Chapter 13

100 Years of Progress in Forecasting and NWP Applications

STANLEY G. BENJAMIN AND JOHN M. BROWN
NOAA/Office of Oceanic and Atmospheric Research/Earth System Research Laboratory
Global Systems Division, Boulder, Colorado

GILBERT BRUNET
Environment and Climate Change Canada, Montreal, Quebec, Canada

PETER Lynch
University College Dublin, Dublin, Ireland

KAZUO SAITO
Atmosphere and Ocean Research Institute, University of Tokyo, Tokyo, Japan

THOMAS W. SCHLATTER
NOAA/Office of Oceanic and Atmospheric Research, Boulder, Colorado

ABSTRACT

Over the past 100 years, the collaborative effort of the international science community, including government weather services and the media, along with the associated proliferation of environmental observations, improved scientific understanding, and growth of technology, has radically transformed weather forecasting into an effective global and regional environmental prediction capability. This chapter traces the evolution of forecasting, starting in 1919 [when the American Meteorological Society (AMS) was founded], over four eras separated by breakpoints at 1939, 1956, and 1985. The current state of forecasting could not have been achieved without essential collaboration within and among countries in pursuing the common weather and Earth-system prediction challenge. AMS itself has had a strong role in enabling this international collaboration.

1. Introduction

Weather forecasting has evolved over the past century from being an empirical, analog-based process with little skill to an increasingly automated high-tech enterprise demonstrating tremendous prediction accuracy. This transformation has been enabled by spectacular advances in our ability to observe the atmosphere, by application of rigorous physical principles to understanding atmospheric phenomena, and by the advent and subsequent enormous increase in digital computing capability. Daily decisions, with outcomes from inconsequential to far-reaching in magnitude made by groups and individuals, are based on the forecast community’s assessment of future environmental conditions—forecasting, and especially, weather forecasting. Vilhelm Bjerknes (Bjerknes 1911) referred to forecasting as the ultimate problem in meteorology.

In this chapter, we trace the progress of weather forecasting and scientific questions leading to that progress.

DOI: 10.1175/AMSMONOGRAPHS-D-18-0020.1
© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
over the last 100 years. This progress is broken down here into four 20–30-yr periods, in each of which, the science and practice of forecasting underwent major shifts. This chapter on forecasting complements scientific developments described in other chapters of this monograph and distills from them. The science of forecasting has followed close on the heels of overall progress in atmospheric science (understanding and numerical representation of physical processes and phenomena described in other chapters) but has often, in turn, accelerated research in these related areas, often following major forecasting failures. Forecasting, in general, and numerical prediction models have aided tremendously in understanding the atmosphere–Earth system (e.g., Randall et al. 2019). Repeated application of the scientific method—asking questions, constructing new hypotheses, and then conducting experiments and collecting observations to test them—is behind the development of weather forecasting and all of the science topics of this monograph.

In each era, we identify, where appropriate, observations used for forecasting and the process of analyzing those observations into a coherent picture (a weather map), growth in the scientific understanding that influenced the practice of forecasting, the community of weather forecast providers from...
As noted by Murphy (1993), weather forecasts only advances in these forecasting components in each era. The overarching focus here is on the evolution of operational forecasting. Within each era, we also describe the quantitative aspects of forecasting, for example, numerical weather prediction (NWP), advances in technology for prediction, how communication of forecasts to users evolved, and comprehension of forecasting by the public. Table 13-1 gives a roadmap through advances in these forecasting components in each era. As noted by Murphy (1993), weather forecasts only realize value when used to make good decisions (see example from 1944 below), and so the overarching focus here is on the evolution of operational forecasting.

This 100-yr period is divided here into these four eras:

1) “Era 1” (1919–39: maps only; observations and extrapolation/advection techniques),
2) “Era 2” (1939–56: increasing science understanding; application especially to aviation; birth of computers),
3) “Era 3” (1956–85: the advent of NWP and dawn of remote sensing), and

This chapter flies lightly over this vast effort, touching down far too briefly or not at all in significant subareas.
Forecasting and its scientific underpinnings were transformed drastically within each of these eras. We focus on weather forecasting, which guides daily human activities, but note the neighboring enterprise of climate forecasting, which uses closely related science and technology, recognizing that seamless weather–climate prediction is now close at hand (e.g., Brunet et al. 2015). Links between weather forecasting and broader Earth prediction are already being forged in many NWP centers as of 2019 and are pivotal to the future of forecasting, as described in the last section of this chapter.

This chapter complements other chapters on Earth-system model development, severe convective storm science research, tropical and extratropical cyclone research, applied meteorology, and observing systems. It describes application to human activities of that science more fully described elsewhere in this volume; forecasting provides a unique science portal into the daily lives of the public and of the forecast providers themselves.

Our forecasting chapter includes the closely related area of NWP applications for real-time prediction. Even in 1919, a form of the equations of motion for NWP already existed from V. Bjerknes in 1904 and Cleveland Abbe in 1901. Numerical solution of these equations had already been attempted by Lewis Richardson as he reported in 1922. Sverre Petterssen (Petterssen 1974) wrote that the Bergen-School-era meteorologists were looking ahead but were unable to quantify forecasting without drastically improved observations and a method to solve the equations of motion. That was not initially accomplished until subsequent work by Carl-Gustaf Rossby (Rossby et al. 1939, and later references), Jule Charney (Charney et al. 1950), and others.

Practical forecasting provided daily tests of atmospheric science hypotheses and daily applications in decision-making. A singular example of this occurred in the latter part of the Second World War. Petterssen [for the (then British) Meteorological Office with Charles Douglas], the British Admiralty, and Irving Krick (representing the U.S. Army Air Force), each independently asked the same question from different locations in southern England in early June of 1944: when might breaks occur in overcast, windy conditions in the English Channel? History-making decisions were based on these forecasts as both sides in the conflict anticipated an Allied military landing onto the coast of France. All were aware of the weighty consequences of these decisions. Forecasting at this time was driven by maps identifying patterns of sea level pressure, wind, and temperature using recent observations mostly at the surface. Changes in cloud and wind conditions and gaps in inclement weather could allow certain weather-sensitive activities, including military coastal landings with air support. Reginald C. Sutcliffe, whose overview (Sutcliffe 1952) summarized forecasting during the middle of the century as well as any, stated, “I can say with some assurance that the techniques of the day (D-day) were virtually innocent of dynamical theory” (Lindzen et al. 1990, p. 205). But the “techniques of the day” did include application of the role of upper-level features on surface weather. This forecast, critical in world history, was severely handicapped when viewed by twentieth-century standards. Observations were extremely limited, but some understanding of fronts and the three-dimensional evolution of air masses and their impact on clouds, precipitation, and winds had been established over the previous 25 years. This understanding had been achieved by scientists from several countries brought together frequently in Bergen, Norway, by V. Bjerknes.

Three surface depictions for 0900 UTC Tuesday 6 June 1944 are shown in Fig. 13-1: one from the Allies [no data are available over continental Europe except not-plotted decrypted data per Lewis (1985)], one from the Germans (no data over the British Isles), and one from a recent reanalysis by the European Centre for Medium-Range Weather Forecasts (ECMWF). All three showed a low in the North Sea, but the Allies had additional upstream observations and had earlier predicted that a gap would occur on 6 June with less cloud cover and wind. Even among the three Allied forecasting subteams (Met Office including Petterssen, Royal Navy, and Krick with the U.S. Army Air Force) coordinated by Group Captain Stagg in the Commanders-in-Chief office, there was intense disagreement about possible windows (a decrease in winds below critical thresholds and reduced cloud cover) on 5 June or none at all (ECMWF 2014; Petterssen 1974, 2001; Stagg 1972; Fleming 2004). Forecasters on both sides were aware of the Norwegian airmass/cyclone/front conceptual model described earlier by Jacob Bjerknes and Halvor Solberg (Bjerknes and Solberg 1921) with their iconic figure (Fig. 13-2) and used it to determine a sea level pressure distribution. Forecasters in 1944 were fully aware of the need to map a coherent 3D structure from synoptic observations [also evident in the German map discussion attached to its surface map in Météo France (2014)]. Out of that structure, they were able to infer possible evolution largely on the basis of advection of the assumed three-dimensional structure. Petterssen (1974) and Stagg (1972) later described that the decision to postpone the invasion from 5 June was the most important decision, as well as the most contentious. The group decision was correct but with a dissenting position from Krick and his subteam using only analog techniques. Petterssen and Douglas assessed forecasts largely using the Norwegian conceptual model with upper-air-to-surface relationships...
refined with local near-surface details. The Allied team agreed subsequently on 4 June that a subtle but sufficient respite from unfavorable conditions would occur on 6 June and issued a momentous forecast. This was based, in part, on a single surface observation from Blacksod in northwest Ireland (Petterssen 1974, 2001). The final decisions each day were made by General Eisenhower, always seeking consensus with his subordinate commanders in chief, on the basis of recommendations and information from Stagg and the forecasting team. German forecasters did not see this weak but sufficient break in the weather, so Rommel returned to Berlin for his wife’s birthday, leaving Normandy just before the actual invasion began.

Weather disasters in different countries led to establishment of forecasting capabilities. The unpredicted Midwest “Children’s” blizzard in January of 1888 (Laskin 2004) and an East Coast blizzard 2 months later led to U.S. congressional investigations (L. Uccellini 2018, personal communication) and the establishment of the U.S. Weather Bureau (USWB) in 1891 (Reichelderfer 1941). Many such defining weather events over the last 100 years were disasters, for example, the 1938 New England hurricane (Avilés 2013), U.S. tornado outbreaks (e.g., 8–9 June 1953 major killer tornadoes in Flint, Michigan, and Worcester, Massachusetts, on consecutive days and an outbreak on 3 April 1974 spreading from Alabama to Ontario, Canada), and Cyclone Nargis (2006) causing 100,000 fatalities in Myanmar. For each of these events, the global weather enterprise (GWE; Thorpe and Rogers 2018) asked itself this question: how could we do better next time? There has been no shortage of hypotheses from each of these events—what better observations, improved NWP solutions, better understanding and more detailed representations of atmospheric processes, and better communication to users would have been required? The ongoing process of weather prediction and its effects on human activity, with increasing successes and less frequent failures, generates new ideas or restatements of old hypotheses. This process has evolved over the last 100 years to a global science development, now with contributions from every nation—a consequence of a single prediction problem for our common planet of residence.

2. The 1919–39 era (Era 1)—The age of maps and graphical methods using advection

a. State of forecasting

The societal need for forecasting increased in this first era, starting in 1919 and ramping up to the start of World War II (WWII). Technology during this period allowed improved observations and communication but not yet a computational solution of any version of the equations of motion.

A day in the life for U.S. forecasters in 1939 is described by Charles Mitchell and Harry Wexler (Mitchell and Wexler 1941, hereinafter MW41) using observations..

FIG. 13-1. Surface pressure analyses valid for the morning (0900 UTC) of 6 Jun 1944: contemporary (produced in real time) analyses from (left) the British (contour interval: 4 hPa) and (right) the Germans (contour interval: 5 hPa), and (center) the corresponding surface pressure (contour interval: 4 hPa) and 10-m winds from a 9-h European Reanalysis of Global Climate Observations (ERA-CLIM) forecast initialized at 0000 UTC 6 Jun. The reconstructed figures are taken from Météo France (2014) and ECMWF (2014). The left panel is British Crown copyright, Met Office, and contains public-sector information licensed under the Open Government Licence v3.0 (http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/).
available via teletype, plotted and analyzed onto horizontal maps for the surface (Fig. 13-3a), 10,000 ft (1 ft = 30.48 cm) above sea level (Fig. 13-3b, valid at the same time), and isentropic surfaces (Fig. 13-3c, to infer a three-dimensional flow structure). Upper-air charts were considered to be an essential addition to produce precipitation forecasts. Forecasters used these maps to identify air masses and boundaries (MW41). The most important technique was simple extrapolation of the past movement and rate of development of weather systems. A daily weather map, a simplified version of the U.S. Weather Bureau surface map (described in 1941 by MW41 as the “principal weather chart”), was provided to newspapers. MW41 indicate that agriculture and transportation were key forecast users. Already, the importance of weather forecasting for aviation interests was growing as the Weather Bureau was directed in 1926 to begin weather services for civilian aviation (Gregg 1935), followed by Charles Lindbergh’s solo flight, nonstop, from New York, New York, to Paris, France, the following year.

