

# 1

---

## Introduction

One of the worst snowstorms in living memory struck the U.S. mid-Atlantic states during 18–19 February 1979. Total snowfall amounts exceeded 0.5 meter (m), and accompanying high winds created drifts of up to 2.5 m. Airports and roads were closed from Atlanta to New York City, and a state of emergency was declared in many areas. This legendary storm, often referred to as the “Presidents’ Day snowstorm,” was notable not only for its severity but also because it was exceptionally poorly forecast.

Media dissemination of weather forecasts to the public is the final step in a complex process. The first step is to collect all the atmospheric observations from the entire globe for a given time. Second, these observations are diagnosed or analyzed to produce a regular, coherent spatial representation of the atmosphere at that time. Third, this analysis becomes the initial condition for the time integration of a numerical weather prediction model based on the governing differential equations of the atmosphere. Finally, the numerical prediction is used by a human forecaster as the basis for the public forecast.

Not surprisingly, the Presidents’ Day snowstorm has been widely studied. In one of these studies, Hollingsworth, Lorenc, Tracton, Arpe, Cats, Uppala, and Kallberg (1985) examined the sensitivity of the numerical forecast to changes in the initial conditions. In midlatitudes, errors tend to propagate eastward in the prevailing westerly flow. The investigators discovered that the predicted evolution of the Presidents’ Day snowstorm in the western Atlantic was extremely sensitive to *small errors in the initial analysis in the northwestern Pacific four days earlier*. In other words, a small localized error in the initial analysis affected the forecast for locations far removed in space and time.

Clearly, accurate analyses of the state of the atmosphere are indispensable for forecasting. They are also invaluable for phenomenological and climatological studies and for the formulation and testing of theoretical models of atmospheric behavior. In

their most popular form, animated full-color analyses are the mainstay of the TV weather show.

In the last century, only observations from local surface stations were available, and these were analyzed using subjective manual techniques. Over the years, as the demand for atmospheric and environmental information grew, increasingly exotic and comprehensive observing systems were deployed for sampling atmospheric parameters. Analysis techniques also had to evolve to provide this vast and heterogeneous real-time data base with spatial and temporal continuity and internal physical consistency.

Analysis problems that are equally complex occur in other scientific disciplines. What distinguishes the atmospheric analysis problem from other problems are the physical properties of the medium, the spatial and temporal characteristics of the observing system, and the applications toward which the analysis is directed. In this chapter we examine all of these characteristics, beginning with the most fundamental, the physical properties of the atmosphere itself.

## 1.1 Atmospheric characteristics

The important components of the earth climate system are

- the atmosphere,
- the hydrosphere (oceans),
- the cryosphere (polar ice fields and sea ice, continental snow cover),
- the lithosphere (earth's surface including hydrology and volcanism), and
- the biosphere (vegetative cover and oceanic flora and fauna).

Each of these components is coupled to the others (Peixoto and Oort 1984). The atmosphere and hydrosphere are strongly coupled through the exchange of energy, momentum, and matter on many space and time scales at the air/sea interface. Changes in continental snow cover affect the surface albedo, and variations in the extent of sea ice affect the heat exchange between ocean and atmosphere. The lakes, rivers, and ground water of the lithosphere are essential elements of the terrestrial branch of the hydrological cycle and are connected to the atmospheric branch by evaporation and precipitation. The vegetative cover of the biosphere affects the surface roughness, albedo, evaporation, precipitation, and moisture capacity of the soil.

The earth climate system displays phenomena on a variety of space and time scales (Figure 1.1).

The atmosphere is the most variable component of the earth climate system in both space and time. The atmosphere system functions to store and distribute the heat received from the sun. Although there is net radiative cooling in the free atmosphere, the earth's surface has a radiative surplus, which is maximum in the tropics and is positive at all latitudes except the winter polar latitudes. This horizontal and vertical radiative distribution forces a circulation in the atmosphere that transports heat upward and poleward.

