presence of a cloud having a certain shape above a certain mountain and the appearance of precipitation in the place under consideration after a certain lapse of time. Beyond any sort of interpretation of the phenomenon in terms of a more or less naive meteorology, the discovery that, on the basis of previous experience, for instance 8 times out of 10 the appearance of that cloud is followed within 3 hours by rain on that area is extremely important and useful. It means, evidently, that there is a statistically significant link between these two phenomena, though it is not possible to determine the causal relationship between them: first of all it is likely that the clouds, that bring rain there, are other ones; moreover the fact that rain does not come in 100% of the cases means that some other phenomenon (that has evidently been overlooked) also affects the appearance of rain in that place.

The concepts of correlation and cross-correlation are by now well established and can be expressed by mathematical formulae. Without going into the details, we will simply state that two variables are positively correlated if algebraically low[10] (high) values in one of them correspond to low (high) values in the other one, at the same instant (correlation) or at a subsequent instant (cross-correlation). The opposite is true of an anticorrelation (or negative correlation). As a rule, if two variables are positively correlated in a very marked manner, either there actually exists a causal relationship between them, or they are both affected by other factors (for instance the trend of a third variable) that "force" them, tuning their behaviours and making them similar. In any case, the analysis of correlations can allow us to discover, in the system under examination, the existence of variables that had previously been overlooked and may turn out to be important.

Confining ourselves to a very limited number of correlations between different variables, we will now analyse some historical series of global temperatures, as we had done at the end of the previous chapter. In particular, we will analyse Figure 2. It shows the well-known anomalies of the global temperatures[11] of the last few decades in the so-called troposphere, that is in the lower layers of the atmosphere, from the

---

[10] Sometimes negative values.

[11] This time with reference to the average of the period 1979–1990.

ground to an altitude of approximately 8 to 18 km (depending on the latitude and season), and the corresponding temperature anomalies in the low stratosphere, right above the troposphere.

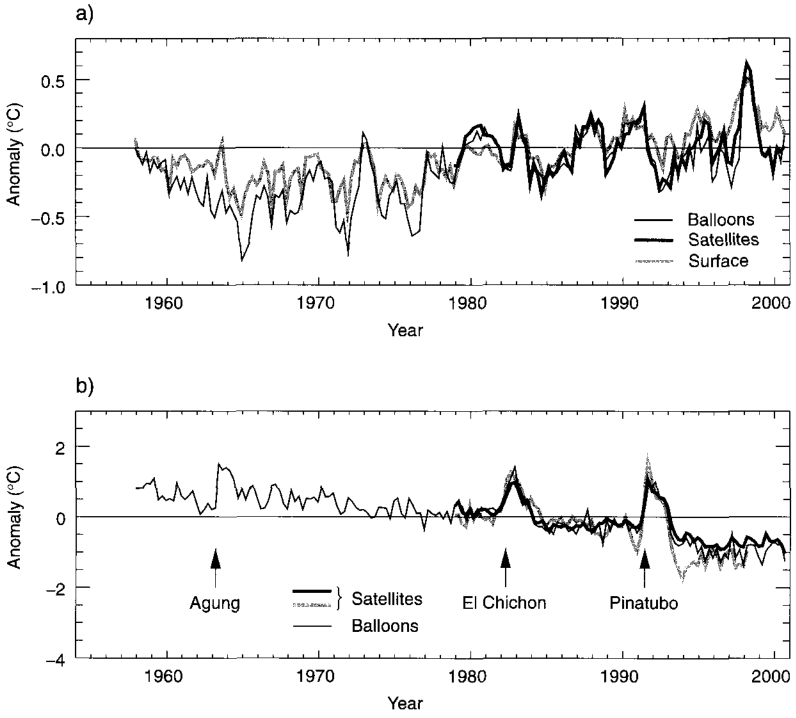


Fig. 2 Curves of global temperature in the troposphere (a) and of global temperature in the low stratosphere (b) (source IPCC).

The first fact to be considered, though it cannot be immediately interpreted here, is that while in the lower layers the temperature tends to increase, in the upper ones it tends to decrease. Apart from this average characteristic of the temperatures, it is interesting to concentrate on the events indicated by the arrows and on the immediately subsequent periods. These events are the three most important volcano eruptions of the last few decades, which introduced a considerable amount of dust in the atmosphere. This dust became more or less uniformly scattered all over the world, and settled in the low part of the stratosphere, where it remained for many months. In correspondence with these events, we can notice an increase in temperature where the dust has settled, and, though less evidently, a subsequent decrease in the temperature of the lower layers. So there seems to be a positive correlation between the amount of dust (not shown) and the temperature in the stratosphere, and a negative correlation between the latter (or the amount of dust) and the temperature in the lower layers.

Beyond the possible interpretation of this evidence within the previously created naive explanatory scheme[12] (which, on the other hand, was found to have some problems in the interpretation of other data), we can state with certainty that the presence of something other than air and clouds between the Sun and the low layers of the atmosphere disturbs the values of a variable such as the temperature. Therefore, if we are to understand the dynamics of the global temperature at a certain altitude, we need to consider the natural events that produce dust.

This observational evidence has a particular significance, because it shows, for the first time, that elements coming from outside, in this case from the lithosphere, can enter into the atmosphere system. So not only the system is not isolated (since its thermal characteristics are constantly changed from the outside by the Sun), but it is also interfaced, at its "boundaries", with other systems that can influence it.

Figure 2 and the discussion that follows its presentation are obviously based on data obtained by means of instruments during the last few decades. In the previous chapter, however, we explained that we also have the so-called proxy data, which we discussed chiefly in relation to the estimates of the global or hemispheric temperature trends during the last 1,000 years. Here we can mention the fact that the core samplings of ice in Antarctica and Greenland make it possible actually to go back to

---

[12] It seems that dust absorbs radiant energy, so it causes an increase in the temperature of the atmospheric layer in which it is present, and at the same times it acts as a screen for the lower layers, preventing a part of the radiation from reaching them.

much remoter periods, though obviously with a lower temporal resolution. We also touched upon the fact that, together with the complicated estimate of the temperature, it is possible to obtain more direct information about the composition of the atmosphere during the periods under examination, by examining the air trapped in tiny interstices in the ice. The examination of the long-term trends of the temperature and of the concentration of the gases that form the atmosphere has revealed some peculiarities, such as those shown in Plate 7.

This diagram reports the estimates, relevant to the last 420,000 years, of the values of the temperature and of the concentration of carbon dioxide (CO_2) and methane (CH_4), as they have been obtained by means of the analysis of core samplings performed at Vostok, in the middle of the Antarctic plateau. The first thing that one notices is the very marked correlation between the three series of data, whose values rise or drop in the course of time in a practically synchronous manner. Another noticeable characteristic is a certain cyclic character of the curves of the three variables.

In the diagram, obviously, it is possible to detect the glacial eras and the interglacial periods, such as the one we have been experiencing during the last few millennia (the last glacial peak, which coincided with a thermal low, occurred approximately 20,000 years ago). The alternation of these different periods and their approximate cyclicity seems to suggest that there is an external forcing factor that determines this pattern. As a matter of fact, at least for temperature, it has been found that this external influence consists chiefly of the intrinsic types of cyclicity of the Earth's orbit and of the precession of the Earth's rotation axis in relation to the plane of the ecliptic ("Milanković's theory"). Some of these effects lead to a different quantity of incident solar radiation, others to a different distribution of it on the surface of the Earth.

By projecting the diagram in Plate 7 into the future and applying Milanković's theory, we should find that now we are at the peak of an interglacial period and are approaching a gradual cooling of the climate. Recent calculations, however, demonstrated that an initial orbital forcing factor needed at least to be amplified in order to account for the changes in temperature found in the past, particularly during the deglaciation stages[13].

Is it possible that carbon dioxide and methane play a "causal" role in the amplification of the orbital effects that we have just mentioned? And, by the way, does the increase in these gases during the deglaciation stages precede, follow or accompany the increase in temperature? Unfortunately the time resolution of the series obtained with core sampling does not make it possible to answer the second question. As regards the first question, in this book we have never discussed the physical and chemical properties of these gases, so we would not know how to reply. There remains the evidence of the marked positive correlation between the three variables examined in Plate 7. This, in any case, may be enough to induce us to consider these three variables as constitutive elements, potentially important in the system under consideration.

Now, taking a step backward to return to the discussion at the end of the previous chapter, we may assert that Plate 7 also gives another contribution: it probably helps us to exclude one of the two hypotheses that had been presented. After having noticed, over the last hundred years, an increment in temperature that exceeds the typical values of the variability of the climate over decades or centuries, we wondered whether this might be due to the effects of a natural variability on a broader scale. The paleoclimatic analysis that has now been carried out shows that we are already near the highest peak in temperature in the last 420,000 years and should be approaching a colder period. This leads us to believe that the influence of a natural broad-time-scale variability characterised by the regularity of astronomic motions (Milanković's theory) is not likely to present unpredictable, opposite-trend aspects and thus to drive the current global warming. Finally, there still remains to be assessed the hypothesis of a disruption of the natural variability due to human activities.

In relation to this matter, and also to the observational evidence obtained by studies such as the core sampling of Vostok, which establishes a close correlation between the trend of the temperature and

---

[13] See, for instance, Petit *et al.* (1999) and references therein.

that of CO_2 and CH_4, we must consider that the concentration of a gas such as CO_2 in the atmosphere depends on complicated mechanisms that involve, for instance, photosynthesis processes and leaf "respiration", and also some mechanisms of oceanic storage. In these cases, the emission or absorption of carbon dioxide is performed chiefly through an "exchange" with other systems that are at the interface of the atmosphere, e.g. the ocean. A result of the fact that CO_2 is emitted also in all combustion processes is that, among these systems, we are led to give full consideration to the biosphere and, in particular, to all the human activities within it that involve combustion (which in most cases originates from the use of fossil fuel in the production of energy for industrial processes, transportation and heating). The deforestation performed by man in order to allow a different use of the soil can also have a certain influence, because it eliminates trees, which "absorb" carbon dioxide.

In this context it is interesting to consider Figure 3, which shows, together with the concentrations, during the last 1,000 years, of CO_2 and CH_4, also those of nitrous oxide (N_2O) and sulphate-containing dust. All the diagrams reveal a considerable increase in the concentrations during the last two centuries. In particular, if we return to Plate 7, a comparison between the data relative to carbon dioxide and methane during the last 420,000 years and the present data shows clearly that the latter are considerably higher than any other value detected in the past. This obviously seems to suggest that the variations are due to a human "perturbation" of the composition of the atmosphere.

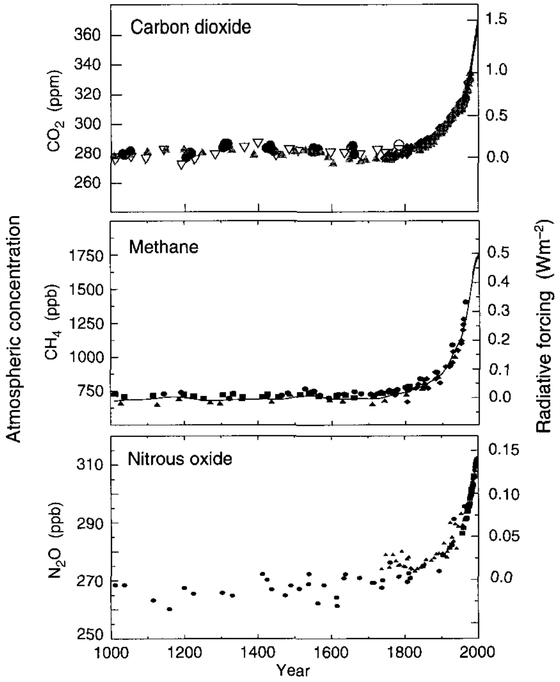
This "impression" is confirmed by some data relevant to anthropogenic emissions[14], which have enormously increased starting from the period of the industrial revolution. For the time being, however, this does not allow us to make any inference about the origin of the warming that has occurred during recent periods, in terms of global temperature.

---

[14] Emissions are something different from concentrations. Whereas an estimate of the former supplies information about the amount of a certain gas or material introduced in the atmosphere, e.g. as a result of human activities, the measurement of the latter gives us an idea of what remains dispersed in the atmosphere after all possible interactions with other systems at the interface of the atmosphere or within it.

Indicators of the human influence on the atmosphere during the Industrial Era

(a) Global atmospheric concentrations of three well mixed greenhouse gases



(b) Sulphate aerosols deposited in Greenland ice

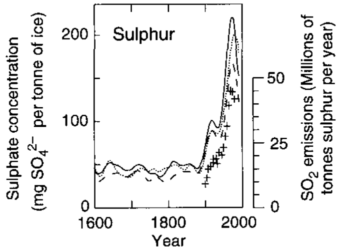


Fig. 3 Concentrations of carbon dioxide, methane, nitrous oxide and sulphate-containing dust during the last 1,000 years, from a combination of proxy data of the past and instrumental data of the last few decades (source IPCC).

Apart from the high correlation found in Plate 7 (and also in the last few decades) between the trend of the global temperature and the concentration of carbon dioxide and methane (a subject that will be taken up again in the next chapter), we do not have, for instance, any cognitive or explicative element that allows us to link the increase in these gases to the increase in temperature.

3.4 Let Us Take Stock of the Situation

In this chapter, on the basis of the data available to us, we covered a path that is typical of a scientific investigation. We saw that these data do not immediately shed light on the operation of the atmosphere system, and were compelled to organise them in an explanatory scheme, albeit a naive-type one. This scheme, however, was found to be in conflict with other observational evidences. At this point, we acted like the experts of local meteorology (farmers and fishermen), who look for what we call correlations and cross-correlations, in order to identify other relevant elements in the system under examination. Now that we have a more complete picture of the system, we need a substantial progress in our theoretical understanding of it. This is what we will endeavour to achieve in the next chapter.

Chapter 4

The Theoretical Framework: Knowledge of Single Phenomena and Complexity of the Earth System

In the previous chapter we mentioned the fact that any type of information becomes usable only if it is fitted into an explanatory and interpretative paradigm. The empirical evidence supplied by meteorological or climatic observations, in particular, is not sufficient, by itself, to reveal the connections among the variables of the system, or to allow us to understand physical phenomena and processes (and even less to predict their future evolution).

4.1 How Can We Read the "Great Book of Nature"?

Adopting a more radical approach, we might argue that perhaps there is no such thing as a pure and simple empirical observation, except in the first attempts to tackle an unknown system. We must consider, as a matter of fact, that as soon as we begin to be acquainted with the physical characteristics of a natural system, within the immense quantity of empirical evidences that Mother Nature makes available to us, the only ones to which we pay attention are the ones that we consider important for understanding the system under examination: only these become real observations[1]. In this perspective, the observations we

---

[1] This is based on a principle that is quite familiar to neurophysiologists, and that here we might call "economy principle". Our brain, because of its finite structure, cannot file all the sensory information it receives (some of it is actually filtered *a priori* by the sensitivity thresholds of our senses). Likewise, because of the finite speed with which we process information, in the solution of a problem we do not consider all the possible

perform voluntarily (e.g. instrumental measurements) are guided by the expectation to increase our theoretical understanding; therefore they are selected within a definite explicative scheme (that they may help to corroborate or disprove). Finally, as we have explained in Chapter 2, the so-called indirect observations that result in proxy data are essentially based on our theoretical knowledge of certain physical phenomena[2].

All this, therefore, indissolubly links observations to the interpretative scheme of the "operation" of the system under consideration. Here we mean to discuss precisely the current theoretical scheme relative to meteorological and climatic phenomena. First, however, we should pose the problem of what this capability of ours to "understand" a certain system consists of. How do we read "the great book of nature"? Maybe, if what we have stated in Note 1 of this chapter is true, natural evolution and, subsequently, cultural evolution have selected an "economical" way of summing up in an interpretative framework all the information we possess about a certain system.

In actual fact it does seem that this is the way things went. While carefully refraining from entering the field of the study of human learning and reasoning (which we leave to cognitive scientists), we will confine ourselves to a brief analysis of how a scientist manages to interpret and sum up the observational data relative to the sector under consideration.

As some authors have pointed out[3], nowadays our scientific image of nature is based on computable functions, and, in actual fact, the intelligibility of the world is due to the fact that we consider it algorithmically compressible. What does all this mean? On the one hand

---

hypotheses. Apparently this is how the human brain works. If, for instance, we analyse the recent wins (and final draw) of the Russian chess player Vladimir Kramnik in his matches with the super-computer Deep Fritz, whose processing and storing capacity is enormously greater than the human one (for instance, it can analyse 3.5 million moves per second), this suggests that natural evolution has been selecting an option that is economical and still successful.

[2] Proxy data are only an extreme, very evident case. In actual fact, the operation of any sensor is based on theoretical knowledge: for instance, in a mercury thermometer, the temperature estimate is achieved by exploiting a phenomenon that is theoretically understood, the expansion of mercury due to changes in temperature.

[3] See, for instance, Barrow (1988).

it shows that (as had already been indicated very incisively by Galileo Galilei) science uses the mathematical language in order to decipher the great book of nature: in particular here we cite computable functions, in which a variable depends on other variables (and possibly on time), from which its value can be calculated, analytically or by means of numerical methods using a computer. On the other hand, the comprehensibility of the physical world is strongly linked to the hypothesis that, by means of mathematics, it is possible to describe the relationships between the variables present in it, in a condensed and economical way. This way, for instance, if we are in possession of observational evidence relative to the values of a pair of variables in a certain system (at the same instant in time or at different instants), once we have found a "physical law" that binds them mathematically, we have acquired a condensed, economical way of describing their relationship. Among other things, this allows us to "predict" the value of one of them on the basis of the value of the other, even in the case of values that have never appeared before in the system.

The example mentioned above is an example of algorithmic compressibility. The theory of any physical process or phenomenon (with its diagnostic or evolutionary equations)[4] supplies precisely this algorithmic compression. Vice versa, in a world where algorithmic compression is not possible, phenomena appear to be random, and the characteristics of that world can be described only by a long list of sequences of observed phenomena[5].

Obviously, it is worthwhile to remind the reader that the concept of causality is always essential when the physical laws of a system that is evolving in the course of time are being sought: it turns out actually to be the pivot of the scientific explanation of any phenomenon or process. We must, however, mention the fact that there exist some balance laws (also

---

[4] Diagnostic equations interlink two or more variables at the same instant in time; evolutionary ones express the evolution, in the course of time, of a variable, which may depend not only on its value at a certain instant in the past and on time, but also on other variables.

[5] We will not advance any further in this analysis. Obviously it would be interesting to discuss whether the compressibility of the physical world depends on an intrinsic characteristic of that world (whatever this may imply) or is an "impression" of ours due to the use of mathematics, understood as a creation of the human mind.

called coexistence laws), when two or more variables are connected by a mathematical relation at the same instant in time. These laws cannot be clearly traced back to cause-effect relationships.

An example of this is Boyle's law for perfect gases, which asserts that, in a transformation in which the temperature does not change, the product of the multiplication of the pressure by the volume of the gas (at the same instant in time) remains constant. Moreover, both in diagnostic and in evolutionary laws, there are relations that connect the values of the various variables in a correlative (or statistical) manner: their contribution to the algorithmic compression of a system is evident, though the fact that we know only these relationships (which are valid only from a statistical point of view, i.e. impaired by an uncertainty that sometimes can be quantified) is often attributed to incompleteness in the description of the system under examination, to the extent that they are unlikely to be regarded as capable of supplying an explanation of its behaviour, at least in classical physics.

Here we cannot proceed any further in this analysis, which would lead us to delve into the concept of determinism. These themes will be taken up again in Chapter 7, within a narrower sphere, but having at our disposal an example of a concrete, realistic study. At present it will be expedient for us to dwell on another "epistemological leaning" of scientists in the study of the physical world, a leaning to which weather and climate scientists are not immune.

4.2 The Local Approach to the Study of a System

As we have already stated in the previous chapter, physicists generally count on the fact that the laws of nature are universal.

This, besides being probably the only way to make it possible to explain phenomena that take place in areas that cannot be investigated directly, naturally leads us to prefer local analyses and the instrumental investigation of small portions of matter. As we have already indicated, the air that is present in my room is the same that is present in the atmosphere: if I count on the fact that the regularities I observe and the laws I discover by analysing its behaviour here and now are the same ones that govern its dynamics in the free atmosphere, I may just as well study its properties in a place where I can use all the available analysing techniques. In this perspective, the "local" approach to the study of a physical system naturally encourages the tendency to carry out a thorough analysis of the basic constituents of a system, for instance air molecules, which can be easily "handled" in a room or laboratory[6].

It is known that this approach to an in-depth study of the constituents of a system has increasingly led us to improve the ideal microscope with which scientists examine nature. In particular, the physics of the twentieth century was characterised by a tendency to study increasingly microscopic entities, up to the so-called "elementary particles" and their constituents, which at present seem to be leptons and quarks. All this research is accompanied by the more or less implicit idea that to know the structure and behaviour of the individual basic constituents of a system (whether they be elementary particles, atoms, molecules or air masses with uniform internal characteristics) means to obtain the key to understanding the "large-scale" behaviour of the system, derived from a "composition" of the elementary processes thus revealed.

In this "rush toward the microscopic", several explanatory theories were gradually developed, often using different languages, i.e. sectors of mathematics, each of which was suitable for describing a single "level of microscopicness". In the current scientific practice, each scientist adopts the theory (with the relevant formalism) that is suitable for explaining the phenomena at the level under his/her examination. However, many people believe that the phenomena pertaining to the most macroscopic levels may be regarded as a composition of more elementary phenomena pertaining to a more microscopic level. If this is true, it will be possible to reconstruct the results and predictions of the theories pertaining to the macroscopic levels by means of the application of a formalism typical of

---

[6] In actual fact, we should pose the problem of what came first: we cannot absolutely take for granted either that the increasingly local or "microscopic" analysis of the constituents of a system was originated by the wave of belief in the universality of the laws of nature, or that, vice versa, this belief actually developed in order to "export" the increasingly detailed knowledge that was being acquired within the microscopic sphere into larger-scale spheres that could not be touched directly by investigation. We will gladly skirt this matter, whose discussion would lead us too far in another direction, deflecting us considerably from the subject of this book.

the lower level, that is by means of theories whose constituents (for instance, the variables considered) are more elementary. This has actually been done in the past. It does not mean, obviously, that the new theories that have thus been developed are easier to use (for instance for prediction activities), or that they replace the old ones in working practice. However, the possibility to explain the macroscopic operation of a system in terms of microscopic variables, regarded as more elementary, is considered a crucial step in the understanding of a system.

The belief that each level may actually be understood through the lower-level constituents and their physical interaction, together with the tendency to develop theories that make this operation possible, is called "reductionism". Extrapolating this tendency and carrying it to extremes, a thorough reductionist believes in the existence of a theory that explains all the phenomenology present in the physical world, and even more in the natural world (including life), starting from its elementary microscopic constituents. The final goal of the reductionist programme is to achieve this result. Now cosmology promotes the "historicisation" of this vision, and there is talk about a "Theory of Everything" that explains the dynamics of the entire known universe, once the physics and, if possible, the initial conditions of the first instants of the universe are known[7].

An example of the partial realisation of this project, which we may consider paradigmatic, is the reduction of thermodynamics to statistical mechanics, where the macroscopic concepts of pressure, density and temperature are translated into microscopic terms and reduced to the average statistical properties of the gas molecules considered in each case[8]. Macroscopic thermodynamic phenomena are therefore described

---

[7] See, e.g., Barrow (1991) and Hawking (2002). Obviously in this vision it is necessary also to evaluate (or upvalue) the roles of determinism and of the so-called anthropic principle (about this, see Barrow and Tipler (1988)).

[8] In order to prevent a misunderstanding, it is expedient to specify that here the adjective "statistical" does not have the meaning previously adopted to define the "uncertain" character in all the cases of a correlation between two variables. Now, in statistical mechanics, the macroscopic variables (e.g. temperature) are determined univocally once the distribution of the corresponding microscopic variables (e.g. the speed of the individual molecules) has been fixed. The use of statistics is necessary only because of the very great number of these variables, and has nothing to do with the uncertainty about

through the properties and interactions of the individual gas molecules. Among other things, this microscopic explanatory vision has made it possible to understand phenomena that were obscure before or were even ignored by thermodynamic texts, for instance the so-called "second-order phase transitions" in crystals[9].

Despite the undeniable success of the reductionist vision in the history of science, I believe that a few words of caution about the project of extreme reductionism are required here. As I have already pointed out, sometimes these microscopic-level theories cannot be applied practically within a macroscopic system, for instance because of the lack of data about the state conditions of the individual microscopic elements, as in the case of the air molecules in the atmosphere. Though on the one hand this is only an impossibility of practical application and not a principle (so it does not impair the cognitive value of the microscopic theory), on the other hand this does not allow a validation of the theory in spheres where there may be phenomena characterised by a particular, macroscopically visible dynamics or self-organisation. We may wonder, for instance, whether a cyclone is (or is not) an emerging entity that can be explained by the molecular dynamics of air and water particles. Maybe it is, but we are not able to verify it.

As we will explain, meteorology took a different direction from the outset, and soon (with a surmise that will be discussed further on in this chapter) actually decided to disregard molecular dynamics in the explanation of some medium- and large-scale meteorological phenomena.

At present, at least in a field of investigation such as that of the atmosphere and Earth system, scientists are basically aware of the importance of studying the microscopic structure and interactions of the elements that form the system. Many people, however, believe that there

---

the determination of the corresponding macroscopic variable. This uncertainty, which does actually exist, is due only to the fact that we do not have a precise knowledge of the individual microscopic variables.

[9] The reader is reminded that classical thermodynamics can be applied also to fluids and solids, besides gases. However, some "anomalous" behaviours, such as the second-order phase transitions, can be explained only in terms of statistical mechanics: an example of these transitions is that of a metal to the superconducting state when there is no magnetic field.

exists a level of complexity and of exchanges among interacting systems that cannot be described in microscopic terms, either as a result of a purely practical impossibility, or as a repercussion of processes and phenomena that actually originate at a macroscopic level. We will return to this topic further on.

For the time being, while remaining within the sphere of reductionist tradition and local instrumental analysis, we will closely examine the elements that may be regarded as the basic constituents of the atmosphere, i.e. the air molecules, and their interaction with the radiation that crosses the atmosphere. Among other things, in the previous chapter various observational clues had suggested the potential importance of carrying out an investigation of this type in order to obtain an explanatory picture that is more realistic and (we hope) does not contradict the observations.

4.3 The Interaction between Radiation and Matter and the Greenhouse Effect

In order to explain the origin of the electromagnetic radiation that crosses the atmosphere, we must return to certain fundamental discoveries of physics in the years before and after 1900, a crucial year for the understanding of some basic radiative properties of bodies.

If we consider the emission of visible light, it is evident that its primary source is the Sun, with a flow on the ground that is more or less variable during the daytime hours, depending on the height of the Sun above the horizon, the presence of clouds, fog, etc. The objects that emit light are not many on the Earth: lightning, lava flows, fire, and obviously man-made light sources. None of these emissions, however, is as continuous and intense as the light coming from the Sun.

Sunlight has been studied thoroughly, starting from the second half of the seventeenth century: that was the period of Newton's experiments with prisms, which revealed the "hidden" colours of white light for the first time. From the viewpoint of theoretical understanding, the disputation between the authors of two antagonistic theories, Newton with his corpuscular theory and Huygens with his wave theory of light, is fairly well known[10]. At the time of their first formulation, both these theories were able to explain the observational data. Then, as time elapsed, evidence in favour of the wave theory accumulated, particularly with the discovery of the phenomena of diffraction and interference, up to the success (which seemed final) achieved with the formulation of Maxwell's equations, by means of which visible light was described within the broader category of electromagnetic waves. This, in particular, led scientists to predict the existence of "non-visible light" such as radio waves, which differ from what is commonly understood as light only because of their longer wavelength[11]. Soon these radio waves were reproduced and revealed instrumentally by Hertz, while Röntgen revealed the waves that are now called X-rays, whose wavelength is shorter than that of visible light. In the meantime, moreover, scientists understood that an individual body that emits visible light usually also emits different-wavelength radiation at the same time.

Without going into technical details, or into historical ones, we can mention the fact that it had already been known for some time that any body in thermal equilibrium emits a radiation whose spectrum (which describes the quantity of energy emitted when the wavelength changes) is determined univocally by the temperature of that body: it is the so-called "blackbody heat radiation". 1900 was a crucial year, because that was when Max Planck determined the correct law of the blackbody spectrum for all wavelengths[12]. A direct consequence of Planck's approach is that the wavelength range in which the maximum quantity of emitted energy falls depends on the temperature of the body. The

---

[10] Newton's theory described the nature of light, its transmission and its interaction with matter in terms of light corpuscles, whereas Huygens's theory contended that the nature of light was that of a wave, with all the phenomena connected to this.

[11] An electromagnetic wave is characterised by its wavelength, λ, which can be represented as the distance between two peaks, or points of maximum value, in the wave train that propagates in space. When an electromagnetic wave falls on a body, it is easier to reveal the so-called wave frequency, ν, i.e. the number of times in which, within a time unit, the wave oscillates in the place where it is detected: this value is inversely proportional to the wavelength ($\nu = c / \lambda$, where c is the speed of light in vacuum).

[12] This, among other things, began a unique thirty-year period in the history of physics, from the viewpoint of theoretical syntheses, leading, in particular, to the formulation of quantum mechanics and the solution of the conundrum relative to the dualism between waves and corpuscles in the interactions between radiation and matter.

immediate application to a body having a surface temperature of approximately 6,000°C (the Sun) and to another one having an average surface temperature of about 15°C (the Earth) shows that the Sun sends us radiation above all in the visible-light range, while the Earth emits radiation above all in the infrared range (non-visible). A graphic representation of the blackbody spectra of the Sun and Earth is presented in Figure 4.

The discovery of a source of radiation other than the Sun is obviously important in view of an analysis of the interactions between the air molecules and the radiation present in the atmosphere, all the more if we consider that the Earth's radiation has characteristics that are different from those of the Sun's radiation, in terms of wavelengths.

Without going into details, let us examine which is the typical effect of the impact of radiation having different wavelengths on the gas molecules that form the atmosphere. The phenomenon on which we will focus our attention is that of the absorption of radiation by the various air molecules, because this "entrapment" of energy in the air contributes to determine its warming[13].

In nature, according to the correct precept of quantum mechanics, the absorption of radiation having a certain frequency (or, similarly, a certain wavelength) by a molecule takes place only if the value of this frequency, multiplied by Planck's constant, h, is equal to the difference in energy between two orbital levels of that molecule. Since each molecule possesses its own characteristic energy levels, it can be inferred that each molecule can absorb only radiations having specific wavelengths.

In semi-classical terms, we can imagine that the effect of the absorption of radiation by a molecule is like the increase in kinetics in the rotational and vibrational motions of that molecule: see Figure 5.

---

[13] Obviously the absorption of radiation is not the only way in which air can be warmed up. Recalling what we have learnt at school about heat transmission, besides irradiation we can cite conduction and convection.

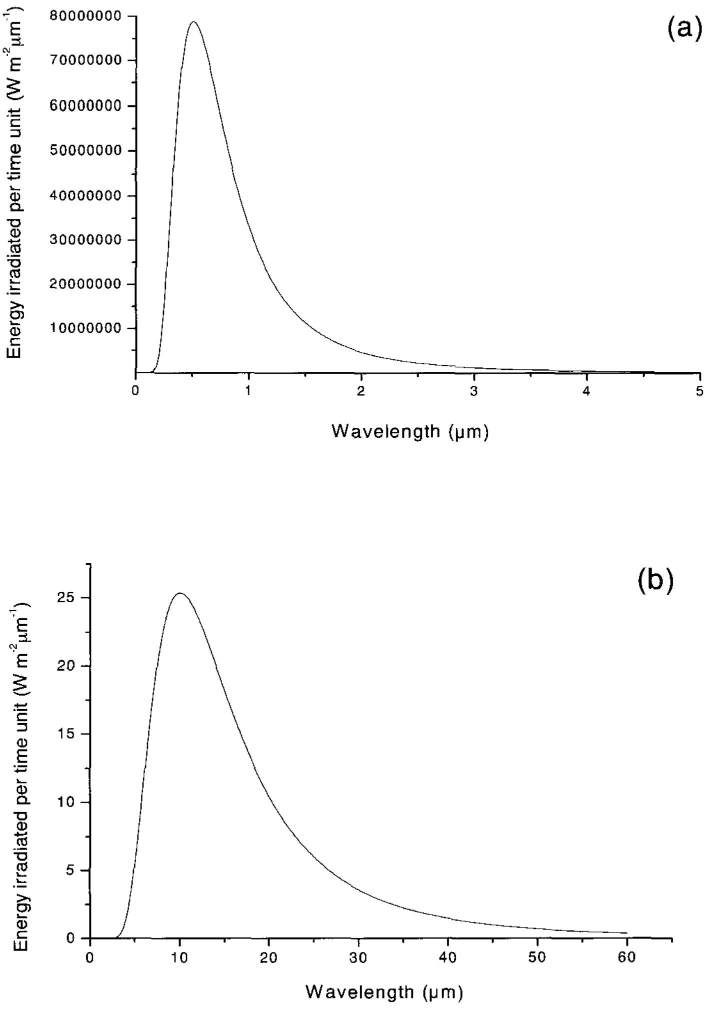


Figure 4. Blackbody spectra of the Sun (a) and Earth (b). The maximum irradiance of the Sun is in the visible range (approximately 0.5 μm); that of the Earth is in the infrared range (approximately 10 μm). Notice also that the total irradiated power always increases when the temperature of the body rises.

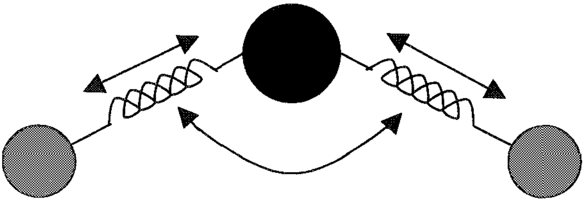


Figure 5. Mechanical model of a molecule: energy absorption creates a greater frequency in the oscillation of the distance between the atoms (along the "springs") and in the oscillations of the angle between the atoms (both indicated by the arrows).

As a rule, ultraviolet solar radiation (UV), conveying a great amount of energy, leads to a molecular dissociation or ionisation[14]. However, this radiation (particularly its most energetic components, UVB and UVC) is almost totally absorbed by ozone (which has the right energetic levels to do this) in the higher atmosphere, warming up the latter. Moreover, the ozone molecule takes part in a rather complex cycle of reactions with molecular oxygen and atomic oxygen[15].

The visible radiation carries less energy than the ultraviolet one, and, in the atmosphere, the typical phenomenon it produces is its diffusion due to impacts with the air molecules[16]; only a very small part of it is absorbed. There is, moreover, a region of the electromagnetic spectrum where the solar radiation and the terrestrial one overlap: it is the so-called "near-infrared" region. Here most of the radiation is absorbed by water vapour and carbon dioxide. Finally, in the region where there is the peak

---

[14] In the former case, the molecules are split into atoms; in the latter, an electron is expelled by the atomic or molecular structure, leaving a positively charged ion.

[15] This balance may be disrupted by chlorinated compounds, which cause the well-known phenomenon of the destruction of the stratospheric ozone layer and the resulting ozone hole detected during the last few years. Here we cannot discuss these topics in detail.

[16] This, as Rayleigh was the first to demonstrate, is why the sky is blue.

of the Earth's emission (called thermal infrared), approximately 80% of the emitted radiation comes out of the atmosphere and reaches the outer space. This range is characterised by some absorption bands due to CO_2 and other minor gases such as methane and nitrous oxide, which — together with the action due to the presence of water vapour and of clouds made up of liquid water and/or ice — lead at present to the absorption of the remaining 20% of the Earth's infrared radiation. The thermal infrared wavelength range is particularly sensitive to a change in the concentration of these gases: in particular, detailed theoretical calculations (forerun by the Swedish chemist Svante A. Arrhenius as far back as 1896[17]) show that their increase leads to an increase in the percentage of absorbed radiation and in the warming of the troposphere (the lowest part of the atmosphere, where there is the greatest quantity of these gases).