For observations available to forecasters in this era, prediction was based largely on surface observations (per MW41), with some pilot balloons. The number of surface meteorological stations in the United States increased from 200 to 350 by 1939 (MW41). Airplanes were dispatched specifically to obtain soundings up to 5 km at 35 sites in the United States by 1935 (Gregg 1935). By 1938, aircraft soundings were replaced in the United States by radiosondes, which, important and in contrast to aircraft soundings, were available in all weather conditions. MW41 describe the importance of the telegraph for transmitting surface observations, enabling subjective analysis of weather maps starting in 1870 by the Signal Service of the U.S. Army. Balloons with meteorographs (similar to subsequent rawinsondes but requiring instrument retrieval) were being launched regularly with data collection in Strasbourg, France, organized by Hugo Hergessel and analysis of the data in Leipzig, Germany, by V. Bjerknes [including for the 1910 case (Bjerknes 1910) used by Richardson]. In this preradiosonde era, these data were not available for real-time prediction but were collected for extensive after-the-fact studies of atmospheric structure and evolution.

Careful analysis of surface observations, particularly from ships at sea, enabled development of the polar front theory of the “Bergen School.” In fact, subjective map analysis was a major component of the forecasting process for over two-thirds of the past 100-yr period. Meteorological technicians were trained to plot observations on maps, and forecasters subsequently identified patterns and subjectively estimated and drew isolines for pressure (isobars), temperature (isotherms), and other variables. Real-time knowledge of the vertical structure of weather systems from direct observations was recognized as critical but remained sketchy, at best, through most of this era, limited by the available technology as noted above. Relationships between the surface and the lower half of the troposphere began to emerge, but there was nothing definitive about conditions higher up. For example, the existence or even the notion of upper-tropospheric jet streaks seems not to have been suspected. Rossby and collaborators (Rossby et al. 1937) found that in the 10,000-ft maps (example in Fig. 13-3b) features would appear and disappear over time, leading him to plot observations on surfaces of constant potential temperature (Shaw 1930) and demonstrate that the 3D airflow with upward and downward motion could be estimated on a single isentropic map. Isentropic analysis was handicapped at that time by the large cancelling terms of the Montgomery streamfunction (solved later by Danielsen 1959) and fell out of favor by forecasters in the subsequent era (1939–56), but was partially revived later (e.g., Bleck 1973).

FIG. 13-2. Idealized polar-front wave [from Bjerknes and Solberg (1921), later quoted in the U.S. forecaster training manual by Haynes (1940)]. In the center is a horizontal view of the distribution of air masses at the ground. The broken line is the boundary (polar front) at the ground between a warm current from the west-southwest (white arrows), displacing to the east a wedge of cold air (black arrows) returning northward from a brief sojourn in southern latitudes. Along the boundary of the receding cold air (warm front) the warm air rises, and its moisture condenses and produces a broad area of rain or snow (shaded area). The upper part of the diagram represents a vertical west–east section north of the center; the lower part is a similar section south of the center. (A Str means altostratus clouds; Ni is nimbus; Ci Str is cirrostratus; Ci is cirrus; A Cu is altocumulus.) The figure is provided through the courtesy of the Norwegian Geophysical Society.
FIG. 13-3. Weather charts for approximately 1200 UTC 29 Mar 1939, from MW41: (a) surface weather chart, (b) upper-air (10,000 ft above sea level) weather chart (the data were gathered by pilot balloons), and (c) the flow pattern for an isentropic surface corresponding to a potential temperature of 301 K.
As advances in radio technology and instrument design made it feasible to send aloft lightweight instrument packages and radio transmitters with balloons, the possibility of obtaining data above the surface and even into the stratosphere on a regular basis was recognized. The first successful flights of a balloon carrying an instrument package that transmitted weather information to the ground were made in France in 1929 and in the Soviet Union in 1930. During the mid-1930s, design was begun in the United States and abroad of a radio meteorograph or radiosonde system that would be capable of making routine upper-air observations (MW41; Dabberdt et al. 2003). From then on, progress was rapid. The first radiosonde in the United States was launched from the Blue Hill Observatory in Milton, Massachusetts, in 1935. In 1936, radiosonde ascents were made by the U.S. Navy using a radiosonde design from the Massachusetts, in 1935. In 1936, radiosonde ascents were made by the U.S. Navy using a radiosonde design from the National Bureau of Standards. Regular radiosonde data became very widespread and essential for forecasting as WWII started, as is discussed in the next section.

b. Scientific understanding enabling forecasting

International science exchange in this era was extensive, especially between Europe and the United States, but it had little effect on day-to-day forecasting yet. The Geofysisk Institutt in Bergen, started by V. Bjerknes in 1917 and led by him afterward, was the initial center for this exchange, including Rossby from Sweden in 1919–20 (Fig. 13-4). Francis Reichelderfer (also Fig. 13-4) from the United States, a Navy pilot during World War I (WWI) and later to be Director of the U.S. Weather Bureau for 25 years (Hughes 1970), was another key visitor to Bergen in 1931, carrying back to the United States his desire to strengthen forecaster training using these new airmass and extratropical cyclone conceptual models. Sverre Petterssen (Fig. 13-4) also spent considerable time in Bergen, and other visitors included Aleksandr Fridmann, the physicist from Russia and Erik Palmén from Finland. A first department of meteorology was established at the Massachusetts Institute of Technology (MIT) by Rossby, a towering figure of this era, bridging the areas of refining conceptual models, for example, isentropic analysis (1937), simplified equations of motion (barotropic vorticity equation), and extremely effective organizational skills (e.g., Lewis 1992). Forecaster training at MIT and, soon after, University of Chicago and the University of California, Los Angeles (UCLA), met an urgent need for the U.S. Army Air Force (changed to U.S. Air Force in 1947). The centrality of international collaborative exchange of science, observations, and predictions, already evident in this era through visits to Bergen and the influence of European scientists toward establishment of U.S. university departments, cannot be overstated, because it spurred the 100 years of progress in forecasting by the GWE (Thorpe 2014). The time was ripe in 1919, as WWI ended, for increasing global cooperation and reducing potential for conflict, including for weather prediction and environmental science as the American Meteorological Society (AMS) and other scientific organizations formed to further open international doors.

This era witnessed codevelopment of conceptual understanding (conceptual “models,”—not numerical) and of equations from fluid dynamics based on fundamental physics from Newton, Maxwell, and others in the eighteenth and nineteenth centuries using Eulerian methods (Pudykiewicz and Brunet 2008, hereinafter PB08). The unifying conceptual model of this era, by far, was the airmass concept, developed largely in Bergen, with fronts representing sloping 3D boundaries between different air masses. The basic cyclone model (Fig. 13-2) and polar-front theory were described by Bjerknes and Solberg (1921), showing that an extratropical cyclone forms within a narrow zone of large concentrated temperature contrast separating adjacent air masses: a front [Fig. 13-2 here from Bjerknes and Solberg (1921), later Palmén and Newton (1951)]. These concepts were not embraced in the United Kingdom or United States until the mid-1930s after extensive visits, especially between workers from Europe and the United States (Petterssen 2001). Rossby (1941) summarized “the scientific basis of modern meteorology” in 1941 in a companion article to the forecasting paper by Mitchell and Wexler (MW41).

Development of mathematical descriptions of processes was a common emphasis across all of the sciences during this era, including atmospheric science. For atmospheric forecasting, however, there were almost no upper-air observations and certainly not any computational capability to solve the governing equations on a real-time basis (Petterssen 2001). Use of mathematics, the universal language of science (Dutton 1976), clarified the description of atmospheric and Earth-system processes. The simplification of the equations of motion to enable a useful large-scale computer solution started with Rossby et al. (1939) in this era and was extended during the next era by Charney (1948; see the photograph in Fig. 13-4) and Eliassen (1949).

A persistent, unanswered question was, Why did cyclogenesis occur? Vilhelm Bjerknes wrote that cyclogenesis occurred with an unstable wave on the polar front. Jacob Bjerknes determined a more thorough structure for linking upper-level airmass structures with surface observations (Bjerknes and Solberg 1921), but a satisfactory, mathematically rigorous theory for cyclogenesis based on the governing equations had to wait for the next era.

c. The “community” of forecast providers

The forecasting community started to coalesce during this era, led by national agencies. Coordinated efforts toward forecasting were established by national
governments after telegraph lines crisscrossed the United States, for example, during the last third of the nineteenth century, making possible rapid transmission of observations to a central location. It was clear that only the government could marshal sufficient resources to gather observations, coordinate analysis and forecasting, and issue forecasts for the benefit of all. Many countries (United States, United Kingdom, Norway, Sweden, Germany, and others) recognized this central role of government for forecasting. Enabled by the invention of telemetry in the late 1800s, in 1891, the U.S. Congress transferred responsibility for weather forecasting from the Signal Corps to the newly created U.S. Weather Bureau (Hughes 1970). In 1938, after the death of USWB Director Willis Gregg, Reichelderfer and Rossby were both recognized as strong candidates as the next director. Reichelderfer was named director (Table 13-2), and Rossby was appointed deputy director—in essence, the USWB chief scientist.

Since weather is no respecter of national boundaries, the benefits of international exchange of data were also recognized early, and the International Meteorological Organization (IMO) was organized in 1873 at the Vienna International Meteorological Congress. One of the missions of the IMO was the international standardization of observing practices to facilitate international exchange of meteorological data. The IMO later became the World Meteorological Organization (WMO) in 1950 under the auspices of the United Nations (UN; WMO 2018). Strong international cooperation has always been a hallmark of meteorology and was essential for effective weather forecasting in
all eras. The International Civil Aviation Organization was established in 1944 in support of safe and efficient international aviation.

d. Forecasting applications

In this era, agriculture, transportation, and public safety were perhaps the most important applications for forecasts: “The object in weather forecasting is to provide the farmer, the city dweller, the shipper of perishable goods, the railroads, the public utilities such as gas, electric, and street-transportation companies, the aviator, and the owners and masters of all types of vessels from fishing and pleasure craft up to the largest passenger liner, and all others interested in the weather, with as accurate weather information as possible and with forecasts as far ahead as practicable…” (MW41, p. 583).

Also in this era, forecasters were faced with unique “Earth-system prediction” challenges with the U.S. Dust Bowl, especially during 1934–36, with its drastic consequences for agriculture and health (Choun 1936; Duncan and Burns 2012). Forecasters used, in essence, short-range surface wind forecasts combined with knowledge of land surface conditions to estimate possible dust-storm conditions. A notable event from that period occurred on 14 April 1935 (Fig. 13-5; National Weather Service 2010). Windy dust-storm days occurred in the context of a larger multiyear drought, exacerbated by poor agricultural practices, a topic that is now studied with today’s Earth-system reanalyses (e.g., Cook et al. 2009) and short-range dust-source and full Earth-system prediction discussed in later sections.