*Atmospheric characteristics*

3

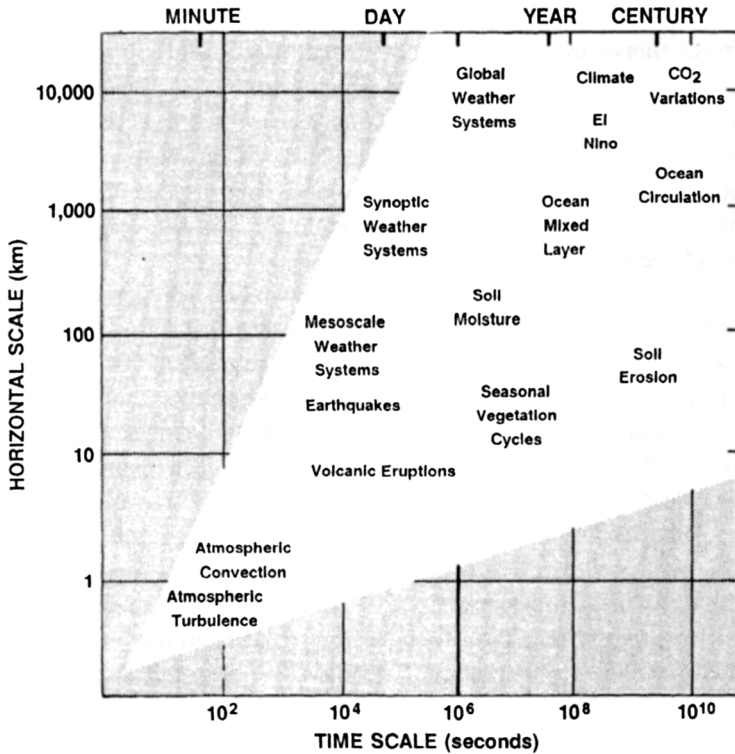


Figure 1.1 Space and time scales of phenomena of the earth climate system. Note that both scales are logarithmic. Phenomena with very long time scales are not included.

In the tropics, the surface air is warm and moist, and the pressure at the surface of the earth is relatively low. As the heated air rises and cools, much of its water vapor condenses out as rain. At high levels, the air travels poleward and sinks at about 30 degrees north and south, creating belts of high pressure. This descending air is heated by (adiabatic) compression. The circulation in this convective cell is completed by an equatorward current at low level, which flows naturally from high to low pressure. The low-level equatorward flow from both hemispheres results in the Intertropical Convergence Zone. The Hadley cell, as this simple toroidal circulation is called, is a direct thermal cell.

As the earth rotates, moving fluid is deflected to the right in the northern hemisphere and to the left in the southern hemisphere (Coriolis force). At low levels in the tropics and subtropics, the Coriolis force deflects equatorward flowing air to the west giving rise to the easterly trade winds. In the same way, the air flowing poleward at high levels must develop a strong westerly component to conserve its absolute angular momentum. Because this westerly current is unstable to small perturbations, the flow at high latitudes is characterized by large-amplitude waves or eddies, which account for the major portion of the extratropical poleward heat transport. The eddies can vary in scale from planetary waves (which are due mainly to orography or thermal

contrast between continent and ocean), through synoptic disturbances or cyclones (which extract energy from the westerly flow through instability mechanisms), right down to individual convective elements. The cyclones have life cycles of about one week and travel along well-defined cyclone tracks that are dictated by the more slowly moving planetary waves. The cyclones have a large rotational wind component and a small divergent wind component. The divergent component, even though small, is important because low-level convergence and upper-level divergence imply upward vertical motion, condensation, and precipitation.

This book concentrates primarily on atmospheric phenomena of global and synoptic space and time scales (see Figure 1.1): space scales from approximately 100 to 40,000 km (the earth circumference) and time scales from a few hours to about one month. However, we do not neglect atmospheric phenomena with different space and time scales and other components of the earth climate system. The shorter time and space scales of mesoscale weather systems and atmospheric convection are discussed in Section 13.7. Time scales longer than one month are considered in Section 1.6. The hydrosphere is discussed in Section 13.8 and the lithosphere in Section 13.6.

In the analysis of the atmospheric state, two aspects are of primary importance: the physical laws that govern the atmospheric circulation and the spatial and temporal spectra of atmospheric phenomena. The physical laws indicate how it might be possible to determine one variable from another, for example, analyzing wind from temperature observations. Knowledge of the spectra of atmospheric variables can be used to determine an acceptable spacing between observations. Spectra are also useful in designing analysis algorithms that successfully separate observational noise from the signal (the true values).