At this point, we must describe the basic mechanism of what is currently known as "greenhouse effect". This phrase originated when climatology began to be popularised, and indicates an analogy between what takes place in a greenhouse and what takes place in the atmosphere. The essential "transparency" of the air to solar radiation (chiefly the visible one) and the absorption of a part of the terrestrial radiation (almost entirely the infrared one) by some gases, called greenhouse gases, is compared to the transparency of the glass of a greenhouse, which allows the solar radiation to come in, but, being opaque to infrared radiation, prevents the long-wave terrestrial radiation from coming out. As we all know, the temperature of the air in a greenhouse is considerably higher than that of the outdoor air: it is fairly common to conclude that this is due to the differential effect of the glass on radiation having different wavelengths. In actual fact, this warming effect does exist, but it accounts only for a small percentage of the warming of the greenhouse, which is mostly due to the mechanical effect of insulation from the external environment, achieved by the glass. So the analogy between the atmosphere and a greenhouse is faulty; however, since the

---

[17] See Arrhenius (1896).

phrase "greenhouse effect" has by now become a part of the media jargon, it still survives.

A fundamental aspect that must be explicitly stressed when dealing with the greenhouse effect is the fact that this effect is essentially due to the presence of some minor gases in the atmosphere and to water vapour: oxygen and nitrogen, which together form approximately 99% of the "dry" atmospheric mixture (i.e. not considering water vapour) do not absorb the infrared radiation. So a constant monitoring of these low-concentration gases is important for keeping track of possible variations in the radiative warming of the troposphere. We must also point out that, contrary to the generally negative connotation of the phrase "greenhouse effect" in common parlance, in actual fact the processes connected to this effect are essential for life on the Earth as it is now: it has been calculated that, if the greenhouse effect were not present, the average temperature on the planet would be lower than the present one by approximately 33°C. In such an extreme condition, the number of known forms of life that might survive is very small.

What has been understood up to now, though obviously not concluding our story, nonetheless is extremely important in view of a global energy balance of the Earth system (formed by land, oceans, ice, biosphere and atmosphere, each of which can be regarded as a subsystem that interacts with the others).

From a physical point of view, as a matter of fact, the Earth is a system upon which there falls an external flow of energy in terms of solar radiation; it responds to this external "forcing factor" by dissipating the energy it has thus received and re-emitting it towards space, again in terms of radiation. Everything that comes into the Earth system and comes out of it takes the form of radiation: the energy balance between the incoming radiation and the outgoing one reveals whether the system is in temperature balance or whether there is an imbalance that promotes its warming or cooling. If, for instance, in the future there should be an increase in the concentration of the greenhouse gases, the latter might entrap a greater quantity of terrestrial long-wave radiation with respect to the present quantity, and it would not be possible for this radiation to reach the outer space; so this would determine a net increase in energy (heat) within the Earth system[18].

4.4 Greenhouse Gases, Clouds and Aerosols

Considering, for the time being, only the atmosphere system, we can perform a concise survey of the gases that play the most important role in the greenhouse effect, i.e. carbon dioxide, methane, nitrous oxide, ozone in the lower layers, and water vapour. The concentration of these molecules in the atmosphere sometimes depends on reactions between the various compounds that are present, as in the case of the complex photochemical reactions that involve ozone near the ground; but more often it results from their emission or absorption by other subsystems of the Earth system that have exchanges with the atmosphere. A typical case is that of water vapour, which is essentially produced by the evaporation of water surfaces (oceans, seas, lakes, rivers) and by the phenomenon of evaporation and transpiration by plants, and is removed by condensation and precipitation phenomena that, as an ultimate consequence, bring it back to the Earth's surface.

A case that is undoubtedly more complex is that of carbon dioxide, whose emission and absorption involve several processes that interconnect various subsystems. The emission of CO_2 is caused, for instance, by volcano eruptions, the breathing of animals, fire in the forests, and the combustion of oil and other types of fossil fuel. CO_2 is absorbed chiefly because it is stored in the oceans and "consumed" in the photosynthesis that characterises the vegetable world. The concentration of carbon dioxide in the atmosphere depends on the balance of this complex cycle, called carbon cycle. The reader should notice, in particular, that the situation becomes the more complicated to describe the more the cycle is affected by human activities: on the one hand, man contributes to the emission of CO_2 through fossil-fuel combustion and arson activities, on the other hand he also contributes to eliminate some "absorbers" of carbon dioxide, for instance through the deforestation

---

[18] Obviously the Earth system would also respond to the increase in energy with other effects, the so-called "feedbacks", which will be discussed further on.

carried out in order to achieve a different use of the soil. Obviously all these anthropic actions tend to promote an increase of CO_2 in the atmosphere.

In order to close the discussion of the global radiation balance, we must now introduce two further elements that are present in the atmosphere and have not been considered yet: clouds and aerosols. Clouds are made up of water or ice, and result from the condensation of water vapour and its possible subsequent freezing. For a detailed description of the interaction between clouds and the radiation present in the atmosphere, we should allow for the specific characteristics of the clouds (composition, shape, height). Simplifying the matter as much as possible, we can assert that the basic effect of clouds on solar radiation in the visible range is that of reflection: they act as a screen. As regards their interaction with the infrared radiation, particularly the terrestrial-origin one, they are able to entrap it between themselves and the ground. The predominance of one of these two effects depends on the specific situation. Simple theoretical calculations show that during the daytime the screening of the solar radiation prevails, and, in comparison with a clear-sky situation, clouds help make the air under them cooler, whereas during the night, when the solar radiation does not fall on the clouds, the only effect that is present is the one on the radiation coming from the ground, which helps to keep the lower layers warmer than they are during starry nights. All this is consistent with the common experience (already mentioned in the previous chapter, taking as an example winter nights in a continental area) of cold nights with a clear sky and warmer nights with a cloudy sky.

As regards aerosols, we must explain first of all that this term indicates a considerable variety of "impurities" present in the atmosphere, ranging from several types of suspended dust particles (produced, for instance, by volcano eruptions or anthropogenic pollutant emissions) to salt crystals coming from the seas and oceans, and to pollen, spores and bacteria. As we have seen in the case of the volcano eruptions discussed in Chapter 2, these types of aerosol often reach a very high altitude in the atmosphere, where the interaction with the radiation refers particularly to the radiation coming from the Sun. Therefore the direct consequences, for instance, of the presence of a dust "cloud" in suspension basically depend on the optical properties of its constituents; these properties in turn depend on the material of which the constituents are formed and on their size (roughly called "particle diameter"). As a rule these "clouds" reflect the visible radiation and act as a screen for the underlying atmospheric layers, helping them to cool down; moreover, particularly if the particles are rich in carbonaceous materials, they lead to a warming of the atmospheric layer in which they are present. This analysis accounts for the observational results presented in Chapter 3, Figure 2.

As we have already suggested, the case of dust in suspension, and more generally that of aerosols, highlight, once again, the exchanges that take place between the atmosphere and the other subsystems of the Earth system, as in the case of the material extracted from the lithosphere and introduced in the atmosphere. And there is also another aspect. This topic allows us to start explaining that any element present in the atmosphere interacts in a complex way with other elements. For instance, the interaction of aerosols with radiation that we have just described is only one of the consequences that these particles have on the radiation balance, the so-called "direct effect". As a matter of fact, there exist at least other two effects on this balance, called "indirect effects". In order to explain them, we must mention the fact that, in the atmosphere, impurities also play a particular role, that of catalysing the condensation of water vapour into droplets of liquid water, facilitating this change of state, which otherwise would have more difficulty in taking place, even in cases of saturation or supersaturation (i.e. when the relative humidity exceeds 100%). To put it in meteorological parlance, aerosols supply condensation nuclei.

The two indirect effects on the radiation balance mentioned above are connected to the consequences of this role. In particular, an increase in the condensation nuclei leads to the presence of a greater quantity of water droplets in the clouds; the diameter of these droplets is smaller, and this changes the optical properties of the clouds, causing them to reflect (i.e. screen) the sunlight more: the net result is a greater contribution to the cooling of the layers below. The other indirect effect is due to the fact that, since the droplets are more numerous and smaller in diameter, it is less easy for them to reach the critical radius and weight above which they fall to the Earth as rain, so the clouds persist as such in the atmosphere and screen out the sunlight for a longer time. In this case too, the cooling of the lower layers is promoted. If the aerosols/condensation nuclei decrease, obviously, the effects are opposite.

This way, the theoretical understanding of the interactions between radiation and the air, which can all be investigated locally and according to the rules of a reductionist vision, led to the discovery of an effect, the greenhouse effect, which plays an essential role in the warming of the lower-layer air. As the reader undoubtedly remembers, radiosonde observations have revealed that the atmospheric layers near the ground are usually the warmest ones; the naive explicative scheme developed in the previous chapter was not able to explain this. Can the greenhouse effect account for this observational situation? In this sense, the fact that understanding the interaction between radiation and other elements present in the atmosphere (specifically clouds and aerosols) can explain some observed phenomena leads us to hope that the correct paradigm for the interpretation of the temperature variations in the atmosphere has been found in the physics of irradiation. In order to answer this question, we will have to broaden the scope of our investigation once again, so as to include other elements and physical processes present in the complex Earth system.

4.5 Approaching a Complete Scheme of Warming from the Bottom

As we have already mentioned (see Note 13 in this chapter), up to now we have discussed only one of the possible modes of heat transmission in the atmosphere: irradiation. Though from a climatic point of view (i.e. averaging values over long periods) heat transmission by irradiation is fundamental, for instance in energy balance calculations, it is reasonable to wonder whether on a shorter time scale it can account for the warming process in the low atmosphere. Here heat is redistributed between the various subsystems of the Earth system, in particular between land, ocean and atmosphere, on time scales that range from the day-night cycle to the seasons.

We have explained that the troposphere is basically transparent to solar radiation, apart from a few absorption bands in the near-infrared range. Moreover, solar radiation in the ultraviolet range helps warm the stratosphere[19], which, however, is basically separated from the troposphere below it (where we live and where all meteorological and climatic phenomena take place). This way we can assert that most of the solar radiation reaches the ground, where it is partly reflected and partly absorbed. On the basis of what we have explained up to now, the reader has understood that the warming of the troposphere takes place from the bottom, by means of the phenomenon of the absorption of long-wave terrestrial radiation by greenhouse gases. However, this statement, which is basically true, reveals only a part of the actual situation: it is true that the warming of the atmosphere takes place from the bottom, but this is not due only to the absorption of radiation.

We can spot a first clue of this by looking at Figure 4 again: we can see here that bodies at different temperatures emit different quantities of energy, which increase as the temperature of the body rises. Now, it is a common experience that the ground warms up during the day and cools down during the night: so we can expect an increase in irradiation during the day and a decrease during the night. In actual fact, though allowing for the fact that in certain situations the approximation to a blackbody may be rough and making the required corrections, a quick calculation allows us to assert that this effect explains only a part of the meteorologically relevant temperature variations in the lower layers of the atmosphere. The absorption of radiation, therefore, is not sufficient to account for short-term temperature variations, for instance during the day-night cycle.

In this chapter we expatiated (some readers may say far too much) on the local analysis of the interaction between radiation and the constituents of the atmosphere; we were led by a reductionist vision, which in turn is inspired by a belief in the universality of the laws of

---

[19] Where, in actual fact, the air is extremely rarefied, and the concept of temperature loses much of its usual meaning.

nature. The same inspiring concept had led us, in the previous chapter, to devise a "domestic" example of the warming of the air in a room by a hot radiator. We had noticed, in particular, that the bubbles of warm air that formed near the radiator moved up as an effect of buoyancy, because they were warmer and less dense (i.e. lighter) than the surrounding air. Soon this created a stratification, in which there was warmer, lighter air near the ceiling and colder, heavier air near the floor. We did not stop at this first aspect, and proceeded to point out that the vertically rising air bubbles cooled down slightly before they reached the ceiling. This led us to conclude that, if the ceiling were removed, the ascending air would cool down further, eventually creating a vertical thermal profile in which the temperature decreases with height, exactly as it does in the troposphere. This simple "domestic" example could therefore "mimic" what happens in the atmosphere, provided there is a source of heat near the ground: in that stage, obviously, the cooling of ascending air was purely an observational assessment, in a situation where pressure decreased with height, without an evident theoretical explanation.

At this point, however, we know that there does exist a source of heat in the low layers of the atmosphere: it is the ground itself, when, as a result of the absorption of a part of the energy conveyed by solar radiation, it reaches temperatures higher than those of the air that is in contact with it. So we can advisedly proceed with the parallelism between what happens in my room and what happens in the troposphere, now utilising some theoretical knowledge that comes from the basic physics of heat transmission and from general thermodynamics.

Though here we have always discussed only irradiation, the best-known mode of heat transmission is undoubtedly conduction, whereby two bodies that touch each other tend to reach the same temperature. Among other things, this is the principle on which the measurement of the temperature of a body by means of a traditional thermometer is based, as when, for instance, fever is measured in a sick person. These measurements are considered reliable if the mass of the thermometer is small and the difference between the initial temperature of the body and that of the thermometer is not too great, that is if the presence of the thermometer does not appreciably affect the temperature of the body. In the case of air that touches a radiator or the ground, it is precisely the air that (over a brief period) takes on the role of an ideal thermometer, because, after a certain time, it ends up by having the same temperature as that of the body it touches, without causing any appreciable change in the temperature of the latter.

This way, for instance, during the daytime the air that directly touches the surface is warmed by conduction. But how are the layers immediately above it warmed? Recalling the mechanical-statistical concept of the temperature of a gas, connected to the average kinetic energy of its molecules, we can surmise that there is a phenomenon (which does actually exist), molecular diffusion, that causes the warmer air molecules (which are faster) to diffuse among the colder-air ones (which are slower) and to transfer a part of their kinetic energy to the latter by means of molecular impacts, with a net warming effect. This does take place, but the molecular "mixing" is rather slow, while in the atmosphere (and in my room) the air is actually mixed up more quickly. The process that seems most likely to be the one that allows this greater rapidity is the previously-described vertical ascent of warm-air bubbles.

This process, which has already been discussed from an observational point of view in an indoor environment, is called "convection". As we have already stated, it can be regarded as a manifestation of the Archimedes' principle, therefore it takes place until the ascending air meets an obstacle (the ceiling of my room), or until, at a certain point in its ascent, it ends up by being surrounded by warmer air. In the latter case, the ascending bubble is now colder and heavier than the surrounding air, and starts undergoing a decided downward thrust, which first slows down its ascent, then inverts its movement, forcing it to descend. We must now explain what determines the previously-mentioned cooling of the air within the ascending bubble.

Here a gravitational effect (to which, in actual fact, also other less-important thermal effects are added) comes into play: it determines the amassment of the air that touches the ground and its increasing rarefaction with the increase in height above sea level. This results in the fact that, if we mark off some small volumes of air at certain heights, a pressure that gradually decreases with height is exerted on their imaginary walls. So if, for some reason, one of these small volumes of air rises in the atmosphere, the fact that it reaches a lower-pressure area causes it to tend to expand (because of the greater internal pressure)[20]. This phenomenon is described well by the first law of thermodynamics.

If we suppose that the "bubble" that rises in the atmosphere is isolated from the surrounding air (like the one that is physically delimited by the balloon that contains the gas), the theoretical determination of the cooling of the bubble on the basis of its expansion is correct, with an excellent approximation. In thermodynamic terms, regarding the bubble as insulated means supposing that it does not have heat exchanges with the outside (in this case the process is called "adiabatic"), and in particular not considering molecular diffusion in the surrounding air. The fact that this approximation (called "adiabatic") leads to excellent results in the calculation of cooling is due to the different time scales that characterise convection and molecular diffusion: the latter is very slow and may be disregarded for up to 12 or 24 hours in the study of certain atmospheric phenomena.

So at this point we have achieved a theoretical justification of the vertical thermal profile characterised by a decrease in temperature with height that has been observed in the atmosphere and that had remained an enigma at the end of the previous chapter. It is due to the combination of the three modes of heat transmission, which co-determine the warming of the atmosphere from the bottom (obviously to all this we must add the horizontal redistribution of heat via oceanic and atmospheric currents, which has not been discussed yet). We can also mention the fact that the validity of the adiabatic hypothesis suggests that sometimes a microscopic description of the system (e.g. in terms of molecular diffusion) is not necessary. Moreover, once we have allowed also for the vertical movements of air masses in which water vapour is present and can condense, the adiabatic hypothesis can be extended to state changes, and considerably helps us to understand the dynamics of clouds. Finally, it also leads us to correctly explain other phenomena that had not been understood previously, such as the warm, dry winds that

---

[20] This fact is clearly displayed in the sad moment in which a balloon escapes from the hand of a child. Rising in the atmosphere, the balloon becomes increasingly large, until the tension on its surface exceeds a critical threshold, causing the balloon to burst.

sometimes descend from the Alps to the Po Valley[21]. Breezes, too, are due to the fact that particular convection cells set in on the territory.

4.6 Nature of the Ground and Air Warming

The theoretical scheme we have just outlined describes a type of atmosphere warming that takes place essentially from the bottom. In this context, it is clear that the nature of the ground that absorbs the solar radiation and helps to warm up the air above it in the ways we have described may be important for a more detailed analysis of this warming.

After all, is it not a common experience that winters are milder in regions of the world that are in close contact with the sea than in continental areas that are quite distant from it? Consider the winters in Rome, Italy, and in St. Louis, Missouri, USA: both these cities are not very high above sea level, and St. Louis is even further south than Rome (so, on the average, it receives a greater amount of incident solar radiation). Yet the winters in these two cities could not be more different: they are mild and with comparatively high temperatures in Rome, and harsh, often snowy and with temperatures that may be extremely low in St. Louis. What determines this difference? Why should the presence or absence of a large water surface nearby have such an influence on the climate?

Once again we must consider the interaction between the incoming solar radiation and the elements that form the Earth system, in particular the Earth's surface. Speaking of this interaction, we stated previously that the surface absorbs a part of the visible radiation coming from the Sun and re-emits radiation essentially in the "thermal infrared" band towards the atmosphere and outer space. However, we never made any distinction between the various types of surfaces: it is now time to do so.

Conduction, convection, and, broadly speaking, the irradiation emitted by a surface depend almost exclusively on the temperature of the surface, independently of its physical nature. However, surface

---

[21] Here we cannot discuss the physics of *föhn* or of other local meteorological phenomena: for this, the reader may consult a text of aeronautic meteorology, because this type of information is usually quite important for aeronautics.

temperature and its changes with time are strongly affected by the mode of absorption of solar radiation by surfaces having different characteristics. More specifically, different materials are characterised by different responses to irradiation: for some, a small quantity of energy is enough to raise their temperature by 1°C, while others require much more energy.

In physics there exists a quantity that measures this "thermal inertia", i.e. that determines the resistance of each body to warming. This quantity is called heat capacity. The energy supplied to a body being equal, the temperature of this body rises more if its heat capacity is low and rises less if its heat capacity is high[22]. As a consequence, in a warming/cooling cycle, a low heat capacity allows the body to respond more quickly to these external forcing factors, with a change in temperature that follows the cycle rather closely, though perhaps with a slight delay. Vice versa, in bodies that have a high heat capacity a great quantity of incident or outgoing energy is needed to change their temperature appreciably; therefore, if the warming/cooling cycle is too quick, their temperature changes only slightly during the times that are characteristic of this cycle.

We will now apply this discussion to the topic of the temperature variations on the Earth's surface, temporarily disregarding the role of ice (which will be considered further on). To begin with, we should remind the reader that the Earth's surface is formed of approximately two thirds of sea and ocean masses, and one third of continental crustal plates. If we disregard the differences that undoubtedly exist between the various types of ground on land areas (deserts, woods, forests, tundras, savannahs, etc.), we can assert that the heat capacity of the seas and

---

[22] In mathematical terms, we can put it this way: the energy (heat) supplied being equal, there is an inverse proportionality between the increase in temperature and the heat capacity of the body under consideration. This can be expressed by a very simple formula: $\Delta T = Q / C$, where Q is the quantity of heat absorbed by the body, ΔT is the difference in temperature between the end and the beginning of its warming, and C is the heat capacity of the body. To make a comparison between the thermal effect of warming on different materials, it is expedient to refer to the same mass unit: in this case it is possible to check the thermal inertia of a specific material in comparison with that of another material, by measuring the so-called "specific heat", c, of the material, which is defined by the formula $c = C / m$, where m is the mass of the body under consideration and C, obviously, its heat capacity.

oceans is much higher than that of the crustal surface above sea level. This does not depend on their greater extension on the Earth: using the terms introduced in Note 22, above, we can assert that it is precisely the specific heat of the water of which they are formed that is always higher than that of the land.

Thus, on the basis of these theoretical considerations, it is easy to predict that the land warms up quickly during the daytime, following the solar cycle, and cools down with equal rapidity (through a loss of radiant energy) during the evening and night. Meanwhile, the temperature of the sea should not be much affected by the day-night cycle, and should undergo only a slight change on this time scale. Instrumental observations, and also common experience, confirm that this does happen. Who has never plunged into the sea during a hot summer day in order to cool down? Who has not enjoyed the typically juvenile experience of a midnight swim? In these cases we have undoubtedly appreciated the temperature of the water, comparatively constant with respect to the more variable one of the sand (and, partly, of the external air).

In the land, the day-night cycle is quite evident in the surface temperature data. It combines with the seasonal cycle, due to the different average height of the Sun on the horizon, whose action on the surface is similar to the one discussed in the previous chapter (with the beams more or less inclined, the amount of radiant energy that reaches an area unit changes). In the seas and oceans, on the contrary, the day-night thermal cycle is only a slight perturbation in the much more evident seasonal cycle, due to the cumulative effect of the enormous amount of energy that, on the average, is absorbed during the summer and dispersed towards the outside during the winter[23]. In any case, the high heat

---

[23] To be more accurate, we should point out that the difference between the thermal behaviour of water and that of land is more marked than that revealed by the calculations made on the basis of the previous theoretical considerations: there actually exists another effect that combines with the one due to the different heat capacities and enhances the difference between the two behaviours. It consists of the fact that, whereas in solid land solar energy is absorbed entirely in the surface layer, in seas and oceans, since they are formed of liquid water, there takes place a certain mixing with deeper layers, so the absorbed energy is distributed in a greater mass, determining a variation in temperature that is smaller than that predicted by the previous considerations.

capacity of the seas and oceans causes their seasonal range of temperature[24], as well as their daily one, to be rather narrow. This leads to the theoretical prediction, corroborated by observational assessments, that, the solar radiation being equal (for instance at the same latitude), on the average solid land is colder than the surface of the sea during the winter and warmer during the summer.

Because of the previously mentioned mechanism of the warming of the lower layers of the atmosphere from the bottom, caused by the Earth's surface, all this leads us to understand how the presence of expanses of water near a location limits the temperature range of the air, tempering the climate, both in its seasonal average level and on a daily time scale. Thus, for instance, at our mid-latitudes, it is possible to find continental areas that are distant from the seas and oceans and have particularly cold winters and sometimes torrid summers, if the temperature values of the air at ground level are compared with those of coastal areas, as in our previous comparison between the winters in Rome and those in St. Louis. In the same way, with reference to the day-night thermal cycle, the presence of the sea leads to a decrease in the daytime maximum temperatures and to an increase in the night-time minimum temperatures, with respect to the values in continental areas at the same latitude[25]. Breezes contribute particularly to this phenomenon, during the day by conveying towards the land air that was previously above the sea, and during the night by blowing from the land towards the sea.

From what we discussed previously, it is evident, in particular, that the seas and oceans act as "climatic dampers", because — like the suspensions of a car cushion the vertical oscillations of a car on rough ground — the masses of water lower the variability rate of the temperature variations in the atmosphere, on several space and time scales. If, more specifically, we consider the global annual averages, we

---

[24] The temperature range is the difference between the maximum temperature and the minimum one within a certain period of time that has a certain cyclic characteristic.

[25] This is why in the previous chapter we chose a continental area for discussing the phenomenon of colder winter nights with a clear sky and warmer nights with a cloudy sky: in these areas of the world the phenomenon is more evident than in other areas.

can see (as we had already mentioned) that the interannual temperature variability, presented in average terms in Plate 5, is also damped down by the contribution of the seas and oceans. Now we can understand that this is due precisely to their characteristics in terms of thermal inertia.

4.7 An Outline of Oceanic and Atmospheric Dynamics

Our discussion, here, of the thermal phenomena in the atmosphere has led us to view the situation from a different angle (that of warming from the bottom) with respect to the angle adopted in the previous chapter, to contradict some naive ideas that had been presented there, and to achieve a theoretical framework in which contradictions with observational experience have practically disappeared. Nevertheless, some temperature data on a regional scale are still unexplained, and the horizontal dynamics of the atmosphere has not been tackled yet. So now, without presuming to do so in a comprehensive manner, we will endeavour to concisely outline some other phenomena, in order to achieve a more detailed treatment of the system under examination and to consolidate our theoretical framework.

It is known that our planet, because of the inclination of its axis, undergoes the phenomenon of the seasons. This is due to the fact that the warming of the Earth's surface by the Sun is greater during the summer than during the intermediate seasons and the winter, because of the different average height of the Sun above the horizon. In any case, if, with a rather low time resolution, we consider the average effects over a year, we can calculate the average energy absorbed by the Earth's surface at various latitudes: its values turn out to be very high at the tropics, lower at the intermediate latitudes, and even lower within the polar circles. The consequent warming of the atmosphere by conduction, convection and irradiation, according to our theoretical pattern, must have a similar trend. Though from a qualitative point of view this is confirmed by observations, from a quantitative point of view there is a discrepancy between the predictions of our scheme and the average temperatures of the air at the various latitudes. In particular, the difference in temperature observed between the equator and the poles is smaller than the one that might be expected theoretically. How can this discrepancy be explained?

A first clue for the solution of this riddle comes from the observation that the difference between the temperature of the oceans at the equator and that at the poles is also smaller than the one estimated by calculating sunlight absorption only. This gives rise to surfaces that are less diversified in temperature, so the resulting influences on the air above them turn out to be more similar to each other than had been predicted on the basis of the previous reasoning. On the one hand we would like to understand how this can happen, but on the other hand we will see that, though from a qualitative point of view this effect goes in the right direction of a mitigation of the differences in the temperature of the air at different latitudes, from a quantitative point of view it is not sufficient to reproduce the correct values of these differences.

In this case, an effect of horizontal transmission of the heat contained in the masses of water is present. Once again, the transmission modes by irradiation and conduction (or molecular diffusion) are not sufficient to explain the phenomenon. Only the action of vast oceanic currents allows such an efficient thermal mixing. Many of us are aware of the existence of the so-called Gulf Stream, which carries masses of water through the Atlantic Ocean from the Gulf of Mexico to the coasts of north-western Europe, mitigating the climate in that area, at least in comparison with the climate on the American shore of the ocean. In actual fact, now oceanographers are aware of the existence of a circulation in the oceans (the "thermohaline circulation") that leads to a global mixing, particularly by means of an enormous underwater "river", the so-called "conveyor belt", whose route covers the entire planet. As a rule, the horizontal route of the ocean currents and their vertical dynamics[26] depend on differences in temperature and salinity[27].

If now we evaluate the surface temperature of the oceans at the various latitudes and include these new data in our explanatory scheme of atmosphere warming, we will find that there is still a discrepancy

---

[26] There are areas in which deep water comes to the surface as a result of ascending movements, and areas in which surface water sinks to the depths of the sea.

[27] It is not possible here to expand on the subject of the general circulation in the oceans. A clear explanation will be found in Lionello (2005).

between the theoretical predictions and the results of observations: to be more precise, from a quantitative point of view there remain a surplus of temperature in the air at the high latitudes and a deficit at the low ones, both of them unexplained. This means that there must be another manner in which heat is exchanged between the tropical areas and the polar ones. This manner does exist, and is present within the atmosphere, in the atmospheric currents of its general circulation. It is not possible, here, to discuss this topic exhaustively. It will be sufficient to mention the fact that, once again, what is quantitatively important is the macroscopic vertical and horizontal motion of air cells, rather than molecular diffusion.

More specifically, it is expedient to point out that, in the description of the atmospheric dynamics, an important role is played by the concept of "air mass", a phrase that denotes a portion of air that is large enough to allow us to disregard molecular diffusion and small enough to allow us to consider it thermodynamically homogeneous, i.e. characterised by univocal values of temperature, density and humidity that are representative of the characteristics of the air of the entire cell. The movements of the air masses in the atmosphere are obviously driven by the forces that act on them, as in the case of what we might call "buoyancy force", which pushes the cells up or down, depending on their temperature and on the Archimedes' principle.

The horizontal motions are affected by the combined action of a force called gradient force (which tends to move the air masses from the high-pressure areas to the low-pressure ones), of an apparent force[28] called Coriolis force (due to the fact that the Earth revolves over a period of 24 hours), of the centrifugal force in non-rectilinear motions, and of a friction force for the air layers near the ground. These forces, together with the fact that it is often assumed that, at a certain height in the atmosphere, air is not compressible[29], lead to a qualitatively and quantitatively satisfactory explanation of the atmospheric movements.

---

[28] It originates as a consequence of the fact that the Earth's reference system is not inertial: within it, a body to which no real force is applied does not remain still or move with a uniform rectilinear motion, but performs more complicated movements.

[29] From a mathematical point of view, this is described by an equation called mass conservation, according to which, once a volume of air has been delimited, the amount of

Let us mention an example. Horizontal and vertical motions may combine, and some may be the cause of the others. In an area where the pressure is low near the ground, the combination of the forces mentioned above tends to cause the air to converge towards the centre of the area; since air is essentially not compressible, this results in its moving up vertically. When the cell rises and cools down, the water vapour present in it tends to condense, forming a cloud. On the contrary, in an area where the pressure is high near the ground, the air tends to diverge from the centre of the area; since air basically does not become more rarefied, this means that a certain quantity of air comes down from the upper layers. Descending, it warms up[30], and the liquid water, if any, that forms clouds evaporates (so the clouds dissolve). Thus we have also achieved a qualitative explanation of the fact that low-pressure areas are often characterised by cloudiness, while high-pressure areas are usually associated to expanses of clear sky.

A last consideration to be made on the atmospheric dynamics is that, because of the slowness of the diffusion process, when two air masses touch each other, they do not mix diffusing into each other, but clash, forming discontinuity surfaces called “fronts”. The cold air usually pushes under the warm one and the warm air slides over the cold one. The mutual diffusion process, and obviously the interaction with the ground below, become important only if two air masses remain stationary next to each other for a long time.

The observational evidence relative to the shifting of air masses having certain characteristics towards latitudes different from the original ones fits perfectly in this explicative framework. These air masses, remaining at a certain latitude for a long time, take on characteristics due to their contact with the ground. More specifically, polar air is cold and

---

air that comes into it must be equal to that of the air that comes out of it. This does not allow the air to be compressed or to become more rarefied. Obviously, this assumption can be made in the atmosphere (the height remaining the same) and not in other physical systems. We all know that, in other conditions, air can undoubtedly be compressed: this happens, for instance, in a bicycle pump.

[30] This mechanism is the inverse of the adiabatic cooling due to ascent in the atmosphere. Now the air cell, descending, undergoes a greater pressure in the low layers. This results in warming.

dry, whereas tropical air is warm and humid[31]. When, because of atmospheric currents, these air masses move towards different latitudes, their thermodynamic characteristics remain almost unchanged for a certain time. In the northern hemisphere, this results in movements of warm air towards the north and of cold air towards the south, which reduce the difference between the temperature of the air in the equator and that in the poles.

From the viewpoint of the physical and mathematical expression of what we have presented qualitatively, in the same way in which we previously showed that vertical dynamics and processes can be understood and described in thermodynamic terms (for instance with the application of the first principle to cases of adiabatic expansion), now we can endeavour to describe the horizontal dynamics of air by means of hydrodynamic concepts and equations. Since, as we have already stated, the two types of motion combine, the two descriptions must be applied simultaneously, in order to achieve an overall vision of the atmospheric motions and processes. From the viewpoint of horizontal motion, air can certainly be regarded as an actual fluid, provided its properties are described by means of averaged concepts like the previously cited one of air mass. It is thus possible to use concepts such as the density and velocity of an air mass, and to apply to them the basic equations of hydrodynamics (known as Navier-Stokes equations) on a rotating reference system and with the forces that act on the air mass.

It seems, then, that the problem has been formulated correctly. However, one of the essential difficulties that make this description very complex is the fact that as a rule these equations cannot be solved analytically.

4.8 Feedbacks and Complexity of System

In this chapter we endeavoured to give the reader a "key" for the theoretical understanding of the phenomena that take place in the atmosphere. We were inevitably compelled to isolate the various phenomena, in order to achieve a better study of their dynamics and of

---

[31] This is due to the greater evaporation of water surfaces at low latitudes.

the causes that produce them. Though we tried to follow the thread of air warming, the picture we obtained has several aspects, so this chapter may seem rather dispersive to some readers. Indeed the picture that has thus been obtained is not simple, so the only way to achieve a more organic vision of what happens in the atmosphere is to understand the mutual cause-effect relationships that each process or phenomenon has with other processes or phenomena occurring in the system under consideration. In the last part of this chapter, therefore, we will endeavour to recover this more organic vision, limiting the treatment to a few explanatory examples. By doing so, we could discover that our system is really "complex".

First of all, in order to "prune" the treatment by eliminating all the elements that can reasonably be disregarded, it is expedient to point out that in the Earth system there exist some processes that are characterised by different evolutionary time scales. For instance, we have already mentioned the fact that the temperature of the seas and oceans varies above all with the seasonal cycle and is only very slightly affected by the day-night cycle. In this respect, therefore, if we wish to study and understand the cause-effect relationships that involve the shorter-time-scale processes, on this scale we can often regard the influence of the hydrosphere roughly as a constant. Likewise, the amount of aerosols present in the atmosphere or the quantity of frozen surfaces on the entire globe can vary only by an extremely small percentage (with respect to their total amount) in the course of a few days; however, as regards the meteorological phenomena based on a small space and time scale, their possible quick local variations (for instance, snowfalls on previously bare ground or the emission of a great quantity of dust or pollutants) should be evaluated for their consequences on the processes that take place on this scale in the atmosphere subsystem of the Earth system.

The approach, obviously, is quite different in the study of climatic phenomena. As a rule, here the specific meteorological phenomena should be regarded as "fluctuations" around certain averages, while the processes with a slower temporal evolution are the ones whose evaluation is most important: they affect these averages, and ultimately guide the changes in climate.

Remaining in the by now familiar field of the discussion of the warming/cooling of the air due to the influence of the terrestrial ground, let us consider the warming produced by the ascent of the Sun above the horizon during the morning. As the hours go by, the ascent of the Sun determines an increasing irradiation that, if it falls on solid ground, leads to an increase in the surface temperature. This rapidly results in an increase in the temperature of the lower layers of the air via the mechanisms that we know well by now (conduction, convection and irradiation). Should everything boil down to this phenomenon, we would be able to accurately predict the temperature at the various hours of the day in a certain location. In actual fact, as we have explained, there are other elements that disturb this situation, for instance clouds. Now, clouds may appear in the sky because they are conveyed from more or less distant regions; but they may also develop or grow locally. In actual fact it is convection that, in the suitable humidity conditions, leads to the cooling of the rising air cells, the condensation of water vapour in the atmosphere and the consequent development of clouds. Among other things, the rise in ground temperature also results in an increase in the evaporation and transpiration of plants, and, to a lesser degree, in the evaporation from water surfaces (whose temperature does not change much). The most evident effect of cloud formation is that of screening the sunlight and causing a smaller amount of it to reach the ground, which consequently tends to warm up less or even to cool down.

We have thus started from a cause — the rise in ground temperature as a result of the incident solar radiation — that gives rise to a chain of effects: rise in air temperature, convection, increase in humidity, condensation. These effects lead to a process, cloud formation, that results in a decrease in the warming of the surface, or even in its cooling, via the screening of sunlight and the reduction of its incidence on the surface. This way we have described a chain of cause-effect interactions whose last link loops back to its beginning: the last effect acts on the cause that has generated the entire chain. There is, in short, a feedback produced by the change in the temperature value on itself. In this case, the final effect is that of lessening the increase in temperature: this is called a negative feedback. In other cases, as we will show, the final effect is that of intensifying the change: this is called a positive feedback.

Figure 6 is a diagrammatic representation of the meteorological interactions that take place in the Earth system. The variables that describe the state of the atmosphere and ground are represented by the rectangles, while the ellipses represent the main meteorological processes. The direction of the arrows indicates whether a process changes the value of a certain variable or whether, vice versa, this value affects a certain process[32]. Obviously it is possible to trace the feedback of the change in ground temperature value on itself (by following the two circular chains: one comprises ground temperature – sensible heat flow – condensation – clouds – radiation – ground temperature, and the other ground temperature – evaporation – humidity – condensation – clouds – radiation – ground temperature). It may also be amusing to count the number of circular interaction loops. And this is only a simplified diagram!