Launched in 1903 with the first flight of a powered aircraft by Orville and Wilbur Wright, commercial aviation started to play a prominent role in forecasting as both a key application and, in the late 1900s, a platform for needed atmospheric observations. Vilhelm Bjerknes summarized this symbiosis with prescience in 1910: “The development of aeronautics will make these [upper-air, or aerological] observations not only possible, but also necessary.” By 1918, only 15 years after the first powered flight, the USWB began to provide aviation forecasts. It became clear that pilots and their passengers needed far more detailed weather forecasts than “the farmer, the seaman, or the average citizen” (Hughes 1970). In 1926, the U.S. Congress passed an Air Commerce Act, landmark legislation for commercial aviation, to establish an official responsibility for USWB to provide these ongoing aviation forecasts (Hughes 1970). By 1933, United Airlines began flying coast to coast in the United States, with takeoff to touchdown taking 20 h (Streever 2016). In their early years, commercial airlines had little weather information, no preplanned flight plans, and suffered increasing accidents, including a Trans World Airlines (TWA) crash in 1935 that killed U.S. Senator Bronson Cutting (from New Mexico). Pan American Airways was the first airline to employ meteorologists, starting in 1927. Six additional airlines took the same step by 1937 (Fahey et al. 2016).

NWP did not begin to affect actual forecasting until the late 1950s, so we defer discussion of this topic until later eras in this chapter, but essential foundations were laid from 1900 to 1939. Vilhelm Bjerknes (1904) laid down a two-step plan for “the rational solution of forecasting” (Lynch 2006, p. 5): 1) a sufficiently accurate knowledge of the state of the atmosphere at the initial time and 2) a sufficiently accurate knowledge of the laws according to which one state of the atmosphere develops from another. Felix Exner (1908) developed a system of simplified equations assuming geostrophic flow and constant-in-time thermal forcing, a precursor to the essential simplifications used later by Rossby and Charney.
Lewis Richardson started to work at the Met Office in 1913, considering V. Bjerknes’s plan for “rational forecasting,” and then attempted a first solution of partial differential equations by finite differencing (Richardson 1922). Richardson’s manuscript was lost during WWI (Lynch 2006; Streever 2016) and retrieved later under “a heap of coal.” Although his forecast was unrealistic, his method was sound and is essentially the method used today. Exner (Exner 1923) reviewed Richardson’s book and doubted that his approach “would ever lead to significant progress in weather prediction. The processes are simply too complicated and fraught with too many small-scale details.” Richardson’s solution to the equations allowed contaminating fast-moving gravity waves and horizontally propagating sound waves (Lamb waves), and a close look into these weaknesses is the central topic of the book by Lynch (2006). Details on NWP history to complement those described in this chapter are provided by PB08. Again, all of this early NWP did not affect daily forecasting during this era.

e. Communication and public perception of forecasts

Communication of weather guidance changed significantly during this era. Time-scales of individual weather phenomena (convective storms, winter storms, and tropical cyclones) have, obviously, not changed detectably through the century. However, the speed of transfer of forecast guidance to decision-makers has changed drastically and repeatedly. At the start of this era, newspapers were the favored medium for transfer of weather-forecast data. Small weather boxes were printed in the New York Times (Henson 2010; Streever 2016). The introduction of radio during this era allowed the start of public AM radio broadcasting in 1920 (especially in the United States—KDKA in Pittsburgh, Pennsylvania—and the United Kingdom). Radio allowed far quicker and more universal transmission of weather forecasts, including to rural areas (Calvert 1921). Since the time scale of hazardous weather events is from days to minutes, rapid communication was always a goal. Dissemination of weather information via radio forged links between USWB forecasters and local broadcasters (Henson 2010).

The public was generally skeptical toward weather forecasting during this period (Henson 2010), with little to no knowledge of its growing scientific maturity. The need to establish professional standards led to the formation of AMS in 1919. The USWB and its leaders at the end of this era, Reichelderfer, Rossby, and Harry Wexler (chief of its Scientific Services Division, 1946–62; Fig. 13-4) were acutely aware of this need within their own staff, a need that mounted in the late 1930s with the increasing likelihood of war. Incorporation of atmospheric science to enable weather forecasting was already being emphasized in every tier from the research community to forecasters, to decision-makers, and even to the public.

3. The 1939–56 era (Era 2)—The age of aviation

a. The state of forecasting

The second era (Table 13-1) saw the advent and growth of technologies that had a massive impact on meteorology in general and forecasting in particular. WWII and its aftermath, the beginning of the “cold war,” were in no small way responsible for this. Weather forecasting at the beginning of this era was very much influenced by the foundational work of the Bergen School 20 years earlier. Surface observations were still the backbone of the observing system, and surface analysis procedures in the USWB tended to follow the principles of surface analysis set forth by the Bergen School: identification of cold, warm, and occluded fronts. Reichelderfer, the new USWB director, drawing upon his exposure to the Bergen School during the last era, was important in requiring that forecasters identify fronts and include them on surface maps (L. Uccellini 2018, personal communication). Avilés (2013) describes a contrast in forecasting (well known to Reichelderfer) for the 1938 New England hurricane between Charles Mitchell and a younger forecaster, Charles Pierce, who, under the influence of the Bergen model for cyclone evolution, argued (correctly) for a landfalling track. Routine upper-air observations by radiosondes were just becoming available at the start of this era, as noted above. Radio direction finding (RDF) developed by the U.S. Signal Corps during WWII allowed radiosondes to add reporting of relatively accurate winds, thus becoming rawinsondes. Thus, little was known about how surface fronts and cyclones might be manifested well above the surface (Rossby 1956); temperature and wind observations above the surface were insufficient to support routine reliable application of the thermal-wind relation to make inferences about upper-level features that might have been useful in a forecast context. Further, despite considerable theoretical work in this area, there was as yet no adequate theory to explain cyclogenesis.

As an example of what was routinely available aloft at that time with which to work, see Fig. 13-3b (from MW41, their Fig. 2). The use of isentropic analysis as developed at MIT during the middle and late 1930s (Namias and Stone 1940) was important in revealing the 3D aspects of weather systems, particularly areas of vertical motion: for this same 29 March 1939 case, Fig. 13-3c shows an isentropic analysis. Note that the solid lines are isobars on the 301-K isentropic surface. The observed winds (approximating
system-relative flow in this nearly stationary system) blowing across the isobars indicate upward motion over the lower Mississippi Valley and eastward into Alabama, Tennessee, and Georgia. Scherhag (1948) devised an ad hoc method to link upper-level flow to surface patterns in the 1930s by advection of 24-h surface pressure tendencies by the 500-hPa wind, a method used in Germany for the following 20 years (R. Bleck 2018, personal communication). In the United States, a central analysis unit was established in 1942 in Washington, D.C., by the USWB to generate such surface and upper-air weather maps, with facsimile dissemination to provide weather information for both civil and military aviation (Reichelderfer 1942).

The “gathering storm” of WWII at the start of Era 2 was a huge stimulus toward improving weather forecasts, since it was recognized early on that, in comparison with WWI, effective use of air power could well be decisive. And, because aircraft operations encounter a wider range of hazards, both on the ground and in the air, they are inherently more sensitive to weather than are sea and land operations. The necessity for reliable, accurate forecasts of weather conditions hazardous to aircraft operations or detrimental to mission therefore provided a strong motivation for research into atmospheric phenomena being revealed by the new upper-air observing capability (rawinsondes), and for development of effective prediction procedures and tools. More information on the history of application of forecasting to aviation is provided in this monograph by S. E. Haupt and her colleagues (Haupt et al. 2019a, their section 3).

Recall that the electronic digital computer did not exist at this time, nor did automated tools for plotting data and analyzing charts. So it was pen for plotting observations, and pencil and eraser for analysis, plus tracing paper and a light table for doing tasks such as subtracting height fields on different pressure surfaces to get thickness values. Typical scenes from the USWB Central Analysis Center are shown in Fig. 13-6 (circa 1950). Quantitative procedures had to be simple and easy to apply, not involving much computation. It was common in both Eras 1 and 2 for practitioners to develop charts and nomograms of various sorts. Many examples of these can be seen in textbooks written during this era that covered forecast techniques, for example, Petterssen (1940, 1956), and Saucier (1955). With rawinsonde data now available, the use of thermodynamic diagrams designed specifically for use by meteorologists such as the pseudoadiabatic chart (or Stüve diagram), the tephigram, and later the skew $T$–$log_p$ diagram came into common use (Fig. 13-7). Techniques of analysis were refined by, for example, applying natural coordinates using analyzed streamlines and isotachs for purposes of computing vertical vorticity and horizontal divergence (e.g., section 2.8 of Petterssen 1956).

Forecasting was based primarily on experience augmented by use of analogs (similar weather patterns), representing an attempt to build on collective experience. Nonetheless, developments in dynamical meteorology began to be applied, particularly in the district (regional)
offices and specialized forecast offices of the USWB. The two biggest, related forecast challenges at the shorter ranges were the motion of weather systems, particularly frontal systems, and the question of development—where could formation of extratropical cyclones be expected, and would existing cyclones on the surface map deepen? Forecast utility was limited to 36–48 h.1 As Petterssen (1974) wrote, the lifetime of extratropical systems was normally about 4–8 days, and entirely new systems (currently unobserved) could form and dominate the weather map within 5 days. The Bergen School, of course, had laid a foundation for this. But in this era, the new, seeming abundance of upper-air data from the rawinsonde network added quite literally a whole new dimension to consider and was a great stimulus to understand more fully the relationship between features aloft and those on the surface map.

While the European front of WWII focused strongly on extratropical forecasting issues, allowing full use of the polar front/airmass conceptual model, the Pacific front, in contrast, forced exposure of many forecasters to the tropics. Riehl (1954, p. v) described his first evening in Puerto Rico in 1943 at a new institute:

On the first evening some of staff walked along the beach and admired the beauty of the trade cumuli in the moonlight. Well schooled in the ice-crystal theory of formation of rain, they had no suspicions about these clouds with tops near 8,000 feet where the temperature is higher than $+10^\circ$C. Suddenly, however, the landscape ahead of them began to dim; then it disappeared; a roar approached as from rain hitting roof tops. When some minutes later they stood on a porch, drenched and shivering, they had realized that cloud tops with temperatures below freezing were not needed for the production of heavy rain from tradewind cumulus. There and then the question arose: How is it with the other theories insofar as they concern the tropics?

We have already noted in section 2 (1900–39) the vulnerability of aviation to adverse weather, in particular, low ceiling and visibility and unanticipated enroute winds. Icing (supercooled liquid water in clouds) and turbulence were also well known as additional weather hazards. As aviation grew in commercial and military importance, the demand for accurate forecasts of adverse flying conditions increased dramatically, prompting the U.S. Congress to pass the Civil Aeronautics Act in 1938. Regulations in the act ensured that air carriers operated in the safest possible manner, in part through increased focus on improved weather information. The aviation demand was also a major motivating factor in the United States for moving the USWB from the U.S. Department of Agriculture to the U.S. Department of Commerce in 1940. The demand for improved aviation forecasts also gave the earlier push for a radiosonde (soon to become rawinsondes with winds added via RDF) network deployed beginning in 1938 (Rossby 1956; cf. MW41, further discussion below). The critical role of military aviation during WWII and the rapid expansion of commercial and general aviation thereafter only increased the prominence of aviation as a driver for improved forecasts worldwide.

b. Observations used for forecasting

Aircraft also became an important source of real-time weather data during this era, albeit initially only in the form of pilot voice reports radioed to the ground. Although aircraft and gliders had been used in meteorological research prior to this era, it was during this period that instrumented aircraft came to be seen as a vital research tool (Emanuel 2019, in this monograph). The advent of the radio-transmitted balloonborne rawinsonde system revolutionized weather forecasting and was accelerated by WWII. In 1938, six USWB airplane

---

1 This translation plus growth/decay strategy was similar to that called nowcasting in the 2000s for much shorter lead times.
stations were converted to radiosonde stations. “Finally, in July 1939, all of the Weather Bureau and 3 of the 9 Navy airplane stations were replaced by radiosonde stations, making a total of about 30 in all.” (MW41, p. 583). Since these balloonborne observations of temperature, pressure, humidity and wind could be made at altitudes at or above the tropopause with some reliability under most weather conditions, this technology opened up a whole new vista for meteorology. See also Stith et al. (2019, in this monograph) for more information. Stickler et al. (2010) summarize the history of radiosonde soundings: 147,000 soundings were taken worldwide from 1938 through 1946, and 10 times as many (now rawinsondes, including winds) were taken from 1948 through 1957, responding to demands of greatly increased commercial aviation after WWII. Ship-based upper-air observations were started during WWII, reaching a peak of 22 fixed ship stations over the Atlantic Ocean and 24 over the Pacific Ocean (Dinsmore 1996). To support transoceanic commercial aviation, a network of 13 ships (Fig. 13-8) was continued after the war over the North Atlantic with additional Pacific ship stations, until satellite data and automated buoys became available in the 1970s (Dinsmore 1996). Surface observations were enhanced, again driven by aviation (both civilian and military), and 30 new airport stations were opened in 1940 at major U.S. terminals (Reichelderfer 1940, p. 6).

