CHAPTER I

INTRODUCTION

1.1. The discovery of the stratosphere

Towards the end of the last century, meteorologists were actively investigating the structure of the atmosphere away from the earth's surface by using kites and balloons carrying thermometers and pressure gauges. It was normally found that the temperature decreased with height at a rate of 5–10° K./km. These results fit in well with the idea of an atmosphere well stirred by winds, since violent mixing in a dry atmosphere should lead to a lapse rate† of 10° K./km. (the dry adiabatic lapse rate). Although the average observed lapse rate is less than the dry adiabatic, the condensation of water vapour would modify the adiabatic lapse rate in just this direction. As a result there was, at this time, a comfortable feeling among meteorologists that there was nothing more to be discovered. Crossley (1934) has attempted to analyse this attitude and has given three reasons why, in those days, it was confidently expected that the temperature of the atmosphere would always decrease with height:

'(*a*) As the temperature of the base of the atmosphere is that of the surface of the earth, and as the outer limit of the atmosphere must approach the absolute zero of temperature, the air temperature must, on the whole, decrease from the surface outwards.

'(*b*) The temperature is observed to decrease with height up to 10 km. or more according to latitude.

'(*c*) There is also a feeling that the temperature should decrease with height because the pressure does.'

Needless to say, these arguments do not bear close examination, but nevertheless the ideas behind them must have had general acceptance.

Despite this apparently satisfactory position the French investigator Teisserenc de Bort started in 1898 to use balloons in an

† According to normal meteorological terminology a 'lapse' is a negative vertical temperature gradient, i.e. temperature decreasing with height. Similarly, an 'inversion' is a positive temperature gradient.

attempt to obtain reliable temperature measurements up to 14 km. Like other workers at this time, he was greatly concerned with the errors of observation, which seemed to fall under three main headings:

(*a*) Errors of pressure, and therefore height, measurement.

(*b*) Lag of the thermometers.

(*c*) Radiation errors, caused by the direct absorption of solar radiation at great heights where the ventilation of the thermometers was inadequate.

It was possible to make estimates of the errors arising from (*a*) and (*b*), but (*c*) was indeterminate, although it was known that it could be large. When it appeared to be indicated in 1898 that above 11 km. the temperature of the atmosphere during daytime ceased to decrease with height and became constant, Teisserenc de Bort naturally concluded that he was faced with radiation errors. Experiments in 1899 showed, however, that this 'isothermal layer', as he named it, still existed at night-time, and two years later enough evidence had been collected for him to be able to announce his discovery as clearly established. By 1904 Teisserenc de Bort was able to publish the results of 581 ascents, 141 of which went to 14 km., and to distinguish some general properties of the layer, e.g. that it is low over depressions but high over anti-cyclones. Fig. 1 shows seasonal average curves constructed from these early results.

Thus, in 1904, the discovery of the isothermal layer had been firmly consolidated, although a number of years were to pass before it was generally accepted. The care with which Teisserenc de Bort unravelled a major discovery from the errors of a very difficult experiment, by means of careful and frequent measurements, makes this research one of the finest in the history of meteorology.

The first man to enter the stratosphere for certain, and to survive, was the Belgian Piccard who ascended to 16 km. in 1931. He might, however, have been preceded. On 5 September 1862 two Englishmen, Glaisher and Coxwell, claimed to have reached a height of 11·3 km. in an open gondola without oxygen apparatus. This may have taken them into the stratosphere, but even if it did, the aeronauts were scarcely fit to make observations, since

Glaisher was unconscious and Coxwell so badly frostbitten that he had to open the release valve with his teeth. This was only one of a long series of ascents which Glaisher and Coxwell made in their 93,000 cu.ft. capacity 'Mammoth' balloon.

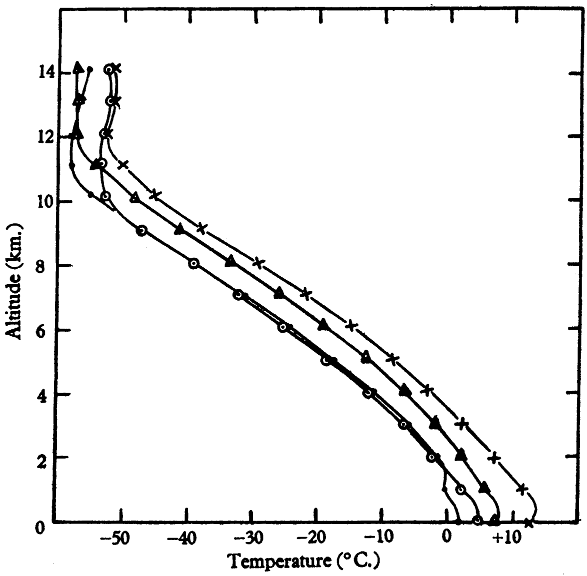

Fig. 1. The first 141 temperature records up to 14 km., grouped according to season. (After Teisserenc de Bort.) (N.B. These do not necessarily correspond exactly with the best modern measurements.) —⊙— spring; —×— summer; —△— autumn; —•— winter.