Now it is much easier to understand our allusion, in the introduction, to the existence of a circular causality, in contrast with the linear causality typical of Newtonian mechanics. Not only the causes do not add up in a linear manner and do not produce a sum of effects, but they actually "act back" on themselves in a non-linear manner. To put it more specifically, if we are studying an evolutionary phenomenon, such as the tendency of the ground temperature and the consequent tendency of the low-layer air temperature, from a mathematical point of view we cannot describe the situation with a single equation[33]. We must analyse the various processes by means of separate equations and their interaction in a system of equations. If, as is often the case, this is not possible (for instance because the system cannot be solved or the dynamic treatment of some processes is not easy), it is possible to identify the basic process that leads to temperature changes and to adjust its values by means of a separate examination of the other processes that affect it. In any case,

---

[32] The diagram shows, among other things, the "sensible heat flow". This phrase indicates the flow of exchanged heat that produces an increase in temperature: it includes the heat transmission processes by convection and by horizontal movements of the air. Though the figure generally shows only interactions between variables and processes, we have drawn a broken line also between the sensible heat flow process and the condensation process, in order to highlight the connection mentioned in the text.

[33] The variable T (ground) cannot be a dependent variable and an independent one at the same time!

there is a quantitative balancing problem in the calculation of the influence of the individual processes on the variable under consideration.

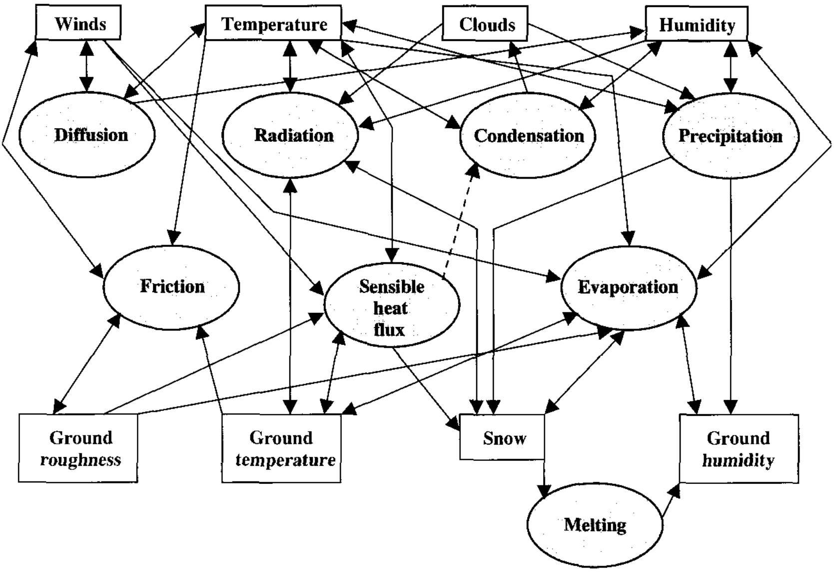


Figure 6. A diagrammatic representation of meteorological interactions and feedbacks.

If now we try to consider the situation from a climatic point of view, that is on a broader time scale, it is not difficult to understand that here too there are many causal loops among the processes that are important in the determination of the climate. Let us reflect, for the time being, on a global level, and suppose that, for some reason, the temperature of the Earth's surface and of the air start increasing. How would the Earth system respond to this increase? Undoubtedly the higher temperature would change the amount of surface covered by ice, causing it to decrease, because conditions characterised by a temperature lower than the freezing point of water would be limited to higher latitudes, or, the latitudes being equal, to greater elevations. This deglaciation effect, however, is not harmless, because it has an important consequence on the radiative balance between the Earth and outer space.

We have not followed up the subject of the properties of ice in its interaction with radiation, but everybody knows that white surfaces reflect most of the visible radiation that falls on them: so do ice and snow. In comparison with them, the solid or liquid surfaces of land and sea reflect much less and absorb a greater quantity of the incoming solar radiation. Therefore the decrease in frozen surfaces results in a greater absorption of solar radiation by the Earth's surface, promoting the warming of the Earth and consequently of the low layers of the air. So we have discovered a very evident positive climatic feedback: deglaciation starts a warming process and encourages a further rise in temperature. Among other things, positive feedbacks are more "dangerous" than negative ones: the latter tend to lessen the effects of a change that is under way and to bring the process back to its starting point, whereas the former tend to increase the departure from the previous situation, leading towards scenarios that are less known, because in many cases they have not been observed previously.

As a matter of fact, a change in a variable in the atmosphere usually affects several cycles. This applies also to an increase in temperature. Another effect, which at present is considered less important than the previously described effect of deglaciation, concerns cloud formation. As we have seen also on a shorter time scale, the intensification of convection and of the evaporation of the seas and oceans tend to cause an increase in the quantity of clouds present in the atmosphere, with a net effect of greater screening of the solar radiation; therefore this is a negative feedback.

Finally, a further factor that should be considered, particularly on a regional scale, is that of the dynamics of the previously mentioned conveyor belt that guides ocean currents. This dynamics depends on the temperature and salinity of the water in the various parts of the globe. Deglaciation, with the consequent flow of fresh water into the oceans, would disrupt this equilibrium. Recent studies[34] demonstrated that an excessive desalination of the northern Atlantic might certainly result in changes in the structure of deep ocean circulation. Some researchers even surmised that the Gulf Stream might disappear, leading obviously

---

[34] See, for instance, Wood *et al.* (1999).

to climatic consequences on the whole of western Europe, where a period of intense cold might begin, with a trend opposite that of the global warming context. Among other things, studies on the dynamics of ocean currents demonstrated that its changes are hardly ever slow and gradual: in most cases they take place within a short period, because this dynamics is affected by the exceeding of certain critical thresholds that determine a sudden change in the circulation[35]. This leads us to partly amend our previous statements about the oceans' "stabilising" role (essentially due to the great heat capacity of water) with respect to the climatic temperature fluctuations. This role, which remains basically active, does not, however, exclude sudden changes in single regions of the globe, resulting from a change in the directions of the ocean currents.

At this point, what should we say about the role of carbon dioxide and other greenhouse gases in the present global warming phenomenon? In Chapter 2 we showed some figures that present a striking positive correlation between the concentrations of CO_2 and the global temperature trend. After the discussion of the interactions between radiation and air molecules, we know that the effect of an increase in the greenhouse gases produces absorption of heat in the troposphere. But how does this effect combine with the other climatic mechanisms and with feedbacks that may be present in the system? Well, if we consider the increase in the concentration of CO_2 in the atmosphere as an observational datum, we can safely assert that it tends to cause an increase in the temperature of the air in the low layers. If we consider the interaction of the air with the ocean surfaces, we will find that in the long run this increase in air temperature will have a warming effect on those surfaces[36]. But now, because of a mechanism that we cannot describe in detail here, the process whereby carbon dioxide is stored in the oceans, and therefore removed from the atmosphere, is gradually becoming less efficient as the

---

[35] From a theoretical point of view, ocean dynamics is said to undergo "bifurcations" and to be partly chaotic. The concept of deterministic chaos will be discussed in depth in Chapter 7.

[36] When we study the cumulative effects of the contact between the air and the ocean surfaces on a large time scale, we can no longer disregard the perturbation it causes in the "measurement" of the temperature of these surfaces, as we had done previously for the sake of simplicity.

temperature of the oceans rises. So this last effect tends to further increase the CO_2 in the atmosphere. We have thus identified a positive feedback.

If we include water vapour among greenhouse "gases", we can assert that the ocean warming obviously leads to a greater evaporation, therefore to a greater presence of water vapour in the atmosphere, and that this further promotes a rise in the temperature of the troposphere (at least until we consider condensation and cloud formation, which, as we have already explained, are elements that lead to a negative feedback).

As the reader has certainly understood on the basis of the information provided in this chapter, the situation of the processes under way in the Earth system, and in its atmosphere subsystem in particular, is extremely complex, if we consider it from the angle of a theoretical understanding. As we have already indicated, in this book it is not possible to delve more deeply into this situation. What we have already explained, however, is sufficient to allow the reader to understand that each variable and each process present in the system interacts in a non-linear manner with other variables and processes, often creating circular chains of cause-effect relationships. Moreover, what occurs in the low layers of the atmosphere depends crucially on the exchanges at the interface with other subsystems, such as land, oceans, ice, and the biosphere (which includes human activities).

As we have previously explained, it is difficult to deal theoretically with this situation in a formal and quantitative way. Though the disciplines involved in the understanding of the various phenomena are all classical and well-established, a classical scientist of a few years ago would not have been able even to quantitatively weigh the various contributions to a typical evolutionary phenomenon such as air warming. Nowadays, on the contrary, this is done. How it is done is one of the main themes we will tackle in the next chapters.

Chapter 5

The Galilean Experimental Method: A Digression?

Summarising in a “skeletal” way the thread of our discourse, we may state that in the previous chapters we explained on which observational data meteorology and climatology are based, illustrated our theoretical knowledge of the individual phenomena and processes occurring within the Earth system, and were led to the final recognition that these phenomena and processes interact in a complex manner. More specifically, under the stimulus of the local study of the properties of matter, a considerable part of the progresses of our theoretical understanding of the system under examination came from reflections on “domestic” observations, in which, for instance, the behaviour of the atmospheric fluid could be observed more accurately. This way, we were able to conduct a direct, “first-hand” investigation into phenomena that normally take place also in the free atmosphere and are an important component of its dynamics, e.g. convection. Incidentally, we should point out that, in small environments that are isolated from the outside, individual processes are probably less influenced by other processes, contrary to the case of the free atmosphere, where a process is usually affected by other processes and sometimes there are strong feedbacks. Thus, for instance, in a closed room convection does not interact directly with sunlight or with the horizontal motions of air.

The approach to a theoretical understanding of the phenomena under examination presented in the previous chapter is based on accurate, sometimes local observations of what occurs spontaneously in nature. The idea that underlies this attempt at a theoretical synthesis is that of gaining knowledge by using both mathematical methods that are increasingly suitable for describing the observational situation (where possible, with more modern, sophisticated techniques), and observations that are increasingly accurate and continuous (in order to obtain as plentiful as possible a sample of different "historical" situations). In this context, obviously, theoretical activity in meteorology and climatology may be negatively conditioned by the fact that the observational results are only those presented by nature in its history. In actual fact, we have seen that some elements of theoretical understanding come from thermodynamics and fluid dynamics, understood as sectors of physics, and here this limitation does no longer exist in scientific practice.

As we shall see, the stepping-stone to a better understanding of nature rested, in some disciplines, on a transition from a tendency to perform increasingly accurate, "first-hand" observations of natural phenomena or processes, to a tendency to force what nature presents to us, practically "manipulating" the system under consideration. In order to examine this breakthrough in physics and in other so-called "hard" sciences, it will be expedient for us first to refer to an analysis of the simple systems that are familiar to all of us and in which the first "experimental" investigations historically took place. Once we have appreciated the advantages of this approach, we will be able to pose the question whether it can be successfully applied also to other more complex systems and to disciplines (such as meteorology and climatology) that have always been historically characterised by purely observational research.

5.1 Aristotelian Physics of Local Motions and the Advent of Galileo Galilei

When we decide to deal with simple systems that are familiar to everybody, we cannot but think of Galilean mechanics. We will therefore analyse (though without going into historical details) some elements of novelty introduced in scientific practice by Galileo Galilei during the seventeenth century: these elements now underlie all experimental scientific disciplines.

During the last few years, Galilei's name appeared rather frequently in the mass media, in relation to the revision process that the Catholic

Church (and Pope John Paul II in particular) has been applying to certain past instances of the ecclesiastic authorities' behaviour. So it is generally known that the Italian scientist, on the basis of his observations through the telescope constructed by him, discovered that celestial bodies are not perfectly spherical[1], studied their motions and eventually got to the point of supporting the Copernican heliocentric theory against the Ptolemaic-origin geocentric one. This was the basic reason for which the Inquisition demanded and obtained Galilei's abjuration.

In actual fact Galilei was not only an attentive observer of the solar system: besides his well-known discovery of Venus's phases, of the existence of Jupiter's satellites and of their motions[2], he also carried out a more "local" activity, studying simple mechanical systems. It is precisely this second activity of his that is regarded by scientists as the most innovative one, because it introduced an investigation method that is largely in use today and is often referred to as "Galilean experimental method".

In the previous chapters we mentioned the fact that pure and simple empirical data, in themselves, do not shed light on the behaviour of matter: they must be fitted into a well-defined theoretical scheme. Only with a "reading" of this type they become usable, both from the viewpoint of the understanding of the system under consideration and from that of possible interventions on it. Galilei's activity in the study of the dynamics of simple mechanical systems revealed, for the first time, the limits of explanations that were regarded as evident and "natural", and were directly induced by the observational data and interpreted in the perspective of the common sense stance of that period. It will be instructive, therefore, to briefly consider how Galilei analysed the problem of the fall of bodies, and, more generally, of all local motions, and how he ended up by introducing a new vision of dynamics in place of the Aristotelian one.

---

[1] Galilei realised, notably, that there are mountains on the Moon. This observation was very important, because it contradicted the Aristotelian-origin assumption that our sublunar world is formed of irregular, imperfect matter, whereas the celestial world must be formed of incorruptible matter, with regular, perfect (spherical) shapes.

[2] All these observations backed up a Copernican vision, or in any case a non-geocentric one.

Most of the Aristotelian description of local motions reminds us of the naive physical image we all developed as children, when we examined these motions and found regularities in them on the basis of the repeated observation of everyday experiences. For instance, noticing that an iron ball falls to the ground more quickly than a wad of cotton, and examining the fall of various objects of different weights and densities, we were led by induction to believe that heavier bodies always fall more quickly than lighter ones. Likewise, when we placed a body on a horizontal surface, we all saw that, if no force is applied, the natural state of motion of a body is that of rest; moreover, in this situation the body certainly moves if one pushes it, but as soon as one leaves off pushing it, it stops, sometimes after having covered a further stretch because it is still propelled by the *impetus* it has received. An important aspect of this description is that it is constantly corroborated by everyday experience, and the actions we perform in harmony with this vision are very useful for practical purposes: for instance, if a heavy object is about to fall on my foot, I know that I must move my foot away very quickly, in order to avoid getting hurt.

In Aristotle's system, obviously, this description that we have called naive was consistently fitted into his vision of the world[3]. From this point of view, for instance, the interpretation of the vertical fall of a stone towards the ground is that, since it is chiefly formed of the earth element, it tends to reach its natural place; likewise the flames of a bonfire, where the fire element is predominant, tend to rise in order to reach their natural place, the sphere of fire. Moreover, Aristotle was certainly not a naive observer: this is testified by the fact that he also allowed for the resistance of the medium in which the motion took place, and went so far as to assert that the velocity, V, of the body whose motion was considered was proportional to the ratio of the imparted force, F, to the resistance of the medium, R; simplifying, and adopting a mathematical

---

[3] In Aristotle's vision of the world, while celestial bodies are formed of aether, all sublunar bodies are formed of a mixture of four elements: earth, water, air and fire. Under the sphere of the Moon, the four elements are arranged in concentric spheres, according to their weight: from the bottom up, first there is the sphere of earth, then that of water, then that of air, and finally that of fire.

form that obviously Aristotle did not use, we may express this with the formula $V = F/R$.

This last assertion is precisely what allows us to understand that Aristotelian physics is not problem-free: everybody knows that if a heavy body is on a rough table and one pushes it delicately it does not budge: a greater force is required to make it move. The proportionality relation that we expressed as $V = F/R$ implies, on the contrary, that the body should move (though slowly) even if a very small force were applied to it. What should we say, finally, of a body to which a force has been applied and that goes on moving even when the force stops being applied? This is what happens when a pen is pushed violently onto a table or an arrow is shot from a bow.

In the Middle Ages some scholars of natural philosophy had already realised that Aristotelian physics had some weak spots. On the whole, however, the solutions proposed remained within the sphere of the Aristotelian principles, at least up to Galilei's time[4]. The Italian scientist's power of analysis and abstraction led him to thoroughly criticise, for the first time, the Aristotelian-type theoretical scheme of local motions and to develop, at the same time, a valid alternative to it.

An attitude undoubtedly peculiar to Galilei, that distinguished him from natural philosophers of the Aristotelian sort, was his approach to natural observations and to their current theoretical interpretation. He often tried to imagine the qualitative results that might be obtained from the analysis of events that had never been observed before, endeavouring to draw inferences from them, even only from the angle of internal

---

[4] To mention an example, in the sixth century after Christ, John Philoponos proposed a change in Aristotle's proportionality relation: it could be expressed as $V = F - R$ and produced a positive velocity only after threshold R had been exceeded. Moreover, the problem of the motion of an arrow was usually solved by imagining that the arrow, moving forward, created a depression on its own butt. Since in Aristotelian physics the case of vacuum was excluded, where the resistance of the medium was null by definition and therefore (because of the relation $V = F/R$) the velocities should have been unrealistically high or even infinite, the continuation of the motion of the arrow was interpreted as a result of the force imparted to it by the air that rushed in to fill the vacuum around its butt. An attempt to solve this problem that broke away from the Aristotelian outlook was that of the theory of *impetus*: in this case the arrow, during its flight, allegedly preserved a sort of memory of the thrust originally imparted to it.

consistency, for the theoretical framework in which the results were being read.

A typical expression of this attitude of Galilei's was the argument with which he "demolished" the Aristotelian theory of the fall of bodies. According to Aristotle, the velocity of the fall of bodies is proportional to their weight, so heavier objects fall to the ground at a greater speed than lighter objects; this seems to be confirmed by natural observations. Galilei tried to imagine a new situation, in which two objects of different weights are joined to form a heavier body, for instance by tying them to each other with a string or even gluing them together. What happens to the new body? According to Aristotle's theory, since its weight is greater than that of each of its two "parts", it should fall to the ground at a velocity higher than that of the fall of the heavier part. On the other hand, in the Aristotelian perspective of a velocity of fall proportional to weight, the lighter part (whose velocity of fall is lower) would slow down the fall of the heavier part, which in turn would speed up the fall of the lighter part: the resulting velocity of the composite body would be intermediate between that of the heavier body and that of the lighter one.

Thus, with two arguments that are both within the Aristotelian logic, we obtain two different qualitative results for the velocity of fall of the composite body. Starting from the assumption of a velocity of fall proportional to weight, we reach a contradiction: this suggests that the assumption is not valid! We do not know whether Galilei ever actually carried out an experiment of this type. In any case, a conceptual contradiction like this one was already sufficient to address him to a different path from the Aristotelian one in the investigation and theoretical interpretation of local motions.

5.2 The Galilean "Style"

So Galilei's first step in the direction of a modern scientific conception was that of imagining the possibility of observing events expressly created by human beings, for instance, events that had never occurred before in nature. In this sense, Galilei spoke of "sensible experiences"[5],

---

[5] Galilei (1967).

both for natural observations and for the "experiences" devised and carried out by him. In the latter case, the "experience", for instance with the construction of the composite object described above, was expressly chosen by the scientist, and its purpose was to test a theoretical explanatory scheme of the phenomenon under examination. Another element that distinguished the Galilean vision from the Aristotelian one was the importance that it gave to the quantitative analysis of the problems by means of an extensive use of mathematics: we have already mentioned the fact that Galilei believed that the book of nature was written in mathematical terms.

The second step (the decisive one) towards a modern scientific conception was taken by Galilei with the method used for choosing and preparing the "sensible experiences" devised and carried out by him, which, from now on, we will call "experiments". On the one hand, at the basis of the *choice* of an experiment there was a theoretical hypothesis about what happens in nature, in the field under examination, for instance the dynamics of moving bodies. On the other hand, at the basis of the *preparation* of an experiment there was the awareness that an individual phenomenon is affected by manifold physical factors, and that it is possible to separate these factors and eliminate or minimise some of them, so as to simplify the system, up to the point of being enabled to study the effect of a single cause on the phenomenon under consideration.

As regards the theoretical hypothesis that underlay the choice of the experiment to be carried out, this hypothesis was usually suggested to Galilei by natural observations and previously conducted "experiences"; however it did not stem directly from them or from the consequent accumulation of data by simple induction. During this stage, the scientist's creative act appeared to be essential. In the scientific biography of Galilei, we can find several ways in which he was led to develop hypotheses for the law of inertia and the fall of bodies. An example derives from the consideration that bodies of different weights fall at different velocities through different media (oil, water and air); by adding to this the fact that the difference between these velocities decreases as the density and viscosity of the medium decreases, Galilei reached, by extrapolation, the hypothesis that in vacuum (where, by definition, density and viscosity are null) all bodies, heavy and light, fall at the same speed, independently of their weight[6].

However, the creative act would have remained barren had it not been united to the ability to prepare an experiment that could reveal something about the validity of the hypothesis thus formulated. Only with this preparation, Galilei's power of abstraction was put to the test. The essential characteristic of the Galilean method of investigation consisted precisely in the combination of the power of abstraction with the ability to "manipulate" physical reality in order to prepare and carry out an actual experiment.

A fundamental idea that Galilei expressed in some of his writings was that the world hides its real mathematical nature behind "spurious" elements that disrupt regularities and thus make it difficult for us to understand it. The Salviati-Galilei of the *Dialogue*[7] explicitly maintained that it was necessary to "deduct the impediments of matter". This programmatic assertion materialised in the preparation of the experiments, when the Italian scientist "released" nature from its impediments, for instance by smoothing off his tables and inclined planes. This way he kept only the essential structure of physical nature, reducing to a minimum what, in his opinion, prevented a simple mathematical reading of the dynamics of local motions, i.e. friction. Thus he discovered the law of inertia (whereby a body that is not subjected to forces persists in its state of rest or uniform rectilinear motion), and found the correct relation between the time and the space covered by non-horizontal motions: this relation eventually led to a direct generalisation in Newton's second law of dynamics.

This Galilean approach was obviously influenced by a vision of reality that was reminiscent of the Platonic one, with the true essence of

---

[6] It may be interesting to point out that Galilei adopted extrapolation as a useful conceptual practice in other occasions as well: the best-known example is that in which he had to defend the veracity of his astronomical observations through the telescope. When someone asserted that the lens had a deforming effect and therefore led to untruthful observations, Galilei replied that observations through a telescope could be personally verified on the Earth (by observing an object at a distance, then going there to check the object). Since this was true for increasingly great distances, it was reasonable to infer that it was true also for distances such as that between the Earth and the Moon.

[7] Galilei (1967).

things situated in an external world (which for Plato was unattainable and for Galilei could somehow be approached through experiments); and this, in fact, was how it was interpreted in the past. Without attempting a philosophical discussion, we wish here to emphasise the methodological novelty of this approach. More specifically, the analysis of the elements of the system (the ones that are considered fundamental and the ones that are considered spurious) and the elimination of some of them (the spurious ones) can be interpreted as a typical causal-influence analysis on a single phenomenon. The elimination of the spurious elements confirms an actual separability of the individual causes that act on the observed phenomenon: this separability is first surmised theoretically, then realised practically with the smoothing off of the tables.

Thus Galilei paved the way for a new "style" in scientific investigations. Scientists were no longer content with carefully observing nature and inferring its regularities by induction, but advanced original hypotheses and put them to the test, with experiments in which all the working conditions were controlled by the experimenter and only rarely corresponded to those found in observational reality. The latter, in many cases, was too complex to be interpreted correctly without a suitable simplification of the conditions in which the observations took place. It was necessary, therefore, to prepare the system so it could be examined in simpler conditions, for instance by removing the spurious elements: this made it possible, more specifically, to determine the effect that a single cause could produce on a certain phenomenon.

5.3 A Galilean Method for Studying the Weather and the Climate?

With Galilei, in short, domestic observations became laboratory experiments, in which the initial state of the system under examination and all the conditions in which the experience took place were decided by the experimenter. In a laboratory we can obtain conditions that would never appear in nature, to the point of making it possible to analyse the behaviour of the system under the influence of a single cause, if necessary repeating the experiment over and over again. Thus the determination of a physical law in simplified conditions can easily be achieved; and a law can be expressed mathematically by means of an equation.

In this process, we have implicitly assumed that the individual concurrent causes that affect a certain phenomenon can be separated from each other; in our case we may even consider eliminating all of them except one (gravity)[8]. This way we get to the point of studying an ideal situation, which would be difficult to find in natural reality. If, however, now our goal is to achieve a correct description precisely of this natural reality in its complexity, the next step may certainly be that of adding another cause in the controlled system and of analysing the combined effect of the previously studied cause and the added one. For instance, we can reintroduce the friction component in the system, and verify how it perturbs the law of inertia or the uniformly accelerated motion in the phenomenon of the fall of bodies. In a simple dynamical system, we may find, for instance, that the effect of friction is that of slowing down motion in comparison with a situation in which there is no friction, and that friction leads to the addition of a term in the equation that describes the motion. Proceeding in this direction, we can attempt to reconstruct natural complexity in the laboratory, where work is carried out in controlled conditions.

These operations of "decomposition and recomposition" of a system in the laboratory, together with the controlled manipulation of all the experimental conditions, have produced enormous results in terms of the description and understanding of systems, starting from classical physics and chemistry, and proceeding to nuclear and subnuclear physics. Our basic knowledge of non-living matter is for the most part due to the application of the Galilean scientific method.

Considering these premises, it is reasonable to suppose that if we could apply this method also to the study of meteorology and climatic

---

[8] In comparison with Galilei's time, nowadays our experimental capability is obviously much more sophisticated, so we can rely less on extrapolations (for instance, we can obtain vacuum minus a very small number of molecules per cm^3). In this context we should point out that, perhaps because of the intrinsic limits of the seventeenth-century experimental apparatus, or more probably because of a cultural inclination of his, Galilei always showed that he preferred "sure demonstrations" to "sensible experiences": it is even possible that he never performed some experiments that have been attributed to him in the past.

changes, our knowledge and our possibility of describing and predicting what takes place in the Earth system would be enormously increased. Is it possible to achieve this theoretical and practical result?

To answer this question, we must first of all emphasise the fact that the purpose of the activity of decomposition of a system in the laboratory is to acquire as detailed as possible a knowledge of the nature of the phenomena and of their responses to individual causes. The subsequent activity of adding other elements to the system is undertaken, on the contrary, in order to achieve, in controlled, repeatable conditions, a more realistic description of what actually takes place in nature, where the various elements of the system interact in a way that is difficult to interpret by simple observation. So the former activity has a purely theoretical value, because it aims at increasing the knowledge of basic phenomena, whereas the latter activity is both theoretically and practically important, because its purpose is both to study the interaction between the various elements of the system, and to attempt to "mimic" the behaviour of the natural system, endeavouring to reconstruct it, somehow, in controlled conditions.

This double operation of decomposition and recomposition of a system in the laboratory sometimes gives rise to problems. More specifically, this activity is usually easy when the phenomena under examination depend on a small number of causes, perhaps linear and easily separated, but is extremely difficult when the phenomena are part of a highly interacting system, where the phenomena "emerge" from a complex, non-linear mixing of concurring causes. As we explained in detail in the previous chapter, this is the case of the Earth system[9].

Another example of the difficulties that are met when attempting to examine the behaviour of the atmosphere in the laboratory is the study of air as a thermodynamic fluid. Obviously in a laboratory it is possible to study small air masses, if necessary using experimental apparatuses (for

---

[9] Here I would like to avoid any misunderstanding: the Galilean experimental method is not applied only to mechanics or classical physics, but also to more "difficult" fields, where investigation requires sophisticated instruments and theories, for instance nuclear physics: here, however, this is enormously facilitated by the alleged linearity of the phenomena, which are described by a theory that is "difficult" but linear, such as quantum mechanics.

instance revolving spheres and cylinders) that force them to perform motions similar to the real ones. In reality, the air masses involved in the atmosphere are much larger. How does this fact affect the "translation" of the experimental results obtained in the laboratory into the macroscopic domain of the real atmosphere? There even exists a theory, called theory of similarity, that deals with the mathematical similarities found between phenomena observed empirically on various scales when they are correctly scaled up or down and that can be applied to the "translation" of these laboratory findings.

What should we say, moreover, about the fact, which has also been already discussed, that some phenomena seem to emerge from complex macroscopic interactions, without having an equivalent in microscopic processes? In the laboratory, we usually study with great attention the elements that form a system (e.g. air molecules) and the processes on this scale (e.g. molecular diffusion), whereas in meteorology the validity of the adiabatic hypothesis, discussed in the previous chapter, suggests that some of these processes may be disregarded, at least in good approximation, while others take on a dominant role.

These last reflections lead us to believe that whereas on the one hand the Galilean experimental method can be applied successfully to the study of the basic elements of the atmosphere and of the Earth system (for instance as has been done actually for the study of the interactions between radiation and matter), on the other hand the complexity of the system seems to elude a local experimental investigation. The ideal thing would be to have a laboratory as large as the Earth itself, in which it were possible to manipulate the system, i.e. to prepare its initial conditions and its "boundary" ones, in order to carry out real experiments and to extract theoretical syntheses from them, in a simplified, repeatable way. Obviously this is not possible: nobody is able to perform such a broad "meteorological experiment" or, even less, such a "climatic experiment", simply because we cannot achieve the control of all the elements of the system. It seems, therefore, that we are not able to fully exploit the Galilean breakthrough, which has been so fruitful in other disciplines.

So up to the nineteen-sixties and nineteen-seventies the study of meteorology and the climate had a purely observational character and had a great difficulty in formulating theoretical syntheses, not so much for the study of the basic elements of the system and of their elementary interactions, to which the Galilean approach could be applied, as for the study of the system in its totality and complexity, for which scientists were compelled to keep to theoretical-empirical rules that had very little descriptive and predictive power.

During those years, however, something changed; and, from that time on, the problem we posed here of the retrieval of the Galilean approach also in meteorology and climatology can be considered in a different perspective. What has changed? Does there exist a technical ploy that makes it possible to get around the problems that had not allowed us to apply the Galilean experimental method completely? Or do we have to achieve a new breakthrough in scientific practice, a change in our cultural attitude, in the broadest sense of the word? We will endeavour to answer all these questions in the next chapter.

Chapter 6

Simulation Models

In the previous chapter we explained how Galileo Galilei usually worked out his own idealised representation of the reality he was studying, first of all formulating hypotheses, then putting them to the test in the experimental "trial". The final result was often expressed in the formulation of mathematical relations that described the phenomena of motion in the controlled conditions of the experiments: Galilei did this, for instance, in the equation that linked the elapsed time to the covered distance in the fall of bodies. For the problem that is being studied here, this equation is a valid example of algorithmic reduction.

This way the mental representation of the course of a certain phenomenon, after it has been studied experimentally in simplified conditions, leads to a mathematical formulation that describes the operation of the simplified system, which is believed to possess the essential characteristics of the systems seen in reality.

To put it in the language of today, Galilei first constructed a "mental model" of the course of motions in basic conditions (i.e. after the spurious elements had been eliminated), then verified this model by means of a "material model" (the inclined plane), and finally worked out a "mathematical model" (the equation of motion) that formally and synthetically described the motions in these simplified conditions.

6.1 How Many Meanings Does the Word "Model" Have?

The concept of "model" is of great importance today in all scientific disciplines, not only in the physical ones. At the same time, as the reader has certainly understood from the various adjectives we have added to it, this term can be used with different meanings in scientific terminology — as it can, actually, also in common language. Any dictionary can show us the great variety of meanings that may be attributed to this noun. A model may be defined as a thing or person that is considered exemplary, and therefore worthy of being imitated; as an original to be reproduced; as a person who sits for painters or sculptors; as an industrial prototype; as a garment that is sewn as a unique specimen according to an original design; and so on.

Proceeding to a more strictly scientific sphere, we will now concisely analyse the various meanings that coexist at present and have developed during the course of the history of science.

The original meaning of the word "model" was probably that of a material reproduction of a physical system on a certain reduced or enlarged scale: for instance, maps of the Earth, or, in more modern times, reconstructions of the structures of crystals (enlarged) and of the atmospheric circulation on revolving spheres (reduced). These reproductions are useful displays of the systems under examination, but, as we stated in the previous chapter, it is not always possible, by changing the scale, to obtain reliable information about the behaviour of the original natural system. Even the inclined planes and the balls used by Galilei cannot be regarded as simplified material models that reproduce, through a simple change of scale, all the characteristics of the motion of stones rolling down the sides of a mountain (velocity, acceleration, space covered, time elapsed). These systems became useful only when their use made it possible to determine some fundamental laws (such as those of the fall of bodies) that could subsequently be applied to systems on different scales. In this sense, modern experimental laboratory reconstructions do not aim at imitating the behaviour of greater systems: their purpose is to reproduce, in controlled conditions, the basic physics of systems on the small scale of the laboratory.

The inability of a simple material reconstruction to achieve an imitation, on a different scale, of the behaviour of a natural system reveals the need for the construction of a mental model in which previous hypotheses or theoretical knowledge coexist with an ideal, simplified representation of the physical system to be studied. This new concept of model actually appears in various examples in the history of physics: the best-known one is perhaps the planetary atom model, in which simple little balls with a negative electrical charge (electrons) revolve around a positively-charged nucleus, under the "thrust" of the Coulomb attraction. The latter is described in the same mathematical form as gravitational attraction, so the analogy between planets and electrons turns out to be cogent.

This concept of model as an ideal display guided by our theoretical knowledge, however, was gradually dropped in the history of physics, undoubtedly in the case of microphysics: the brief history of the planetary atom model illustrates the difficulties inherent in the displaying of microscopic bodies[1]. On the other hand, in classical physics the physical elements that are considered are macroscopic, therefore can be displayed, located and separated. This suggests that it may be possible to construct an ideal model, at least for some phenomena: for instance, we can visually imagine the interaction between two air masses, in a limited environment, under the action of their different thermodynamic characteristics, with the warm air masses sliding above the cold ones and the latter pushing under the former, thus "modelling", at least qualitatively, the frontal interaction we briefly described in Chapter 4.

In any case, both in classical physics and in quantum physics, the real breakthrough in the understanding of a phenomenon takes place when a mathematical model of that phenomenon is worked out, that is when a description of it (and, if possible, also of its future evolution) is achieved by means of mathematically formalised physical laws that represent an algorithmic reduction of the apparent complexity of the real system. The mathematical model is usually formed of one or several equations in which the individual variables represent the properties of the individual elements of the real system under examination (for instance, the temperature and humidity of the air masses that come into contact with each other). So there exist some actual rules of correspondence that

---

[1] Simple calculations of classical electrodynamics demonstrate, for instance, that an electron that performs a circular motion (and therefore undergoes a uniform acceleration) quickly loses energy and soon ends up by falling on the atomic nucleus. The quantum description, moreover, which is the only one to adapt to the properties of microscopic bodies, shows that they cannot be displayed, located or separated in a classical sense.

interrelate the descriptive variables and the properties of the real objects, whose interactions can be studied concretely in the laboratory.

Obviously, when one is in possession of the mathematical model of a phenomenon (obtained, for instance, from the formulation of hypotheses and from the previous execution of experiments), this model can be validated again in new experimental situations in the laboratory. The validation on the real system, though in the simplified conditions of the laboratory, is always necessary if the model is to be regarded as a scientific one.

In practical terms, if, for instance, we are in possession of an evolutionary mathematical model (i.e. a model that is believed to be able to describe the evolution of a certain phenomenon or process in the course of time) and wish to find out whether it is a more or less faithful representation of an experimental reality in the laboratory, we must usually solve the equations of the mathematical model, by fixing an initial instant and a final one, and by determining the values of the variables at the initial instant and all the boundary conditions of the system under examination. If we consider the example of the two air masses that have different thermodynamic characteristics and interact, we will have to introduce into the equations of the mathematical model the various initial values of velocity, temperature, humidity, etc., of the two air masses, and also the temperature of the walls of the room that contains them, both at the initial instant and throughout the experiment. Once the equations with these initial and boundary data have been solved, and once the laboratory experiment has been carried out, we can compare the results predicted by the mathematical model with those obtained through the measurements. If the values predicted by the model for the different variables in the different points of the experimental room are "similar" to the ones that have been detected experimentally (i.e. if they remain within the error bars that characterise these measurements), we may assert that our model is a good description of what takes place in the reality of the laboratory.