Radio detection and ranging (radar) did not find its niche as a critical tool for detecting hydrometeors until after WWII. During this conflict, it was used mostly to detect aircraft, but it was realized that radar could also detect precipitation. After the war, surplus radar equipment became available for civilian use (Rinehart 2004, p. 3). Meteorologists, particularly at the University of Chicago, McGill University (Montreal, Quebec, Canada), and MIT, began seriously investing in it as a tool for weather research. Particularly noteworthy here was the Thunderstorm Project (Byers and Braham 1948), a groundbreaking study of convective storms using radar. By the end of this era in the mid-1950s, radar was being used to study synoptic-, meso, and convective-scale weather systems at several institutions. Forecasters recognized its potential as a tool for detection and short-term (by extrapolation) forecasting of precipitation and severe local storms (Whiton et al. 1998).

Without question, the technology of electronic digital computers has revolutionized weather forecasting far beyond any other technology. Invented during this era, computers had the potential to fulfill Richardson’s
dream of using the basic governing equations of atmospheric motion to predict the weather based on physical principles laid out by V. Bjerknes (Bjerknes 1904) and Abbe (Abbe 1901; Willis and Hooke 2006). This topic will be discussed more later.

c. Science understanding enabling forecasting

This era also saw a revolution in meteorological education, which underpinned the expansion of atmospheric science understanding and the training of forecasters. At the beginning in 1939, only a few schools in the United States offered specialization in meteorology, and much of this was due to the influence of Rossby. He joined the MIT faculty in the Department of Aeronautics in 1928, founded the MIT Department of Meteorology (the old Course XIX) in 1932, and then later moved to the University of Chicago (1940), where he gathered and trained a small cadre of meteorologists including Palmén and George Cressman, which later formed the nucleus of the Department of Meteorology there. As the storm clouds of war gathered in Europe, J. Bjerknes found a welcome refuge at UCLA and, with Rossby’s encouragement, founded the department there in 1940.

After the Pearl Harbor attack in Hawaii in 1941, the acute need for weather forecasters to support the U.S. war effort was obvious. Extensive training programs for weather officers were organized at MIT, UCLA, the University of Chicago, and New York University. The programs gave Fred Sanders, Edward N. Lorenz, and many other foundational atmospheric scientists of later eras their initial training in meteorology. Along with this educational effort, thousands of enlisted personnel got a crash course in meteorology at Chanute Air Force Base, Illinois, during WWII. Many of them ended up working in offices of the USWB after the war ended. The WMO, formed in 1950, greatly aided coordination among national forecasting services and helped to meet the needs of the growing international aviation industry.

The work of Rossby et al. (1939), particularly the development of the famous formula for the speed of waves in a uniform zonal current based on conservation of absolute vorticity,

\[ c = \frac{\beta L^2}{4 \pi^2}, \quad (1) \]

provided a rational foundation for forecasting the motion of weather systems. This formula relates the phase velocity \( c \) to the mean zonal flow \( \mathbf{U} \) and the wavelength \( L \). Sutcliffe et al. (1951) verified the formula using upper-air charts, and Petterssen (1956) modified the formula by incorporating lateral shear across the current.

At issue was the level at which the formula should be applied, since it was quickly recognized with the advent of the upper-air network that the flow near the tropopause was much stronger than that at lower levels in the typical open wave, whereas vertical cross sections through such waves showed them to generally remain coherent, with a westward tilt with height. A second important relation from Rossby et al. (1939) indicated that, under zero horizontal divergence, purely horizontal flow, and barotropic conditions, absolute vorticity was conserved:

\[ d(\zeta + f)/dt = 0, \quad (2) \]

where \( \zeta \) is the vertical component of relative vorticity and \( f \) is the Coriolis parameter. Cressman (1948) and Sutcliffe et al. (1951) argued for 500–600 hPa as giving the overall best results for application of Eq. (2).

Rossby had demonstrated the importance of group velocity, showing that a packet of planetary waves could propagate in a direction contrary to the individual waves. This was applied to the phenomenon of downstream development in a qualitative manner, including by Ernest Hovmöller in his iconic time–longitude diagrams (Fig. 13-9; Hovmöller 1949). Hovmöller diagrams have become widely used into the current time for conveying time–space-varying patterns, as described by Persson (2017).

The closely related concept of constant absolute vorticity trajectories (CAVTs, also developed by Rossby; see Petterssen 1956, his chapter 8) also gained some currency with forecasters during the immediate postwar period. These were computed by noting the latitude of the core of the jet at an inflection point in the flow between a trough and a ridge (identifying absolute vorticity at this location) and calculating the trajectory of air from that point forward in time assuming constant absolute vorticity with the wind speed held constant and continued zero cross-flow shear. Such hypothetical trajectories originating between an upstream trough and downstream ridge would curve anticyclonically so long as they remained poleward of their starting position, and cyclonically if they remained equatorward. CAVTs were used to gauge the potential for discontinuous retrogression of a long-wave trough in the westerlies or the possibility that a new long-wave trough would form between two existing ones. This is an example of an attempt to incorporate fundamental dynamical considerations into the forecast process in a manner that was at least marginally tractable with use of the calculators that were available at the time.

Related to these considerations, additional fundamental appreciation of the dynamics of midlatitude mobile weather systems became more firmly established.
That the flow in the upper troposphere was systematically stronger than that at low levels, together with simple considerations based on the vorticity equation, argued that there must be horizontal divergence in the upper troposphere ahead of troughs and convergence behind. This naturally argued for convergence ahead at low levels and divergence behind, with upward motion in advance and descent behind in the middle troposphere. Thus, the typical spatial relationship between precipitation and weather features was put on a firmer foundation. One of the authors (Brown) recalls an impromptu one-on-one tutoring on these points by J. Bjerknes in the UCLA map room in the early 1960s.

During this era, observations confirmed that the airflow above the surface on the scale of the mobile midlatitude cyclones and anticyclones was rotationally dominant, that is to say, the winds aloft were approximately geostrophic, and that the vorticity of the flow was typically much larger in magnitude than the divergence. This suggested that simplifications of the basic governing equations for forecasting applications were possible, as illustrated by the examples mentioned above. But it remained for Charney (1947) to show mathematically that, given a sufficient horizontal temperature gradient, a simplified set of equations incorporating these concepts yielded unstable solutions possessing the time and space scales of observed cyclones and anticyclones. Charney (1948) then went on to formalize the quasigeostrophic (QG) system through a rigorous scale analysis of the basic governing equations. This simplified system has the advantage of eliminating the fast gravity wave (and acoustic) modes yet maintaining the dynamics that produces the dominant weather systems of the midlatitudes. The timing of this was extremely fortuitous in that it coincided with the invention of the electronic digital computer, for the first time making it possible to do extensive mathematical calculations quickly and accurately. So, at this point in time, there appeared both the computing capability (limited as it was by current standards) and a set of equations to which this capability could be applied much more easily than to the set used by Richardson one-quarter century earlier.

d. Numerical weather prediction becomes possible

These coincidental developments were not lost on USWB leadership at the time. John von Neumann, the famous mathematician and key architect of the first computers, which were used initially for weapons research, recognized that the nonlinearities associated with transport in fluid flows required intensive computational power for solution and that meteorology presented an opportunity for peacetime application. He approached decision-makers including Reichelderfer, USWB Director since 1938, and the result was a decision in 1947 to form a small core of scientists at the Institute for Advanced Study at Princeton University to explore the possibilities. Charney was tapped to lead this group, which included Philip Thompson, who came from UCLA, George Platzman from the University of Chicago, and Arnt Eliassen from Norway. Eliassen left to go to UCLA in 1949, and Ragnar Fjørtoft came from Norway to replace him. In 1951, Norman Phillips came from the University of Chicago to join this effort.2

2 Interested readers may wish to find more information on this period from Lynch (2006), Harper (2008), and Lindzen et al. (1990), including on the tensions in this assembled group.
The chosen strategy was to start with the simplest form of the equations governing the synoptic-scale flow in midlatitudes first, namely, the conservation of absolute vorticity [Eq. (2)]. Held (2019) states that Rossby did not see the extension at that time to a predictive equation based on potential vorticity conservation. Wiin-Nielsen (1991, p. 43) relates that Charney made the same claim, even though Platzman disputed this in a famous interview (Lindzen et al. 1990). The plan was that, once a satisfactory means of solving this equation was developed, tested, and validated, attempts would be made to introduce additional computational layers to allow representation of observed baroclinic structure but still within the QG context so that smaller time steps would not be necessary. As well, it was planned to explore generalizations of the quasigeostrophic set that relaxed the geostrophic constraint on the advection of horizontal wind such that curvature effects were taken into account. The choice of a domain of integration was also important; using the concept of group velocity for Rossby waves, Charney and Eliassen (1949) reasoned that a domain covering roughly one-half of the hemisphere would allow forecasts of up to 36 h or so without excessive contamination by lateral boundaries.

By early 1950, the details of the computational procedure had been worked out, and over a month’s time four 24-h forecasts were successfully made on the Electronic Numerical Integrator and Computer (ENIAC) machine at Aberdeen Proving Grounds, Maryland. These forecasts proved to be surprisingly accurate, considering the simple dynamics, pointing out the importance of the nonlinear advection of vorticity. This work is described in another landmark paper (Charney et al. 1950) in which the Meteorology Group at Princeton established working modeling guidelines, for example, the equivalent barotropic assumption (Charney 1949), allowing the 500-hPa level to represent the troposphere through an assumption commonly used in barotropic model forecasts: that the horizontal wind increases in speed with height without change of direction.

The work of the Princeton group garnered considerable attention such that, by sometime in the early 1950s, there was widespread recognition that computer-produced forecasts that are based on the governing equations, filtered so as to remove the gravity and acoustic modes, represented a promising new approach to the forecast problem, including that of baroclinic development. Using a three-level QG model, the Meteorology Group next tested forecasts of the powerful November 1950 U.S. East Coast storm (Smith 1950) and other cases of cyclogenesis. These attempts were not uniformly successful and led to consideration of more accurate versions of the filtered equations and to more serious consideration of the hydrostatic primitive equations despite their admission of gravity wave modes (e.g., Charney 1955). The USWB meanwhile decided to pursue implementation of the equivalent barotropic model. A new Joint Numerical Weather Prediction Unit (JNWP) was formed under USWB auspices in Washington, D.C., in 1954 with a dedicated IBM, Inc., 701 computer to support the work and a mandate to develop operational NWP capability that would be available for both civilian and military forecast application. The Princeton Meteorology Group disbanded in 1956, Charney and Phillips went to MIT, and Fjørtoft had earlier returned to Norway. Cressman became the JNWP director, and Frederick Shuman and Joseph Smagorinsky were two other prominent members.