The governing equations of the atmosphere can be written in terms of the independent variables (three spatial variables and time) and the dependent variables (such as mass, temperature, the three components of motion, humidity, chemical species, and cloudwater). In general, these equations are nonlinear partial differential equations of which the most important are the equations of motion, the first law of thermodynamics, and the mass and humidity conservation equations.

Atmospheric phenomena can be characterized by their aspect ratio  $L_z/L_H$ , where  $L_H$  is a characteristic horizontal scale and  $L_z$  a characteristic vertical scale. A reasonable value for  $L_z$  is about 10 km; so in the global and synoptic scales,  $L_z \ll L_H$ . Under these conditions, the atmospheric flow is, to a very high degree of approximation, in hydrostatic balance, and vertical motions are much weaker than horizontal motions. In mid and high latitudes, the synoptic and to some extent the global scales exhibit a substantial degree of geostrophic balance. That is, the Coriolis forces are approximately balanced by the pressure gradient forces.

Space/time spectra of atmospheric flow are shown in Figure 1.2 for time scales of days and horizontal space scales of  $10^3$  kms. The contours are isolines of 500 millibar (mb) geopotential variance and are spaced logarithmically. It is evident

*Atmospheric characteristics*

5

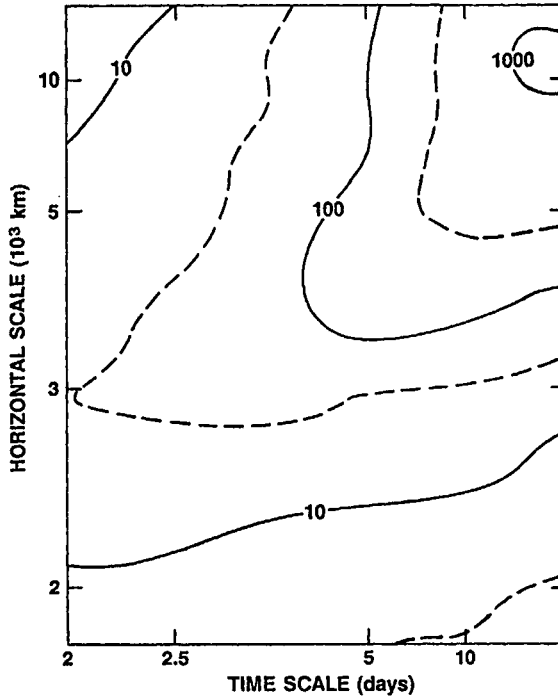


Figure 1.2 Space/time spectrum of atmospheric variance. (From Pratt, *J. Atmos. Sci.* 36: 1681, 1979. The American Meteorological Society.)

that the bulk of the variance is at relatively larger time and space scales and that shorter time scales are associated with shorter space scales.

The temporal spectrum of atmospheric kinetic energy shown in Figure 1.3 is logarithmic in time and linear in kinetic energy. This diagram is somewhat schematic and so does not show units for the kinetic energy. Four obvious energy peaks appear. The peaks at one day and one year are the diurnal and annual cycles, respectively. The peak that occurs between one day and one month is associated with the baroclinic eddies or cyclones in the midlatitude westerlies; the peak at about one minute is associated with atmospheric turbulence and convection. A relative minimum occurs at time scales of about one hour.

A spatial spectrum of vertically integrated atmospheric kinetic energy at horizontal scales of 40,000 to 200 km is shown in Figure 1.4 as a log/log plot. The global wavenumber  $k$  is defined such that wavenumber 1 has a wavelength equal to the earth's circumference. Curve S (stationary) indicates all variance associated with time scales longer than about one month, and curve T (transient) indicates all variance associated with the remaining time scales. As would be expected from Figure 1.2, curve S has most of the variance at very long waves. The transient curve T has a maximum at wavenumbers 8–10 (several thousand kilometers) associated with the baroclinic eddies.