What we have stated up to now makes it evident that the mathematical model represents an element of theoretical knowledge of the system under examination. This knowledge may be limited to a very simple real system, which is decomposed in the laboratory up to the point of considering a single cause acting on the phenomenon under investigation (as in the cases of the motions considered in the previous chapter); or this knowledge may be extended to a mixing of concauses that act on the phenomenon. In any case, the activity of decomposition and recomposition of the real system present in the laboratory has an analogue in the mathematical model, which, for instance in the case of a recomposition, may be formed of a system of coupled equations. And at any rate the purpose of the elaboration of a mathematical model nowadays is very pragmatic: it usually aims at mathematically describing and reconstructing the experimental reality within a very limited domain.

If a more complete theoretical construct is sought, we have to use the term "theory", and the range of validity of a theory must be broader, at least enough to include all the experiences known in the field under consideration; in a theory all arguments may spring from axioms (i.e. from basic postulates regarded as certain), and the aim of a theory may even be that of grasping the essence of the investigated reality (which in Galilei's view meant approaching the Platonic perfect world). Once a theory has been adopted, it is possible to work out models that use its equations to describe particular phenomena or processes: as a rule, when a laboratory "experience" is available, only a comparison between the results of the model and the experimental reality can lead to the corroboration or refutation of the individual equations.

6.2 The Simulation Approach

As we have explained, when the system under examination is a complex one like the atmosphere or the whole Earth system, the Galilean experimental method, and the consequent fruitful interchange between mathematical models and laboratory experiments, can be applied only locally for studying the basic properties of the individual constituents of the system and a few simple interactions between them. This is true for any system characterised by strong feedbacks and manifold concauses that interact in a non-linear manner on a single phenomenon, in other words for any complex system. Must we therefore lose all the fruitfulness that springs from the interaction between theory and experiment through the concept of mathematical model?

To answer this question, it is expedient to shift our attention temporarily from the experimental difficulties we have just mentioned to the theoretical ones that underlie non-linear mathematical models. With the usual recomposition technique, when we wish to describe a complex system in which two or more causes produce the change of a certain quantity in a circular manner (i.e. with a feedback), we consider a system of two or more non-linear equations, whose solution, with certain initial and boundary conditions, supplies the model's prediction for the changes in the variable under examination. Now, the crucial fact here is that these systems of equations make it possible to achieve analytic solutions only in very particular, simplified cases that are hardly ever realistic. Up to a few decades ago, this limitation was added to the difficulties of experimental reconstruction, and made it utterly impossible to tackle the problem of working out a mathematical model for a complex system, drawing from it theoretical prescriptions about the system's behaviour and evaluating its effectiveness through an experimental comparison.

Today, at least from the angle of theoretical description, the situation has changed. The introduction of computers has made it possible to solve these systems of non-linear equations in all physically substantial and relevant cases. Obviously these solutions are not analytic, but numerical: they are found by software programmes by solving these systems step by step, supplying a spatially and temporally discrete evolution of the variables contained in them.

All in all, we can consider a real system that evolves in the course of time, and we can describe its behaviour by means of a mathematical model whose equations are solved numerically by a computer[2]. If the variables present in these equations are linked to properties of the objects of the real system, what we obtain is a "simulation" of the real system's behaviour on the computer. So we will use the expression "simulation models" to indicate those mathematical models whose equations depend on time and on variables that correspond to some actual properties of the

---

[2] If the equations are differential, this is called "numerical integration".

system under examination, and are solved numerically by means of a computer.

In the "virtual" world of the simulation model, it is possible to follow, step by step, the evolution of the variables in space and time, if necessary displaying their values in a graphic form: in this sense, at least in classical physics, the simulation models make it possible to meet the need for visualisation that is particularly felt by scientists and that characterised the ideal models we discussed previously.

As we have explained, before the introduction of simulation models the comparison between a mathematical model and a real experiment depended on the ability of the individual scientists to carry out analytic calculations in strictly predetermined and simplified conditions, and on the reconstruction of these situations in the laboratory. Now the possibilities of experimental verification are becoming more extensive. And that is not the whole story. Even in cases where it is not possible to reconstruct in the laboratory the complexity of a macroscopic real system, a simulation model allows us to consider the variables and interactions that are regarded as fundamental and to simulate the evolution of a simplified system, subsequently comparing its behaviour (simulated in the model) with the one observed in reality. In this case, the goal may be to verify the correctness of a choice of variables and interactions that aims at reproducing (more or less faithfully) the phenomena and processes found in observational reality.

As our explanation so far has made clear, the strong point of the simulation approach consists chiefly in the fact that it makes it possible to tackle complex problems from a theoretical point of view. Simple systems, with only one or a few causes that interact in a linear manner, can be described by means of equations characterised by analytic solutions. On the contrary, the "recomposition" of a complex system starting from the causes and interactions we regard as fundamental — an activity that, as we have explained in the previous chapter, is extremely arduous in the real world of the laboratory — can be undertaken more easily in the virtual world of the simulation model. This world can be manipulated and controlled at will, and within it we can study how certain macroscopic phenomena emerge from the interaction of basic processes.

More specifically, this reconstruction activity can be made increasingly realistic by introducing in the simulation model some theoretical elements relative to interactions or processes that are regarded as less and less important. In this case, obviously, the experimental validation is not achieved by means of a comparison with the results of laboratory experiments, but from the observation of a real system.

In the light of what we have just explained, it is clear that the Galilean experimental method discussed in the previous chapter can be applied directly only within the local sphere and to the basic properties and interactions of the constituents of a complex macroscopic system. Should we wish to reconstruct the complexity of a real system of this type, this would be possible not in the real world of a laboratory, but in the virtual one of a simulation model. In the latter, we would start from basic theoretical elements, the individual equations of the model (the only elements to be validated by means of a classical experimentation activity in a real laboratory), and compound them, in order to simulate the complex interactions that take place in the real system. Though on the one hand the real world eludes the researcher's control, on the other hand the simulation model can be extensively manipulated, and there the scientist can carry out "virtual experiments". In the world of the model we can determine, *a priori*, the conditions in which the interactions are to be studied, and at any time we can change both the equations and their coupling (this allows us, for instance, to describe the intensity of the feedbacks theoretically), and also the initial and boundary conditions that affect the phenomena under examination.

So at this point we have explained the advantages of the application of the Galilean experimental method to simple systems, we have illustrated the difficulties of a purely observational approach to the study of meteorology and science of climate, and we have concisely presented the simulation approach to the study of complex systems. The conclusions we have drawn are that, though it is impossible to reconstruct in the laboratory the complexity of macroscopic systems such as the atmosphere, it is possible to "transfer" this reconstruction into a different sphere, that of the simulation model. Within the latter we can retrieve our theoretical knowledge (through the model's equations), the description of the system's physical state (through the initial conditions) and the interaction with other systems or with the external environment (through the evolution of the boundary conditions). This way, experimental activity becomes possible in the world of the model, and its validation is achieved by means of a comparison with the behaviour of the real system that is being simulated, albeit in a simplified manner. Once this change of perspective has been accepted, it seems natural to regard the simulation approach as the "heir" of the Galilean approach to the study of reality.

6.3 Conceptual Novelties in the Simulation Method and in Its Use

Given the conceptual importance of the simulation method in modern scientific practice, before we proceed, in the next chapters, to discuss its specific use in meteorology and science of climate and to stress the peculiar, original aspects of these applications, it is worthwhile to dwell a little more on the aspects of novelty (shared by all the areas of application) of this approach[3].

Summarising concisely what we have already partly discussed in this chapter, and adding a few elements of further clarification, we may list the main characteristics of a simulation model, as follows.

- There is a one-to-one correspondence between the quantities of the real system and the variables of the simulation model: what evolves in reality has an equivalent in what evolves in the model. It is possible to display graphically and follow over a period of time the behaviour of these basic variables: thus we obtain a concrete (and dynamical) equivalent of the ideal models worked out in the past by physicists.
- The real system is assumed to be decomposable into basic processes and interactions that can be described by single equations, "pieces" of theory that achieve the algorithmic reduction of these processes and interactions. The purpose of the simulation model is not to corroborate or falsify these equations, but to reconstruct the complexity of the system by means of these equations, to simulate

---

[3] The remarks that follow in this chapter are considerably influenced by some discussions I had with D. Parisi, whose clarity of mind and depth of thought were a great help to me.

the operation of the real system, and to make it possible to validate the model by means of a comparison with the large-scale behaviour of the system. Underlying this way of proceeding, there is the fundamental idea that, whereas for studying and understanding simple systems it is essential to decompose the system and analyse its individual causes in the laboratory, for understanding a complex system it is necessary for the individual components of the model to interact so as to correctly reproduce the behaviour of the real system. In this sense, to reproduce correctly means to grasp (both qualitatively and quantitatively) the complex non-linear dynamics of the system under examination.

- The simulation of the evolution of the characteristics of the real system is achieved by means of a numerical solution of the equations of the simulation model whereby the values of the corresponding variables in space and time are obtained in a discrete manner[4].
- With simulation models, the behaviour of a system is studied as a whole and in its complexity, without isolating it from the environment with which it interacts. The action of the latter can be described dynamically, for instance by means of an equation that outlines its evolutionary behaviour and interacts, in a non-linear manner, with the other equations, or only as an external forcing factor that is not affected by feedbacks from the system.
- In the world of the simulation model, scientists are completely in command of the virtual system that simulates the real one. They can carry out numerical experiments and repeat them at will, with an extreme ease in changing the theoretical elements of the model and the situations of the experiments (for instance, by changing the values of the variables). More specifically, as after Galilei scientists began to manipulate reality in order to obtain experimental situations that had never appeared in nature, now in the simulation model scientists can create possible worlds that have never existed in

---

[4] The solution of the equations of the simulation model takes place in a discrete manner: there are no continuous solutions like those of the analytic solution. The reader who wishes to understand more concretely how this can occur is referred to the examples in the next chapter.

natural history. As we will explain in Chapters 7 and 8, this possibility is extraordinarily important for understanding future environmental scenarios.

- Referring again to the fact that in a simulation model it is so easy to manipulate the studying conditions, in the same way in which it is possible to hypothesise future scenarios and observe the relative behaviour of the system, in "historical" sciences it is possible to reconstruct the past (for instance, different environmental conditions) and to study the behaviour of the virtual system in the model under these conditions. This activity is extremely important, because only for the past we can achieve a validation of the model by means of a comparison with the behaviour of the real system that is being simulated.
- In the same way, by changing the equations of the model or their coupling parameters (which define the feedback values), scientists can easily evaluate the empirical effectiveness of several theoretical schemes in the evolution of a complex system.

Having said this, I would like to conclude this short *excursus* in the world of simulation models with a few more basic remarks that are particularly relevant to what we will discuss in the next chapters.

It is known that in present-day science specialisation is carried to extremes. As early as the first half of the twentieth century, the great Danish physicist Niels Bohr humorously expressed this conviction by remarking that "an expert is a person who has made all possible mistakes within a very narrow field". Well, when we deal with complex systems, we frequently happen to meet with different theoretical descriptions of individual components of a system, in each of which the language typical of the specific discipline that studies it has been adopted. If we wish to reconstruct the behaviour of the system as a whole, we must then "link up" these languages. In this sense, the working out of a simulation model for a complex system compels the individual experts to speak the same language, in particular to devise a unified, interdisciplinary formalism to be introduced in the model.

The emergence and increasing development of simulation models, therefore, essentially promoted the tendency towards a unified, interdisciplinary vision of reality that has been appearing in the scientific world during the last decades. But the introduction of simulation models had other consequences as well!

In actual fact, the use of these models in natural sciences and in sciences of mankind has led to changes in scientific practice. The ease with which scientists could manipulate the elements of the model led them to regard simulation models as virtual experimental laboratories where they could carry out "experiments" that were not possible in reality. We have already explained, for instance, how difficult it is to reduce the complexity of a system such as the atmosphere to the narrow space of a laboratory; we should also consider that it is impossible to carry out real experiments for very long periods, like those that are characteristic of climatology. Well, in the virtual laboratory supplied by a simulation model, space and time can be expanded at will: the speed of calculation of a computer makes it possible to simulate experiments that in reality would have to be prolonged for tens or hundreds of years.

If we add to this the capability of a simulation model to perform a "synthesis" through the recomposition of the phenomena — i.e. the fact that it can realistically account for the complexity and manifold interactions and feedbacks of a complex system — we can understand why these models are becoming essential tools in modern scientific practice.

Finally, our remarks up to now have made it clear that simulation models are useful for the study of complex systems. The structure of simulation models is usually influenced by a tendency that is extensively present in the history of science and that we have already discussed: reliance on a reconstruction of the behaviour of a macroscopic system that is based on the composition of elements and interactions that are regarded as fundamental and have been validated separately in the experimental activity of the real laboratory. If these basic elements belong to a different (lower) level with respect to the phenomena to be studied, this amounts to a relapse into a form of reductionism; otherwise, it is nothing but reliance on the possibility of reconstructing the complexity of reality on the basis of simpler phenomena.

In the next two chapters we will examine some concrete applications of these simulation models within the scope of this book. We will realise, however, that from these applications there emerge some important conceptual aspects that need to be discussed. In the last chapter we will concisely return to an examination of the methods of present-day model making and of possible alternative approaches that may be conceptually interesting and practically useful within the sphere of the disciplines that study complex systems.

Plates

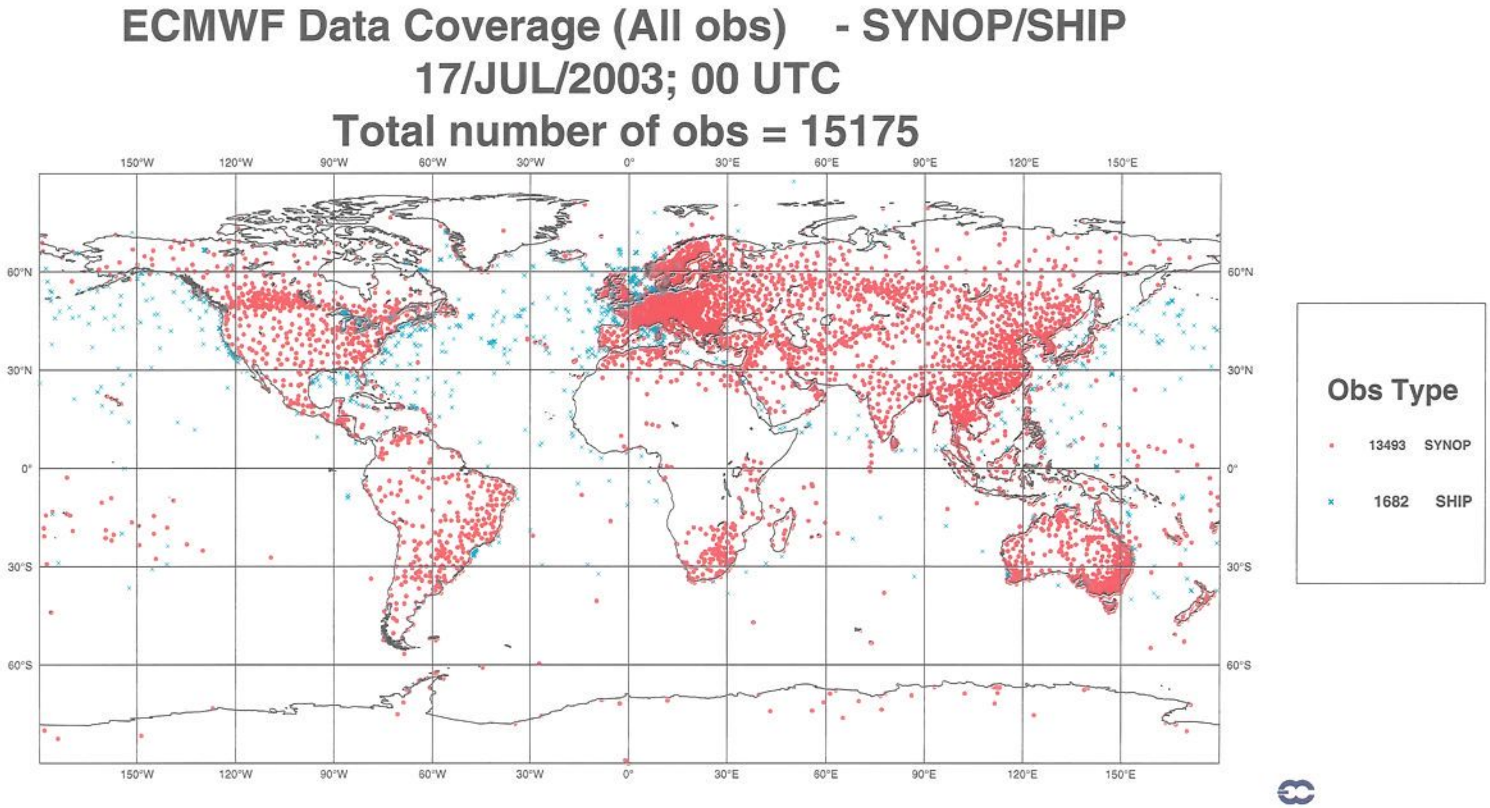


Plate 1. Global coverage of ground level observations as they reach an international forecast centre on a randomly selected day (source ECMWF).

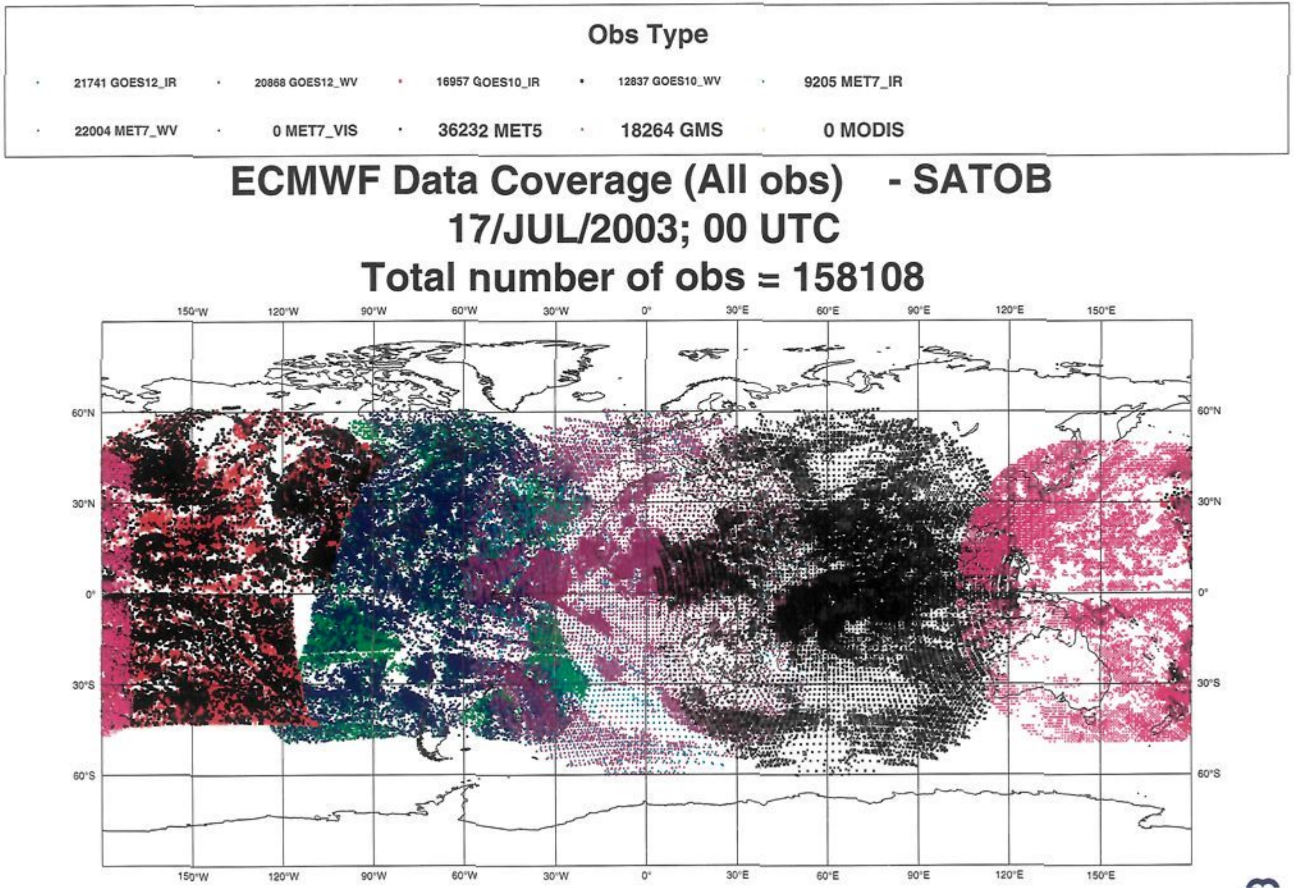


Plate 2. Coverage of observations by geostationary satellites on a typical day (source ECMWF).

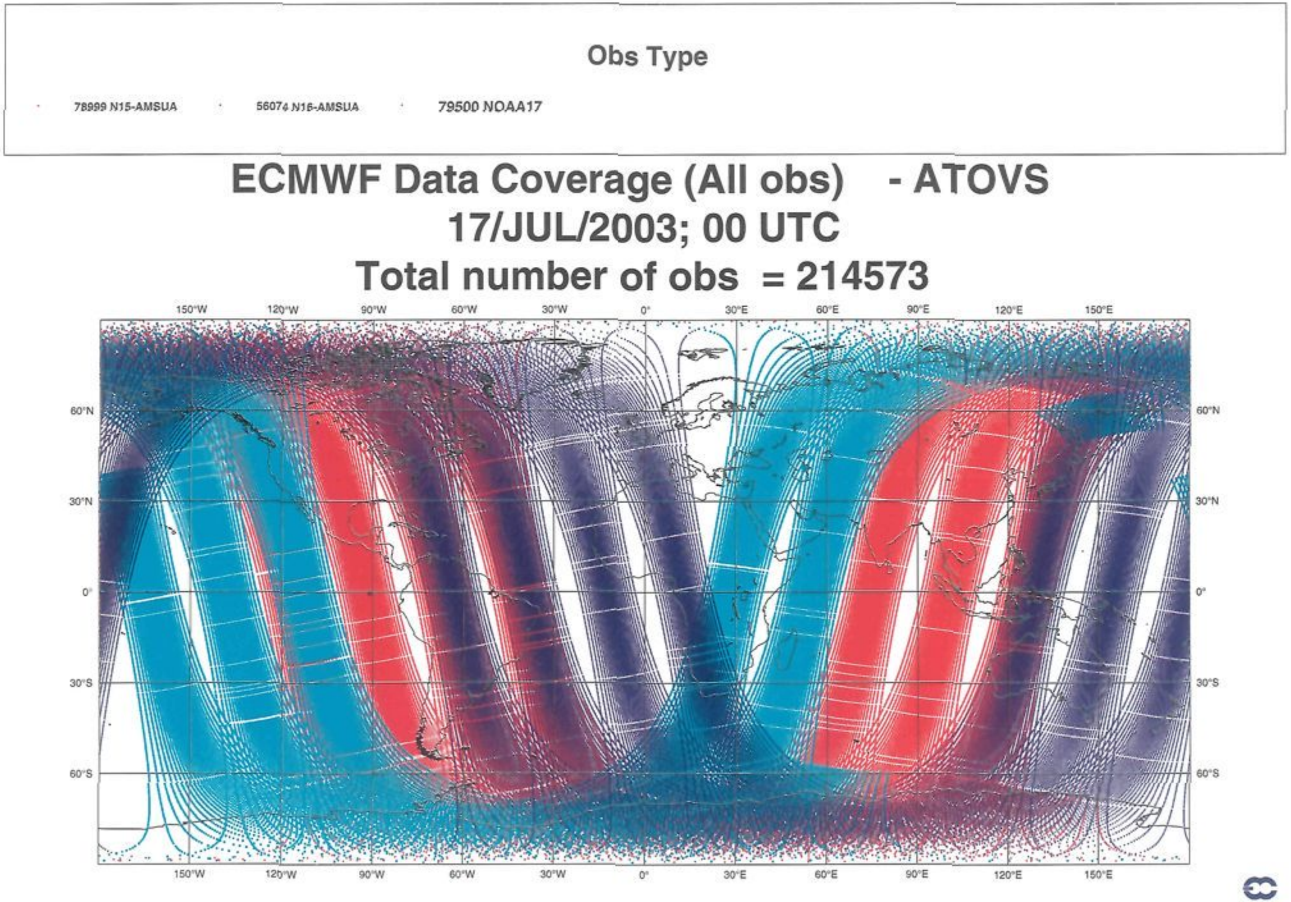


Plate 3. Coverage of observations obtained by the TOVS and relevant to the period 21–03 GMT of a typical day (source ECMWF).

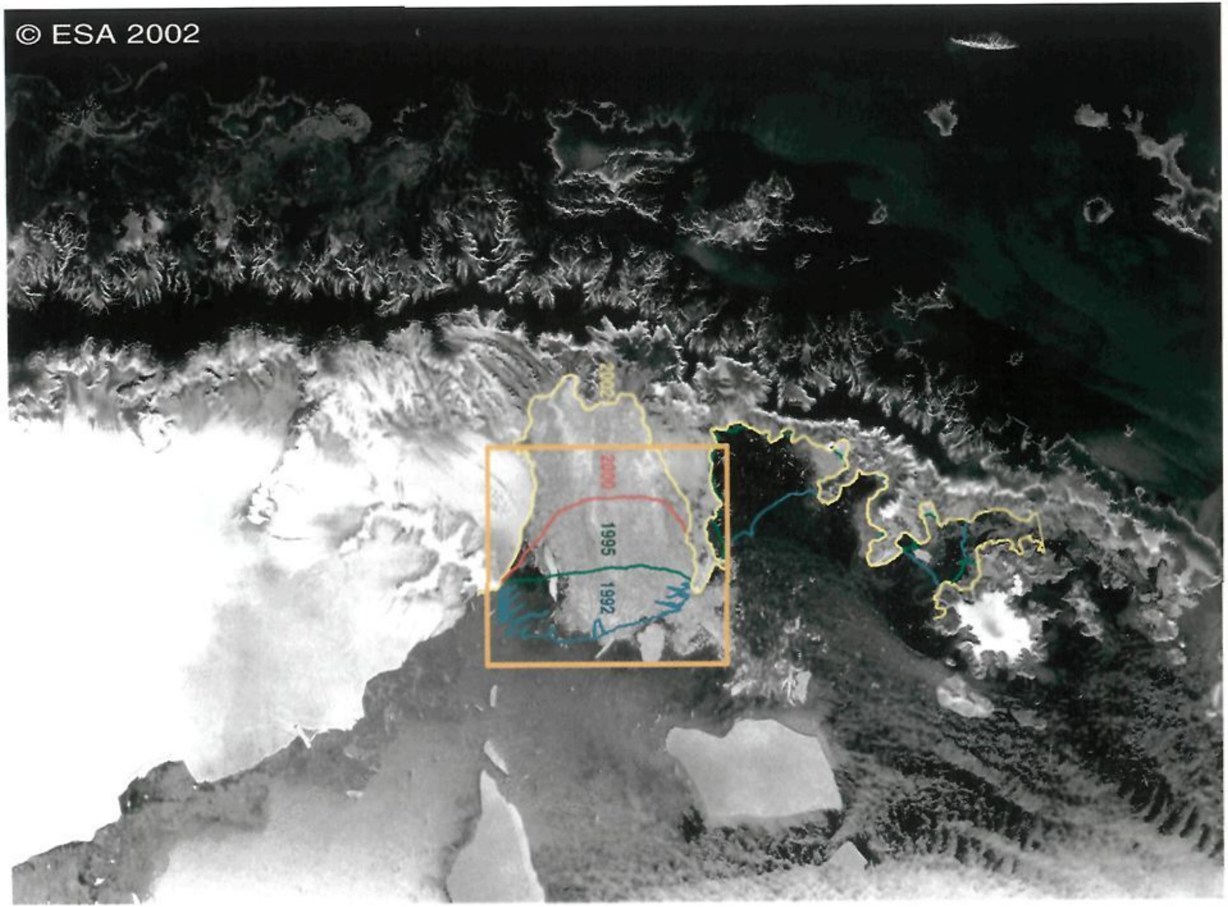


Plate 4. An example of the disgregation of the Antarctic pack: ASAR looks at the Antarctic Peninsula (source ESA).

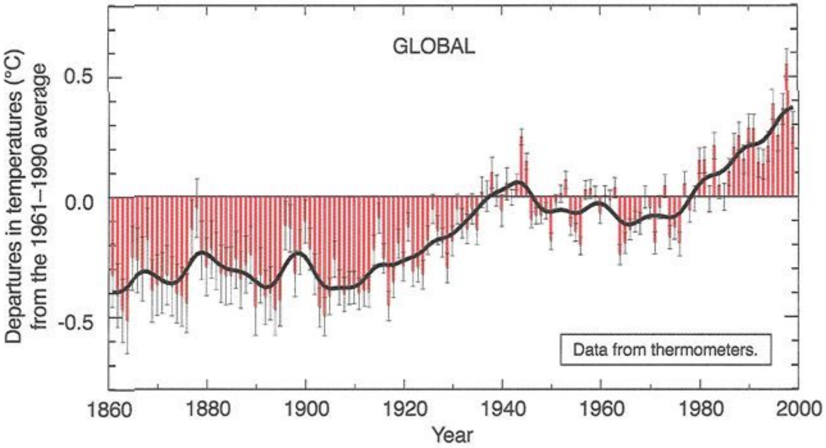


Plate 5. Anomalies in the global temperatures of the last 140 years (source IPCC).

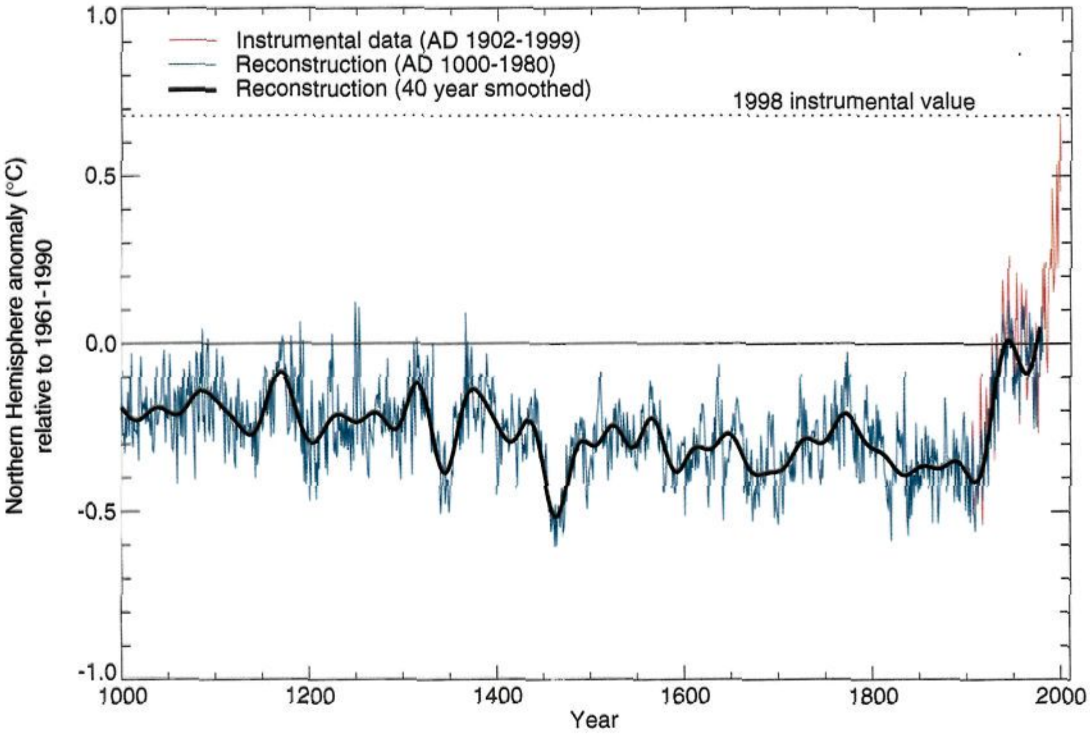


Plate 6. Temperature anomalies during the last millennium (source IPCC). The estimated error bars are shown in gray.

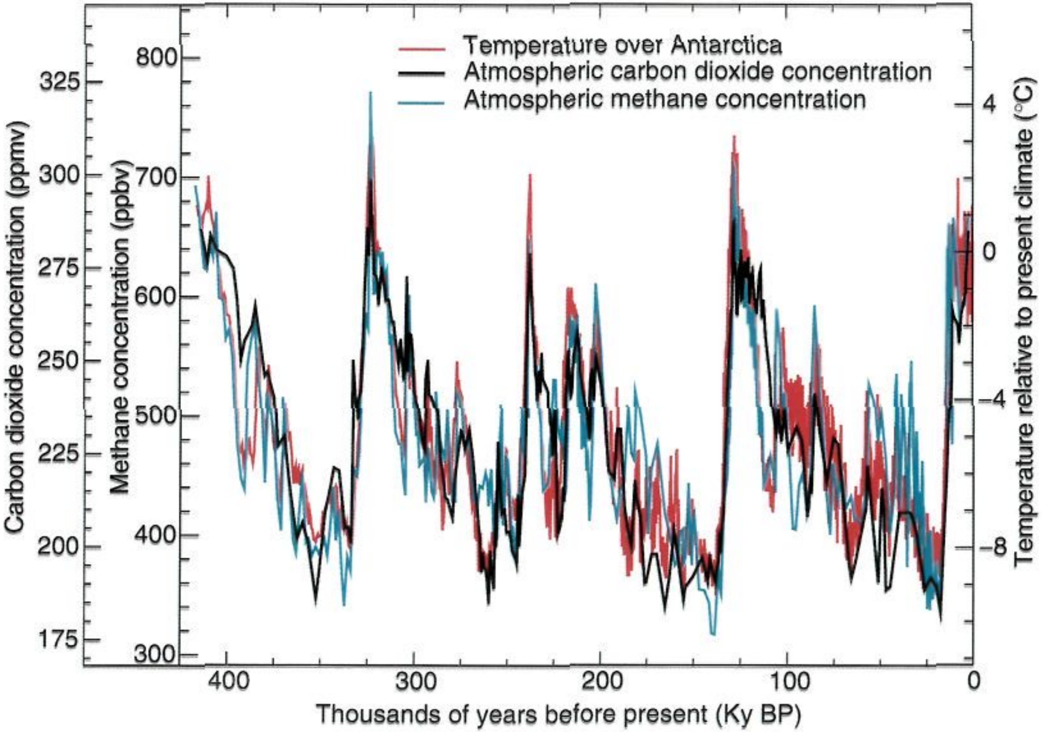


Plate 7. Comparison among the historical series of temperature, concentration of CO_2 and concentration of CH_4 (source IPCC).

Monday 3 Feb 2003 00 GMT
ECMWF Forecast T+36 VT: Tuesday 4 Feb 2003 12 GMT
Geopotential height 500 hPa + Temperature 500 hPa

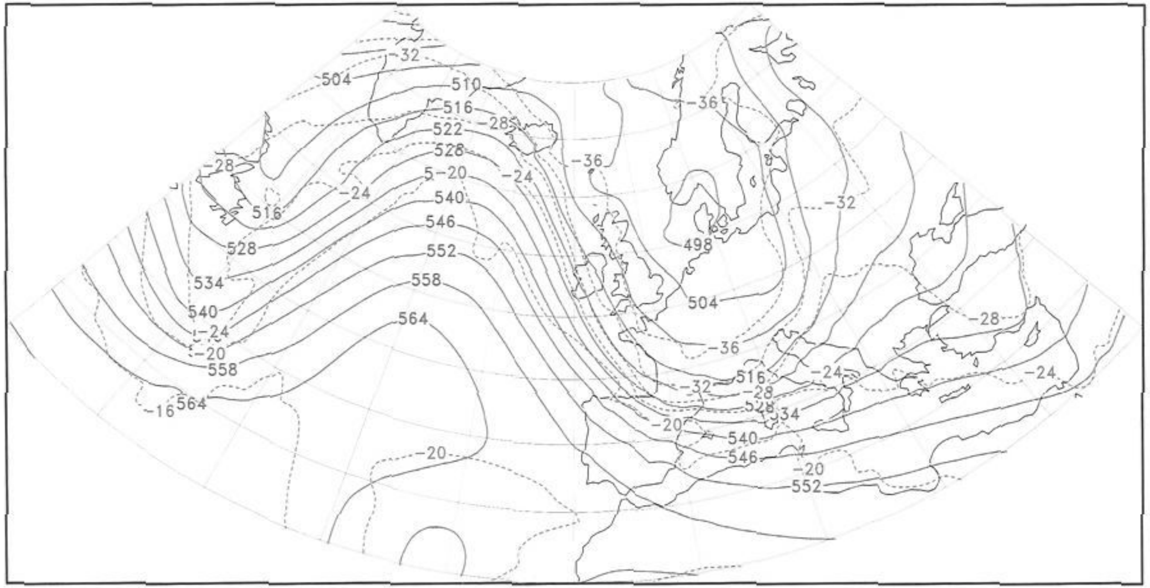


Plate 8a-d. Prognostic charts for a weather forecast (source of the data ECMWF).