The United States was not the first nation to actually produce daily, real-time NWP forecasts. This honor belongs to Sweden, where in Stockholm under the leadership of Bert Bolin and others, including Rossby who had returned to Sweden [Swedish Meteorological and Hydrological Institute (SMHI)], daily integrations with a barotropic model were begun in 1954 using a computer of Swedish design (Persson 2005) several months before the JNWPUs. We discuss the impact of these initial NWP forecasts in the next section.

e. Communication of forecasts

Era 2 was important in the history of forecasting in respects other than observations and NWP. With an improved scientific basis of forecasting that emerged, daily weather forecasts began to have a bit more credibility with the general public. The success of the D-day forecast undoubtedly helped.) Radio and newspapers continued as the principal media through much of this era, but other media gained in importance. An automatic telephone forecast service was introduced by the USWB in some major cities by 1945 and became increasingly popular (NOAA 2006). Television became available in the early 1950s (from 10 million sets in the United States in 1950 up to 22 million two years later), and television news figures became part of the larger “community of forecast providers” (Henson 2010), in some cases becoming highly recognizable local personalities. Weather maps could be shown directly on television, tying the audience with what forecasters themselves used, together with a degree of animation depending on the gesturing skills of the television presenter. The role of television is discussed more in the next section.

In the years following WWII, some forecasters who had been trained in the newer methods involving use of routinely available upper-air data chose to go into private business, seizing on opportunities to provide
customized service to particular industries. They started private forecasting services for individual clients or took forecasting positions with industry. Aviation was an obvious source of employment for meteorologists, and several of the airlines formed weather units. Among the better known were Pan American, TWA, Northwest (later part of Delta), American, and United. By 1942, 17 airlines operated weather offices in North America, also including Braniff, Capital, Chicago and Southern, Continental, Eastern, Mid-Continent, Northeast, Panagra, Trans Canada (now Air Canada), and Western (Fahey et al. 2016). Airline meteorologists were tightly integrated into day-to-day flight planning and scheduling. Others advised the construction industry by advising on anticipated conditions for weather-sensitive activities, such as pouring concrete. Utilities needing to better anticipate weather-caused changes in demand contracted with weather-forecast providers or hired meteorologists to work in house. With its legacy from the U.S. Department of Agriculture, the USWB continued to provide detailed short-term agricultural forecasts during this era—for example, frost and freeze warnings for areas where frost-sensitive produce was grown. Private-sector meteorologists claiming to have skill at forecasting beyond 3–5 days ahead garnered a following, and their forecasts affected the commodities markets.

The popular U.S. weekly news magazine Time featured Carl-Gustaf Rossby on its front cover in December 1956, thereby acknowledging that weather forecasting was an emerging science that was profoundly affecting human activity. As noted by many (e.g., Petterssen 1974; Phillips 1998; Fleming 2016), Rossby’s scientific vision, organizational skill, and personal drive were unparalleled. As a Swede and an American as well, Rossby was a worthy choice to represent progress in forecasting by the international community. Rossby died from a heart attack the following year at the age of 59.

4. The 1956–85 era (Era 3)—Forecasting with a mix of maps, graphics, and NWP

a. State of forecasting

In the mid-1950s, forecasting was largely subjective, based upon advection of upper-level patterns with data from the now well-established global rawinsonde network and upon Bergen-School concepts about the evolution of extratropical upper-level waves and their links to surface pressure patterns. Sutcliffe (1952) described how experienced forecasters could successfully anticipate precipitation, clouds, and even surface winds. The first NWP forecasts became available to forecasters in Sweden in 1956 (Persson 2005) and in the United States in 1958 via facsimile machines. Thus began a long-term trend toward improved forecast accuracy as a consequence of increasing skill in NWP [Fig. 13-10 shows National Centers for Environmental Prediction (NCEP) 500-hPa forecast accuracy, with important NCEP computer upgrades indicated by the arrows]. Although the initial numerical predictions of the 500-hPa field had skill, forecasters had great difficulty in drawing reliable conclusions about conditions on the ground. This problem persisted for many years until surface mesoscale features could be represented in regional models.

Not until the 1980s did NWP become the backbone for weather forecasting in the United States (L. Uccellini 2018, personal communication). Throughout this 1956–85 era, there was broad recognition among forecasters about the limitations of operational NWP models, and model impact on actual forecasting was advisory at best (R. McPherson 2018, personal communication). Although upper-air skill showed a steady improvement starting in the 1950s (see again Fig. 13-10), Ramage (1978) stated that skill improvement was very little for surface fields, including precipitation and especially in the tropics. Early in the NWP era, expectations outpaced performance. Concern about the relative roles of humans and automated guidance in forecasting was already evident in Snellman’s (Snellman 1977) article describing “meteorological cancer.” This paper described how weather forecasting and the role of weather forecasters had to adapt quickly and continually to rapid development of technology and science. As noted by Hoke (2004), the role of forecasters evolved strongly into model evaluation and determining modeling system deficiencies. Meteorologists presided increasingly over the forecast loop by making forecast adjustments based on recognized systematic model errors. This role continues into the late 2010s with the U.S. National Weather Service (NWS), a consequence of forecasters viewing observations daily and evaluating models in light of observations as a way of ascertaining to what extent the model forecasts in a particular situation can be trusted. Forecasters now compare observations and models on a day-to-day basis, more so than any other participants in the “forecast enterprise,” including model developers or verification specialists.

Era 2 laid the groundwork for NWP to become a major player in the forecasting process. But NWP had yet to prove itself in the real day-in-and-day-out world of forecasting, even in the most obvious matter of providing guidance for the motion of synoptic-scale weather systems. As operational NWP came of age in the first half of Era 3, tensions arose between forecasters
schooled in the methods employed in WWII and taught in immediate postwar meteorology departments and numerical modelers versed in the new field of computers and numerical methods and with a strong foundation in the theoretical meteorology of the day. A gradual resolution of these tensions as NWP skill steadily improved was a defining feature of this era (1956–85) and the following one. Requirements for improved forecasting continued to proliferate during this period, including for the general public and certainly for aviation. The first commercial trans-Atlantic jet flight occurred in 1958 (International Air Traffic Association 2018), and the spread of transoceanic commercial aviation was accelerated in part by the introduction of the Boeing 707. Aviation safety became an increasing concern as commercial and general-aviation activity continued to grow rapidly and occasional major weather-related crashes caused many deaths. Aviation fuel costs also increased sharply along with world oil prices in the late 1970s, increasing the need for more accurate short-range upper-level wind forecasts.

By 1975, forecasters used knowledge of cloud physics, planetary boundary layer processes, radiation, orographic effects, and land surface interaction, much of it empirical, to fill in the details based on the synoptic-scale dynamic solution from the models. Green et al. (1966), Browning (1971) and Harrold (1973) introduced “conveyor belt” concepts (using an isentropic relative-flow perspective) to describe how both stable and convective precipitation could be understood within a baroclinic disturbance. By this time, Browning and Harrold (1969) and Austin and Houze (1972) had characterized mesoscale precipitation structures with the help of radar and three-dimensional airflow observations. These transport structures, refinements of the original Norwegian polar-front model as summarized by Palmén and Newton (1969), were noted in satellite images: filaments of water vapor transport later labeled as “atmospheric rivers” (Newell et al. 1992). A related airflow-based conceptual model explaining production of very high convective instability from an elevated mixed-layer capping cooler moist air was described by Carlson et al. (1983). These conceptual models were often confirmed, in part, by use of three-dimensional trajectories to determine source regions and elevation (e.g., Palmén and Newton 1951). Improved accuracy from use of isentropic coordinates for trajectories was demonstrated by Danielsen (1961).
for transport of water vapor, ozone, and, especially at that time, radioactivity.

b. Observations used for forecasting

Remotely sensed observations from satellite and radar became routinely available during this era and greatly transformed short-range subjective forecasting and atmospheric research, particularly on the mesoscale. The launch of Sputnik in 1957 precipitated the international space race and helped to define the start of Era 3. Since the 1960s, weather satellites have surveyed the entire globe daily. Images from TIROS-I, launched in 1960 by the United States, generated great excitement. By 1980, TIROS-N was providing clear-air vertical temperature and moisture profiles every 12 h at 250-km spacing with global coverage (Smith et al. 1981). In the mid-1970s, geostationary satellite imagery became available. It provided high-frequency, real-time monitoring of clouds, a major observational advance. Satellite cloud imagery gave, literally, a new perspective on atmospheric structures, leading, in turn, to revised conceptual models about three-dimensional airflow patterns upon which forecasters relied heavily. This advent of satellite images transformed forecasting, especially in the very short range.

Global rawinsonde data were available every 12 h during this era from fixed locations on land and sea; these data were the basis for regional and global upper-air analyses and NWP initial conditions. Rawinsonde data became the backbone of forecasting during Era 2, and this situation continued in Era 3. The rawinsonde network changed little except that soundings from the 13 ships over the Atlantic and Pacific were terminated in the 1970s as satellite and automated buoy data became available (Dinsmore 1996). Surface observations remained fundamental and essential for diagnosing current conditions and forecasting out to 2 days, as described by Sutcliffe (1952) and others. The WMO mandated surface “synoptic” observations every 6 h, but aviation needs gradually drove many countries to make surface observations every hour. Hourly observations better portrayed how the weather situation was evolving relative to the current forecast. They also highlighted, as never before, mesoscale features such as nonlinear wind shifts and isallobaric patterns. The number of global observations reaching the U.S. National Meteorological Center (NMC) doubled in the late 1970s (McPherson et al. 1979) because of the First Global Atmospheric Research Program (GARP) Global Experiment (FGGE). Augmented observations from land and sea, satellites and aircraft, were responsible.

After WWII, surface and upper-air charts were essential for improved characterization of the preconvective environment when severe local storms were considered possible, an area of inquiry led by the United States. Radar data became available (Doviak and Zrnić 1993; Atlas 1990) during this period, in the United States initially from refurbished military radars. The potential of radar as a hurricane-monitoring tool was realized with the first images of U.S. East Coast hurricanes in 1954–55 (Carol, Diane, and Hazel). The absence of radar information was keenly felt when Hurricane Audrey struck Cameron, Louisiana, with unexpected intensity in 1957, causing a disastrous storm surge (550 fatalities; Dunn and Miller 1960). In 1959 the USWB deployed the first radar specifically designed to detect and monitor weather targets: the Weather Surveillance Radar–1957 (WSR-57; Whiton et al. 1998). Initial deployment was along the East and Gulf coasts for improved hurricane tracking and warning. Radar summary charts began to be produced, consolidating radar data from civilian and military radars in 1956, and distribution via facsimile every 3 h began in 1960 (S. Weiss, NOAA, 2018, personal communication). During the 1960s, network coverage expanded to include much of the central and eastern United States that is prone to severe weather. For each WSR-57, meteorologists constructed and transmitted once per hour a radar coded message that gave location, movement, and tops of significant weather echoes detected by radar. The Weather Bureau’s Severe Local Storm Warning Center (SELS) in Kansas City, Missouri, composited these messages and forwarded them to local forecast offices for monitoring convective weather.

c. Science understanding enabling forecasting

Tetsuya (Ted) Fujita from the University of Chicago tutored a new crop of scientists during this era using observations of opportunity (often including home movies obtained from the public) and often poststorm surveys to improve understanding of convective storms. Fujita (e.g., Fujita 1970) meticulously documented evidence for hitherto little-known substorm phenomena including microbursts, downbursts, and multiple vortices within tornadoes. Several U.S. universities not already mentioned established meteorology departments or atmospheric science programs from 1946 to 1969 including The Pennsylvania State University, the University of Wisconsin, the University of Washington, the University of Oklahoma, Texas A&M University, Florida State University, the University of Illinois, the University of Michigan, and others, usually within larger

3From an unpublished report by J. Galway on the use of the radar summary chart at the National Severe Storms Forecast Center in 1965.
Earth sciences departments. Formation of these departments was prompted in part by early concern that human activity was causing environmental degradation. The government began to regulate the emission of atmospheric pollutants such as sulfur dioxide and oxides of nitrogen. This gave private-sector meteorologists the opportunity to predict air pollution episodes and created a demand for meteorology graduates skilled in air chemistry and trajectory forecasting under stable conditions.