1.2. Nomenclature

It was soon discovered that the 'isothermal layer' is hardly ever exactly isothermal, and the name dropped out of general use after a decade. The most striking property of the layer was then recognized to be its great hydrostatic stability, suggesting that it might be stratified as opposed to mixed (this being the definition of the meteorological term 'stratification'). This led Teisserenc de Bort himself to suggest the word 'stratosphere' for his discovery and the word 'troposphere' for the underlying atmosphere.† Since much interest would obviously attach itself to the boundary between these regions, and since this boundary appeared to be a

† Greek *tropos* = turn; troposphere = turning or mixing sphere.

rather definite feature, it was not long before a name was found for it, the 'tropopause', suggested by Hawke and popularized by Napier Shaw.

The problem of upper-atmosphere nomenclature is now one of importance and difficulty. The definition of terms depends upon the definition of measurable features, but unfortunately most upper-atmosphere measurements are made on the limits of observation and frequently can be interpreted in more than one way. Moreover, the atmosphere is not in a steady state and measurements will never be exactly repeatable. Finally, such measurements as exist—even the voluminous records of meteorological services—are at the best only a few spot measurements in a complicated three-dimensional field. Under these circumstances care has to be exercised in the naming of a feature, since merely giving a name may imply a physical significance which later turns out to be illusory.

The difficulties of definition stem from the many different types of feature which may be measured and therefore named: thermal structure, dynamical factors (wind, etc.), concentration of major constituents (oxygen, nitrogen), concentration of minor constituents (water vapour, ozone), visible phenomena (water and ice clouds, aurorae, nightglow emissions) and electron density. There is little reason to expect detailed correspondence between these features, and fortunately there is no reason to give them all equal weight. The cornerstone of studies in atmospheric physics will always be the transformations of energy which take place, and the dynamical and thermal structures of the atmosphere are most indicative of these transformations. Dynamical factors, however, such as wind, are far too variable to be of use in general discussion, and the thermal structure is therefore the obvious choice for the purposes of nomenclature.

In fig. 2 is shown the approximate temperature distribution in the atmosphere up to 120 km. in temperate latitudes; methods of measurement will be discussed in the next chapter. Nearest to the earth's surface is the layer [*a*], where temperature decreases with height at about 6·5° K./km.; this is universally known as the 'troposphere'. The troposphere ends at the 'tropopause', [1], or if doubts exist as to whether this really is a well-defined feature the term 'tropopause layer' may be introduced. We will adhere to the use of the word 'tropopause', implying thereby a belief that a rather sudden change of temperature gradient at this point is the normal state of the atmosphere. The word 'stratosphere' was originally introduced to indicate the approximately isothermal region above the tropopause. We shall use the word in a wider sense, to mean that part of the atmosphere between the tropopause

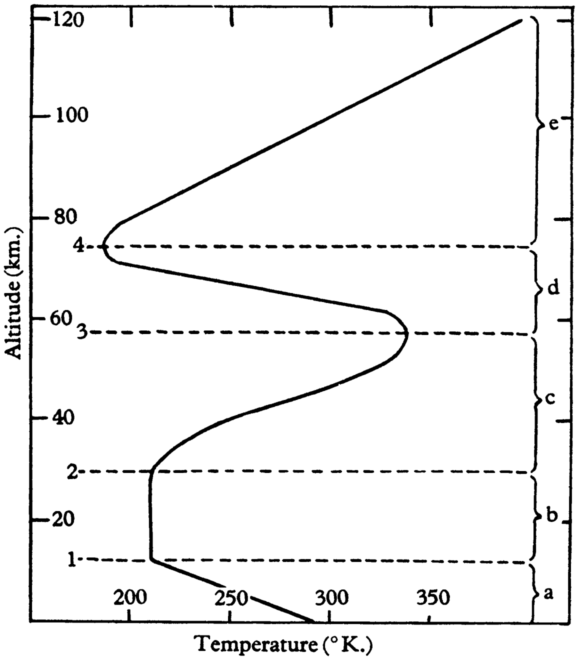

Fig. 2. Approximate vertical temperature distribution in the atmosphere in temperate latitudes.

and the inversion which starts near 75 km., [4]. By analogy with the tropopause, the feature [4] has been called the 'stratopause'. The stratosphere can conveniently be divided into the approximately isothermal region [*b*], or 'lower stratosphere', and the rest [*c*+*d*], or the 'upper stratosphere'. Above the stratopause lies a region whose thermal properties are not well understood, and which has mainly been investigated by means of radio techniques which have established the existence of large electron densities. It seems reasonable, therefore, that at present this region should be named after the various ionized layers which have been discovered (viz. D, E, F_1 and F_2), and that the whole region up to at least 400 km. should be called the 'ionosphere'. It has also been suggested that the upper layers of the atmosphere above 600 km., where collisions are extremely rare, should be called the 'exosphere'.