Monday 3 Feb 2003 00 GMT
ECMWF Forecast T+36 VT: Tuesday 4 Feb 2003 12 GMT
Mean Sea Level Pressure + Temperature 850 hPa

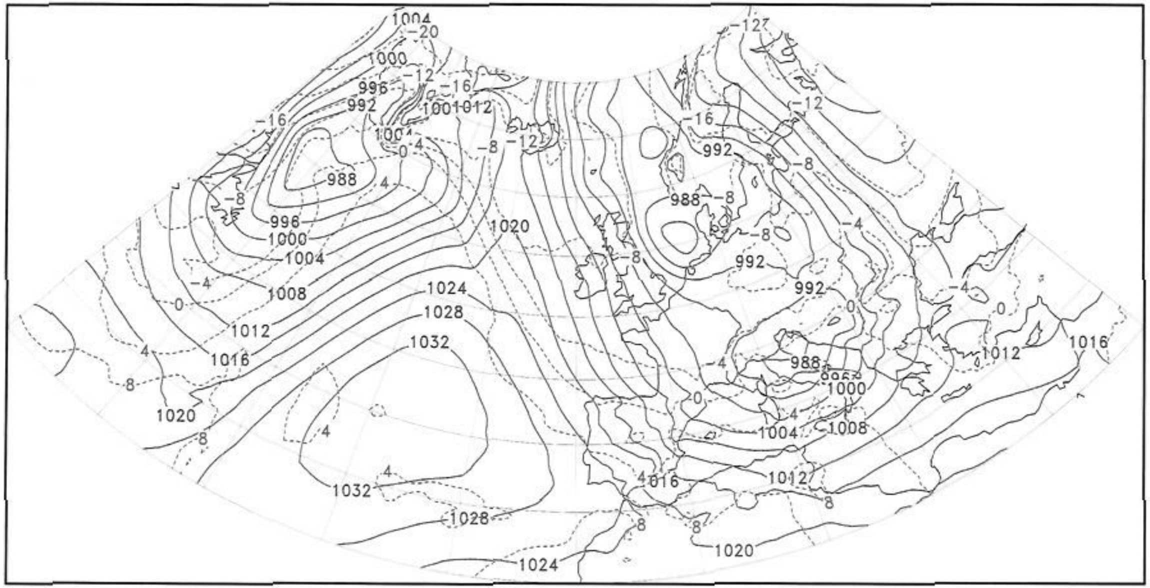


Plate 8b. (*Continued*).

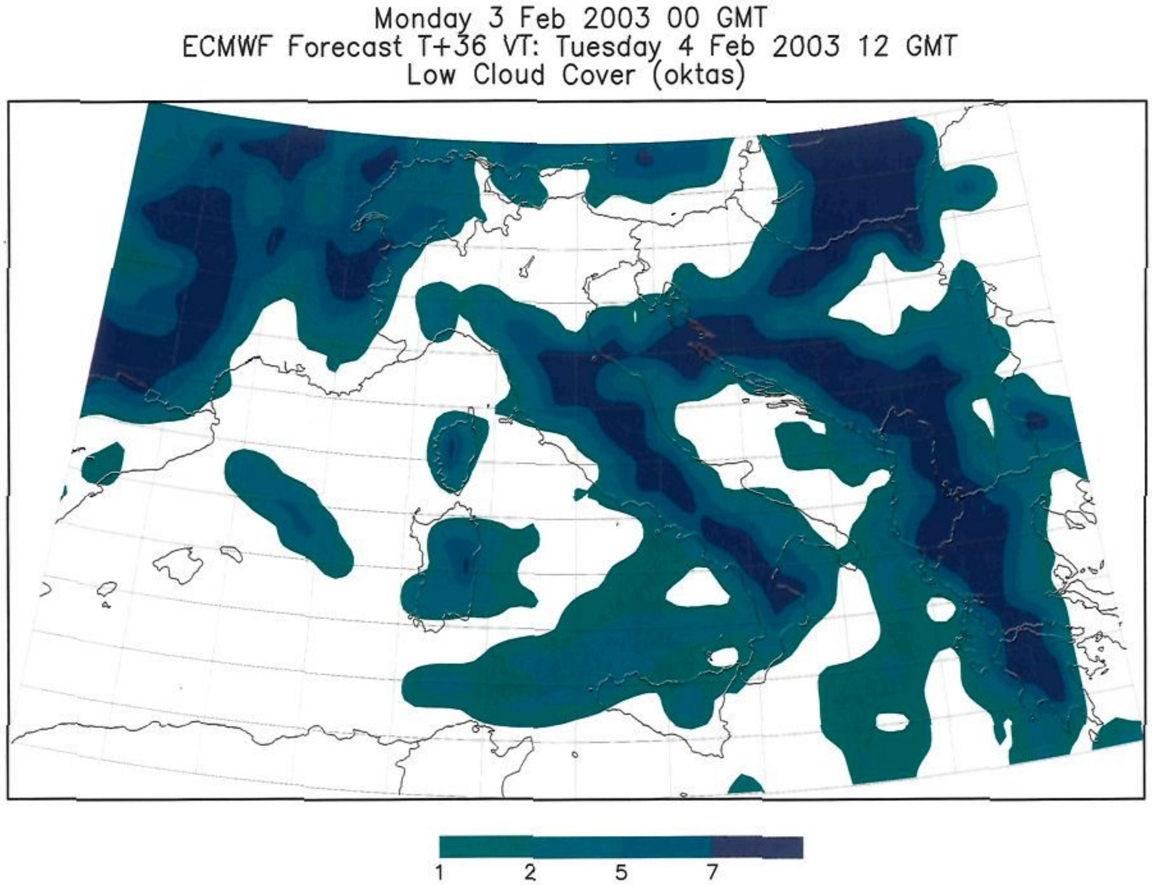


Plate 8c. (*Continued*).

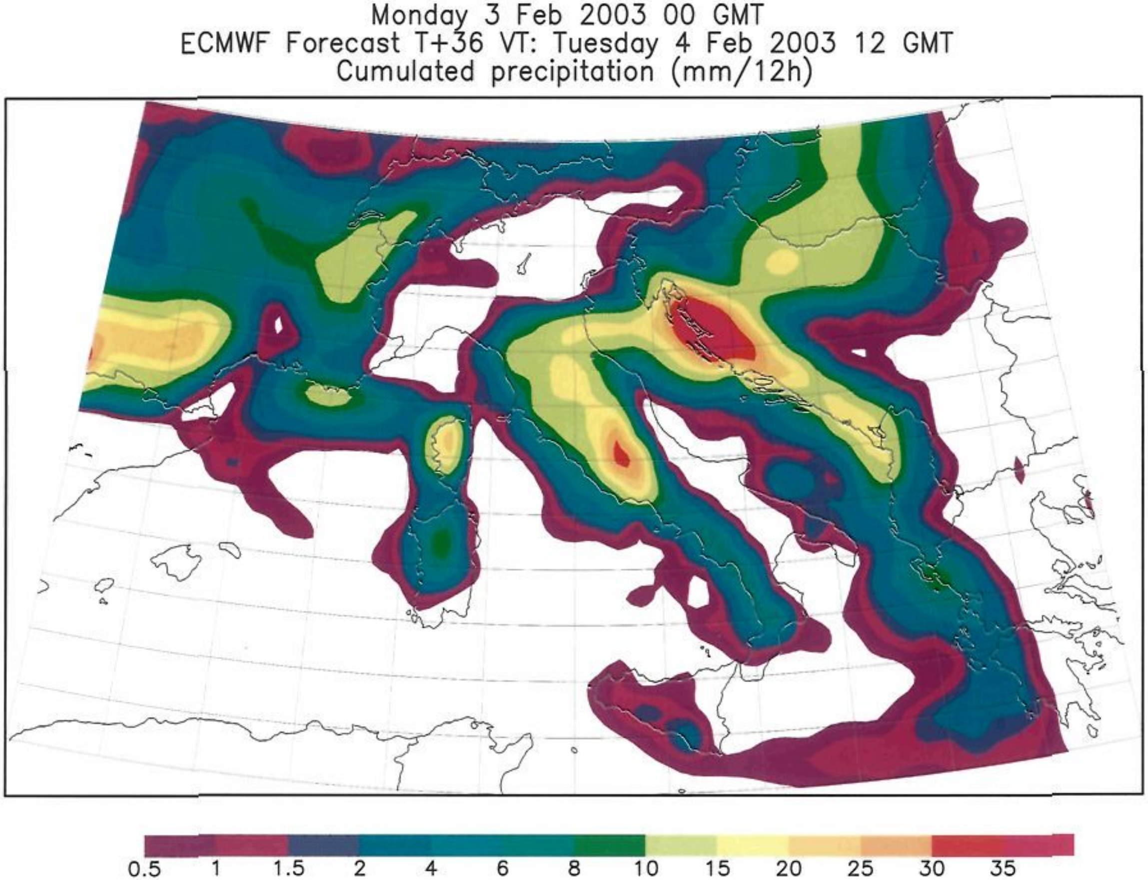


Plate 8d. (*Continued*).

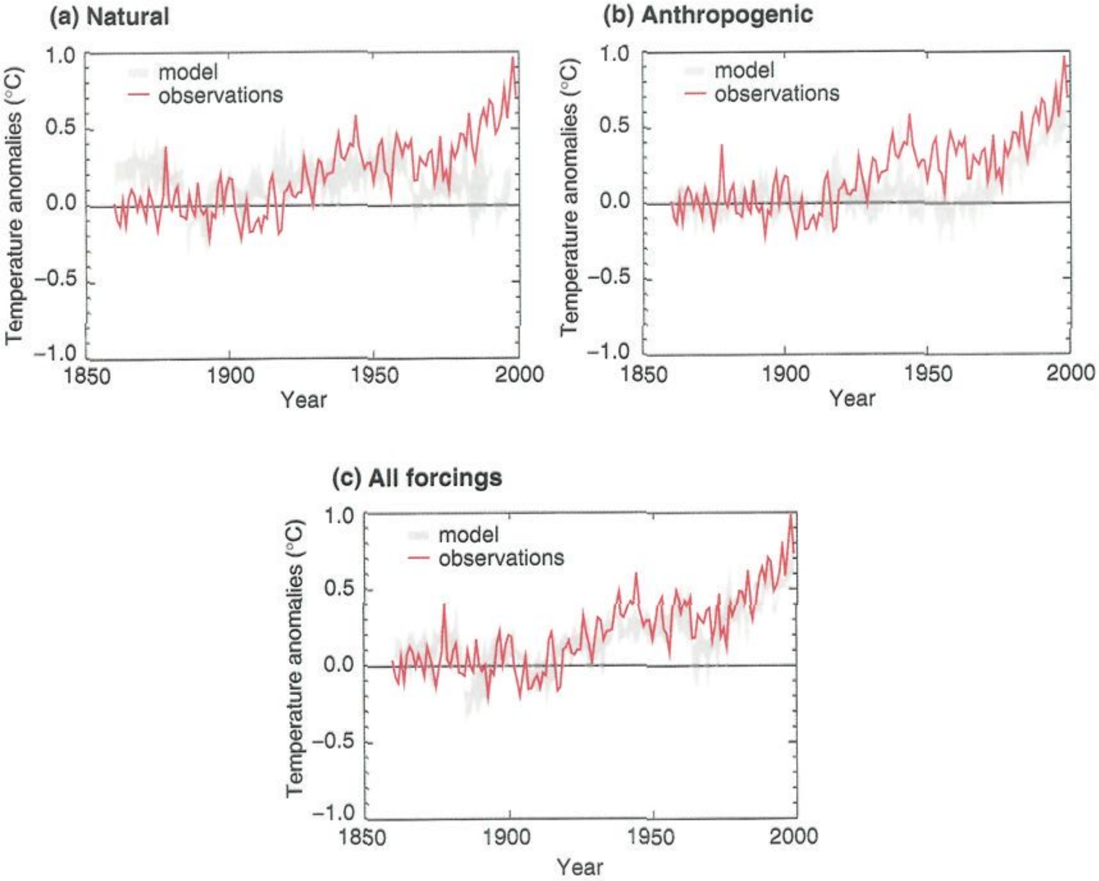


Plate 9. Reconstruction of the trend of the global temperature subject to the various forcing factors considered in an AOGCM (source IPCC).

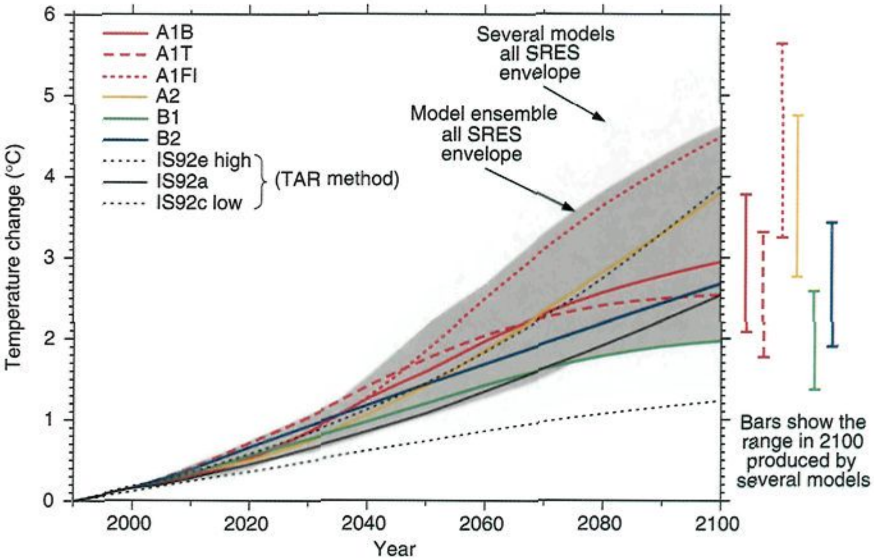


Plate 10. Evolution of global temperature during the present century, as forecast by simplified climate models under the influence of various emission scenarios (source IPCC).

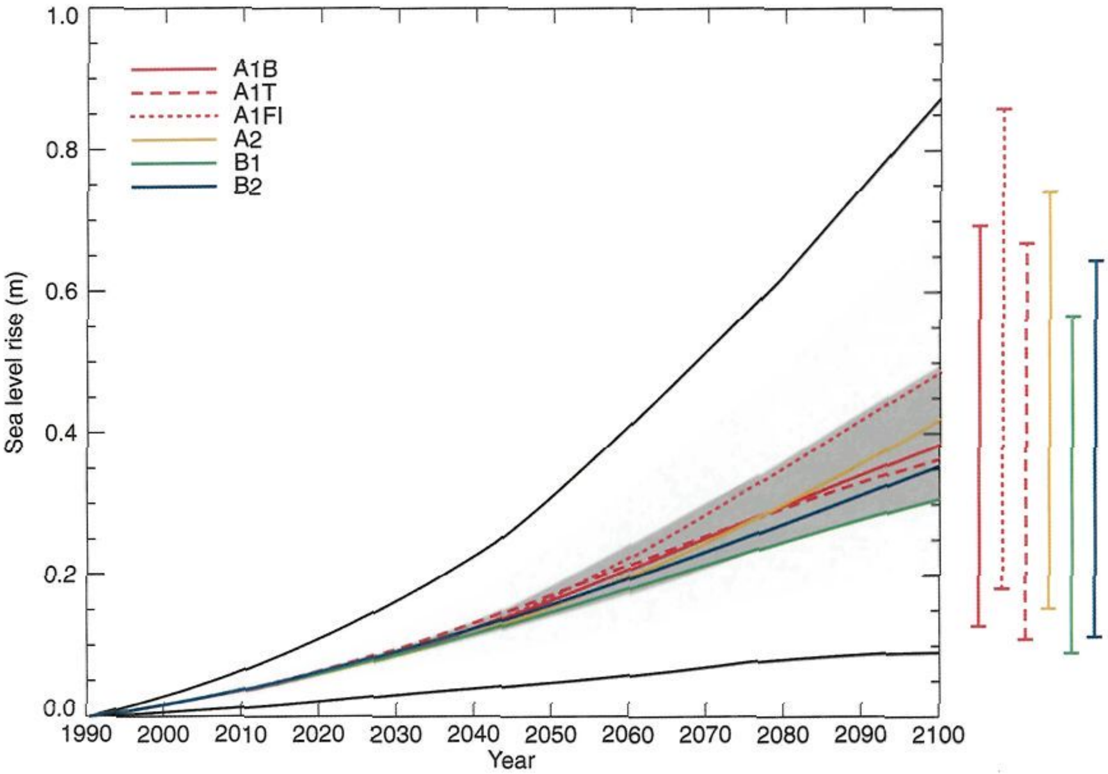


Plate 11. Sea level rise forecasts (source IPCC).

Chapter 7

Meteorological Models

In the previous chapter, we introduced the concept of simulation model. We particularly underlined the peculiarities that suggest that the use of these models in disciplines characterised by the study of complex systems leads to a change of methodological paradigm in the cognitive approach to these systems. Now we will move on to a more concrete level of discussion, analysing the structure of the models devised for weather forecasting, and examining the impact of their application.

At the same time, however, within this very cogent and particular applicative framework, we will see that there emerge some problems (and relative attempts at a solution) that have a much more general significance and will enable us to discuss further changes of paradigm in the scientific approach to the study of complex systems.

7.1 The "Perception" of the Weather Forecasting Activity

In order to concisely evaluate how important the weather is in everyday life, it may be useful to notice how much this subject is discussed by ordinary people, that is by people who do not have any specific professional interest in weather data and forecasts. From this point of view — apart from clichés, such as the one that in the British Isles this is everybody's main topic of conversation — it is evident that in the countries of Northern Europe and the United States "meteorological culture" is much more developed than in the temperate-climate countries of the Mediterranean area. This is due to the fact that in the former countries the weather has a greater impact on everyday life, because there is a more frequent occurrence of intense cold, snow storms, frost, prolonged rain, and, in the United States, of violent phenomena such as hurricanes and tornadoes.

In the present-day society of information, a more quantitative index can be supplied by the presence of weather information on television. In all the northern countries this information is usually given quite often and in a very invasive way (rather like commercials). There even exists a channel, the Weather Channel, that is completely dedicated to world-wide weather information. On the contrary, in the countries of the Mediterranean area, including Italy, apart from occasions of particular cold or heat waves, in which the newscasts take care to underline the minimum or maximum temperatures in certain locations and to supply forecasts for the next few days, information about the weather is restricted to the regular forecast features.

Despite this limited presence of meteorologists on television, during the last few years, in the countries of the Mediterranean area, there has been an increase in the popularity of these features: their audience and viewing figures have risen almost to the levels of those of Saturday-night shows featuring popular showgirls. It is not clear why this has happened: some people believe that the cause is modern society's greater need to plan leisure time, others the fact that during the last few years the climate seems to have changed and there has been an increase in the frequency of extreme phenomena — such as cold or heat waves, violent storms and floods — from which people have to defend themselves. Other people believe that the meteorological features today are watched more because the weather forecasts have at last acquired a high degree of reliability.

On the one hand, the recognition of a high statistical reliability of the weather forecasts shows that there has been an increase in the credit of meteorology as a scientific discipline that has been making advances during the last few years[1]. On the other hand, this increase in credit is not always combined with an increase in meteorological culture. If one asks a person why the quality of the weather forecasts has improved, in most

---

[1] In this context it is significant to notice that up to a few years ago, for instance, in the makeup of Italian daily newspapers, the weather forecasts were always placed near the horoscopes, in an obnoxious divinatory couple, often combined also with the lottery draw news. Now at last some newspapers have made the decision of giving the weather column its own space, raising it to a rank of greater credibility.

cases one will obtain a reply like "Because now with satellites everything is easier...". Many people actually believe that the forecasts are produced by the satellites. The reader who has patiently reached this point in the book knows, on the contrary, that satellites are an instrument of observation, whose data can be used for a better definition of the state of the atmosphere system at a certain instant, but that they do not have any value for prediction purposes, at least on the time scales involved in the forecasts for the general public. It is clear that the misunderstanding in which laymen have fallen stems from what they see on television, where sequences of images from a satellite are often shown: they reproduce the movements of clouds during the last few hours over a certain territory. Though over a very short period (a few hours) these movements can be projected into the future, a forecast of the evolution of these clouds (formation of new ones or dissolution of old ones) and of their further movements requires knowledge of the evolutionary laws of the system and their interaction, therefore a model.

7.2 The Heart of a Meteorological Model: Primitive Equations and Their Numerical Solution

Here too, as in the previous chapters, we will refrain from adopting a historical perspective on the development of models in the past. We will only mention that the first theoretical forecast (which was worked out by discretising the meteorological variables at an initial instant and some basic equations, which were then solved "by hand" step by step) was published by Lewis F. Richardson in 1922: it was a 24-hour forecast, which turned out to be decidedly wrong. Almost 30 years elapsed before this early attempt was resumed with some hope of obtaining a more encouraging result, and with more reasonable calculation times, since in the meantime the American ENIAC, the first computer on which the calculations of a meteorological model were performed, had become available. Richardson's unsuccessful attempt allowed scientists to understand that it was not sufficient to know the basic equations that govern the physics of the atmosphere, because it was necessary to simplify them, in particular by eliminating the phenomena, such as the propagation of sound waves, that could impair the correctness of the forecast of the future behaviour of the atmospheric flow. So, for at least twenty years, very simplified models that used extremely approximate equations were worked out. Finally, starting from the nineteen-seventies, there appeared the first prototypes of the present-day models, the so-called "primitive-equation models". In this book we will describe only these models.

As we have already explained in the previous chapter, the central core of a simulation model consists of the equations that represent our theoretical knowledge about the system we are going to study, in some cases also in its temporal evolution. In order to describe highly interacting systems such as the atmosphere, where various causes concur in producing a single effect and there are feedbacks on the causes, it is necessary to couple several equations in a single system to be solved numerically. The variables that are contained in the equations, and whose future values are to be forecast, are physical quantities that can be measured in the real system. If, as in meteorological models, the equations are based on partial derivatives and include temporal evolution, the second element that is needed is the determination of the variables at the initial instant in the whole spatial domain considered by the model. Finally, the boundary conditions (e.g. the state of the ground in relation to time) and the external forcing factors (e.g. the solar radiation cycle) temporarily complete the ingredients of this recipe for a simulation. In actual fact, we will see later on that, in order to hope to achieve a realistic simulation of the physical evolution of the atmosphere, it is necessary to add at least another fundamental element.

A preliminary, essential remark must be made at once: standard meteorological models, at least the ones that aim at producing a correct medium-range forecast (up to 7-10 days), are characterised by a dynamical treatment only of the atmosphere subsystem within the broader Earth system discussed in Chapter 4. In practical terms, this means that the equations refer only to the dynamics of the atmosphere, and that the interaction with the other subsystems is expressed as an interaction with the external environment through the consideration of forcing factors and boundary conditions, even if they are evolving and partly depend on what takes place in the atmosphere system[2]. This is due to the fact that, as a rule, the evolutionary dynamics of the other subsystems that interact with the atmosphere produces changes in the latter that are slow in comparison with the time scales involved here. In a model whose purpose is to deal dynamically with the changes that each subsystem of the Earth system causes in the other subsystems, and with the resulting feedbacks, it is necessary to express the individual dynamics in terms of equations, and then to couple these equations in systems that can be solved numerically. We will briefly return to this topic in the next chapter, because in climatic models a more decidedly dynamical approach is required.

For the time being, it is interesting to concisely explain the basic equations (now usually called primitive equations) that are introduced into a meteorological model for a weather forecast. Adopting the principle of system decomposability, which leads to the determination of the individual equations, the complexity of the atmosphere system is usually reconstructed by means of six primitive equations: the first two, called "diagnostic equations", are laws like those we have called balance or coexistence laws in Chapter 4, and supply the relationship between different variables at the same instant; the other four equations, called "prognostic equations", depend on time and supply the evolution of the values of the involved variables. Let us briefly examine these equations.

- Equation of state of gases: supplies the link between pressure, density and temperature in an air mass at the same instant.
- Hydrostatic equation: this equation is obtained, by means of a scale analysis[3], from the equation of vertical motion, and supplies the approximate relationship between the density of the air and the change of pressure in relation to height. As we shall see, it is

---

[2] For instance, if in the world of the model there occurs a snowfall on a previously bare ground, from that moment on, and for a period that depends on the amount of snow that has fallen and on the temperature of the air, the boundary condition that determines the state of the ground is changed, so as to allow the model to correctly simulate the decrease in absorption of solar radiation on that part of the territory.

[3] The scale analysis method makes it possible to determine the relative importance (in terms of orders of magnitude) of the individual terms of which the equation is formed. In actual fact, as a result of a scale analysis, some of these terms are disregarded and the situation is simplified.

considered valid for models where the spatial resolution is not very high.

- Continuity equation: it ensures the conservation of mass. Once a volume that delimits a portion of air has been determined, for instance in the case of the hypothesis of the perfect non-compressibility of the fluid, this equation ensures that if a certain amount of air gets into the volume under consideration, the same amount gets out.
- Navier-Stokes equation: this is the equation of motion in fluid dynamics for the horizontal components of wind: from the vertical part of the complete three-dimensional equation, after a scale analysis, the hydrostatic equation is obtained.
- First law of thermodynamics: it is an evolutionary equation for temperature that allows for the thermodynamic processes in the atmosphere, such as the adiabatic warming or cooling of the air due to vertical movements, the release of latent heat in changes of state, etc.
- Continuity equation for the water (liquid, solid or vapour) contained in the atmosphere: it is an evolutionary equation that obviously allows for all the processes connected to the changes of state, i.e. evaporation, condensation, fusion, solidification, sublimation and rime formation.

As we have already mentioned, the computer-aided solution of equations such as those we have just described involves numerical techniques of step-by-step solution of the equations. If there existed analytical techniques for arriving at general solutions or even only for solving significant specific cases, it might be possible, perhaps, to use the computer with a method more similar to that of a classical mathematician: nowadays symbolic-calculus software packages that solve certain equations analytically are available. Unfortunately the mathematical knowledge about the properties of these equations is still limited, and the dream of an analytical solution is still far off: to this day there does not exist a general theorem of existence and uniqueness for the solutions of the Navier-Stokes equation.

In a situation like this, the successful strategy seems to be the one that had been outlined by Richardson as early as 1922, that is the replacement, in the equations, of the derivatives with the finite differences between points at a certain spatial distance (in the case of derivatives with respect to space) or at a certain time lapse (in the case of derivatives with respect to time). The concept of the derivative of a certain quantity is a sort of generalisation precisely of the concept of the increment or decrement of that quantity in a time or space unit, because it determines the rate of difference of the quantity for infinitely small spatial distances or time lapses. Thus, by replacing the derivatives with finite differences, we achieve a "discretisation" of the equations under consideration. This means, in particular, that these values will have to be calculated on a finite number of points in space and on a sequence of discrete temporal steps. Thus, in the world of the model, the space continuum and the time continuum that characterise the description of spatial and temporal evolution in physics are replaced by a three-dimensional spatial "lattice" and a sequence of discrete temporal steps.

So, in order to concretely bring about this "discretisation" of the equations, it is necessary to define the so-called "grid", i.e. the spatial lattice on whose points the values of the variables present in the equation are to be calculated in the immediate future (at the next discrete time step). Proceeding step by step, it is then possible to obtain the values of these variables on all the points of the grid for the forecasting period to be considered[4].

Obviously there exist several types of grids: a natural grid, for instance, is one that is defined by the points of intersection between meridians and parallels on the Earth's surface; though appearing to be completely natural, this grid has the negative characteristic that its points are very close to each other towards the poles, further apart at medium latitudes and even more sparse near the equator. The grids that are chosen usually have a more uniform, equidistant distribution of points.

---

[4] For the sake of completeness, and only for the interested reader, it is worthwhile to mention the fact that this scheme based on finite differences is not the only one that allows a discretisation of the equations: there exist also some methods, called "spectral methods", in which the spatial changes in the variables are expressed as components in the Fourier space or in spherical harmonics. The results of the two methods are quite comparable.

Moreover, the reader should notice that the grids are not two-dimensional, but three-dimensional, since they extend upwards in the atmosphere, with several vertical levels and a series of concentric two-dimensional grids around the Earth.

Likewise there exist several techniques for the numerical solution of the equations on the grid that has been chosen. These techniques tend to cut down to a minimum the approximations inherent in the discretisation, and to elude certain problems stemming from the numerical solution of the equations. Nowadays there exists an entire branch of mathematics, numerical analysis, that studies these problems. Here, obviously, we cannot dwell on this topic[5].

Returning to our grid, we must remark, that, as a rule, the thicker it is, the closer the finite differences come to the values of the derivatives of the space-time continuum. The problem that arises at once, however, is that of the calculation times: if a three-dimensional grid exceeds a certain degree of thickness, the number of calculations to be carried out by the model increases excessively. On the other hand, our most pressing need when using a meteorological model is to obtain results in a time that is practical for their use: obviously if a model takes two days to produce a forecast for the day after, the result, even if it is correct, becomes utterly useless. Therefore the lower limit for the distance between the meshes of the grid is shifted, basically following the performance of the new computers, which are increasingly swift. Nowadays, for a global model (whose domain of numerical integration is the whole Earth), we stop at a typical distance between two points of the grid (the so-called "grid spacing") that is about 50 Km horizontally, while vertically there are about 60 levels in the first 60 Km of the atmosphere (but these levels are not equidistant, because they are closer to each other in the lowest layers, less and less close as the altitude increases, and very distant from each

---

[5] There are no texts in which the techniques of numerical integration, and the dynamic and physical structure of meteorological models, are explained in a thorough but accessible way. For the reader who is interested just the same, we recommend Krishnamurti and Bounoua (1996).

other beyond the tropopause)[6]. Because of this discretisation, each point of the grid essentially identifies a cell centred on that point, for instance a parallelepiped whose horizontal sides are 50 Km and whose vertical sides are a few hundreds of metres longer.

Since the physical quantities of the future are calculated only at the intersections of the lattice, the grid spacing essentially determines the resolution of the model: the variations that take place in any variable between two adjacent points of the grid cannot be represented by the model if their values exceed the range of the values present on those two points. It is as if we tried to reconstruct a sinusoid $y = \sin x$ with a 2π period by means of a number of discrete points whose distance between each other exceeds π on axis x: in this case it is not possible to correctly represent the shape of the sinusoid. In other words, the discrete numerical solution of a model necessarily "averages" the real physical behaviour of the system under examination.

The present limitations in the spatial resolution of a global model are partly overcome if we focus on a smaller portion of territory: there exist some models, called "limited-area models", in which the numerical solution of the equations is limited to a less extensive area, for instance the Euro-Atlantic region of the northern hemisphere. In these cases, it is possible to obtain a thicker grid while preserving the same order of magnitude in the number of grid points. Obviously these models need the conditions at the boundaries of their area of interest, and these conditions can be supplied only by the evolution forecast through a global model: for this reason, limited-area models cannot be regarded as self-sufficient.

In these models it is possible to achieve a horizontal spatial resolution even only of a few Km. Beneath 10 Km, particularly in cases where the ground (which is the lower boundary of the atmosphere) is characterised by complex orographical features, the validity of the hydrostatic approximation disappears, so it is necessary to use a Navier-Stokes

---

[6] The reader is reminded that the tropopause is the area where the temperature, which decreases with height in the areas below, starts increasing with height. The tropopause is situated at an altitude that depends on the latitude, season and meteorological configuration: at medium latitudes it is, on the average, between 11 and 14 Km above sea level.

equation in the model also for the vertical component. These models are called “non-hydrostatic models”, and will not be examined in this book.

7.3 Physical Parameterisations

The previous remarks led us to understand that the discrete numerical solution of the equations of a model based on a grid and characterised by a certain grid spacing necessarily results in our obtaining an averaged description of the real evolution of the atmosphere system. This is a serious problem when there is the need to draw a distinction between the weather in a certain point and that in another one separated from it by a horizontal distance smaller than the grid spacing; this difficulty, however, is obviously present in any model that simulates reality in a discretised manner on a computer. The real physical problem that must be taken into account is a result, instead, of the fact that in meteorology there exist phenomena on very small scales — smaller than the grid spacing — that heavily influence the value of the quantities present in the equations solved by the model on the points of the grid. Basically these phenomena are not “seen” by the system of discretised equations, so, if everything were limited to a solution of this system on the given grid, the description of the physical atmosphere system would turn out to be inadequate, and the validity of the forecast would be greatly impaired.

In order to allow the reader to understand that these small-scale phenomena do not pertain to marginal aspects of the weather as it is usually perceived, it is sufficient to point out that one of them is the presence of thunderstorm clouds: the diameter of a thunderstorm cloud leading to a typical summer thunder shower is always less than 50 Km, and often even less than the 10 Km that have been taken as the lower limit for the grid spacing of a non-hydrostatic model. A model that does not “see” these clouds obviously would result in a completely wrong forecast of the amount of precipitation: it is known that the most intense precipitation comes precisely from convective phenomena.

But this is not the whole story! There is a further complication: the presence of thunderstorm clouds, which are very tall, obviously also perturbs the value of other variables “handled” by the model with its dynamic equations. For instance, the quantity of water (in its three states of aggregation) present in the cells that contain the thunderstorm cloud and are identified by the neighbourhood of the relative grid points in the vertical direction is considerably altered on the entire column. Moreover, the strong upward currents present within a cloud lead to a vertical conveyance of matter and humidity and to a vertical redistribution of the temperature. More specifically, there appears a vertical thermal profile typical of an air mass that rises adiabatically in the atmosphere, with phase transitions within it from water vapour to liquid water and ice: the latent heat that is released in these transitions causes the temperature within the cloud to be usually higher than that of the surrounding air. Finally, the presence of a thunderstorm cloud results in a decided change in the balance between the incoming radiation and the outgoing one in the cells under consideration: sunlight is intercepted, and so is the long-wave radiation coming out of the ground (this too affects the temperature values).

The example mentioned here of the possible presence of thunderstorm clouds highlights the existence of an intense evolutionary phenomenon on a scale that is not solved by the equations of the model with the selected grid spacing. We have already explained that this deficiency is serious from the viewpoint of forecasting, both because of its immediate consequences (unforeseen strong precipitation) and because of the perturbations in the other variables that the model causes to evolve independently on the grid through the solution of the equations.

We will not dwell on other small-scale phenomena that have a similar effect (though perhaps in a less quantitatively evident way, as in the case of the mechanical turbulence in the low layers); it is clear, in any case, that there is the need to express the action due to these small-scale processes in relation to the quantities on the scale of the grid that are dynamically treated by the model.

The strategy adopted for this purpose is that of creating routines, program modules called "physical parameterisation modules", that describe the evolution of these small-scale phenomena. Usually we have separate but interacting modules for convection, radiation, turbulence, etc.; they are used, every now and then, to "adjust" the values of the variables, which, in the meantime, evolve separately through the solution of the equations. The use of this interacting set of dynamical equations and physical parameterisations is the only method with which we can hope to achieve a correct forecast of the evolution of the characteristics of the atmosphere system.

7.4 Determination of Initial State and Analysis Procedure

By considering the primitive equations and the physical parameterisation schemes, both solved in a discrete manner on a three-dimensional grid, and with a discrete time step, we have introduced our theoretical knowledge of the atmosphere system into a model that simulates it. We have already pointed out, moreover, that external forcing factors and evolutionary boundary conditions make it possible to describe the relationships between this system and the other systems that define its external environment (oceans, lithosphere, biosphere, etc.). After the theoretical framework has been thus defined, the model can be "linked" to the real system by means of the variables present in the equations, which have a one-to-one correspondence with certain physical quantities in the atmosphere.