At McGill University, a radar meteorology group from the physics department and an Arctic meteorology group from geography united in 1959 to form the Department of Meteorology. McGill has been a leader in the training of many distinguished atmospheric scientists including André Robert, who developed the first global spectral numerical model and integration in 1965. At the University of Tokyo in Japan, T. Okada and S. Fujiwhara built the basis of the modern Japanese meteorological science/service. The Geophysical Institute (Division of Meteorology) of the University of Tokyo was established in 1941, the year that Japan entered WWII. Notable meteorologists, many important later in U.S. atmospheric science (e.g., Y. Ogura, A. Arakawa, K. Gambo, A. Kasahara, Y. Sasaki, K. Ooyama, K. Miyakoda, S. Manabe, Y. Kurihara, T. Murakami, T. Nitta, M. Yanai, T. Matsuno, I. Hirota, and T. Takeda), studied under Professor Sigekata Syono, the first Japanese honorary member of AMS (AMS 1961). Their scientific strength was in dynamics and numerical computation, and their activities led to the 1960 NWP Symposium in Tokyo (Fig. 13-12, described below). Their experiences between the 1940s and 1960s, including difficulties during WWII and following their emigration to the United States, have been reviewed by Lewis (1993) and remembered by Kasahara (2015).

Two events in the 1970s spurred research on mesoscale and convective-scale phenomena: 1) the arrival on the research landscape of a sizable cadre of young, enthusiastic scientists with an interest in subsynoptic (mesoscale) phenomena, and 2) a number of weather disasters having strong convective-scale character, for example, the 9–10 June 1972 Rapid City, South Dakota, flash flood; the 3–4 April 1974 tornado outbreak; the Big Thompson, Colorado, flash flood on 31 July 1976; and the Johnstown, Pennsylvania, flood of 19 July 1977. This research advanced the theory of thunderstorms with rotating updrafts (e.g., Rotunno and Klemp 1982, 1985) and had major applications in forecasting [e.g., Maddox (1980) for mesoscale convective complexes]. Between 1977 and 1990, much of this activity centered in the National Oceanic and Atmospheric Administration (NOAA)–National Center for Atmospheric Research (NCAR) complex at the East Campus of the University of Colorado Boulder (in Boulder, Colorado), which housed a fortuitous mix of scientists with theoretical and practical forecasting interests.

Aviation-focused research, motivated by major commercial airline accidents, was part of this emphasis. For example, the Joint Airport Weather Studies (JAWS) project investigated thunderstorm microbursts (UCAR 1982). JAWS and other studies led to stronger recognition of the roles that mesoscale organization and orographic forcing play in flash floods. They also identified synoptic patterns that suggest flash-flood potential in the next 36–72 h. The end of this era saw planning and execution of field studies relating to midlatitude mesoscale convective systems, the prime example being the Oklahoma–Kansas Preliminary Regional Experiment for STORM-Central in May–June 1985 (Cunning 1986). Research during this period motivated plans for a major NWS reorganization, discussed below.

d. Community of forecast providers

The year 1957 marked the beginning of the space age and fresh emphasis on Earth science across many countries. The International Geophysical Year started in 1957. Dwight D. Eisenhower convened a Presidential Science Advisory Committee. Through this committee, scientists from 14 U.S. universities formed the University Corporation for Atmospheric Research (UCAR) and then, with support from the National Science Foundation (NSF), formed NCAR. In collaboration with UCAR-member universities, NCAR expanded the depth and quality of atmospheric science and contributed mightily toward long-term improvement of weather forecasting [more details are in Fleming (2016)]. The most important event in European forecasting over the last 60 years was the establishment of ECMWF in 1975 with 17 European member states, pooling European resources toward improved weather forecasts and climate data (Lynch 2006). ECMWF started to produce daily global forecasts in 1980. Three of the senior leaders at ECMWF (Lennart Bengtsson, Daniel Soderman, and Aksel Wiin-Nielsen) were from Scandinavia, all trained during the initial successful NWP effort in Sweden a few decades earlier (Persson 2005).

After the Sputnik launch in 1957 and with the subsequent impetus for increased science in the United States, President Kennedy noted in 1961 the international flavor of meteorology as a noncontroversial area for contributions from all countries for the benefit of the world as a whole. Kennedy proposed then to the United Nations a joint goal for cooperative weather prediction.
The UN passed a resolution later that year to advance atmospheric science and develop common forecasting capabilities across UN countries, including regional meteorological centers. Springing from this were the International Atmospheric Science Program (IASP, which led to GARP) and the World Weather Watch, which coordinates and promotes common weather observations and telecommunications. The latter are still today a backbone for NWP and forecasting (Petterssen 2001, p. 298). Numerous national weather services also began to provide precipitation forecasts with increasing quantification. The WMO now terms these organizations as national meteorological and hydrological services (NMHSs).

The U.S. NWS progressed to a higher level of scientific maturity and professionalism during this era. The United States was unique among countries in its distributed, local-emphasis approach to forecasting with 280 local offices in the 1960s. Reichelderfer continued his 25-yr post as USWB director until 1963. Wexler (see the photograph in Fig. 13-4) played a key role as USWB Chief Scientist under Reichelderfer. Starting in 1946, he strongly influenced USWB efforts on air pollution, radar, weather satellites, early consideration of climate change, and international collaboration (Fleming 2016). Thus, Reichelderfer had unmatched chief scientists during his tenure: Rossby and Wexler. Collaboration between civilian and military weather agencies waxed and waned through this 100-yr period, and the JNWPU split in 1960 into three organizations: the NMC (part of the USWB), a U.S. Air Force component (Air Weather Service), and a U.S. Navy component (Fleet Numerical Oceanography Center) (Shuman 1989). Smagorinsky, who started the Geophysical Fluid Dynamics Laboratory (GFDL) in Princeton, New Jersey, spoke of the early difficulty in the mixes of operational and research staff and also of civilian and military staff (A. E. MacDonald 2018, personal communication). Wexler died in 1962 at only 51 years old. Two years after Reichelderfer stepped down in 1963 after 25 years as USWB Director (Table 13-2), Cressman stepped in as the new director for the next 14 years.

In 1970, the USWB was renamed the U.S. NWS, as a component of the newly created NOAA, retaining Cressman as NWS director. Hughes (1970) provides an excellent history up to this transition point, and key U.S. NWS milestones over a longer period from 1800 to the present day are identified in a NOAA timeline (https://www.weather.gov/timeline). Richard Hallgren succeeded Cressman in 1979. Despite firm directions from Reichelderfer that forecasters adopt the Norwegian air mass analysis and from Cressman that all USWB staff receive 20 h of training including dynamics, the pace of change lagged even while the pace of science clearly accelerated in the 1960s and 1970s (R. McPherson 2018, personal communication). Cressman and Phillips had visited Princeton in 1956, and Cressman became convinced that NWP was the wave of the future (A. E. MacDonald 2018, personal communication). In 1978, the NWS introduced the first computer workstation into its offices. It was called Automation of Field Operations and Services (AFOS) and marked a giant step toward modernization. AFOS displayed data and simple maps. Although limited, it replaced the previous teletype and facsimile.

Design of a broad NWS modernization, featuring new observations (Doppler radars and wind profilers; Weber et al. 1990; Benjamin et al. 2004c) and better computer-based tools for forecasters began later in this era under Hallgren. Both NWS and the NOAA/Office of Oceanic and Atmospheric Research contributed to this modernization, including development of a far more powerful workstation called the Advanced Weather Interactive Processing System (AWIPS, not operationally available until the next era after 1985). The NWS modernization was realized in the 1990s and is described more in this chapter under the next era. Hallgren (Table 13-2; NWS director for 1979–88) led the push for the NWS modernization, including reversing NWS staffing from predominantly technicians to predominantly meteorologists with stronger scientific training (A. E. MacDonald 2018, personal communication). Hallgren was also instrumental in establishing the UCAR–NWS-cosponsored Cooperative Program for Operational Meteorology, Education, and Training (COMET) training center in Colorado in 1989. He also created a new Science and Operations Officer (SOO) position in the 1980s to accelerate incorporation of new technology and science into NWS Forecast Offices. Changes in the NWS field structure under Hallgren included expansion of the number of full-service Forecast Offices from 52 to 123 (including collocation with universities where possible) and the closure of about 171 smaller, less capable offices.

During the 1950s and 1960s, the USWB established special centers for forecasting tropical cyclones and severe local storms. Hurricane prediction was recognized as a major public need at the beginning of this chapter (e.g., Galveston, Texas, in 1906 and New England in 1938; Avilés 2013). Hurricane prediction for the tropical Atlantic Ocean and Caribbean Sea was established in Miami, Florida, at the National Hurricane Center in 1965. Kerry Emanuel (Emanuel 2019) describes the history of hurricane prediction in another chapter of this monograph.

Early in 1948, U.S. Air Force officers Ernest Fawbush and Robert Miller (Fawbush and Miller 1952, 1954; also
see Doswell et al. 1993) identified a set of conditions that could lead to tornadic thunderstorms in the central United States. On a particular day, that later spring, with those conditions evident, they issued what would now be called a tornado watch. Incredibly, a tornado struck their duty station, Tinker Air Force Base, Oklahoma, in early evening. Not by accident, in 1952 the USWB followed up by forming its own severe weather unit in Washington, D.C., in coordination with the U.S. Army and U.S. Navy [Weather Bureau–Army–Navy (WBAN)]. This unit was renamed SELS in June of 1953 after the devastating Flint, Waco (Texas), and Worcester tornadoes earlier that month. By 1954, SELS expanded and moved to Kansas City. By 1957, SELS was issuing both public forecasts and forecasts tailored to aviation for severe thunderstorms (Fahey et al. 2016). In 1966, SELS acquired a new name: National Severe Storms Forecast Center (NSSFC). In 1997, NSSFC was again renamed the Storm Prediction Center (SPC) and moved to Norman, Oklahoma, as part of the NWS modernization. A brief history of NWS severe weather prediction is provided by Corfidi (1999) and SPC (http://www.spc.noaa.gov/history/early.html). Harold Brooks and his coauthors (Brooks et al. 2019) provide additional details on severe storm research and prediction in the United States, South America, Europe, and Australia in this monograph.

By the late 1950s, dedicated positions focused on aviation weather had been established at USWB district forecast offices. These offices issued Forecast Advisories and what are now called Terminal Aerodrome Forecasts (TAFs) for major airports. An additional specialized office for international aviation was operated at LaGuardia Airport in New York City in the late 1940s. The crashes of Eastern Airlines Flight 66 in June 1975 at John F. Kennedy International Airport at New York City and Southern Airways Flight 242 in April 1977 in Georgia, both in thunderstorms, resulted in further reorganization of forecasting support for aviation within the NWS. The Convective Signet Unit was established, collocated with NSSFC, and NWS Center Weather Service Units (CWSUs) staffed by NWS forecasters were set up within each of the Federal Aviation Administration (FAA) Air Route Traffic Control Centers to advise controllers of adverse enroute weather conditions (turbulence, icing, and convection). During this era, many NMHSs sprang up in countries that did not already have them, with WMO providing international coordination for both observations and prediction; much of this was driven by aviation (as also mentioned under the previous era).

In 1953, the USWB granted teletype and facsimile (Fig. 13-21, described below) data access to private forecasters (Henson 2010), a major step toward broadening the community of forecast providers. By the late 1950s, a “weather wire” was used to transfer data from the USWB to many stations, and data were available to all to avoid any exclusivity. This substantially increased the amount of weather information available to the public via radio and television. In 1970, the AMS Award for Outstanding Services to Meteorology by a Corporation was given to a group of vanguard private consulting companies—A. H. Glenn and Associates; Murray and Trettel, Inc.; North American Weather Consultants; Northeast Weather Service; and Weather Corporation of America—belonging to the larger community of forecast providers at this point.