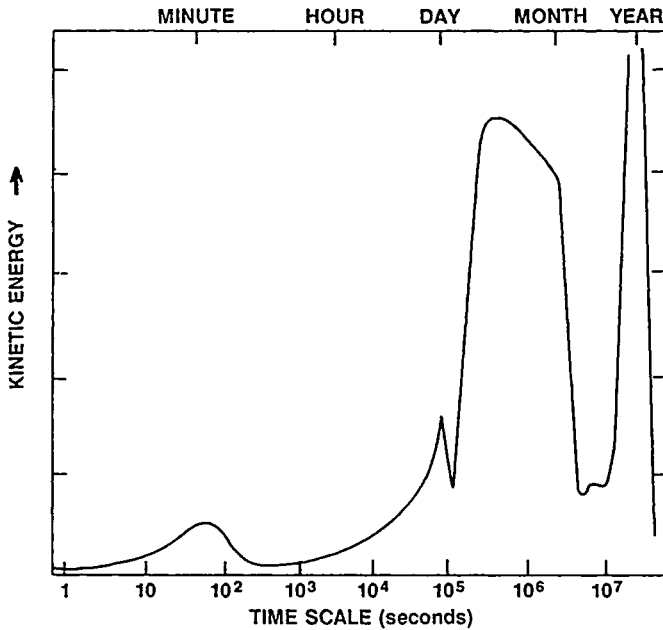


Figure 1.3 Temporal spectrum of atmospheric kinetic energy. (After Vinnichenko 1970)

Atmospheric turbulence theories suggest that inertial subranges should exist for which the kinetic energy spectra obey a  $k^{-5/3}$  power law for a three-dimensional fluid and a  $k^{-3}$  power law for a two-dimensional fluid. The small aspect ratio in the synoptic and global scales suggests a largely two-dimensional flow, which should become increasingly three-dimensional as  $k$  increases. The slopes of the  $k^{-5/3}$  and  $k^{-3}$  power laws are indicated on Figure 1.4. The observed spectra are not very reliable for the shorter scales (dashed line) but seem to fall somewhere between  $k^{-5/3}$  and  $k^{-3}$ .

As noted earlier, the atmospheric circulation is governed by physical laws. The formulation of these governing laws in the nineteenth and twentieth centuries put meteorology on a scientific basis and led to an important conceptual breakthrough.

## 1.2 The ultimate problem in meteorology

The main impetus for the study of the atmospheric circulation has always been people's desire to forecast the weather (predict future atmospheric states). Bjerknes (1911) referred to this as the ultimate problem in meteorology and outlined an approach for tackling it. According to this approach, two conditions must be satisfied to successfully predict future atmospheric states:

- I The present state of the atmosphere must be characterized as accurately as possible.
- II The intrinsic laws, according to which the subsequent states develop out of the preceding ones, must be known.