At this point, if we wish to achieve a simulation of the behaviour of the atmosphere in a situation that is actually present in nature, all we need is the determination of the initial state of the atmosphere through the measurement of the physical quantities introduced in the model as variables. Obviously, it will be necessary to define the value of these variables at the initial instant in all the points of the grid that characterises the model, so as to supply a complete initial condition for the numerical solution of the theoretical scheme. If we consider that the meteorological measurements performed on the planet are far from being arranged on a three-dimensional grid[7], we can understand that the problem of supplying the model with a complete initial condition is far from easy to solve. In actual fact, any reputable meteorological office employs a considerable part of its researchers in the optimisation of the

---

[7] The reader will recall, for instance, Plate 1, discussed in Chapter 2, which showed the highly unhomogeneous horizontal distribution of conventional surface-based observations and the shortage of observations on the oceans and African continent.

methods (and operating process) for obtaining an estimate of the initial state of the atmosphere in the form of initial data on the grid of the model.

The first solution that enters our mind when we are looking for the value of a variable in a point where its measurement is not available is to obtain this value by means of the interpolation of the known surrounding values. Actually, given a spatial distribution of values of a certain quantity (acquired, for instance, by means of measurements), there now exist some software packages that can interpolate the value of this quantity onto a point where it is unknown, provided the number of available measurements is sufficient to ensure the reliability of this reconstruction. Can we apply this method to the determination of the values of all the variables of the model on all the points of the grid at the initial instant of the simulation? Certainly! But if we did it, we would immediately realise that the results are quite meagre from the physical point of view.

The key for understanding the reason for this lies precisely in the physical nature of our data, which are representative of a system characterised by a certain balance or coexistence between several quantities at the same instant and by a certain consistency between spatially close points. In fact, though the data that come from a meteorological office with gross coding or transmission errors can be corrected by a quality check that might be introduced also in a simple interpolation software package, discovering systematic errors and more subtle ones (instrumental or due to other causes), or obtaining a physically consistent interpolation requires a different treatment of the data — a treatment that allows for the physical relationships that interlink different quantities.

Moreover, we must consider that the amount of available data is often not sufficient for a satisfactory interpolation: while for conventional observations we have just mentioned the presence of gaps in the observational network, for satellite observations we have stressed (in Chapter 2) the shortage of points in the vertical direction in the SATOB messages and a similar problem also in the SATEM ones, because of the highly averaged measurements that are characteristic of the vertical soundings performed by the TOVS. Furthermore, the data are actually completed by those coming from the polar satellites, but the latter are usually not synchronised with the conventional data or with those coming from geostationary satellites. In order to be able to somehow compare these data with those of the synoptic hours, we will inevitably need a temporal "consistency", i.e. evolution equations.

The situation we have thus described leads us, therefore, to drop the idea of a simple mathematical interpolation of the data, quantity by quantity, and to prefer a more dynamical way of dealing with this problem. An approach that might seem natural is that of correcting the results of the interpolation by using the physical constraints we know. Let us now briefly discuss what might be a typical strategy for tackling the problem of determining the values of the variables at the initial instant on all the points of the grid.

On the whole, if we consider the importance of the physical constraints in the reconstruction of the discrete spatial distribution on the points of the grid, we are led to completely overturn the apparently natural approach to the problem.

In actual fact, instead of starting from a mathematical interpolation and correcting its results by applying known physical laws, we prefer to start from the values of the variables, fixed by means of the forecast results coming from the previous run of the model (6 or 12 hours before). This way, we obtain a first guess for the fields[8] of all the variables at the initial instant of the new run of the model. Only at this point the observations come into play, obviously in the role of correctors for the fields that have been determined with the first guess.

This rather original approach makes it possible to allow fully for the physical laws of the atmosphere system, both the balance ones and the evolutionary ones, which are all included in the model. Moreover it makes it possible to obtain more reliable initial conditions on the areas where there are few observations and where their simple mathematical interpolation would give rise to serious problems. Finally, since the

---

[8] Without referring specifically to the strict mathematical definition, within our context of discrete model making, a field is a set of values of variables on all the points of the grid: it is a scalar field if the variable under consideration is a quantity that is defined by a single numerical value (a scalar), or a vectorial field if the field variable is defined by a vector.

obtainment of the first guess is based on the use of an evolutionary model, which theoretically can supply the fields of the variables at each temporal step, it is possible also to retrieve the correction contribution relative to the observational data that have not been read precisely at instant $t_0 = 0$ of the new run of the model, but have been included, for instance, in the interval $t_0 - 3$ hours $< t < t_0 + 3$ hours.

The procedure we have thus briefly defined is called "analysis". More specifically, a very recent achievement of research in this area is the possibility of including in the analysis observational data coming from measurements that are not synchronous with the initial instant. In this case, rather sophisticated mathematical techniques are used[9], and the analysis is called "four-dimensional", because, besides the three spatial dimensions, the temporal one is also included, in order to account for the inflow of data relative to instants different from t_0.

7.5 The Products of a Meteorological Model

Now that we have discussed all the ingredients of the "recipe" for a weather forecast simulation, we will proceed to the executive stage of this recipe. The analysis procedure supplies the initial data to the model, which starts the simulated evolution with time. After the very first period (a few hours of simulated time), during which, despite the physical approach of the analysis, there may still be the influence of a few settling-down problems in the balancing between the fields of the various quantities[10], we will begin to obtain the first reliable, consistent results. The values of the variables provided at all the points of the grid are stored in suitable memory areas of the computer in which the model is "running", at fixed intervals and throughout the simulated-time predefined for the run of the model (usually 10 days for a global model and 3 days for a limited-area model).

---

[9] For the reader who is interested, we can summarise by stating that it is essentially a matter of solving minimisation problems by means of variational calculus and of the construction of an adjoint representation of the linearised model.

[10] This problem is called "spin-up effect".

The files containing the values of all these variables on all the vertical levels of the model remain usually in the archive of the data processing centre of the meteorological service or institution that develops and operates the model, and are available, if necessary, for subsequent studies. What reaches the users is something different. The technical users (i.e. the people who need these data in order to process them further, both for scientific and for purely applicative purposes) receive digital, coded data[11] on certain vertical levels, for the required areas and forecast times (usually every 6 hours of simulated time). The other users normally receive graphic elaborations that supply a selection of the same pieces of information, but in a form that can be directly understood by anybody, because they are usually displayed in weather maps that can be interpreted more or less readily.

Plate 8 shows an example of these weather maps, also called "prognostic charts", drawn from the global-circulation model of the European Centre for Medium-range Weather Forecasts (ECMWF). This figure represents the forecasts for 4 February 2003 at 12 GMT, obtained by the operational run of the model after 36 hours of simulated time from the date of the initial analysis. This forecast is illustrated here by displaying the situation predicted for that date and time, by means of an upper-level chart (a), a surface chart (b), a cloudiness chart (c) and a precipitation chart (d). We will now concisely analyse some traits of these products, which obviously result from graphic-display elaborations of the data provided by the model and distributed in a discrete manner on the grid.

The first chart, shown in Plate 8a, describes the situation forecast at an altitude approximately at the middle of the troposphere. In actual fact, the processing of upper-level charts is influenced by an inheritance due to the fact that these charts were originally produced for flight aiding purposes. Since the altimeters of aircrafts are basically barometers, i.e. instruments that measure the atmospheric pressure, stabilising the height of an aircraft during the cruising stage means flying the aircraft at a fixed pressure and not at a certain altitude above sea level. So in upper-level charts, instead of placing ourselves at a certain altitude and ascertaining

---

[11] In order to minimise file sizes and transmission times.

what pressure there is as we move horizontally, we place ourselves at a certain pressure and display the altitude at which this pressure is present over the area under consideration[12]. So each blue line in this figure represents a line that joins the points where the pressure of 500 hPa is at the same altitude (represented by a number that indicates tens of metres of height). Likewise, the red broken lines identify the points that are characterised by the same temperature at the pressure of 500 hPa. A vast altitude minimum is visible: it is forecast to be centred over Denmark.

The next chart, shown in Plate 8b, represents the atmospheric pressure read at sea level (by means of blue lines that connect the equal pressure points, called isobars). The forecast field of the temperature at 850 hPa (about 1,500 metres) is also displayed by means of red broken lines, called isotherms, that connect points of equal temperature. There is a forecast area of low pressure that involves the whole of central Europe and extends towards Italy, where the pressure is particularly low in the central and northern parts of the country.

These first two charts are a graphic representation of the fields forecast for basic variables that are supplied to the model as initial conditions, then treated dynamically by the model. The actual use of these charts requires a certain forecasting experience. There are, moreover, charts that refer to variables that have been reconstructed or somehow drawn from the basic ones. For instance, from the temporal trend of the liquid water or ice content it is possible to infer the presence or absence of clouds on the territory (if necessary, also determining their altitude): see the example in Plate 8c, which zooms onto Italy and the neighbouring countries, and represents the cloudiness forecast in terms of low cloud cover[13]. Moreover, the amount of liquid or frozen water that falls to the ground from clouds supplies a precipitation forecast: see

---

[12] This is facilitated by the fact that in meteorological models the vertical coordinate is often expressed in terms of pressure, because this results in a simplification of the form of the equations.

[13] In Plate 8c the forecast low-cloud cover is indicated in the oktas scale, at whose ends 0 oktas means clear sky and 8 oktas means completely overcast sky. Low clouds are characterised by the fact that their base is at an altitude of less than 2 Km and are very significant from a meteorological point of view: for instance, they often lead to precipitation.

Plate 8d. Contrary to the first three charts, which represent "snapshots" of the fields expected at the time of the forecast, these charts represent the distribution of the precipitation expected to fall to the ground in the 12 hours of simulated time that precede the time limit of 12 GMT of 4 February 2003. Obviously these last two charts are much more self-explanatory than the previous ones, and can therefore be utilised more directly (though, unlike the layman, a careful forecaster who is acquainted with the characteristics of the model also knows how to correct possible defects in the forecast of these quantities).

7.6 The Emergence of Deterministic Chaos and Ensemble Integrations

We have just shown an example of the graphic display of the fields provided by a meteorological model. Obviously a model is regarded as valid if it correctly forecasts what will eventually take place in the reality of the atmosphere; this validation can be carried out only by means of a comparison with *a posteriori* analyses of data obtained through the observation of the real system. In practice, checking the performance of a model is an important aspect of the activity of a meteorological office, because it enables the researchers to understand the effects of possible changes in the model, in terms of the accuracy of the simulation and therefore of the validity of the forecast. This validation takes place both as a daily operation carried out on the new forecasts that are produced and in an analysis of the behaviour of the model in case studies characterised by particularly significant meteorological situations.

As we already stated in the previous chapter, when we leave the controlled conditions of the laboratory and attempt to reconstruct the complexity of reality, we are inevitably compelled to choose the theoretical elements we regard as fundamental and to leave out the ones we regard as secondary and less important for the dynamics of the system under consideration. Once we have worked out a simplified model, obviously it can be improved by the addition of further theoretical elements, so we may be led to believe that, through successive improvements, it is possible to eventually achieve a reproduction of all the details of the behaviour of the real system. Obviously a project that is so ambitious can be based only on confidence in the validity of the theoretical scheme, which in our case is formed of diagnostic and prognostic differential equations and parameterisation schemes and on the possibility of rendering our theoretical scheme complete and univocal.

We may be doubtful about the possibility of achieving the univocality of the scheme, because in our case it would be a matter of having to determine in a univocal manner (obviously on the basis of theoretical considerations) the value of the parameters that currently determine the coupling of the various parameterisation schemes in the models and therefore the intensity of the feedbacks. At present these parameters are fixed by means of an activity that might be called artisanal, balancing them, i.e. mutually "fine tuning" them, as we do when we fine tune a television set in order to centre the frequency of a channel and obtain optimum vision. In any case, at least for the time being, we will pretend that this problem does not exist, and assume that it is possible to work out a complete and univocal model, a "perfect" model that can accurately reproduce the future behaviour of the atmosphere system.

If we could have a model whose theoretical scheme perfectly reproduces the physics of the atmosphere system, the only element of uncertainty in the future forecasts would stem from approximations and errors we might make in the determination of the initial state. An error here would obviously propagate, step by step, during the discrete numerical integration. It remains to be explained how this takes place.

Present-day models, like the earlier ones, are characterised by forecasting errors in the various estimated quantities, and these errors are usually the greater the longer the simulation time. All this is rather reasonable, because, since the model is not perfect, it leads to inaccuracies in the forecast, which, step by step in the discrete numerical integration, will propagate and get amplified in the course of time. These errors cannot be distinguished *a priori* from the ones that stem from inaccuracies in the determination of the initial conditions. In this framework, however, it is interesting to point out that the amplification of the errors is not linear (i.e. practically constant), but is affected by sudden upsurges that, after a certain time lapse, cause the behaviour forecast by the model to be completely different from the real behaviour of the atmosphere system. On what does this fact depend?

In order to study this simulated behaviour, and to hope to be able to identify the elements that are important for understanding it, we must separate the contributions due to the model from those due to the initial situation. More specifically, if we temporarily assume that our model perfectly simulates the behaviour of the atmosphere, a way of imitating the propagation of errors due to an incorrect estimate of the initial situation is that of starting two runs of the model with two slightly different initial conditions: one of the runs simulates the behaviour of the atmosphere on the basis of a correct estimate of initial conditions, the other does the same thing on the basis of a somehow incorrect estimate.

What this simulation experiment reveals is that the error thus introduced in the initial condition is amplified in a manner that is not linear or slow and gradual, but "exponential", with a sudden upsurge after a certain time lapse, exactly as in the comparison between our working simulations and the behaviour of the real atmosphere. On the one hand this observation may suggest that our way of representing the evolution of the atmosphere is basically correct, but on the other hand it shows that our representation of the atmosphere system by means of a system of non-linear differential equations and physical parameterisations is extremely sensitive to initial conditions. In these theoretical schemes, if we start from two initial conditions that are even highly similar, after a certain simulated time lapse the solutions will diverge: this is a phenomenon called "deterministic chaos"[14].

Since, because of the discrete, unhomogeneous and partly discontinuous nature of atmosphere monitoring, the initial condition for the runs of the model is always identified (through the analysis process) in an approximate way, we can endeavour to minimise this error but cannot ever eliminate it. So even if we had a perfect model, the approximate estimate of the initial state would lead, after a certain time lapse, to the production of forecasts that considerably differ from the

---

[14] Though the emergence of deterministic chaos is typical, as a rule, of systems of non-linear equations, it was discovered precisely in the meteorological field, by Edward N. Lorenz. He reported his findings in a famous article of 1963 and more recently produced an interesting account in Lorenz (1994).

behaviour of the real system. This limit, which obviously is intrinsic to the theoretical scheme that has been adopted, leads to the recognition of a maximum theoretical predictability period beyond which it is not possible to produce forecasts with this type of model: the period is approximately 10 to 15 days, though in actual fact the model starts performing rather badly even before.

Once this limitation, which is completely inherent in our theoretical scheme, has been recognised, the next goal is to find the way to reduce its damage to the forecasting activity. This will be the topic of the next pages. First, however, it is worthwhile to dwell somewhat on the absolutely fundamental significance of this discovery.

Classical physics is entirely based on a deterministic description of the evolution of dynamical systems, inasmuch as its time-dependent equations can be solved in the future and can yield a univocal solution if the initial conditions from which they start are univocal. Some philosophers, whose most paradigmatic representative was Laplace, adopted the philosophical stance of defending the most absolute determinism, meaning by this that "if an intellect, at a certain instant, knew all the forces that animate nature and the mutual positions of the beings that compose it, and if it were so vast as to submit these data to an analysis, it would condense into a single formula the movements of the largest bodies of the Universe and that of the lightest atoms: nothing would be uncertain for it, and both the future and the past would be present before its eyes"[15]. Laplace clearly went beyond determinism when he asserted that all natural phenomena could be reduced to the (deterministic) laws of mechanics.

In a stance like this one, the concept of determinism leads directly to that of the univocal prediction of the behaviour of a system. Apart from previous evidences that had already undermined this concept — such as the discovery, made by Poincaré, that the three-body problem cannot be solved exactly, and the discovery of quantum phenomena — nowadays, with the emergence of chaos in systems of non-linear deterministic equations of classical physics, it is being shown that in systems that are sensitive to initial conditions a deterministic approach cannot ensure a

---

[15] Laplace (1820).

reliable evolutionary uniqueness in the future. So will we have to adopt a different forecasting strategy? After all, Laplace himself, being aware of the finiteness of human knowledge and of the difficulty of completing his own deterministic-mechanistic programme, was the creator of the probability theory. Let us now endeavour to picture our situation with the help of classical methods.

If we consider a system of interacting particles (without any internal structure), the state of this system at a certain instant is defined when we know the position and velocity of all its particles. If, in a very simplified situation, our system were formed of a single particle constrained to move in a definite direction, we might display its state at a certain instant as a point in a Cartesian coordinate system with axes x and v_x. Though obviously for a particle that moves on 2 or 3 spatial dimensions (and even more for a system with several particles) it is not possible to graphically display the situation of its state, this conceptual representation remains valid: the state of the system at a certain instant is represented by a point in a multidimensional space (called "phase space"). Though now there is an uncertainty about the determination of the state (e.g. at initial instant t_0), we can represent the situation by considering not a point but a hypervolume, whose hypersurface delimits the zone where there is the initial state we have not been able to determine univocally. In the case of the single particle constrained to move along axis x, this new situation can be easily displayed as shown in Figure 7.

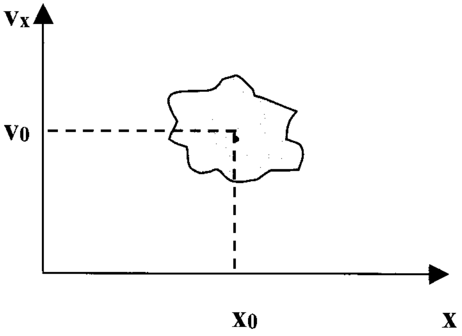


Figure 7. Display of uncertainty about the initial state of a particle constrained to move along axis x.

Returning now to the world of our meteorological model, here its state at a certain instant is defined if the values of all the variables at that instant on all the points of the grid are defined. The total number of these values is very great: for a global model it is somewhere around several tens of millions of data. In perfect analogy with what we have just discussed about particle systems, we can define a "state space" in which each axis is relative to a single variable on a single point of the grid. This way, the initial state, i.e. the initial condition determined by the analysis, turns out to be a point in this multidimensional space. Likewise, the uncertainty in the determination of this initial condition can be represented as a hypervolume, more or less extended around this point.

In this perspective, the meteorological model, by predicting the value of all the variables on all the points of the grid through the numerical integration of the equations and parameterisation schemes, does nothing other than determine, in successive time steps, the evolution of this point in the state space. But obviously the analysis-based initial state is approximate: it does not represent the real initial state of the atmosphere exactly. This, however, is likely to remain within the hypervolume that represents the uncertainty region in the determination of the initial state of the model.

At this point, obviously, we have no means to univocally determine the initial state within the uncertainty region and thus to supply the model with an initial condition that is quite similar to the real one: its best estimate for us is the one given by the results of the analysis. However, the evidence that two initial states that are very similar may lead to very different developments after a certain time lapse seems to suggest that we should examine the evolution of the entire volume that determines the uncertainty. This is done in systems of more limited dimensions, where a function that measures the probability of finding the state in the various points within the volume is also associated to this volume[16]. In certain systems described by non-linear equations we find that, beyond certain temporal thresholds, this volume becomes stretched out and greatly distorted, causing the individual points within it to become even very distant from each other. This means that initial states

---

[16] This function is called "probability density function".

that are similar to each other evolve into final states that are quite different.

It is obvious that the purpose of the study of the temporal evolution of this hypervolume is not to determine the initial state that results in a correct evolution (this would be possible only *a posteriori*), but to identify the moment in which this volume starts becoming so distorted as to lead to very different developments of the various states included in it. After this moment, because of this uncertainty about the initial condition, the deterministic forecast (which starts from a single point) might no longer be reliable.

In the world of a meteorological model it is not possible to determine a probability density function and study its evolution in the course of time. Nevertheless, the researchers of various meteorological offices posed the problem of studying the reliability of deterministic forecasts and identifying the time limits after which the evolution of various initial states begin to diverge. The idea underlying these studies is that of perturbing the results of the initial analysis, to the point of determining a certain number of initial states (approximately 50) that are representative of a volume of uncertainty; then these initial states are caused to evolve by means of the model, with a number of runs equal to that of the initial states under consideration. These operations are called "ensemble integrations"[17].

If now we examine the forecast results for the variables on the points of the grid of a certain zone, we can find out when the results of the model are no longer reliable for that zone. We will discover that the reliability interval depends on the situation of the weather: there are some situations in which the developments forecast by the runs of the models based on the various initial conditions remain close to each other for a long time, and others in which they begin to diverge after a few days. Figure 8 (from Pasini and Pelino (2000)) shows the results of a series of ensemble integrations for the forecast of the wind speed at surface level on the Italian weather station of Brindisi. The reader will

---

[17] Because of computer-time problems, all these additional runs are usually carried out on grids whose resolution is lower than that of the operational model. Here we obviously cannot dwell on these methods. The reader is referred to a review article on this subject (Buizza, 2001).

notice that the trajectories (therefore the wind speed forecasts) remain rather close to each other during the first three days of simulated time, but diverge broadly after 72 hours.

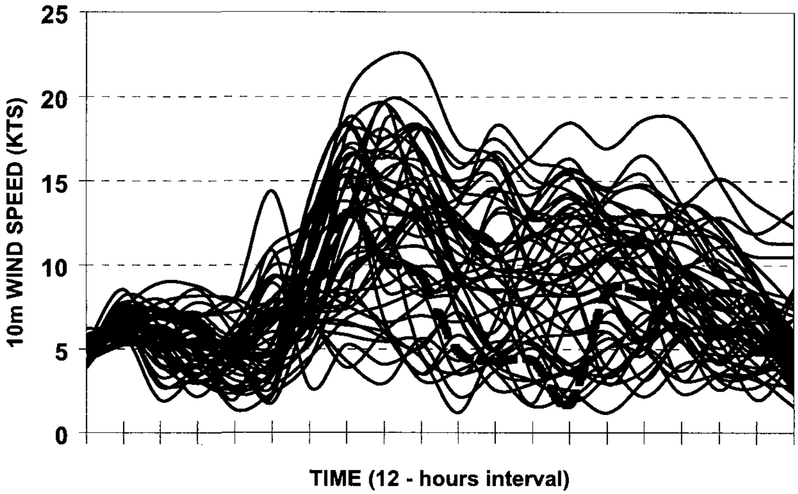


Figure 8. Results of ensemble integrations for a surface-level wind forecast at Brindisi, Italy, obtained on the basis of perturbations of the analysis of 13 June 1998 at 12 GMT. The broken line represents the forecast of the operational model (figure reprinted with permission from Elsevier).

We will overlook the complex statistical processing carried out on the data yielded by ensemble integrations, and mention only the fact that the most immediate possible use of these data is to attribute to the forecast of the deterministic operational model (the one that starts from the non-perturbed analysis data) an index of forecasting reliability that depends on the time limit of the forecast[18]. In situations of high predictability,

---

[18] Further theoretical aspects of the evolution of the state of a complex system in the state space will be discussed in the next chapter.

such as those that evolve very little (due, for instance, to "blocking situations", when a high-pressure condition at all levels persists for many days over a vast area), this index turns out to be high, up to a distant forecasting time limit, so the forecasters may commit themselves up to that date. In opposite situations, where reliability decreases quite soon, the forecasters will have to limit themselves to a forecast of a few days, because the reliability attributed to the operational model decreases quickly after that time limit.

7.7 A Few Conceptual Remarks

In this chapter we analysed the structure and performance of meteorological models, highlighting a very concrete applicative activity, and also some aspects that are important for the development of scientific thought, such as those relative to the emergence of deterministic chaos and the use of ensemble integrations. Now it will be expedient to dwell a little on the peculiarities that these models show in the sphere of the simulation paradigm discussed in the previous chapter: these peculiarities reveal some conceptual difficulties of modern scientific practice in the study of complex systems.

To begin with, weather forecasting models closely follow the structure of a typical simulation model, since they are formed of a core of differential equations based on partial derivatives that represent our theoretical knowledge of the basic processes and phenomena of the atmosphere. These equations have been essentially corroborated by fluid dynamics and thermodynamics researches with laboratory experiments, so the purpose of the meteorological model is not to verify their validity. The fact that we start from individual equations, combine them in a system and use them together with parameterisation schemes means that we rely on the possibility of reconstructing the complexity of a system such as the atmosphere through the recomposition of some basic elements and interactions, in the form in which they have been understood theoretically after having been studied separately in the laboratory.

In this recomposition activity, though the primitive equations that are considered are practically the same for all models, the parameterisation schemes are different and undergo a constant development, whose main purpose is to improve the description of the processes that take place at a scale smaller than the grid spacing, including previously overlooked details and phenomena. In this regard we can assert, from a historical point of view, that the path towards the correct simulation of the future behaviour of the atmosphere has been passing through the examination of constantly new theoretical elements, which had previously been regarded as secondary and were subsequently added to the scheme of the models.

Though here we will not deal with the history of meteorological modelling, it is worthwhile to point out that the first models based on primitive equations were dry adiabatic ones, i.e. they regarded the atmosphere as a system where there is no heat exchange with the external environment, and air as a compound devoid of humidity or water in its three states of aggregation. Afterwards, the continuity equation for water was added: it also supplied the precipitation for non-convective phenomena. Then a parameterisation scheme was added for convection, but it only led to an adjustment of the vertical thermal profile of the involved cells and to the production of convective precipitation (in the model, all the condensed water fell to the ground and there was no cloud formation). Later on, convection schemes with cloud formation were introduced, and it became possible to consider parameterisation routines also for the radiative exchanges between the Earth and the free atmosphere, and for the influence of the ground[19]. We may therefore assert that, in meteorology, the modellers' ability to simulate the behaviour of the atmosphere has really improved, owing to a recomposition activity that led them to consider an increasing number of new elements of theoretical description, with particular reference to parameterisations and the relative feedback cycles acting on the variables of the system.

---

[19] Several of these transitions led to a very considerable improvement in the forecasts. A presentation of some of these different theoretical schemes will be found, for instance, in Krishnamurti and Bounoua (1996).

The models thus obtained are now able to forecast the future characteristics of the atmosphere, at least within the predictability period that can be estimated on the basis of the courses of the ensemble integrations. Above all, they correctly reconstruct the development of medium- and large-scale systems, such as the low and high atmospheric pressure configurations and the fronts with the cloudiness associated to them. The limits of these models consist in the fact that forecasts on very limited areas (which are sometimes also characterised by complex orographical features) cannot be solved correctly because of the finite grid spacing. So sometimes in the world of a model it is not possible to tell the windward side of a small range of mountains from the leeward one, where the weather may be quite different.

The recognition that the macroscopic dynamics is captured well by the meteorological models supports the idea that the complex non-linear mixture of theoretical elements introduced in the models is sufficient to account for the complex emergence of large-scale phenomena in the atmosphere. If we consider, for instance, the fact that the equations of the model are applied to air cells whose horizontal sides are about 50 Km long, this is not surprising, because here we are not describing a macroscopic system in terms of elements belonging to a different (smaller) scale: in the latter case the model would be a reductionistic one, but this is not the situation of the meteorological models. Here the macroscopic concept of air mass, discussed in Chapter 4, can safely be applied to simulated evolution.

From this point of view, the success of these models (which are basically the only instrument we have for reliably forecasting the weather for a period longer than 24 hours) may reassure us by confirming that the method adopted (i.e. that of unravelling the skein of the complexity of the atmosphere by means of experiments in a fluid-dynamics and thermodynamics laboratory and of subsequently recombining these elements in a model) has been quite fruitful. This is undoubtedly due to the fact that, in order to reconstruct the behaviour of the system, we have placed ourselves at its same macroscopic level. However, there do exist some macroscopic phenomena, such as the outbreak and evolution of tropical cyclones and hurricanes, that are not captured well by the models. This appears to be due to two factors: on the one hand the causes that prime these phenomena seem to be on a small scale; on the other hand, at least in their mature stage, they can be regarded as phenomena of macroscopic self-organisation that perhaps represent a complexity that is emerging at a macroscopic level and cannot be reduced completely to the basic laws of the system.

We have seen that the results of these weather forecasting models are supplied in digital form — and therefore can be objectively evaluated *a posteriori* by means of statistical indices that quantify the validity of the forecast — or transmitted as prognostic charts or sometimes graphs like the one we showed for the results of the ensemble integrations on a station. Weather maps, which allow us to follow the forecasts of the various fields in successive temporal steps, represent just the evolution of the ideal model we discussed in the previous chapter, because they make it possible to display (dynamically and not statically) the behaviour of the system under examination, in this case the atmosphere.

This immediate display enables us to comprehend the evolution of the system “at a glance”, even without consulting the data. Sometimes simply comparing these charts with the analysis obtained subsequently at the time limit provided for the model is sufficient to allow us to understand whether in that case the model has yielded satisfactory results or not, even without having to carry out laborious *a posteriori* statistical analyses. This is very useful when we analyse the characteristics of the model and of possible versions of it that have been modified on the basis of past case studies that were particularly significant from the meteorological point of view. In this regard, the possibility of having at our disposal a graphic display of the behaviour of a complex system such as the atmosphere facilitates the evaluation of the results, and substantially changes scientific practice in the analysis of the behaviour of this complex system. Evaluations are often performed on the ideal model whose graphic evolution is available: a check is carried out to ascertain in what position the various models place a certain atmospheric low, what values they forecast for its minimum value, etc.

These last examples help us to understand that we can change the physics of a model and control its forecasting performance on case studies of the past. However, the manageability and the flexibility of models as an instrument allow us to do other things as well: for instance, to construct possible worlds that do not exist currently in nature. This may be interesting when we wish or need to act on the natural environment in order to change it. For instance, let us suppose that we have decided to build a dam with an artificial lake, for the production of electric energy: if we have a reliable meteorological model (i.e. a model that has been validated on the basis of past situations or of everyday work), we can evaluate, *a priori*, the environmental impact of the lake in terms of the changes it may cause in the weather of the surrounding area. Obviously we can perform tests on the past years for which we have the meteorological data relative to the area under consideration. First we must run the model with the state of the ground as it is in reality (these runs of the model are called "control runs"), then we must change the state of the ground, creating a lake in the world of the model; at this point, we can run the model with these new boundary conditions. The fundamentals of meteorology explained in Chapter 4 lead us to reckon that the presence of this expanse of water will essentially affect the radiative exchange between the Earth and the free atmosphere, the transmission of heat, and the atmospheric humidity content. Basically, the most evident changes we expect are in the temperature of the air at surface level, in the precipitation regime, and in the phenomena connected to them. An analysis of these results may advise us to build the dam or not to build it, or to change the project in a certain way. In actual fact, if we perform these tests over a period of several years, the simulations will become climatic ones, and for these the reader is referred to the next chapter.

To round off this discussion, we must return briefly to the emergence of deterministic chaos that has led to the need of developing and using the method of ensemble integrations. We have already explained that temporal evolution may move apart two points that were initially quite near in the state space, causing them to become even considerably distant from each other after a certain time lapse[20]. Moreover, their way of moving away from each other is rather characteristic: after a certain instant the points diverge quickly (in an exponential manner): this is called a bifurcation. Where possible, the method for amending this flaw

---

[20] Obviously this is a distance in a multidimensional space.

in the deterministic treatment of a non-linear dynamical system has consisted in studying the evolution of the probability density function associated to the volume that quantifies the uncertainty about the initial state. In the case of meteorological models, an alternative method has been adopted: that of studying an extensive range of evolutionary paths, one for each point considered at the initial instant.

Whatever alternative we are compelled to choose, these examples show that the concrete forecast of the future state of a complex system cannot be achieved univocally with a deterministic prediction. Both the methods highlight the need to adopt a probabilistic point of view in a domain — systems of non-linear differential equations — that was considered the undisputed realm of determinism. The concrete example of the resort to ensemble integrations (which also has a share in changing the scientific and applicative practice in the world of meteorology) particularly reinforces this vision. In any case, it seems that Laplace's dream has vanished for good.

This page is intentionally left blank

Chapter 8

Climatic Models

In the previous chapter, we explained how it is possible to work out a simulation model of the atmosphere that can supply weather forecasts for a certain time lapse. Here, reminding the reader of the definition of climate given in Chapter 2, we will consider whether it is possible to work out simulation models that account for the average weather and its variability over a few decades, both validating them for the past and applying them to a forecasting of the future climate. As we have always done in this book, here we will not only evaluate the applicative results of these models, but also analyse their aspects of conceptual novelty as a means of investigation that by now has become a part of present-day scientific practice in the study of a complex system like the Earth.

8.1 From Weather Forecasting to Climate Forecasting: What Changes?

In the course of Chapter 4, we endeavoured to construct a theoretical vision, at least a qualitative one, of some phenomena and processes that take place in the Earth system, particularly highlighting the characteristics of some feedback cycles, and thus revealing the complexity of the system. When in Chapter 7 we proceeded to work out some simulation models for weather forecasting, we focused on the atmosphere subsystem, supplying an algorithmic reduction of its fluid dynamic and thermodynamic characteristic features by means of primitive equations and parameterisation schemes. In doing this, we dealt dynamically only with the atmosphere subsystem: the representation of its interactions with the other subsystems of the Earth system was confined to the status of a series of influences of an environment outside the atmosphere, and these influences were expressed in the model as boundary conditions and external forcing factors. More specifically, in this non-dynamical treatment of the relations between the various subsystems, the influence of the external environment is usually a one-way one, i.e. it is not subject to feedbacks, except in an artificial and, in any case, non-dynamical way[1].

What made it possible to choose this approach was, notably, the fact that the changes that take place in the subsystems situated at the interface with the atmosphere are usually slow in comparison with the meteorological evolution, on the time scales that are considered (up to 10 days). If we decide to follow a path that will lead us to the elaboration of models for studying the evolution of the climate, therefore at much larger time scales, the evolution in time of these subsystems (with their influences on the atmosphere and relative feedbacks) becomes no longer negligible. From this point of view, we should fully consider the dynamics of these subsystems in their interaction with the atmosphere, through the elaboration of a coupled system of diagnostic and evolutionary equations. Among other things, the fact that systems such as the oceans are characterised by inertia factors greater than those relative to the atmosphere may suggest that they play a stabilising role, thus promoting the predictability of the characteristics of the overall Earth system.

A further element that emerged in our discussion of meteorological models and is important both from a conceptual point of view and from an applicative one, is the recognition of the existence of a limited predictability time lapse for forecasts. In this perspective, if for the weather we cannot produce forecasts over more than 10 days, how can we hope to achieve a forecast of the climate over a future period that may extend to several decades?

The key for answering this question lies in the statistical formulation of the definition of climate and in the consequent concept of climatic forecast. The weather is defined when the values of certain atmospheric variables that approximately determine the state of the atmosphere are

---

[1] See the particular case described in Note 2 of Chapter 7.

known; weather forecasts aim at predicting the evolution of this state with time in an accurate, deterministic way (with the obvious limits revealed by the emergence of deterministic chaos and the consequent use of ensemble integrations). When we speak about the climate, on the contrary, we mean to define a certain period of time by means of average values and of the variability of quantities that are important in the atmosphere; climatic forecasts, therefore, do not aim at determining the precise appearance of various atmospheric states, but at revealing future "scenarios" in which these quantities take on certain average values and a certain variability, and in which it is possible, if necessary, to determine whether one or several "classes" of atmospheric states are more frequent than others[2].

8.2 The Concept of "Attractor" and Climatic Simulations

As a rule, the interest in the average behaviour of a system and in its statistical variations over a fairly long period of time, in comparison with the wish to determine the precise behaviour of a system, leads to lower requirements during the forecasting stage. This type of forecasting, therefore, is easier to carry out, particularly if there is the possibility of determining the statistical laws followed by the system. An extreme example, familiar to all of us, is the situation that arises when we repeatedly roll a dice. Though we cannot foretell the sequence of the sides that will come up, over a great number of throws we can predict the average frequency (1/6 for each side), and also the variability in the number of times each side comes up.

Obviously, for the system we are discussing in this book we do not know any well-defined statistical laws such as the law of large numbers for the distribution of random events: for instance, it is still under debate whether the course of the climate is driven by a complex, chaotic dynamics or by a sequence of random events[3]. In any case, we can adopt the depiction of the evolution of the system, which is represented by a

---

[2] The possibility of determining these classes will be briefly discussed further on.

[3] The interested reader is referred to the article by Pasini *et al.* (1997) and references therein.

point in the multidimensional state space, as a trajectory of this point in the course of time. Temporarily disregarding the problem of the univocal determination of the initial state, knowledge of the deterministic laws of the evolution of the system offers the possibility to determine the curve that the point that represents its state traces within this multidimensional space in the course of time. Vice versa, a statistical knowledge of the system would make it possible to know the average distribution of these points in the state space, their variability and the shape of the "geometric figure" (if any) formed by them.