Aviation passenger miles grew by a factor of 2 between 1970 and 1985 (https://data.worldbank.org/indicator/IS.AIR.PSGR?view=chart). Exposure to hazardous weather conditions also increased, as already mentioned. Air pollution was a particularly important new application during this era, with initial air-quality standards being applied. Boundary layer meteorology also had a special application area in this era: nuclear energy regulation and safety. [See Walker and Wellock (2010) for relevant history, including the incidents from Three-Mile Island in Pennsylvania in 1979 and Chernobyl in (now) Ukraine in 1986.] Transport of possible nuclear fallout was a major concern and became a specific application for forecasting in this era. Energy applications of weather prediction were primarily for load forecasting in this era and not yet for weather-dependent generation from wind and solar energy. Haupt et al. (2019a,b,c) describe the evolution of many of these atmospheric science applications in this monograph.

### e. Numerical weather prediction in this era

The first NWP model output became available during this era. Expectations from models were modest: accurately show baroclinic wave development and improve prediction of extratropical cyclones. A nagging question still lingered in 1956: what was the cause of midlatitude cyclogenesis? The theoretical work of Charney (1947) and Eady (1949) was definitive in identifying baroclinic instability as the fundamental source of cyclogenesis. In practice, however, this did not readily translate into a usable formula like Rossby’s wave-speed formula [Eq. (1)]. The important contribution of latent heating to rapid oceanic cyclogenesis was not fully appreciated until late in this era, when higher-resolution models with improved representation of latent heating were able to successfully simulate such cases.

Regarding the equations of motion—what could be used at that time? As noted earlier, it was appreciated
during Era 2 that the full equations with only one (hydrostatic) approximation were more accurate, but they also required a shorter time step for numerical stability. The JNWPU started with the three-layer QG model (Table 13-3), which had the potential to predict cyclogenesis resulting from baroclinic instability but obtained better overall results with the barotropic vorticity equation, a single-level model (Lynch 2006). Spurious retrogression of long waves in the equivalent barotropic model was solved by allowing effects of divergence (Cressman 1958; Bolin 1956). Initial resolution at the JNWPU was 381 km [a resolution curiously forced by line-printer resolution at that time as described by Lynch (2006)] on a Cartesian grid (Table 13-3). Figure 13-11 shows a sample of line-printer output, enhanced manually, of a real-time 24-h forecast from the JNWPU and verification 1000-hPa heights for a case from May of 1956. No diabatic or viscid processes were addressed in these initial models.

The first real-time NWP forecasts were produced in Sweden in late 1954 and then by the U.S. JNWPU in early 1955. NWP-based maps were not provided to forecasters in the United States until 1958. As is true today, early operational models were constrained by available computer power, so developers had to consider trade-offs among resolution (horizontal and vertical, linked to the corresponding time step), duration of forecast, domain size, time of forecast delivery, and, later, sophistication in representation of subgrid-scale processes. As computer power increased, all of these could be treated in more detail to improve forecast accuracy and timeliness. In later eras but not in this one, data assimilation took its place to occupy an increasing proportion of the NWP computing resource.

The USWB eventually replaced the single-layer barotropic model with an improved three-layer QG baroclinic model (Cressman 1963) and then introduced the baroclinic six-layer primitive equation (PE) model (Shuman and Hovermale 1968) in 1966, which covered much of the Northern Hemisphere at 381-km horizontal resolution (Table 13-3). The six-layer PE model was the first U.S. model to include parameterizations of diabatic and frictional processes, with a frictional coefficient related to terrain height, and simple radiation processes, both applied in the coarse lowest layer.

The international collaborative effort to develop operational weather models was exhibited at the International Symposium for Numerical Weather Prediction in Frankfurt in 1956 with over 50 participants from the United States, Japan, and 11 European countries. Development of NWP numerical methods was under way in the United States, United Kingdom, Sweden, Germany, and Belgium [Platzman survey reported in Lynch (2006)]. The next NWP symposium four years later in Tokyo (Fig. 13-12) had 131 participants from many countries. The Japan Meteorological Agency started its first real-time NWP runs that same year in 1960, the third country to do so, with a single-level barotropic model (Nitta and Saito 2004).

PE models were pursued in other countries. The German Weather Service [Deutscher Wetterdienst (DWD)] introduced one for operational forecasting in 1966, the same year that the USWB introduced its six-layer PE model. The Japan Meteorological Agency introduced a Northern Hemisphere PE model in 1975 with a horizontal resolution of 381 km and four layers. Introduction of diabatic processes, especially driven by deep convection, led to problems: spurious “grid-point storms” (Kasahara 2000; Lynch 2006). In 1972, The Met Office introduced a 10-layer PE model (Bushby and Timpson 1967).

By the middle of this era, multilevel NWP models incorporated components to describe diabatic and viscid effects via early versions of parameterizations of subgrid-scale atmospheric processes including water vapor and its change of phase, convection, radiation, turbulence, and exchange of heat, moisture, and momentum at Earth’s surface. Richardson (1922) recognized very early that successful weather prediction would have to include these additional processes and spent 98 of the 236 pages in his book on them (clouds and water—12, energy and entropy—8, radiation—19, turbulence—36, and surface/soil/sea—23; Lynch 2006). The 1972 Met Office PE model included a fuller set of physical parameterizations, ambitious for its time, including atmosphere–surface exchange, topographic forcing, subgrid-scale convection, and lateral diffusion. The United States introduced in 1971 a higher-resolution (191 km) “local forecast model” [Limited-Area Fine Mesh (LFM); Howcroft 1971], which was a smaller-domain version of its six-layer PE model. It covered North America and had the same simpler physics described by Shuman and Hovermale (1968). Condensation physics was introduced experimentally into models in late 1950s. It was quickly realized (e.g., Kasahara 1961; Lilly 1962) that in a moist, conditionally unstable atmosphere the smallest horizontal scales grow the fastest. A representation of the subgrid impact of latent heat release needed under these conditions led to initial convection parameterizations (Charney and Eliassen 1964; Ooyama 1966; Kuo 1965—based on horizontal moisture convergence and other parameters) and fostered a vast development area that continues today. Readers can find a much more thorough summary of treatments of cumulus and stratocumulus in the monograph chapter on Earth-system models by
<table>
<thead>
<tr>
<th>Year</th>
<th>Dynamic core</th>
<th>Resolution (“L” = no. of vertical levels)</th>
<th>Domain</th>
<th>Data assimilation</th>
<th>Physics/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1958</td>
<td>Barotropic</td>
<td>381 km L1</td>
<td>NH</td>
<td>Cressman 1959</td>
<td>No physics</td>
</tr>
<tr>
<td>1962</td>
<td>Quasigeostrophic</td>
<td>381 km L3</td>
<td>NH</td>
<td></td>
<td>No physics</td>
</tr>
<tr>
<td>1966</td>
<td>Primitive equation</td>
<td>381 km L6</td>
<td>NH</td>
<td></td>
<td>Simple diffusion; radiation</td>
</tr>
<tr>
<td>1974</td>
<td>Primitive equation</td>
<td>2.5° lat/lon L9</td>
<td>Global</td>
<td>Hough analysis (Flattery 1971)</td>
<td>See Stackpole 1978</td>
</tr>
<tr>
<td>1985–86</td>
<td>Global spectral model (GSM)</td>
<td>R40 (300 km) L18</td>
<td>Global</td>
<td>Statistical interpolation</td>
<td>GFDL physics; interactive clouds (Kanamitsu et al. 1991)</td>
</tr>
<tr>
<td>1993</td>
<td>GSM</td>
<td>T126 L28</td>
<td>Global</td>
<td></td>
<td>Pan and Wu 1995</td>
</tr>
<tr>
<td>2005</td>
<td>GSM</td>
<td>T382 (38 km) L64</td>
<td>Global</td>
<td>Gridpoint statistical interpolation</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>GSM</td>
<td>T574 (23 km)—2010</td>
<td>Global</td>
<td>Hybrid 3D-ensemble variational (3DEnVar)</td>
<td></td>
</tr>
<tr>
<td>2012 May</td>
<td>GSM</td>
<td>T574 (23 km)—2010</td>
<td>Global</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015 January</td>
<td>GSM; semi- Lagrangian advection</td>
<td>T1534 (1km) L64</td>
<td>Global</td>
<td>May 2016—4DEnVar</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>FV3 gridpoint cubed sphere</td>
<td>C768 (13 km) L64</td>
<td>Global</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIG. 13-11. Early primitive graphics from a line printer (with contours added manually): (top) 24-h forecast for 1000-hPa heights from gridpoint “contouring,” valid 0300 UTC 24 May 1956, and (bottom) the verifying analysis, also valid at 0300 UTC 24 May 1956. The figure is from the JNWPU and is provided through the courtesy of the NOAA/NCEP Weather Prediction Center and G. DiMego.
FIG. 13-12. Participants of the International Symposium on Numerical Weather Prediction in Tokyo in November 1960. The figure is from Nitta and Saito (2004, their Fig. 3), provided through the courtesy of the Meteorological Society of Japan. (The original figure was inside the front cover of the Proceedings of The International Symposium on Numerical Weather Prediction in Tokyo: November 7–13, 1960; Syono et al. 1962.)
D. Randall and his coauthors (Randall et al. 2019, their section 4a). Representations of boundary layer turbulence using Monin–Obukhov similarity theory began toward the end of this era, for example, by Louis (1979). Margaret Lemone and her coauthors (Lemone et al. 2019) provide a full chapter in this monograph on boundary layer meteorology, including representation of subgrid-scale processes.

These parameterizations in multilevel NWP models allowed direct prediction of fields important to the public, including precipitation and surface (screen level) temperature. NWP centers needed to consider what atmospheric processes could be captured in them, as the decades passed, for improved model prediction. The sophistication of these parameterizations was also limited, of course, by computer power, their effectiveness toward improving skill, and incomplete understanding of physical processes.

Operational NWP took a leap forward with the introduction of global models in the 1970s, the first being an expansion in 1972 (Table 13-3) of the U.S. six-layer PE model to a nine-layer 2.5° latitude/longitude version (Stackpole 1978). ECMWF introduced a similar model, its first operational model, in 1979 with a 1.875° grid and 15 levels. Using efficient Fourier transforms, Robert (1969) developed spectral models. Hemispheric spectral models became operational in Canada and Australia in 1976. The United States introduced a global spectral model (Sela 1980) in 1980 using spectral representation of meteorological fields with rhomboidal truncation at wavenumber 30 (R30) and 12 vertical levels; Eliassen had earlier suggested such a model (PB08). Hemispheric spectral models became operational in Canada and Australia in 1976 (PB08). ECMWF moved to a triangular truncation representation (T106) in 1983 with its Integrated Forecast System. Transformation from spectral space to physical (grid point) space was accomplished by efficient Fourier transforms, allowing physical processes operating in vertical columns (e.g., convection, turbulence, and radiation) to be applied within spectral models. Regional models allowed higher-resolution representation of atmospheric processes. Phillips (1979) tested one by the end of this era, the U.S. Nested Grid Model (NGM). It was implemented in 1985 (Table 13-4). In this era, all models used the hydrostatic simplification for pressure, defining it as the weight of air in the column above without any vertical acceleration.

f. Data assimilation—A central component of NWP

All forecasts, and especially NWP forecasts, are initial-value problems. The process to obtain a best estimate of the three-dimensional, multivariate (temperature, wind, moisture, etc.) conditions on a grid at a given time to initialize an NWP model is a minimization problem known as data assimilation (DA). Vilhelm Bjerknes (1911) gave the three steps for prediction: 1) acquisition of observations, 2) diagnosis or analysis, and 3) prognosis. Today, the second process, DA, integrates all knowledge (observations, physical relationships, and statistics) to obtain a single best estimate of the true three-dimensional state of the atmosphere.