In Table I the nomenclature adopted here is compared with suggestions which have been made by Chapman (1950) and Flöhn and Penndorf (1950). Chapman gives three forms, based upon thermal properties, homogeneity (the atmosphere is believed to be thoroughly mixed below the stratopause, see Chapter III)

TABLE I. *Atmospheric nomenclature*

Reference	This book	Flöhn and Penndorf	Chapman I	Chapman II	Chapman III
a	Troposphere	Advection layer	Troposphere	Homosphere	Neutrosphere
1	Tropopause	Tropopause layer	Tropopause		
b	Lower (Stratosphere)	Isothermal layer (Stratosphere)	Stratosphere		
2			Stratopause		
c	Upper (Stratosphere)	Warm layer (Stratosphere)	Meso-incline (Mesosphere)		
3		Ozonopause (Stratosphere)	Meso-peak (Mesosphere)		
d		Upper mixing layer (Stratosphere)	Meso-decline (Mesosphere)		
4	Stratopause	Upper tropopause	Mesopause	Homopause	Neutropause
e	Ionosphere	Ionosphere	Thermosphere	Heterosphere	Ionosphere

and electrical properties. He distinguishes therefore between a 'mesopause', a 'homopause' and a 'neutropause'; but since experiment cannot distinguish clearly between the heights of these features there seems little object at present in introducing this degree of complication.

The nomenclature of Flöhn and Penndorf, although less complicated than Chapman's, is not very appropriate. The 'isothermal layer', as we have seen, is hardly ever isothermal. The 'warm layer' includes only half of the feature which might deserve this name. Finally, experiment indicates slow mixing throughout the stratosphere, and therefore the name 'upper mixing layer' conveys no useful meaning.

It is not suggested that the nomenclature adopted in this monograph has any particular physical significance.† Its main merits are that it introduces no unusual terms, so that those familiar with

† In fact Chapman's use of the word 'stratosphere' is correct physically.

the literature will recognize the meanings of terms without formal definition, and that it conveys no more than the experimental evidence justifies.

1.3. Scope of the monograph

In the following chapters we will discuss briefly the main physical properties of the stratosphere, only considering the troposphere and the ionosphere where it is essential to preserve continuity. Work upon this region of the atmosphere has mainly been concerned with measurements and calculation of the temperature, the composition, the dynamical structure and the radiation balance. With regard to temperature, we require to know not only the general vertical structure but also the variations which take place daily, seasonally, and from one latitude to another, for such information gives a valuable guide to underlying processes. Atmospheric composition is of interest from many points of view: most methods of measuring temperature depend upon a knowledge of the mean molecular mass of air; the extent to which gases separate out under gravity indicates the degree of mixing which is taking place; the radiation balance at levels below the stratopause, which creates the sources and sinks of energy necessary to maintain the atmospheric motions, is almost entirely controlled by the concentrations of three minor atmospheric constituents, viz. ozone, water vapour and carbon dioxide, which present many interesting problems of measurement; ozone is created and destroyed *in situ*, and its concentration is so intimately bound up with atmospheric conditions that the study of this gas is almost a branch of geophysics in its own right. Finally, the dynamical structure of the atmosphere is one of the clearest indications of the transformations of energy which continually take place, and the theory of radiation transfer is required to link together such studies with information upon temperatures and composition.

By limiting discussion to the stratosphere we shall not be considering many important atmospheric phenomena. As far as is known, a number of changes take place just above the stratopause. The atmosphere becomes ionized and therefore conducting, and there have been many interesting studies upon the reflexion of radio waves and short-period variations of the earth's magnetic field. Also, approximately at this level, most molecules except nitrogen begin to be dissociated by the ultra-violet radiation from the sun, leading to uncertainties in our knowledge of the mean molecular mass for air, and therefore to uncertainties in temperature measurement (see Chapter II). Further, atomic and molecular mean free paths become so great that diffusive transfer begins to predominate over mixing processes. Aurorae practically never penetrate into the stratosphere, and although some night-glow emissions are now considered to be located near 70 km., they are still generally associated with higher levels.