Though we do not possess a detailed knowledge of all the laws that determine the evolution of the system, or an in-depth statistical knowledge of the system on the basis of purely theoretical explanations, the fact that there exist some coexistence or evolutionary laws interlinking the variables indicates that not all combinations of their values are possible[4]. In our case, this leads us to presume that, in the evolutionary history of the system under consideration, its state cannot cover all the points in the state space (as would be possible if there were no functional links between the variables), but must be confined in a subset belonging to it. The geometric figure that identifies this subset is called "attractor" and determines the points that are physically possible for the state of the system, given certain external conditions (essentially forcing factors and boundary conditions).

Statistical knowledge of the system may therefore consist merely in knowledge of the characteristics of the attractor: the barycentre of this geometric figure may represent the average state of the atmosphere, and the extension along the various axes may represent its variability. If, moreover, we were to identify some parts of the attractor that are particularly thick with points, i.e. are highly "frequented" by the state of the system, this would testify to the existence of particularly probable state classes.

The concepts of state space and attractor are obviously rather abstract and, because of the high dimensionality of our system, not easy to visualise. So we will resort to a two-dimensional mechanical analogy.

---

[4] For instance, in the simple case of two variables, x and y, that are interlinked by a linear law, only the points belonging to this line are actually possible on plane xy.

Let us consider a straight, horizontal little track that can be used as a track for marbles. If we bend this track downwards and upwards, we obtain a piece of a little roller-coaster like the one shown in Figure 9.

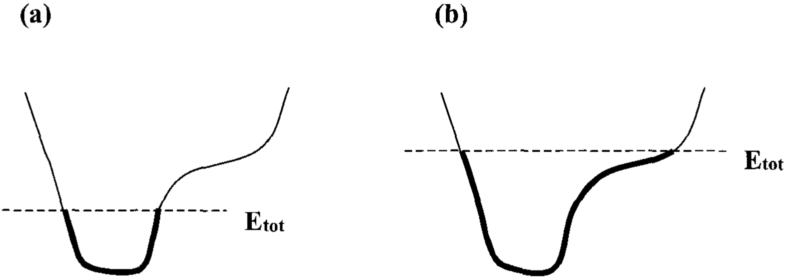


Figure 9. Mechanical analogy for the concept of attractor.

Now, if we imagine a condition of ideal frictionless motion, when we place a marble at a certain point of this track and let it start from a rest condition, it will run down the track, then run up again to the same height on the other side; subsequently its motion will be inverted, and it will return, after a certain period, to its starting point. The path covered by the marble (indicated by the thicker line in the figure) is determined by the height from which we have started the motion: as we know from the principles of mechanics we have studied at school, this determines its potential energy (which, since the marble starts from the rest condition, coincides with the total mechanical energy of the system, E_{tot}). At this point, if we change the height from which we release the marble, i.e. if we change the value of E_{tot}, the states that are possible for the marble change as well, as shown in the two parts of Figure 9.

The curves drawn with a thick line are thus the analogue of the attractor in the state space: as you can see, when the energy of the system changes, the extension and shape of the trajectory change as well. Among other things, in this simple mechanical example we are also able to determine which are the "most-frequented" stretches of this trajectory, i.e. the point where we are most likely to find the marble if we look at the system at any instant: if we apply the law of conservation of total energy (understood as the sum of kinetic energy and potential energy), we will find that the velocity of the marble is lower in the highest parts of the trajectory and higher in the lowest parts. So the stretches where the marble stays for a longer time are the highest ones: this would probably be particularly noticeable in the stretch of less steep slope to the right in Figure 9b.

Up to now, nobody has been able to determine the characteristics of the attractor of the atmosphere or, even less, of that of the Earth system. The main cause of this deficiency is the fact that there are few points available in the state space (for the brief period of extensive monitoring of the system: to be optimistic, the last century) with respect to the high dimensionality of the system: in nature we must be content with historically observed states, which, however, are only a small subset of the actually possible ones. We must add to this that, as the years went by, the forcing factors outside the system also changed, making the problem even more difficult to tackle from a theoretical point of view.

At this point, it may be natural to think that in this situation the models, since they allow us to reconstruct possible worlds and to repeat the simulations over an extremely long period of simulated time and with different initial states, may be able to contribute to the determination of the attractor of a system. If this is possible, although a precise determination of the evolution of the system is not obtained, the statistical properties of this evolutionary system can be recognised. Since these properties are precisely the ones that characterise the definition of a climate, the application of certain models may turn out to be essential for our purposes.

Once the model of a certain physical system has been worked out, for instance by means of a system of differential equations, in this dynamical model the attractor is the figure that determines the "asymptotic" behaviour of the system, i.e. the behaviour that appears after a transient stage of imbalance of the variables (such as the spin-up stage discussed in relation to meteorological models)[5]. Figure 10 shows the attractor for Lorenz's three-dimensional system, the first system in which the phenomenon of deterministic chaos was revealed. Notice the point at the centre of the figure, from which the motion starts: after a short stretch of trajectory, the asymptotic behaviour is reached on the attractor. This attractor is characterised by two wings, which express different classes, or regimes of motion.

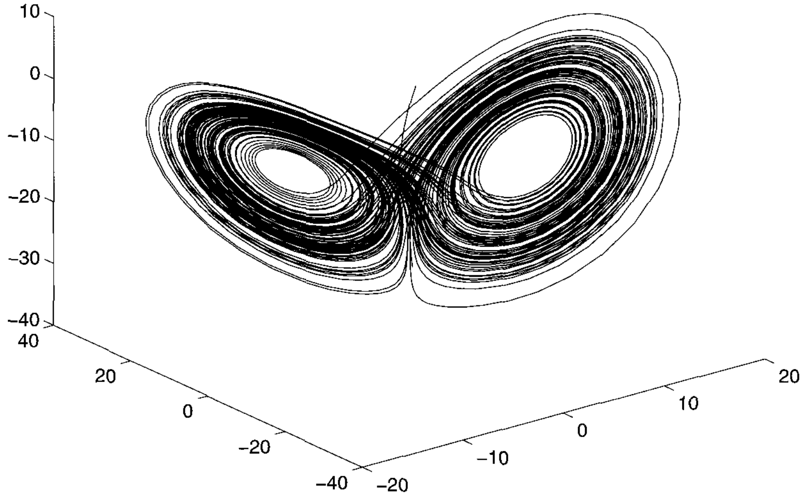


Fig. 10. The Lorenz attractor (figure reprinted from Pasini and Pelino (2000) with permission from Elsevier).

Incidentally, we would like to point out that the existence of attractors such as Lorenz's one contributed to the emergence of a "fractal"

---

[5] To return briefly to the mechanical analogy presented above, here the asymptotic behaviour of the system coincides with the real behaviour, because we imagined an ideal system without friction. If we introduce this slowing-down component as well, we will see that the velocity of the marble is gradually lessened (transient stage), and the marble ends up by stopping at the lowest point: this fixed point may therefore be the real attractor for motion in a real system with friction.

geometry, typical of objects characterised by a non-integer dimensionality[6].

8.3 Approaching the Description of a Coupled and Highly Interacting Climate System

In Chapter 4 we examined some phenomena and processes that take place in the atmosphere, and highlighted the influence (and relative feedbacks) to which they are subjected by what occurs in the subsystems at the interface with the atmosphere. Whereas for the short time limits characteristic of weather forecasts some of these interactions may be disregarded, or parameterised in a simple manner, on climatic time scales this is no longer possible: the simulation model would completely lose its effectiveness in relation to observational reality. More specifically, on these time scales the evolution of the subsystems at the interface with the atmosphere can no longer be defined *a priori* (through a given evolutionary law or by means of a representation of its internal dynamics alone), because it is affected by feedbacks due to changes that have taken place in the atmosphere in the meantime. Therefore it is necessary for the dynamics of the various subsystems to be coupled in a single, highly interacting dynamical system, where what occurs in a subsystem affects the dynamics of the other subsystems and is affected by feedbacks from them.

In the history of science, and above all in the course of the recent development of science, characterised by an extreme specialisation, the various subsystems of the Earth system were studied within the domains of separate disciplines. In each of these disciplines, theoretical knowledge has been advancing, and in these sectors, as in the study of the atmosphere, during the last few decades there has been an intensive use of simulation models. In some cases, like those of oceanic circulation models (though with the obvious differences due to the fluid medium and to the particular role of salinity), the formalism that is used is very similar to the one adopted in atmospheric models. In other cases,

---

[6] Obviously here we cannot dwell on fractal geometry. The reader is referred to a book written by the creator of this discipline (Mandelbrot (1987)).

particularly when study is addressed to the contributions to the dynamics of a particular subsystem that are also caused by biological organisms (as in the development of vegetation, or in the part of the carbon cycle due to emission or absorption by plants and algae), specialistic studies lead to very detailed representations of the processes on a local level, but require further development in order to fully consider and adequately size up the contribution of these factors in the climatic system on a regional or global level.

If we wish to make a list of the main components of the climatic system, we may consider the following subsystems, cycles and processes of the Earth system, for each of which we possess an individual theoretical treatment, as it has evolved in the course of the history of scientific knowledge.

- *Atmosphere*. This is a system we know well, and whose modelling was the subject of the previous chapter: obviously this knowledge is crucial for the problem we are considering, because it is within this system that climatic phenomena are detected, through the change in the values of its variables within a given time lapse.
- *Oceans*. The extension of the sea interface, the characteristics of its thermal capacity and contribution of humidity to the atmosphere, together with the dynamics of ocean currents and cycles such as that of *El Niño*, cause the ocean to be a fundamental factor in the interactions between the atmosphere and the external environment. Oceanography is a rather ancient discipline; at present it supplies a theoretical description of this system at the same macroscopic level as that of the description available for the atmosphere.
- *Continental surfaces*. Like that of the oceans, the behaviour of the solid interface is essential for understanding the flows of heat, humidity, etc., exchanged by the ground and the atmosphere. Here, too, great attention must be given to the changes in the state of the ground that affect the radiative balance. The reader should notice, however, that some of the changes in the ground are to be attributed not to natural dynamics, but to human activities (e.g. the deforestation or the establishment of new crops).
- *Cryosphere*. This name indicates the subsystem formed of the sea ice and land ice, whose evolution is now described dynamically. The

dynamics of the formation or melting of ice obviously affects the radiative exchange between the ground and the atmosphere, makes it possible to better define the capability of absorption and reflection of solar radiation in the course of time, and also affects the salinity rate of the various parts of the oceans.

- *Aerosols*. We already explained that in certain conditions the lithosphere introduces in the atmosphere dust of various origins (for instance, volcanic ash from eruptions), and we described the possible primary, direct effects of this on the radiative balance, the secondary effects on cloud formation, and the consequences of these effects. The duration of the persistence of these aerosols in the atmosphere depends on several factors, including their chemical composition. Now we are able to describe the life cycle of aerosols formed of sulphates and also to schematically describe the behaviour of aerosols of different compositions. It is important to point out, however, that the emission of these aerosols depends chiefly on natural events that may be hard to predict, or on anthropogenic processes.
- *Carbon cycle*. This is the complex cycle that accounts for the presence and amount of carbon dioxide in the atmosphere. The presence of CO_2 depends on mechanisms of emission, absorption and storage that involve both physical and biological phenomena in the oceans and on land: it is known, for instance, that in photosynthesis CO_2 is absorbed, and its carbon is stored in the wood that forms the trees. As in the previous case of aerosols, now most of the CO_2 emissions have an anthropic origin.
- *Vegetation*. All the vegetable world affects the climate and its changes, and is affected by them: on the one hand its presence (with different species) influences the albedo[7], the humidity introduced in the atmosphere with the evaporation/transpiration process, and the absorption of carbon dioxide; on the other hand its growth is driven by weather and climate factors that it contributes to change. During

---

[7] The reader is reminded that the albedo is defined as the ratio of the energy reflected in space by the Earth, clouds and atmosphere to the total incident energy (coming from the Sun).

the last few years, there have been some attempts at a dynamic modelling of the vegetation. Here too, however, it is necessary to point out that these attempts must include, as data, the changes in vegetation caused by mankind — first of all the deforestation activity.

- *Atmospheric chemistry.* The chemicals present in the atmosphere, most of which are emitted as a result of human activities, interact in a complex way through physical phenomena (such as diffusion or deposition on the ground) and chemical reactions. All this is interesting for the study of possible climatic changes, because the presence or absence of certain chemical species affects the radiative balance and other phenomena in the atmosphere. On the other hand, the changes may give rise to feedbacks on the atmospheric chemistry: one of several possible examples is the fact that the occurrence or non-occurrence of certain reactions and their speed depend on the temperature. A theoretical treatment, necessarily simplified, is now available, and attempts are being made to introduce it as an interacting model component within a climatic model.
- *Mankind.* As we have repeatedly stated, in almost all the points of this list of the subsystems that form the Earth system, the presence of mankind affects various important cycles of the system, to the point that it is reasonable to start thinking of the possibility of beginning to produce models of human activities in their interactions with the other subsystems. In this sense, we mean to interpret the presence of mankind as an integral, fundamental part of the Earth system: though on the one hand a "human dynamics" is obviously hard to define, as the so-called "human sciences" have shown, on the other hand considering the human element within the Earth system may be the only way to reveal the feedbacks of possible climatic changes on human activities, and to understand, in an integrated way, the sustainability and limits of the system.

Under the present circumstances of climatic modelling, a model that is (so to speak) complete, that integrates in a single dynamical system all the components we have just discussed, does not actually exist. In particular, the human element is still outside the model-system that is being studied dynamically, and acts on the model as a factor that changes some of the forcing factors of the system. Moreover, as regards the aspects relative to the atmospheric chemistry and, partly, to the dynamics of vegetation, the integrated modelling treatment of these subsystems is still rather problematic, both because of "connection" problems of the formalisms in the model and because of numerical problems often due to the different scales of development of the relative phenomena with respect to the atmospheric and oceanic processes in particular.

In this picture, the "hard core" of present-day climatic models consists of a strong coupling of a meteorological model with an oceanographic one, in which the aspects relative to the continental surfaces and cryosphere are included as an integral part. To this base there are added some interacting modules relative to the dynamical behaviour of aerosols (both those containing sulphates and others having a different composition) and of the carbon cycle. The models thus obtained are called AOGCMs (Atmosphere-Ocean General Circulation Models).

The need to carry out simulations at least over several decades obviously leads to a reduction in the resolution of these models, in comparison with that of the purely meteorological models discussed in the previous chapter. At present, the state-of-the-art values for the various coupled models are, in a meteorological model, an average distance between the grid points of 250 Km horizontally and 1 Km vertically (where, however, in many cases an increased resolution is kept in the lower layers). In an oceanic module, the range is usually from 125 to 250 Km for horizontal resolution and from 200 to 400 m for vertical resolution. Obviously, as we have already seen for weather forecasting models, here too parameterisation schemes are used for the processes that take place on a spatial scale smaller than that of the grid, such as the convection of clouds in the atmosphere and oceanic convection.

In spite of the difficulty of including the elements that are more distant from a purely physical treatment, we can safely state that modern coupled models are an important theoretical training-ground for the multidisciplinary representation of complex phenomena like those that take place in the Earth system. More specifically, the need to work out a model (the only method with which we can hope to attain an understanding of the phenomena of this highly interacting system), compels experts of various disciplines to speak the same language, that is to work out a unitary formalism for obtaining software to be run on a computer (everybody knows that computers are stupid and can carry out instructions only if they are self-consistent and written in a univocal language).

8.4 Experiments for Validation and Sensitivity Testing of a Climatic Model

In the course of the brief description of coupled models that we have just completed, we duly highlighted the importance of regarding the various subsystems of the Earth system as interacting dynamically in a single model, particularly in order to avoid one-way influences of the external environment on the model-system under consideration. This makes it possible to account for the mutual interactions between these subsystems, including the feedbacks, therefore to achieve a more consistent and correct simulation of the evolution of the whole system.

In actual fact, there also exists another motivation for dynamically joining the various subsystems in a single model. In fact (returning to the concepts introduced in Section 8.2), if we believe that the model correctly simulates the meteorological and climatic behaviour of the Earth system, we can think that the climate is represented by information that can be extracted from the attractor of the model, in terms of average state, variability and most frequent states. Strictly speaking, however, the attractor of a system is a "static figure", that does not change in the course of time[8], only if the external forcing factors and boundary conditions are fixed: in this case the system is called "autonomous"[9]. The

---

[8] And that therefore is outlined increasingly well by observing the state of the system over increasingly long periods.

[9] It is fairly simple to observe how the statistical properties of an autonomous system are driven by the values of external constraints or forcing factors. If, for instance, we consider an isolated system formed of a gas in a container to which a certain amount of energy has been previously supplied (possibly in the form of heat), the velocity of the particles within a statistical distribution (called Maxwell-Boltzmann distribution) is determined by this amount of energy. If we supply more heat to the system, then isolate it

most immediate way of not losing anything of the wealth of knowledge gained from the concept of attractor is therefore that of rendering the model-system as autonomous as possible, by reducing the evolution of the external environment to a minimum, i.e. by dynamically including all the evolutionary subsystems in the model[10].

Now that we have acquired this theoretical background and this vision of climatic models as dynamical systems, we will endeavour to apply them concretely. The first thing we require of a climatic model, obviously, is a correct reconstruction of the climate of the past. Let us therefore consider a period in which the external forcings and the boundary conditions can be regarded as practically constant, and let us proceed to simulate the climate over this period.

The first problem we must tackle is obviously that of the sensitivity to initial conditions: if we start from an initial condition determined by a (necessarily approximate) analysis at a given instant, after a certain time lapse the simulated evolution will be quite different from the real evolution of the system under examination. This problem was explained for meteorological models in the previous chapter, and now reappears in the model-system that is extended also to the other subsystems. Now the problem actually seems even more serious, both because for the distant past no accurate analyses are available, and because the lower resolution of a climatic model in comparison with a meteorological one renders the analysis even more approximate.

In actual fact, as we mentioned previously, the statistical character of the definition of climate makes it possible to get around this problem: in particular we can consider the properties of the attractor under the

---

again, once the internal equilibrium has been reached, the velocity of the particles has changed (and its average value has increased).

[10] It is evident that in any case there will remain some constraints or forcings outside the Earth system: the reader should consider the amount of incident solar radiation, which is precisely what determines the energy supplied to the system. Obviously it is not possible to establish a perfect parallelism with the case described in the previous note, because here we are in a non-linear dissipative system with many interacting components. However, should we suppose that there is a radical change in the so-called solar constant, we might expect that the system would respond to this different value of the forcing with a change in its average state, that is with a change in the barycentre of its attractor (which, we must point out, is not defined only by the temperature values, but also by the values of the other variables), and possibly also with changes in its shape.

influence of the specific boundary conditions and external forcings, thus evaluating the average climatic properties and the variability on the basis of the statistics of the states of the system in the period under consideration. Obviously, because of the high dimensionality of the state space of the model, we will extract only the variables we regard as fundamental for the reconstruction of the climate.

At this point, however, it is evident that the single run of the model supplies only an evolutionary trajectory in the state space, a sort of "sampling" of the attractor, just like the evolution of the real system is a particular realisation of the possible physical states. So if we have no hope of being able to reconstruct this real trajectory, because of the emergence of deterministic chaos, we may just as well increase the statistics and run the model repeatedly with different initial conditions. This way we can explore the state space more thoroughly and reconstruct the attractor with a greater abundance of details[11]. If we carry out a considerable number of runs (which here too we can call "ensemble integrations"), we can determine the average values and possible climatic variability over the period under consideration, and we can also understand whether the individual, particular realisation of the climate extracted from the data of the real system falls into the class of climatic values simulated by the model. If it does, the model is corroborated by the real data: this is the validation stage[12].

Once the state space has been explored and the climatic variability has been evaluated by means of the analysis of the sensitivity to initial conditions we have just described, if a validated and physically self-consistent model is available, it is possible to evaluate the climatic

---

[11] It is worthwhile to remind the reader that the attractor is the figure that determines the asymptotic behaviour of the model-system, which appears after the transient stage due to the initial unbalancing problems of the fields. In this regard, we would like to point out that, in modelling practice, work often starts from all sorts of initial conditions (e.g. a condition of atmosphere at rest), beginning to consider the data produced by the model only after the end of this transient phase, which in a climatic model may even be several years, because of the unbalancing of the hydrological part of the module relative to the continental surfaces.

[12] During the validation of one or several models, it is also possible to carry out studies on the sensitivity to changes in the modelling representation of the various processes described by means of equations or parameterisations, if necessary with expressly-made projects of intercomparison between different models.

response of the model-system to changes in the boundary conditions and external forcing factors. Assuming that our theoretical knowledge expressed in the equations and parameterisation schemes has an extensive domain of validity, this is done in order to analyse the behaviour of the Earth system thus simulated, once a new balance has been reached in extreme conditions (doubling or quadruplication of the concentration of CO_2, total deforestation, conditions of "nuclear winter" due to a thermonuclear war, etc.). Since, as we have seen, the Earth system is very complex, there actually exist countless factors that affect the climate, both directly and indirectly, through feedback mechanisms. In this framework, studies on the sensitivity to boundary conditions and external forcings may make it possible to distinguish the factors that have a greater impact on climate change.

This way we can also begin to explore the scenarios that may appear in the future. As a rule, some runs of the model (called control runs) are carried out with the values of the forcing factors and boundary conditions typical of the present period or of a reference sample period. Then the model is run with some of these parameters changed (usually by bringing them to extreme values).

Since the model is practically autonomous, once some ensemble integrations have been carried out it is possible to evaluate the statistical changes in its average state and in its variability, through the variations in the two attractors, as they are outlined, respectively, by the control runs and by the runs with altered forcing factors or boundary conditions. The parallelism with the situation described in Note 10, i.e. the gas system to which different amounts of energy are supplied, is rather cogent. A situation of this type — returning for an instant to the case of the repeated rolling of a dice — can be obtained by applying a tiny weight to a side of the dice, which thus becomes a loaded dice. Now, though the temporal sequence of the draws remains unpredictable, the frequency of the draws, obtained empirically after the dice has been rolled for a considerable number of times, will turn out to be different, and, from now on, predictable for that particular dice.

The analysis performed by means of climatic models, therefore, is basically statistical; as in other systems, this treatment supplies results that otherwise would be unattainable. More specifically, in the analysis of the sensitivity of a climatic model with perturbation of the forcing factors and boundary conditions, the first thing to be evaluated is the significancy of the climatic changes that are found. Since the first runs determine the averages and the natural variability of the quantities that are important for the definition of the climate of the control period, it is necessary to ascertain whether the climate outlined by the runs with perturbed parameters falls into the range of climatic values of the control period, or whether it does not. In the latter case, the changes that are found are statistically significant, and we can recognise an impact of the change in the values of the forcing factors and boundary conditions on the climate; in the former case, these values fall into the natural variability of the climate, so there does not exist a clear signal of an influence of these external changes on the dynamical system.

8.5 Evolutionary Validation and Climatic Forecasts

The analysis of climatic models carried out up to now was particularly focused on the evolution of the state of the model-system in the state space and on the concept of attractor, which, as we explained, summarises the statistical characteristics of the system. Our discussion reveals the close connection between the method for studying the climate on the Earth and the methods adopted for analysing complex systems that are characterised by smaller dimensions and a greater manageability (possibly systems of which we have an experimental counterpart in the laboratory). This attributes to modelling the difficult task of verifying advanced theoretical physics methods in an extremely concrete and applicative field, characterised by impact consequences that can be verified in a very cogent way.

Despite the recognition of this noble origin of the structure and investigation methods of climatic models, which issue from the theoretical physics of complex systems, the rigour of theory (as usually happens in applicative disciplines) must sometimes be "softened" because of the working conditions, which often require the achievement of results (albeit partial and approximate) within a reasonable time. Thus, for instance, in the validation and sensitivity testing experiments that are actually carried out with present-day models, because of the prohibitive computer times required, the use of ensemble integrations was rather reduced, and was limited in most cases to a small number of runs. A discrete, rather restricted sampling of the attractor was thus obtained, so it is reasonable to point out that this sampling may turn out to be insufficient for determining the desired climatic statistics with due accuracy.

In the theoretical context of the representation of the Earth system by means of equations and parameterisation schemes, there has been substantial progress during the last few years. Up to a short time ago, for instance, in the physical coupling between the oceans and the atmosphere it was necessary to introduce an exchange of flows of heat and water whose nature was not physical (i.e. they were not observed in nature), in order to reconstruct the present climate in a satisfactory manner: this was called "flow adjustment". Nowadays the improved representation of the exchanges at the interface between the oceans and the atmosphere makes it possible to dispense with this artificial adjustment. In any case, we must remember that, like the meteorological models discussed in the previous chapter, climatic models contain some parameters that in any case must be fixed (without losing track of their physical meaning) with a somewhat artisanal activity of numerical experimentation.

These remarks reveal how difficult it is to preserve an extremely high degree of theoretical rigour in the application of climatic models to the validation and sensitivity studies presented in the previous section. In these studies, the approach to the problem is theoretically correct in any case, because the boundary conditions and external forcings are regarded as practically constant (the model is autonomous). This means that we are examining a situation of equilibrium, where, after the transient stage due to the effects of the unbalancing of the initial fields, the trajectory of the state of the model-system lies on its own attractor.

The problem that comes up at once, if we consider the data of the external forcing factors and boundary conditions in the last century, is that these factors and conditions cannot at all be regarded as almost constant throughout the period. Among other things, there is some evidence of the fact that they have a certain influence on the value of some quantities that have a share in defining the state of the system, such as the quantity shown in Figure 2 of Chapter 2, relative to the changes in temperature detected after volcano eruptions. If we add to this the fact that the influences of human activities, in particular anthropogenic emissions of greenhouse gases, have been increasing on the whole[13], we can understand that, in order to simulate the behaviour of the climate during the last century, validating a model on these real data, we cannot postulate conditions of equilibrium within an autonomous model-system.

At this point, however, the whole theoretical construction based on the concept of the attractor of an autonomous system begins to "totter": the system can no longer be studied in a state of equilibrium, but must be studied in evolutionary conditions, i.e. transient ones. The calm, reassuring situation where, at least in theory, we could run the model for a long simulated time with the same forcing factors and boundary conditions, thus determining with great richness of detail the shape of the attractor that determined its climatic statistics, is no longer permissible. If we really wish to validate a model with respect to real observations, we must cause it to reconstruct the climatic statistics of its evolutionary history: we are in the conditions of an evolutionary validation.

In these conditions it is no longer possible to speak about an attractor as a static figure in the state space. In actual fact, the problem due to the fact that forcings and boundary conditions depend on time has been studied in dynamic systems characterised by small dimensions, and it has been noticed that the trajectory of the model-system tends to shift progressively towards attractors having a different barycentre, extension and shape, though obviously it never lies on them: the system is in transition and not in equilibrium. This shift takes place sometimes gradually and sometimes more suddenly, and sometimes depends on the position in which the point that determines the state of the system is located in the state space. In essence, it has been demonstrated that an evolutionary problem of this type can be studied, though the extensive statistics that in the stationary, equilibrium-based problem sprang from the fact that the trajectory lay repeatedly on the static attractor must be

---

[13] We remind the reader that in climatic models the mankind subsystem has not yet been dynamically coupled with the other subsystems.

obtained here by further increasing the use of repeated runs of the model with different initial conditions.

The outcomes of the evolutionary validation of these coupled models during the twentieth century, if we include the real data of the concentrations of greenhouse gases, changes in the soil, and so on, lead to satisfactory results where averages and global or hemispheric variability are concerned, while they show discrepancies (even considerable ones in some cases) on individual regions of the globe. In particular, the variability with time, on the scale of a few years or a decade, is caught better if real variations in Sun irradiation and volcanic aerosol emission are included as forcing factors.

When we have worked out a climatic model and have validated it in an evolutionary manner over a historical period for which we possess real data, we can rely on the validity of its theoretical scheme. We are ready, at this point, to apply the model to the simulation of the future evolution of the Earth system, in order to forecast possible climatic changes in a more or less near future, possibly in an accurate, reliable way.

But here comes the difficult part! Since the system is non-autonomous, above all because the human influences are an external factor that is not dynamically coupled with the other subsystems of the Earth system, the behaviour of these influences in the course of time (supplied as forcings or boundary conditions in the model-system) are affected by forecasts, made outside the model, that can be produced only by the so-called “human sciences”. This fact introduces so many important factors of uncertainty that it is absolutely necessary to hypothesise various future scenarios.

Since variability in natural forcings is rather unpredictable and is evaluated statistically, the essential forcing factors that come into the model-system from the outside are basically the concentrations of greenhouse gases and anthropogenic aerosols, and the changes in the use of the soil. Both are driven by the scenario that can be surmised for the global and regional socio-economic development (which notably determines the emission of these gases), and are determined after having considered the feedbacks of the system, which may lead, for instance, to the absorption of a part of the emissions. Socio-economic development, in turn, follows market laws, but is also affected by interactions with the political forces. This situation is summarised in Figure 11, which shows the cascade of factors that influence the data to be included in the climatic model — these data, in turn, obviously influence the results of the forecasts. The figure also surmises that there is a feedback of the results of the models on the political forces: considering the situation of international negotiations, at present this seems to be an excessively optimistic surmise[14].

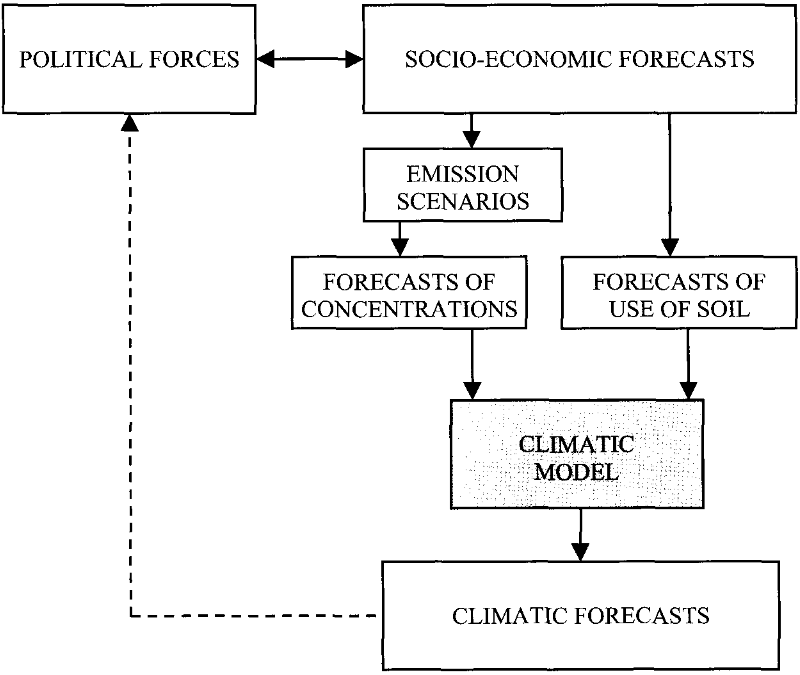


Figure 11. The socio-economic forcing factors for a climatic model.

---

[14] We will return to this topic in Chapter 9.

Of late, the scientific community that deals with climatic models has acquired a range of scenarios of socio-economic and emission developments. The climatic simulations refer more and more often to runs of an individual model on the basis of the data of these scenarios, so as to give out a similar range of climatic developments, in which each forecast is relative to a scenario for economy and emissions that has been surmised before starting the run. Some examples of these forecasting results will be discussed in Section 8.7.

8.6 Simplified Models and Regional-Scale Models

As the reader has undoubtedly understood from what we have explained up to now, the computer power required for carrying out simulations with coupled atmospheric-oceanic models that dynamically involve the other subsystems as well (the so-called AOGCMs) is very considerable, even with a limited resolution. It is so considerable that it turns out to be extremely difficult to run these models for more than a few decades and with a number of ensemble integrations sufficient for determining correct statistics.

In this situation, on the one hand scientists felt the need to have simplified models that could lead to longer integrations and also facilitate the testing of the sensitivity to changes in fundamental mechanisms (these tests are particularly onerous from a numerical point of view). On the other hand, the low horizontal resolution and unsatisfactory results in the reconstruction of the real climate on a regional scale (though extended to subcontinental areas, such as the Mediterranean area, northern Europe or Saharan Africa) led the scientists to consider models that were more expressly dedicated to catching details on this scale.

The class of simplified climatic models includes several types of models. We can only describe them concisely, in the list below.

- Energy Balance Models (EBMs). These are the models with the highest degree of simplification, because they regard the Earth's atmosphere as a single point, and, in the course of time, evaluate the global radiative balance (i.e. the balance between the incoming solar radiation and the outgoing one), under the influence of changes in greenhouse gases and sometimes also aerosols. A slightly more advanced form of these models also evaluates the conveyance of energy between different latitudes: in this case the evolution of the temperature for each latitude band is obtained.

- Radiative-Convective Models (R-CMs). These models consider the radiative transformations that take place when energy is absorbed, emitted or diffused, and the role of convection, which is basically understood as a factor that modifies the vertical thermal profile of the atmosphere, with a mechanism similar to the one mentioned in the previous chapter (Section 7.7), which we may call "convective adjustment". In R-CMs, it is possible either to regard the atmosphere as a single vertical column or to introduce a horizontal spatial distinction (though a rather rough one).
- Statistical-Dynamical Models (SDMs). These models usually combine the horizontal energy transfer of EBMs with the radiative and convective approach of R-CMs. However, the transfer of energy from the equator to the poles takes place in a slightly more sophisticated way, on the basis of theoretical and empirical relations in the flow between different latitudes. Parameters such as the speed and direction of the wind are estimated by statistical relations, while for obtaining an estimate of the horizontal diffusion of energy the laws of motion are used.
- Earth-system Models of Intermediate Complexity (EMIC). This is a rather diversified class of models, in which the individual models all have the characteristic of being an attempt to bridge the gap between AOGCMs and the previously-described simplified models. They are usually dynamical models characterised by certain simplifications with respect to the AOGCMs, particularly in their atmospheric and oceanic modules; but, just the same, they offer the possibility of introducing a series of bio-geo-chemical cycles, such as the carbon one, and of evaluating the model's sensitivity to changes in the forcing factors and theoretical schemes much more easily and for a longer simulated time, in comparison with what is possible with AOGCMs.

As a rule, all these models are used for testing, in a simplified way, new theoretical schemes and previously disregarded feedbacks (possibly examining the various components of the Earth system one at a time), and for obtaining scenarios over time lapses that are so prolonged as to make it impossible to carry out simulations with fully coupled models. In particular, once one of these simplified models has been validated, its results can be rendered quantitatively consistent with the averaged results of an AOGCM, by adjusting certain parameters with a fine tuning operation[15]. This makes it possible, for instance, sometimes to use a simplified model in place of a more complete AOGCM in the long-range forecasting of climatic scenarios, at least in the cases in which interest is focused on averaged global- or hemispheric-scale climatic values.

In the opposite case, i.e. when averaged data are not sufficient and more precise information is needed, on a scale even smaller than that solved by an AOGCM, it is necessary to adopt "regionalisation" techniques that make it possible to carry out a more accurate, reliable investigation on these scales. On a regional scale (by "region" here we mean a part of a continent that is characterised by a comparative homogeneity of climate), the climate is affected not only by the average global values of certain variables and by the oceanic and atmospheric circulation on a global or hemispheric scale, but also by some more local forcing factors, for instance the presence of complex orographical features, great lakes or practically closed sea basins such as the Mediterranean Sea, the characteristics of the use of the soil, etc. Moreover, the climate of a particular region may be heavily influenced by cycles that take place elsewhere in the ocean or atmosphere: this is the case of the far-reaching influence of the phenomena resulting from *El Niño* and the North Atlantic Oscillation (NAO).

By studying these regionalisation techniques, we can hope to obtain a more detailed simulation of the spatial structure of the temperature and precipitation, a smaller-scale description of the atmospheric circulation (e.g. convective systems, tropical cyclones, breeze circulation), and a representation of the processes whose frequency in time is higher than

---

[15] This term already appeared in the previous chapter (Section 7.6): it was used for highlighting that part of artisanal activity present in meteorological models. Now it can also be used for climatic ones.

that of processes that can be simulated by an AOGCM, e.g. frequency distributions and intensity of precipitation.