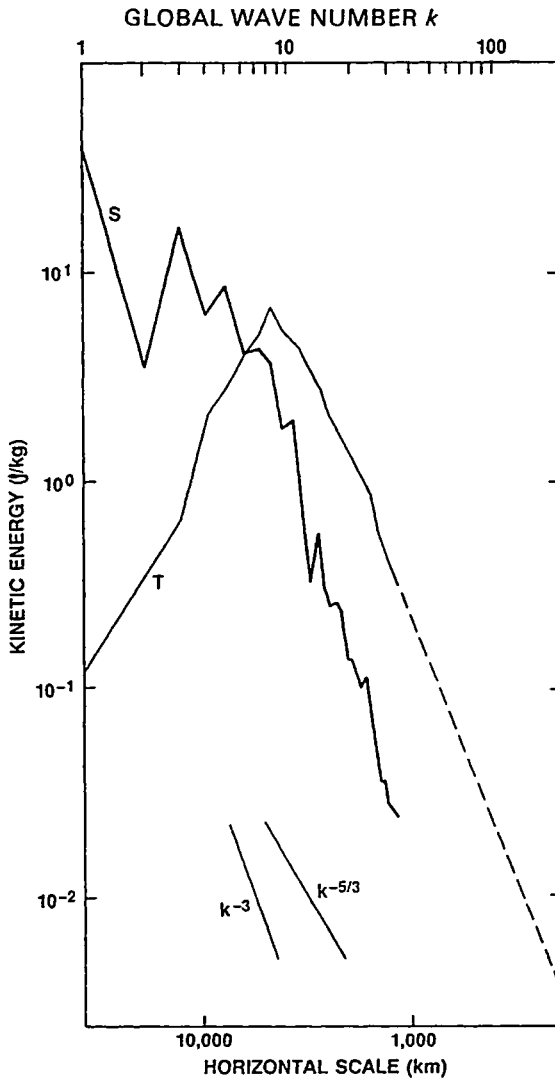


Figure 1.4 Spatial spectrum of atmospheric kinetic energy. (Curves S and T are from Boer and Shepherd, *J. Atmos. Sci.* 40: 164, 1983. The American Meteorological Society, with dashed portion of T curve inferred from Brown and Robinson, *J. Atmos. Sci.* 36: 270, 1979. The American Meteorological Society.)

Conditions I and II define weather prediction as an initial-value problem. In today's terminology, Bjerknes' approach would be called deterministic because future states of the atmosphere are assumed to be completely determined from the present state. For the practical implementation of the initial-value approach, Bjerknes outlined a program that was subdivided into three partial problems or components:

- 1 The observation component



- 2 The diagnostic or analysis component
- 3 The prognostic component

Components 1 and 2 are related to the characterization of the present state (condition I); component 3 is related to condition II.

Component 1 requires an observation network distributed throughout the atmosphere. At the observing points, the dependent variables (mass, wind, etc.) are measured. The distances between observation stations and the time between observations should be sufficiently small to adequately resolve the space and time scales of the phenomena of interest. Bjerknes also suggested that all observations be taken at approximately the same time (the principle of simultaneity), but this restriction is no longer felt to be practical.

In the diagnostic component (2), the observations are analyzed in a way that produces regular spatial representations of the dependent variables at fixed times. Such representations could consist of regularly spaced points on a grid or coefficients of a functional expansion. The density of the gridpoints or the number of terms in the functional expansion should be sufficient to adequately represent each of the dependent variables. The diagnostic function should consider not only the observations themselves but also the intrinsic relations between different dependent variables (mass and wind, for example).

In the prognostic component (3), the governing equations of the atmosphere are used to predict future states from the present state. The governing equations are highly implicit and have to be rewritten so that they can be used to *integrate* or *march* the dependent variables forward in time from their initial values. Bjerknes' era lacked the resources and knowledge to carry out his program successfully, but now all three components are performed routinely.

We now consider all three functions in more detail in order to establish the context in which the diagnostic component (the subject of this book) is exercised.

### 1.3 The observing system

An existing meteorological station in the British Isles has been either an outgrowth from an astronomical or magnetic observatory, or it has adjoined the house of an enthusiast who lived there for reasons unconnected with meteorology, or it has been pushed out to the confines of the islands to grasp as much weather as possible, or it has been placed in charge of coastguards because they are on duty at night, or it has been set on a mountain to test the upper air. Excellent practical reasons all these, but it is remarkable that the properties of the atmosphere, which are expressed by its dynamical equations and its equation of continuity, appear to have no influence on the selection. (Richardson 1922, p. 217)

Richardson's remarks, written in 1922, indicate quite clearly his frustration with the British meteorological observing system of the time. His complaint was essentially of the heterogeneity of the observing system, which was designed for human convenience rather than sampling the atmosphere in a manner consistent with underlying physical principles. As will be seen in the following historical review, the



heterogeneity of the observing system has its roots deep in history (Khrigian 1970).

The present-day global meteorological observing system has been evolving for 300 years. The basic meteorological instruments (thermometer, barometer, hygrometer, and anemometer) had been invented by the middle of the eighteenth century. Meteorological observatories such as the Paris Observatory commenced taking regular observations in the late 1600s. Short-lived meteorological observing networks were set up in Britain (Royal Society, 1724–1735) and Russia (Great Northern Expedition, 1730–1745). The observations collected by these networks were dignosed long after the fact due to the poor communications of the times. The Palatine Academy of Sciences and Letters in Mannheim, Germany organized the first international observing network, in which regular observations were collected during the 1780s and 1790s from as far away as the Ural Mountains and Cambridge, Massachusetts.