All these extremely interesting topics will be only briefly mentioned where they lead to information about the atmosphere near the stratopause. The reader who is interested in these problems is, however, fortunate in the recent publication of a number of books and reviews dealing with them, which are listed at the end of this monograph.

Going to the other extreme, the stratosphere is just above the levels involved in the important and complicated phenomena of weather. There is a tendency now for meteorologists to include the lower stratosphere in their dynamical considerations, although marked reactions of stratospheric conditions upon weather have yet to be demonstrated. The extremely important subject of cloud and rain formation is almost entirely restricted to the troposphere, though the highest layers of cirrus and cumulo-nimbus lie only just below the tropopause.

The remaining portion of the atmosphere, which will be described in this monograph, is neither so close that mistakes in the computation of minor perturbations are a subject for discussion in the press and on the radio, nor yet so remote that it is necessary to work with techniques akin to astronomy. It is a region whose main features can be measured and explained by fairly straightforward physical methods, and it is the application of these methods to a field problem of great complication which will be the main concern of this book.

Finally, there will only be passing reference to the important subject of atmospheric oscillations, not because it is inappropriate, but because it has already been discussed in a monograph in this series (Wilkes, 1949).

1.4. Tools of research

Much of our knowledge of the upper atmosphere is based upon measurements made at ground level, and owing to the care and accuracy possible in the laboratory such measurements will always be of great importance. There is little in common between various ground-level experiments, except perhaps that most make use of some section of the electromagnetic spectrum, and they all need to be described individually in their context in later chapters.

In recent years it has become more and more common to attempt to make direct measurements by transporting instruments into the stratosphere on various types of vehicle. The limitations of such experiments are often imposed by the characteristics of the vehicle itself, and therefore it is worth while to discuss the important features of the three possibilities available, viz. balloons, aircraft and rockets.

1.41. *Balloons*

The first attempts to make measurements of the upper atmosphere employed kites and balloons. Balloons soon proved their superiority by ascending to much greater heights. The techniques of manned and unmanned balloon ascents differ very greatly, and the former have now fallen into disuse, though in the past much useful information has been gained by this method of investigation.

The main characteristic of the manned balloon ascent is the great weight that has to be carried. The observer is a considerable load, particularly since he must be protected against the low pressures of altitudes greater than 10 km., and such ascents are justifiable only if a quantity of apparatus is carried. The last great ascent was by the American helium-filled Explorer II (1938). It reached a record height of 22 km. In order to lift the weights involved, the capacity of the balloon was 3,750,000 cu.ft., although only a small fraction of this could be filled with helium at ground level since the pressure at the ceiling was only $\frac{1}{28}$ atmosphere, and a really elastic balloon is not acceptable for a manned ascent. The Explorer II had a spherical aluminium gondola in which three men could be carried. The total flight time was 8 hours 13 minutes, during which time many observations were taken, including cosmic-ray intensities at various zenith angles, the height distribution of ozone by spectroscopic methods, atmospheric ionization, the composition of the air with particular reference to oxygen, carbon dioxide and helium, the presence of spores and moulds, and air photographs for geological purposes.

Manned balloon ascents are unlikely to be of future service to research. They are difficult and costly to organize, considerable danger to human life is involved (one Russian and one American ascent have ended in disaster), and the advance of techniques has now made it possible to obtain all important results with much cheaper unmanned balloons.

The main difficulties with unmanned balloons are the light, reliable and automatic equipment required and the recovery of data after the flight. Data may be obtained either by the recovery of the whole balloon or by the radio transmission of instrument readings to the ground. The radio sonde, now universally used for routine observations, employs this second method. For research purposes the apparatus may be too complicated for this technique to be used, and the usual procedure is to attempt to recover the balloon and the attached equipment. In this country, with its dense population, a very good recovery rate, as high as 80 %, may be obtained by attaching to the gondola a label with the promise of a small reward, and since the other 20 % are lost partly in the sea or other inaccessible areas, the choice of favourable occasions can lead to an even better recovery rate. Alternatively, the study of upper winds and the preliminary release of a pilot balloon, together with tracking of the main balloon by telescope and radar, can give a good estimate of the point of landing, where the investigator may reasonably hope to proceed by car with a pair of binoculars and thus be present at the time of landing.

Before the last war, rubber balloons were in general use. The maximum height attained for certain with such balloons was about 30 km., but such heights were only rarely reached, 20 km. being more usual. Long experience is required to obtain good results with reasonably cheap balloons, and much of the progress has been made by a few individual investigators—for example, E. Regener, of Stuttgart, who has described some of his earlier experiences in an interesting article (Regener, 1935), to which the