In short, there are three types of techniques that make it possible to obtain results on a regional scale: application of atmospheric models with a high resolution that can be modified, if required; regional climate models (RCMs); and downscaling statistical techniques. Whereas the use of high-resolution global atmospheric circulation models for a few decades of simulated time is still in the experimental stage and quite expensive in terms of computer time, the use of RCMs and downscaling statistical methods seems to be more established, and leads to rather generalised improvements in the results.

Regional climate models are basically limited-area models "nested" within an AOGCM. This means that the integration is carried out on a limited region of the globe, with a high-resolution grid, and is driven by the boundary conditions supplied by a global model. This way it is possible to perform a detailed study of the climatic evolution of a certain zone. To this day, the large-scale information given by the boundary conditions is usually not affected by feedbacks due to the smaller-scale evolution described by the RCM; however, experiments have been started on a dynamical nesting in which the feedbacks of the regional scale on the global one lead to the interactive running of the global model together with the regional one, allowing for the feedbacks due to the evolution described by the latter.

The statistical downscaling methods are based on the fundamental assumption that the regional-scale climate is conditioned by two main factors: the large-scale climatic state, and the regional or local characteristics of the territory (orographic features, distribution of land/sea boundaries, soil cover, etc.). At this point, if we manage to develop a statistical model that links some large-scale characteristic features to some regional- or local-scale variables over a reference period, we will be able, over a certain period in the future, to use the climatic forecasts of the large-scale characteristic features obtained by means of an AOGCM and to find the values of the regional or local-scale variables by means of the application of this statistical model. This way we will be able to obtain smaller-scale climatic forecasts. To do this, we can use several techniques, ranging from a basic multiple linear regression to more sophisticated artificial-intelligence methods such as neural network models.

8.7 Simulation Results

Now that we have established that climatic models are dynamical systems with attractors and certain properties relative to trajectories in the state space, and once we have outlined their structure, we can finally proceed to analyse the main results yielded by their concrete application. In doing this, we will present only some results that are common to all currently existing models and are regarded as well-established by the international scientific community. Moreover, the compactness of this book and the fact that it is focused on the study of climatic changes in the present and in the immediate future compel us to overlook some difficult and prolonged simulations relative to paleoclimatic studies.

First of all, we would like the application of climatic models to help us to shed light on a point that has appeared several times in this book: the possibility of determining the causes of the climatic change that has taken place during the last century, as revealed by the data reported in Chapter 2. In the subsequent chapters our knowledge of the Earth system has increased; we have become aware of the various factors that affect the climate; and we have understood that the system is extremely complex and that therefore it is not possible to carry out a linear causality analysis, as is done, for instance, on a simple Galilean mechanical system. This complexity was what led to the development of simulation models, first for meteorology, then for the investigation of climate. These models, starting from laboratory-validated basic laws, essentially constitute an instrument for unravelling such an intricate skein.

Well, we have said that the first stage of the testing of a model consists of the validation of the model in the reconstruction of the past climate, either as an independent system (if it is possible to consider a period in which the boundary conditions and external forcings can be regarded as almost constant), or in a transient phase (if the boundary conditions and external forcings are in evolution). As a matter of fact, during the last century the changes in the natural and anthropogenic influences that affect the Earth system have been so important as to put it in evolutionary conditions: therefore, using the observational data for these boundary conditions and forcing factors, it is possible to attempt to reconstruct the climate by means of simulation models. As we have seen, in this situation the statistics of the model that is applied is somehow reduced (the trajectories in the state space no longer lie on a static attractor), and it will be basically necessary to take ensemble integrations into account.

With reference to Plate 9, which presents some reconstructions of the global-scale temperature, we will now briefly examine the results of this causality analysis. In the three parts of the figure, the red line represents the anomalies observed in the global temperature with reference to the average of the forty years between 1880 and 1920; the gray bands represent the results of the reconstruction of this temperature that have been obtained by means of four ensemble integrations of an AOGCM. Part (a) refers to runs of the model in which only the purely natural forcing factors, i.e. solar radiation and volcanic activity, have been caused to change (according to actually observed values), while the other forcings have been left at a constant value. In Part (b) the situation has been inverted: values that are variable and have been actually observed are considered for the partly or chiefly anthropogenic forcing factors, such as greenhouse gases and sulphate atmospheric aerosols, while the solar radiation and volcanic activity are kept at a constant value. In Part (c), all the previously discussed forcing factors have been caused to change, according to the values that have been observed.

The results presented in Plate 9 are particularly interesting. More specifically, the anthropogenic forcing factors appear to be necessary for a correct reconstruction of the global temperature data relative to the last three decades. We must point out, however, that the reconstruction of the whole series requires the introduction of the variability relative to both types of forcings. The evolution of the global temperature is thus reconstructed in a satisfactory way, and the use of climatic models also leads us to infer that the anthropogenic forcings are a primary cause of the change in the global temperature.

Now that we have seen how an AOGCM can reconstruct the global temperature values of the last 140 years, we may pose the problem of what happens in the reconstruction of the temperature in individual zones of the Earth. If we limit ourselves to considering the temperature of the air in large continental or oceanic masses, the situation is fairly satisfactory; but as soon as we pass to the regional (i.e. sub-continental) scale, the errors of this type of model become considerable: there are typical errors of ±3-4°C in the temperature and between −40% and +80% in the amount of precipitation, in comparison with the observed values. In this field, we are helped by the regionalisation techniques, which lead to a considerable reduction of the errors, particularly those in the reconstruction of the temperature, which now remain within a range of ±2°C.

Once the models have been validated in the reconstruction of the past climate, it is possible to tackle the problem of forecasting future climatic scenarios on the basis of the available socio-economic scenarios, as shown in Figure 11. Because of the variety of socio-economic and emission scenarios, in practice it has not yet been possible to run the climatic models that are most expensive in terms of required computer resources (i.e. AOGCMs) for all these scenarios. In order to evaluate the prevailing tendency, over the next 100 years, of average quantities such as the global temperature for all the available scenarios, simplified models have been applied to this forecast, obviously only after having optimised them with reference to an AOGCM, by means of fine tuning operations, as described in the previous section.

An example of the results thus obtained is presented in Plate 10, which indicates the global temperature forecast up to the year 2100. The graph explicitly shows the curves relative to the evolutions of the temperature, forecast on the basis of nine of these scenarios, with the uncertainty bars relative to the first six. The dark-gray band includes the set of simulations relative to a single model, with data from no less than 35 different scenarios, whereas the light-gray band includes the set of results obtained by 7 different models with the data of all the scenarios. It is evident that all the scenarios suggest a more or less marked rise in the global temperature in the models. This analysis leads to the conclusion that, with a reasonable degree of reliability, in the year 2100 we may expect a rise in temperature ranging from 1.4°C to 5.8°C with respect to the value of 1990.

Though these results are extremely significant, it would be desirable to obtain more, particularly as regards the values of other quantities and phenomena whose degree of intensity and more or less frequent appearance contribute to a change in the climate of a certain zone. For this purpose, obviously, we must resort to the results of simulations performed by means of AOGCMs. So, though the currently available computer power has not allowed a wide-spectrum analysis (on all the scenarios that have been surmised) like the one we have just presented for the global temperature, it will be worthwhile in any case to briefly summarise the evaluations of the future climate obtained by means of simulations with fully coupled models. We will condense these forecasts into the following basic points:

- the temperature of the air near the surface (both of the land and of the sea) will rise;
- the temperature of the air should rise more during the night than during the daytime, so there should be a drop in the temperature range;
- the temperature of the sea will rise;
- the whole troposphere will undergo a warming process;
- the low stratosphere will tend to cool down;
- the amount of sea ice will decrease;
- the snow cover on the continents will also decrease;
- the water vapour in the troposphere will increase;
- the so-called "heat index"[16] on the continental zones will tend to rise, and this will lead to worse conditions of meteorological and climatic comfort for mankind, at least at the low and medium latitudes;
- the precipitation at the low and medium-low latitudes will decrease (this, combined with the rise in temperature, will lead to the risk of an increase in the desertification of certain areas);
- the precipitation at the high and medium-high latitudes will increase;
- some models forecast an increase in the frequency and intensity of tropical cyclones;
- some models forecast an increase (albeit a slight one) in the frequency and intensity of heavy precipitation at mid-latitudes.

---

[16] This is an index that is obtained by combining temperature with humidity.

The forecasts of the AOGCMs have been refined, of late, by the use of regionalisation techniques. These recent studies make it possible to evaluate the changes in temperature and precipitation regimes in individual subcontinental regions. For instance, in these simulations, Alaska, northern Canada and Greenland undergo an above-the-average warming during the winter, while precipitation decreases in a rather marked manner in the Mediterranean area during the summer. Recent studies have also revealed that the melting of ice may contribute a great quantity of fresh water to the northern Atlantic Ocean, and may deviate or lessen the thermohaline circulation there; so, in a context of general warming, northern Europe may undergo a regional-scale cooling[17].

So, besides the practically univocal forecasts of a rise in temperature, changes in other important parameters are also forecast. For instance, though these results should be regarded only as a general indication, in the situation of the Mediterranean area (and in Italy in particular), we can surmise a change in the precipitation regime, with precipitation whose total amount will perhaps be less abundant, but whose appearance will be concentrated in single, more violent episodes, with non-beneficial effects.

To conclude this overview of results, we would like to point out that the forecasts we have concisely summarised have consequences also on phenomena or processes that are usually not perceived as strictly climatic. For instance, the rise in temperature and the decrease of the expanses of ice both in the sea and on the continents directly affect the sea level to be expected during the next 100 years. Still referring to the variety of scenarios surmised for the economic situation and the emissions, Plate 11 shows some forecasts of the sea level, which will be clearly rising up to 2100. The six lines are relative to six different scenarios, while the dark-gray band refers to the averages of the runs of some AOGCMs, and the light-gray band indicates the scatter of the individual runs of the models.

Obviously what has been presented in this section is a summary of the forecasts that are believed to be rather well-established. It is clear that in a discipline like the study of climatic changes, characterised by an

---

[17] The reader is reminded of the previously cited Wood *et al.* (1999).

"explosion" of scientific studies, almost every day an article appears in some specialised journal, bringing further contributions to the understanding of phenomena and to their forecasting in the future. In particular, some phenomena that had been previously neglected or dealt with in a perfunctory manner in the models are now analysed more carefully, particularly with reference to their importance in the determination of future climatic scenarios. Nowadays, for example, the need is felt for a more thorough study of the consequences of the water cycle and of the use of the soil, particularly in relation to the indirect effects of atmospheric aerosols on clouds.

8.8 Further Remarks about Climate Change and Its Study

What we have presented in this chapter has made it possible at last to tackle the problem of the study of climate change by means of an instrument suitable for exploring the complexity of the Earth system. Some signs of the fact that certain influencing factors might turn out to be fundamental in determining climatic changes had already been pointed out starting from Chapter 3. Then the theoretical remarks in Chapter 4 contributed, at least partly, to the determination of the action of some of these single factors, considered individually. Only now, however, within a climatic model, we have been able to evaluate the complex, non-linear mixture of these diverse concauses of climatic changes. When this was being done, the anthropogenic influences on the evolution of the climate were revealed in all their importance: we saw, in particular, that these influences account for the net rise in the global temperature during the last thirty years.

Have we thus identified the increase in anthropogenic emissions of greenhouse gases as the main cause of the changes observed during the last few decades? Undoubtedly the physical interaction between these gases is well known, and their role as an important influence on the climate does not seem to be very controversial. Obviously, as we have seen, the elaboration of all-embracing climatic models is still a rather distant goal, since there still exist some processes that need to be understood better and evaluated more carefully in the models. This means that, in the future, the inclusion of new feedback mechanisms may slightly reduce the importance of this role of human activities. We should point out, however, that the main anthropogenic feedbacks are all positive: an example is the deforestation activity, which tends to eliminate absorbers of CO_2 and leads to a further increase in the concentration of CO_2 in the atmosphere, with the relative consequences on trapped heat and temperature.

During the last few years, some of these less-known mechanisms have been investigated better: we should mention the role played by the oscillations of ENSO on the interannual variability of the climate, or the more accurate analysis of the carbon cycle on land and in the oceans. At the present stage of climatic modelling researches, two points that have not been clarified completely yet are the role of clouds in the feedback cycles and how cloud formation is influenced by the indirect effects of changes in aerosol concentration.

In short, the system we are dealing with is complex, and this cannot be concealed. The future is quite likely to bring further progress in knowledge and simulation, and this may partly change our vision of climatic processes. However, what we know at present seems sufficient to allow us to assert that the influence of human activities on climate change is considerable, and that impact studies (which are consequent upon the forecast climatic scenarios, but will not be discussed in this book) show that it is necessary to act at once in order to avoid the most objectionable consequences of climatic changes, such as an excessive rise in sea level that might force the population to leave some coastal territories, with mass departures from the most threatened zones.

In this context, one of the efforts modellers should make is to improve the theoretical understanding of the instrument they are using, endeavouring to come as near as possible to the rigour of the science of complex systems. This would lead, in particular, to a greater awareness of the intrinsic qualities and limits of climatic models. In this perspective, for instance, modellers should delve further into the "chaoticity" of the system (i.e. the sensitivity of its future evolution to errors in the determination of its initial state), by analysing the zones of the state space where two neighbouring states suddenly diverge in an exponential manner (the so-called "bifurcation points"). While in meteorology — where the goal is a forecast, as deterministic as possible, of the evolution of a state in the future — these bifurcations lead to an actual unpredictability after a certain time lapse, in climate studies they lead the system to classes of possible states, whose knowledge determines the statistics of the system. It is reasonable, therefore, to assert that while in meteorology modellers "suffer" bifurcations, in climate studies they can "exploit" them, in order to determine more extensive climatic statistics over a certain time lapse.

In particular, the change in the frequency (and therefore probability) of the occurrence of certain classes of states in comparison with others may be due to a change in the external forcings: in the past, rather sudden climatic changes occurred several times (the reader is reminded of Plate 7, which shows the comparative swiftness of the transitions from glacial eras to interglacial periods). In a non-autonomous system like this, an analysis of the behaviour of the evolving trajectories (on a no longer static attractor) may allow us to understand these sudden transitions.

In the panorama we have thus outlined, the resort to ensemble integrations on the runs of an individual model is essential. At present, in actual fact, ensemble integrations are also beginning to be carried out starting not only from different initial states, but also from different models: this makes it possible to test the reliability of the results obtained and, in some cases, to increase the confidence that can be placed in these results. The ideal objective at which efforts would be aimed if this were possible is that of achieving a forecast of the distribution of probabilities for the states of the system in the future; but this is obviously impossible, because of the multidimensionality of the system. In this context, then, a result that appears to be attainable is an accurate determination of the classes (or scenarios) of possible climates under the influence of forcing factors and boundary conditions of the system. The key to all this is in the statistical analysis of the dynamical treatment carried out with climatic models.

After having emphasised that climatic models possess to the utmost degree the characteristics of simulation models that integrate the knowledge coming from different disciplines and make it possible to carry out an experimental activity of construction of possible future worlds (scenarios), at this point we would like to conclude this chapter with a few more remarks of a conceptual rather than practical nature.

As we have already stated, models that simulate the behaviour of complex systems do not have the purpose of corroborating or falsifying the equations of the model: this activity is left to laboratory experiments in controlled and simplified conditions. The purpose of these models is to reconstruct the complexity of reality and to simulate phenomena and processes on a macroscopic scale. As concerns us, then, what does validating a climatic model mean? Is it possible to falsify a model of this type?

Whereas in the realm of experimental laboratory activity there is the possibility of carrying out crucial experiments that — through a comparison with the behaviour of the real system — make it possible, for instance, to draw a distinction between alternative theories or models, in numerical experimentation, tests can be carried out in order to evaluate the physical consistency of several models in extreme situations, though these situations do not usually appear in nature.

In the validation of climatic models, in particular, besides checking the physical self-consistency of the scheme, an attempt is usually made to understand the nature of the interactions in the model. The comparison with the data observed in reality, however, is usually carried out in a "filtered" manner, that is, for instance, through a comparison with the meteorological and climatic analyses, which, as we have explained in the previous chapter, stem from a combination of data and results of forecasting models. So there remains a difficulty in falsifying a model in the "Popperian" sense of the term[18].

The situation, therefore, is not as clear as in real laboratory experimentation, which, however, is possible in very simplified conditions and for studying individual phenomena or processes. So the example of the study of the climate in a complex system like that of the Earth is also the paradigm of a new approach to science — a science that

---

[18] The reader is reminded that a theory or model is falsified if even only one particular situation is found in which the behaviour prescriptions stemming from the theory or model clash with the data obtained from natural observations or laboratory experiments. On this subject, consult Popper (2002), where the possibility of falsification is actually used as a criterion for a boundary between science and non-science.

probably would not be regarded as science by Popper: see note 18. In particular, the choice among the various models cannot be made simply on the basis of the possibility of their falsification.

This last remark encourages us even more to feel that, with the analysis of climatic models, we have reached the threshold of a modern science, where, in order to understand the complexity of the natural environment, we will inevitably be forced to relinquish a classical approach to science.

Chapter 9

Conclusions and Prospects

Now that we have reached the end of this journey in the complex world of the atmosphere and Earth system, it is expedient for us to take stock of what we have discovered and to evaluate some prospects of the future development of scientific knowledge in this field.

The route followed in this book showed the complexity of the natural environment in which we live, at least as regards the part relative to meteorological and climatic phenomena and processes. At the same time, we became aware of the diversity of the systems under consideration, in comparison with the ones we were able to deal with in a school laboratory and whose operation we could easily understand, often thanks to the application of the Galilean experimental method. Understanding the behaviour of the atmosphere and of the Earth system required a breakthrough similar to the one that took place in the seventeenth century with the adoption of the method devised by Galilei: today, simulation methods are the key for compounding phenomena and elementary processes, and for thus reconstructing the complexity of reality in the virtual, controllable world of computers.

In our domain, we have seen that the application of meteorological and climatic models has led to immensely important practical results, in particular to forecasts of considerable climatic changes over the next hundred years.

Moreover, in this *excursus* on the science of weather and climate we made several forays into the domain of theoretical physics. In doing so, we had the chance to recognise, for instance, that the concept of univocal deterministic forecast can no longer exist for a complex system, and that a probabilistic description has rightfully come into the (once unchallenged) realm of determinism. An important concept like that of attractor was similarly added to the statistical definition of a climate, and turned out to be extremely useful for the determination of the latter. On the whole, besides changing scientific practice in a substantial way, simulation models (particularly meteorological and climatic ones) also contributed to change our outlook on nature.

In this book, we have endeavoured to outline a conceptually meaningful vision of the scientific approach to the study of the weather and climate. In doing so, on the one hand we explained some scientific findings and pieces of knowledge, while on the other hand we discussed conceptual and epistemological issues[1] relative to the nature of the systems under examination, to the methods used for studying them, and, ultimately, to problems inherent to the intelligibility of nature.

At this point, in order to complete the picture, we should perhaps present some conclusions and discuss the prospects of future development with reference to two important points: the results of climatic models, with the consequent actions that can be undertaken; and the present and future developments of the modelling paradigm applied to the study of the weather and the climate.

9.1 The Results of Climatic Models and "What Should We Do?"

In the previous chapter, we discussed climatic models, and, in particular, presented the basic results attained through their application to the forecasting of future climatic scenarios. Without going again into the details of these results, we may mention the following: the consistency with which these models forecast a more or less marked rise in the global temperature over the next hundred years; the consequences, on a continental or regional scale, of the possibility of an increase in extreme events (ranging from prolonged drought to conditions of heavy precipitation); and the rise in the sea level.

We already stated that the AOGCMs are still in the development stage and do not yet consider all the processes that act in the Earth

---

[1] See Note 5 in the introductory Chapter 1.

system: this will be briefly discussed in the second part of the present chapter. However, AOGCMs show the influence of human activities on the climate[2]. In this perspective, does anyone intend to do anything? In particular, since the problem is a global one, has an international negotiation been activated?

The answer to these questions is obviously "yes". Under the aegis of the United Nations, in December 1997 a document was drawn up in Japan, the (by now well-known) Kyoto Protocol, which establishes an international-level commitment for the reduction of some greenhouse gases, including CO_2. The rationale behind this document obviously stems from the acknowledgement of the contribution of these gases to the global warming and of the consequent need to limit their emission, compatibly with the development requirements of the poorest countries in the world. In practice, once the increase in greenhouse gases has been identified as a cause of climatic changes, an effort is made to stop or slow down the warming process by acting on this cause: it is an attempt at a "mitigation" of the effects.

As a matter of fact, the actual situation of the international negotiation has reached a deadlock: the Kyoto Protocol has come into force on 16 February 2005, after its ratification by Russia. With the entry of this country, the last of the thresholds laid down for the Protocol to become legally valid has been passed: now the sum of the emissions of all the countries that have ratified the Protocol is more than 55% of the total emissions detected in the reference year 1990. However, it seems that the stance of the Bush government against the ratification of the Protocol by the United States is thwarting the international efforts: the United States are currently emitting more than one fourth of the annual quantity of greenhouse gases emitted in all the world. We must also add that many scientists consider the Kyoto Protocol important more as an expression of international will than as an actual measure for mitigating the phenomenon of global warming: it has been pointed out by several experts that more drastic measures would be necessary.

---

[2] This is underscored not only by the last report of the IPCC (Houghton *et al.* (2001)), but also by a more recent account of the National Academy of Sciences of the U.S.A. for President Bush (see National Academy of Sciences (2001)).

In this situation, some researchers propose engineering solutions for confining carbon dioxide underground or in the oceans: for the time being, these solutions are not very feasible or safe, and are financially too exacting. The American government, on their part, have initiated a vast programme of studies on the impact both of anthropic activities and natural influences on climatic changes, maybe in the secret hope of demonstrating that natural elements and not mankind are the main causes of global warming. If this were true, the present pattern of economical development might be able to grow further without the hindrance of negative effects on the ecosystem, at least as regards the climate.

What we have explained in the previous paragraph essentially allows us to understand that human activities lead to an increase in greenhouse gases and therefore promote a global-scale rise in temperature, creating feedbacks that are all intrinsically positive. Undoubtedly the magnitude of this effect in comparison with that of other changes in natural forcings is open to question, but a shred of common sense and a "precautionary principle" (advocated by many people) should lead us in any case to reduce the extent of these anthropogenic causes, also because the Earth system, being highly non-linear, does not always respond gradually to changes in forcing factors.

From this angle, the "wait and study" tactics of the American government causes perplexity. Many people feel that, instead of spending energies and resources in studies for mitigation projects that might turn out to be practically unfeasible, it would be better to concentrate on "adaptation" studies, i.e. on the evaluation of how to reduce the impacts of climatic changes (by now considered inevitable) on the territory and population. In actual fact, now international efforts are focused particularly on studies of this type.

This subject, of course, is partly beyond the scope of this book. It is interesting, however, to discuss it, and to show, for instance, that there is a connection between political stances, with their models of socio-economic development, and scientific climatic activity, which, like all sectors of modern science, increasingly depends on the financing of research activities.

9.2 The Future of Models for Studying the Weather and Climate

Returning to the main theme of this book and to the modelling approach we have presented for analysing and forecasting the meteorological and climatic behaviour of the Earth system, we can now concisely summarise the importance of this paradigm, highlighting its strong points and weaknesses, and then proceed to discuss the prospects of future development of research in this field.

From a theoretical point of view, we have repeatedly emphasised the conceptual importance of the simulation paradigm applied to systems such as the atmosphere or Earth system, its capability to reconstruct reality in a virtual, controllable laboratory, and all the advantages that can be associated to this in terms of possibilities of experimentation for understanding complex phenomena and processes.

From a practical point of view, as regards the study of the weather, we have seen that meteorological models are the only instrument we have for obtaining detailed forecasts beyond 24 hours. Their deterministic structure, combined with the use of ensemble integrations, allows us to evaluate their accuracy and reliability at the various forecasting time limits. The drawbacks of these models are revealed in the following circumstances:

- in long-range forecasts, when the theoretical predictability limit (which is intrinsic to a non-linear model and marks the onset of deterministic chaos) is reached;
- in very short-range forecasts, when the problems due to the initial unbalancing of the fields (the previously mentioned spin-up effect) and the long computer time required for the analysis and the model do not make it possible to obtain reliable forecasts that can be used within an operational time;
- on a local scale, because of the limited resolution of the models[3];
- when there is the need to obtain forecasts of some variables that the model does not supply directly, such as the meteorological visibility (important, for instance, for determining fog situations).

---

[3] For instance, this problem is particularly felt in aeronautics, where forecasts on a particular site, i.e. the runway of an airport, are required.

As regards climatic models, we would like to emphasise that in this sphere we appreciate, even more, the capability of the simulation method to interconnect phenomena and processes that arise within different, mutually interacting systems. From this point of view, it is obvious that the greatest limit that can be found at present in climatic models is precisely that of being still rather far from having completed the reconstruction of the real system: climatic models currently include in their schemes only the interactions that are considered most important. Other limits are obviously due to the low horizontal resolution of these models: this problem has been only partly solved by the regionalisation techniques.

Considering the excellent results obtained with the application of the simulation method to weather forecasts and the study of the climate, the meteorological and climatic research centres scattered all over the world persist in applying the simulation paradigm, in quest of constant improvements, which also involve the elimination of the drawbacks we have listed above. An example that refers to meteorology is the elaboration of coupled atmosphere-ocean models, which, combined with the use of ensemble integrations, is now leading to an extension of the predictability time lapse, to the point of obtaining seasonal forecasts in terms of scenarios for some quantities such as temperature and precipitation. Climatically important examples are the attention that is being given of late to previously-neglected interactions, such as those relative to the influence of cosmic rays and solar variability on cloud formation (an attempt is being made to include them in climatic models), and the development of projects (such as the Japanese Earth Simulator) in which new-generation supercomputers and high-resolution models, as all-embracing as possible, will be used for a more complete study of the Earth system.

In this attempt at working out increasingly complex and sophisticated models, the scientific enterprise that involves meteorology and climate science has acquired the standing of a "Big Science". This term indicates a science whose development requires very substantial funding (sometimes unaffordable for a single country), which in our case is used basically for managing major computer centres and for supporting large teams of researchers and technicians.

In my opinion, when a level of applicative development like this is reached, there begins a process of science policy and sociology whereby the same type of research tends to be self-nourished in the future. To explain this more clearly: since supporting these centres for the researching and operational development of models is quite expensive, the people who finance them and those who manage them must guarantee a constantly improving final product; therefore, once a paradigm has been found that works (and ensures a constant, though maybe slight, progress), it is unlikely that this paradigm will be abandoned in order to investigate less-explored paths. Moreover, for the individual researchers (whose career is linked to the number, and, partly, to the quality of their publications on international scientific journals) it is easier to publish researches performed with standard and extensively accepted methods. This fuels what Thomas S. Kuhn calls "normal science"[4], to the detriment of investigations that are more innovative but are also more at risk of being unsuccessful.

I would like, obviously, to prevent any misunderstanding: the simulation paradigm we have discussed has so many important positive aspects, both from the conceptual point of view and the applicative one, that the constant improvement of meteorological and climatic models can only be looked on with favour. The effect we have just described does nothing but further amplify the success of these models.

In spite of this, some attempts have been made recently to tackle certain meteorological or climatic problems outside this paradigm, and, in actual fact, some techniques for bridging certain gaps in these models are emerging. In many cases, the authors of these works are researchers who do not belong to big research centres but enjoy a greater scientific independence and can explore more original, though more risky, paths[5].

---

[4] Kuhn (1996).

[5] For instance, most Italian scientists, including the author, are in this situation. The fact that the persistent shortage of funds in the Italian world of research does not make it frequently possible to compete with big foreign research groups leads the individual scientists to deal with original subjects or methods, which allow them to carry out high-level (and perhaps more creative) researches despite the fact that their resources are scanty.

In actual fact, there exist some domains of meteorology and climatology where the models we have described are in difficulty at present, and presumably will be so in the future as well. I refer, for example, to very short-range local meteorological forecasts. Moreover, some phenomena of remote correlation, the so-called "teleconnections" (for instance, between the NAO index in the Atlantic Ocean and the winter temperatures and precipitation in Europe), suggest that there exist some macroscopic variables that determine the climate on a certain zone of the globe.

On the other hand, the method for the reconstruction of reality in a simulation model is extremely sensitive if it is applied to a non-linear system characterised by a great number of feedbacks: we have already mentioned the somewhat artisanal activity with which the various parameterisation schemes are balanced. This set of problems cannot be solved in a simple way, and leads to uncertainties that are often hard to quantify.

All these considerations have led some researchers to choose a different path from that of the classical models we have described: a path that does not presume to replace these models, but endeavours to bridge their gaps and to catch the evolution of the system under examination in a more direct manner, without getting "bogged down" in the meanders of the balancing of the various parameterisation schemes and of the many feedback cycles. For instance, these researchers use artificial-intelligence methods such as neural network models, non-linear artificial systems that mimic some simple functions of the human brain[6].

With these models, the simulation paradigm is abandoned: here there undoubtedly exist some elements that have a one-to-one correspondence with the real system (for instance, the meteorological variables), but they are not what evolves with time. Learning on the basis of previous experiences supplied by the researcher, and readjusting the connections between their own artificial neurons, these models manage to find non-

---

[6] Some examples of applications of this type can be found in regionalisation techniques, in very short-range forecasts (see, e.g., Pasini *et al.* (2001)), in the ENSO forecast (see, e.g., Tangang *et al.* (1998)), and in the analysis of climatological data (see Pasini *et al.* (2005)).

linear correlation laws among the various states of the real system and to account for its characteristics (and possibly also for its evolution).

This opens the way to a different chapter in the history of modelling, no less fascinating than the one we have dealt with here. For instance, imagine a little artificial brain that finds diagnostic or evolutionary laws in cases that are so complex as to remain unintelligible to us even after the application of our theoretical knowledge of the problem[7]. At the same time, the neural network model may also indicate the fundamental variables for a dynamical description of the real system under examination.

At this point, however, the subject becomes excessively vast, and I should start writing another book. Perhaps I may manage to do so some day.

---

[7] A case of this type is described in Pasini and Ameli (2003).

Bibliography

Alley, R.B. (2002). *The Two-Mile Time Machine: Ices Cores, Abrupt Climate Change, and Our Future*, Princeton University Press.

Arrhenius, S.A. (1896). On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground, *Phil. Mag.*, 41, pp. 237-276.

Barrow, J.D. (1988). *The World Within the World*, Oxford University Press.

Barrow, J.D. (1991). *Theories of Everything: The Quest for Ultimate Explanation*, Clarendon Press.

Barrow, J.D. and Tipler, F.J. (1986). *The Anthropic Cosmological Principle*, Oxford University Press.

Bateson, G. (1980). *Mind and Nature. A Necessary Unity*, Bantam Books.

Buizza, R. (2001). Chaos and Weather Prediction — A Review of Recent Advances in Numerical Weather Prediction: Ensemble Forecasting and Adaptive Observation Targeting, *Il Nuovo Cimento*, 24C, pp. 273-301.

Einstein, A. and Infeld, L. (1967). *Evolution of Physics*, Free Press.

Fermi, E. (1937). *Termodynamics*, Dover Publications.

Galilei, G. (1967). *Dialogue Concerning the Two Chief World Systems: Ptolemaic and Copernican*, University of California Press.

Ghirardi, G.C. (2003). *Sneaking a Look at God's Cards: Unraveling the Mysteries of Quantum Mechanics*, Princeton University Press.

Greco, J. (1998). Introduction: What Is Epistemology?, *The Blackwell Guide to Epistemology* (J. Greco and E. Sosa eds.), Blackwell.

Hamblyn, R. (2001). *The Invention of Clouds: How an Amateur Meteorologist Forged the Language of the Skies*, Farrar Straus Giraux.

Hawking, S. (2002). *The Theory of Everything: The Origin and Fate of the Universe*, New Millennium Press.

Hodell, D.A., Curtis, J.H. and Brenner, M. (1995). Possible Role of Climate in the Collapse of the Classic Maya Civilization, *Nature*, 375, pp. 391-394.

Hood, B.M. (2004). Children's Understanding of the Physical World, *The Oxford Companion to the Mind* (R.L. Gregory ed.), 2nd edition, Oxford University Press.

Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K. And Johnson, C.A. (eds.) (2001). *Climate Change 2001: The Scientific Basis*, Cambridge University Press.

Krishnamurti, T.N. and Bounoua, L. (1996). *An Introduction to Numerical Weather Prediction Techniques*, CRC Press.

Kuhn, T.S. (1996). *The Structure of Scientific Revolutions*, 3rd edition, University of Chicago Press.

Laplace, P.S. (1820). *Theorie Analitique des Probabilités*, Courcier.

Lionello, P. (2005). *Oceans: Their Motions and Role in Climate*, Foxwell and Davies.

Lorenz, E.N. (1994). *The Essence of Chaos*, University of Washington Press.

Mandelbrot, B. (1982). *The Fractal Geometry of Nature*, Freeman.

National Academy of Sciences (2001). *Climate Change Science: An Analysis of Some Key Questions*, National Academy Press.

Pasini, A. and Ameli, F. (2003). Radon Short Range Forecasting through Time Series Preprocessing and Neural Network Modeling, *Geophys. Res. Lett.*, 30 (7), 1386.

Pasini, A., Lorè, M. and Ameli, F. (2005). Neural Network Modelling for the Analysis of Forcings/Temperatures Relationships at Different Scales in the Climate System, *Ecol. Mod.* (in press).

Pasini, A. and Pelino, V. (2000). A Unified View of Kolmogorov and Lorenz Systems, *Phys. Lett.*, A275, pp. 435-446.

Pasini, A., Pelino, V. and Potestà, S. (1997). Evidence of Structured Brownian Dynamics from Temperature Time Series Analysis, *Nonlin. Proc. Geophys.*, 4, pp. 251-254.

Pasini, A., Pelino, V. and Potestà, S. (2001). A Neural Network Model for Visibility Nowcasting from Surface Observations: Results and Sensitivity to Physical Input Variables, *J. Geophys. Res.*, 106 (D14), pp. 14,951-14,959.

Pease, C.B. (1994). *Satellite Imaging Instruments*, Wiley & Sons.

Petit, J.R. *et al.* (1999). Climate and Atmospheric History of the Past 420,000 Years from the Vostok Ice Core, Antarctica, *Nature*, 399, pp. 429-436.

Philander, S.G. (2004). *Our Affair with El Niño: How We Transformed an Enchanting Peruvian Current into a Global Climate Hazard*, Princeton University Press.

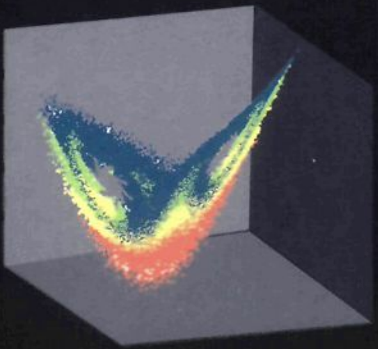
Popper, K.R. (2002). *The Logic of Scientific Discovery*, 15th edition, Routledge.

Tangang, F.T., Tang, B., Monahan, A.H. and Hsieh, W.W. (1998). Forecasting ENSO Events: A Neural Network-Extended EOF Approach, *J. Clim.*, 11, pp. 29-41.

Wood, R.A., Keen, A.B., Mitchell, J.F.B. and Gregory, J.M. (1999). Changing Spatial Structure of the Thermohaline Circulation in Response to Atmospheric CO_2 Forcing in a Climate Model, *Nature*, 399, pp. 572-575.

Yourgrau, W. and Mandelstam, S. (1979). *Variational Principles in Dynamics and Quantum Theory*, Dover Publications.

Index



FROM OBSERVATIONS TO SIMULATIONS

From Observations to Simulations leads us on a fascinating journey through the methods used for the scientific analysis of complex systems such as the atmosphere and the Earth system: from meteorology and climatology, as observational sciences, to the development of models and the use of computers as virtual laboratories. In plain, accessible language, avoiding technicalities, but highlighting the conceptually meaningful aspects, the book describes this "Copernican revolution" in meteorology and climatology, a change in methodological paradigm that rigorously tests the definition of some classical concepts, such as "causality" and "prediction." This is the first book that guides the general public (and sets the specialists thinking) through research on complex systems which is contributing to a change in our outlook on nature.

World Scientific
www.worldscientific.com
5930 hc

SEPS
SEGRETARIATO EUROPEO PER LE PUBBLICAZIONI SCIENTIFICHE

ISBN 981-256-475-6
9 789812 564757