The idea of a global real-time observing system first became a reality in the nineteenth century. A number of technological and organizational developments occurred after 1800 that made a global network possible. On the technical side, meteorological instrumentation substantially improved. More important, however, was the invention of the telegraph, which allowed meteorological observations to be communicated rapidly.

On 14 November 1854, a severe storm in the Black Sea destroyed the French fleet at Balaklava. Because the storm had been observed the previous day over the Mediterranean, the French government asked Urbain LeVerrier, director of the Paris Astronomical Observatory, to study the circumstances surrounding this phenomenon. He obtained weather observations from around Europe and was able to trace the path of the storm. In 1855 he presented to Napoleon III a plan for a great meteorological network, designed to warn mariners of impending storms. Thus was born the first permanent observing network and the first national weather service. Other nations quickly followed suit.

Meteorological phenomena, of course, transcend national boundaries, and it was realized that the observation network would require extensive international cooperation. In particular, observing practices, units, and observation times had to be standardized. Increasing international cooperation and consultation took place throughout the latter half of the nineteenth century, but the International Meteorological Conference of Vienna in 1873 placed this cooperation on a formal diplomatic basis, and a permanent international committee was established shortly thereafter. One of its main responsibilities was the standardization of meteorological observations. This was no mean task; even the general adoption of metric units by English-speaking meteorologists took another 75 years.

Thus, by the beginning of the twentieth century, a global real-time observing system was in place. It was primarily a surface network, mainly confined to land and distributed rather arbitrarily, as noted by Richardson. Sporadic attempts to sound the free atmosphere with balloons, kites, and aircraft occurred in the early years of this century, but not until the second world war was an international upper air network established. This network employed the balloon-borne radiosonde, which

measured directly the temperature, pressure, and humidity and transmitted encoded observations to nearby ground stations by radio. The radiosonde balloons could be tracked by theodolite or radar to yield wind measurements.

By 1950, the upper air network provided good coverage over land, but there were still large oceanic areas from which observations were rarely received. In fact, Spilhaus (1951) reported that over large areas of the southern oceans less than 500 meteorological observations (of any kind) had been made in half a century. In the later 1960s, satellite-borne radiometers were developed and launched. With the deployment of these instruments came the hope of obtaining uniform observational coverage over the globe and thus finally filling in the large oceanic data voids. The invention of the electronic computer in the late 1940s made possible the real-time processing of all these new data.

In 1979, the nations of the world organized the Global Weather Experiment (GWE) to observe the atmosphere as systematically as possible using the most advanced technology then available. In addition to the operational observing network, many special observing systems were deployed, such as drifting buoys to measure surface variables over the oceans, dropwindsondes ejected from aircraft, and constant-level balloons to observe the upper circulation. The experiment ran from December 1978 to November 1979 and provided a comprehensive global meteorological data set.

The international meteorological observing system today is known as the World Weather Watch (WWW). It is supervised by the World Meteorological Organization (WMO), which is a United Nations agency. The WWW has three components. The Global Observing System (GOS) consists of the basic surface and radiosonde networks, aircraft and satellite systems run by the national meteorological services. The Global Telecommunications System (GTS) consists of telecommunication facilities and arrangements for rapid transmission of observations and processed information. The Global Data Processing System (GDPS) has world meteorological centers in Melbourne, Moscow, and Washington D.C. plus national meteorological centers that collect, store, process, archive, and disseminate the observational data in real time.

Weather observations from the surface and radiosonde networks are processed in the following way. At the observing station, the observer processes the raw information, making appropriate corrections for local effects and checking the internal consistency of the observations. The observations are then coded in a form suitable for transmission, using an international code called SYNOP (for land surface observations), SHIP (for sea surface observations), and TEMP (for upper air observations). This coded information is transmitted to regional centers and then to national and world meteorological centers. The information from all stations is collected, decoded, checked for errors, archived, and disseminated. Unfortunately, this process takes time: a complete set of global observations is not usually available until several hours after observation time.

Surface and radiosonde observations are taken regularly at *synoptic* times. These are 00 and 12 GMT (Greenwich mean time) for radiosondes and 00, 03, . . . , 21, 24