To obtain initial conditions for his early numerical prediction experiment, Richardson (1922) interpolated hand-analyzed maps from an earlier case onto a grid. These upper-air maps had been produced previously by the Leipzig Geophysical Institute under the direction of V. Bjerknes using data from a balloon with self-registering instruments to produce the first isobaric analyses as described by Lynch (2006, his section 6). The U.S. JNWPU used similar methods for initial NWP tests, requiring hand interpolation to a grid from previous hand-analyzed subjective analyses of observations such as shown in Fig. 13-1 (left and right panels) and Fig. 13-3. The United States switched initially to an objective, programmable technique by Gilchrist and Cressman (1954). This technique optimized the fit of a biquadratic function to the observations, thus defining analytically a two-dimensional initial field (e.g., 500-hPa height). Bergthorsson and Dööö (1955) from Sweden were the first to use prior information from a previous forecast. A short-term forecast carries forward in time information from previous observations. This gridded forecast is called a first guess or background. It is a starting point for the current analysis. Bergthorsson and Dööö (1955) interpolated current observations to the model grid and used observation-minus-background differences to produce an increment field of far smaller magnitude than the background, which was then added to the background. This incremental DA design was adopted for U.S. models by Cressman (1959). Sweden (SMHI) used a 24-h forecast of 500-hPa height as a background for an analysis of current conditions. Persson (2005) states that the emergence of objective analysis (replacing hand interpolation) “aroused stronger emotions from the meteorological community [including Tor Bergeron] than the introduction of objective [numerical] forecasts.”

The adaptation of forecasters to changes in automation continued.

Data assimilation came to be considered a problem in maximum likelihood estimation, weighing both observation error and background (prior forecast) error. Daley (1991, his section 1.6) summarizes the goal of the DA cycle: use prior information to improve analysis quality and subsequent NWP accuracy while incorporating improved observation quality control. A formally correct minimization formulation for DA, statistical
<table>
<thead>
<tr>
<th>Year</th>
<th>Model</th>
<th>Resolution and vertical coordinate</th>
<th>Domain</th>
<th>Frequency update/data assimilation</th>
<th>Physics/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>LFM model</td>
<td>191 km L6</td>
<td>North America</td>
<td>12 h</td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>7L PE</td>
<td>191 km L7</td>
<td>NH</td>
<td>12 h</td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>NGM/Regional Analysis and Forecast System</td>
<td>80 km 16L; sigma coordinates</td>
<td>CONUS</td>
<td>12 h; OI—DiMego 1988</td>
<td>Replaces LFM</td>
</tr>
<tr>
<td>1993</td>
<td>Eta</td>
<td>80 km 18L; eta coordinates</td>
<td>North America</td>
<td>12 h; OI</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>RUC</td>
<td>60 km 25L; isentropic-sigma hybrid</td>
<td>Regional-CONUS</td>
<td>3 h; isentropic OI</td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>Meso-Eta</td>
<td>29 km 50L</td>
<td>North America</td>
<td>12 h (0300, 1500 UTC initialization); OI</td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>Meso-Eta</td>
<td>29 km 45L</td>
<td>North America</td>
<td>1 h, isentropic OI; digital filter initialization</td>
<td>Reisner et al. mixed-phase multispecies cloud; 6L soil model</td>
</tr>
<tr>
<td>2002</td>
<td>Eta</td>
<td>12 km 60L</td>
<td>North America</td>
<td>6 h</td>
<td>Ferrier et al. cloud</td>
</tr>
<tr>
<td>2005</td>
<td>RUC</td>
<td>13 km 50L</td>
<td>CONUS</td>
<td>6 h; 3D-var</td>
<td>Thompson et al. (2004) mixed-phase cloud</td>
</tr>
<tr>
<td>2006</td>
<td>NMM replacing Eta (NAM)</td>
<td>12 km 60L; sigma–pressure hybrid</td>
<td>North America</td>
<td>6 h</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>RUC</td>
<td>13 km 50L; isentropic-sigma hybrid</td>
<td>CONUS</td>
<td>1 h; radar reflectivity assimilation</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>NMM</td>
<td>4 km 60L nest inside NAM</td>
<td>CONUS</td>
<td>6 h</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Rapid Refresh—WRF-ARW dynamic core</td>
<td>13 km 50L</td>
<td>North America</td>
<td>1 h; radar reflectivity assimilation</td>
<td>Similar physics to RUC</td>
</tr>
<tr>
<td>2014</td>
<td>HRRR</td>
<td>3 km L50; based on WRF-ARW</td>
<td>CONUS</td>
<td>1 h; gridpoint statistical interpolation with radar reflectivity latent heating and cloud nonvariational DA</td>
<td>Thompson et al. (2008); see Benjamin et al. (2016a)</td>
</tr>
</tbody>
</table>
interpolation (Eliassen 1954; Gandin 1965), contributed to improved NWP forecast accuracy from the late 1970s into the 1980s.

Statistical interpolation [sometimes called *optimal interpolation* (OI)] allowed objective use of information on observation errors, forecast errors, and spatial correlation of both. In particular, methods to estimate the spatial covariance of short-range forecast errors were found to be particularly important for improving analysis quality and forecast accuracy. In the United States, Bergman (1979) used simple geostrophic relationships to model the spatial correlations between wind and height errors in the background field.

Application of an initialization step (e.g., normal mode initialization; Williamson and Temperton 1981) to modify three-dimensional analysis grids by forcing initial balance was found to be effective in avoiding excessive gravity wave generation in the beginning of the forecast that rendered it unusable (Lynch 2006). The United States used a Hough analysis initialization (Flattery 1967) to force balance between wind and geopotential fields by zonal and meridional wavenumbers (Stackpole 1978). Other initialization methods are discussed in the section on the next era. Further information on development of DA during this era is provided by Daley (1991) and Kalnay (2003).

Rawinsondes were the key observations assimilated into NWP analyses in this era. Polar-orbiter satellite data were available, and it was thought in the United States in 1969 that retrieved height profiles would suffice for global assimilation (R. McPherson 2018, personal communication). This perception launched the National Environmental Satellite, Data, and Information Service into a multidecadal effort to produce retrieved soundings from radiances, but in the 1990s direct assimilation of radiances was found to be a far more effective use of satellite data, as discussed in the next era.

Although no ensemble forecasting with multiple different predictions valid at the same time actually took place during the 1956–85 era, the table was inadvertently set by numerical precision experiments performed by Lorenz in 1960 (Lorenz 1960, 1963, 1993) as described by Kalnay (2003, pp. 205–208). Using a simple dynamical model, Lorenz found that truncating the last significant digit in real-number initial conditions caused the forecast to diverge from one that included all significant digits in the initial conditions. The two forecasts were completely uncorrelated after two weeks of integration. Lorenz (1963) later wrote about the limits of predictability in long-range forecasts, which, in turn, spawned ideas about ensemble forecasting. Operational global ensemble NWP forecasts by the United States and ECMWF both started in the same year, 1991.

The skill of large-scale NWP models was first assessed by comparing predicted and observed height gradients on the 500-hPa pressure surface by means of the S1 score [Teweles and Wobus (1954), described further by Kalnay (2003, pp. 2–3)]. A steady increase in accuracy of 500-hPa wave patterns during this era from 1956 to 1985 is evident in S1 scores for NCEP 36- and 72-h forecasts that are shown in Fig. 13-10. NCEP also tracked precipitation skill with the threat score (Fig. 13-13) from its NWP-augmented and forecaster-issued forecasts over 24-h periods from days 1 and 2 from 1964 onward. The threat score was calculated as $C/(F + O - C)$, where $F$ is the total area enclosed by 1-in. (1 in. = 2.54 cm) predicted isohyets, $O$ is the total area enclosed by 1-in. observed isohyets, and $C$ is the total area included within both (the area where the forecast for 1 in. or more precipitation was correct). This metric shows a modest increase starting in the mid-1970s and a sharper increase for day-3 precipitation forecasts after their assessment began in 1999.

During this era, the forecasting community asked this question, “What is a good forecast?” This question could be considered at two tiers—first, a verification problem for forecasters and, second, at a broader level: does the forecast help the user? These tiers correspond to “goodness types” 2 and 3 in the essay on this topic by Murphy (1993). Public forecasts increasingly presented probability of precipitation, and investigators showed that the percentage of events occurring should roughly match the forecast probability, or reliability, summarized by Sanders (1963), Murphy and Winkler (1977), and others. At the same time, it was shown that the sharpness of those forecasts (distinguishing between yes and no events) was another important forecast attribute: using the same forecast of climatology each day might have perfect reliability but would be useless (i.e., essentially zero sharpness). Partly to help to instill these concepts into students with a synoptic bent and to give them an opportunity to put classroom learning into practice, several university meteorology departments, including MIT and The Pennsylvania State University, began informal forecast contests.

Initial statistical postprocessing using NWP output, trained by multiple years of observation-model pairs, was introduced as “model output statistics” (MOS) in the United States (Glahn and Lowry 1972; Klein and Glahn 1974). Evaluation of probabilistic forecast skill was important in the 1970s, because statistical postprocessing, as an add-on to NWP, forced continual reassessment of how the human–machine mix could produce the best forecast (e.g., Snellman 1977). Introduction of MOS resulted in enhanced concern about the role of the forecaster in the United States.
At the broader level of assessing forecast value for users, Myers (1971) examined objective methods for decision-making from probabilistic forecasts by examining the gain/loss for correct/incorrect forecasts in specific applications such as protection of materials for a construction company. Private forecasters now produce such statistically postprocessed guidance for a vast number of decision scenarios, but this was less common in the early part of this era. Forecast prediction in this period emphasized large-scale variables (e.g., 500-hPa height anomaly correlation) and precipitation. Although forecast and NWP improvements have been based on hypotheses with subsequent tests as mentioned at the beginning of this chapter, statistical comparisons did not then and still do not commonly show significance testing, as noted back in 1968 (Panofsky and Brier 1968).

**g. Communication and public perception of forecasts**

Public perception of meteorology as an important science affecting everyday life grew during this period. Initial television images from the TIROS spacecraft gave the first visual evidence of the complexity of cloud and weather patterns and therefore, hinted at the challenges faced by meteorologists. Henson (2010) describes the key role of television in transferring weather information to the public. The level of scientific content in broadcast meteorology slowly increased as gimmickry receded. The public desired effective communication of weather information to plan their day or week ahead. AMS took steps to increase professionalism in broadcast meteorology (Seitter et al. 2019). Many new private weather companies came into existence and grew in the United States; they used publicly available data (e.g., NWS observations and maps) to tailor forecasts for specific users. These included AccuWeather (started in 1970 and focused initially on newspapers and radio), WSI, Baron Services, Weather Central, and many others. (Henson 2010). The availability of cable television led to the start of The Weather Channel, a weather-only cable channel, in 1982. A similar Weather Network in Canada began television broadcasts in 1988. The weather information available to the public proliferated with these new venues, but the information was not always consistent, nor was the source always known.

**5. 1985–2018 (Era 4)—The modern era**

**a. State of forecasting**

In this most recent era, NWP unquestionably became dominant in the forecast process. The role of
forecasters relative to NWP evolved, and, by the 2000s with the steady improvement of NWP models, forecasters spent less time modifying forecast grids but concentrated more on communicating and interpreting NWP guidance to nonmeteorologists for decision support. The role of the private sector as communicators also increased substantially, as did the application of environmental prediction to human activities.

The key transformative events from the mid-1980s to the mid-1990s included the introduction of greatly improved DA methods and more effective global models. For global models in particular, the introduction in the 1990s of direct assimilation of satellite radiances instead of rawinsonde-like retrieved soundings was a major leap forward. Far more effective regional models became available, aided by improved high-frequency observations, including radar and satellites, innovative DA techniques to use those observations, and development and refinement of nonhydrostatic models with sophisticated cloud and boundary layer physics. Forecasters witnessed the end of widespread use of paper maps for forecasting and the introduction of computer-based displays. As mentioned above, NWS offices in the United States procured their first AFOS workstation around 1980. AFOS was central to forecast operations until about 2000. All of these milestones will be discussed below in greater detail.

Early progress during this era in global model accuracy was evident in a successful 5-day forecast of a rapid intense cyclogenesis, resulting in a major March 1993 U.S. East Coast snowstorm, unobserved 5 days earlier, as described by Uccellini et al. (1995). The U.S. global model, along with the ECMWF and Met Office global models, all showed potential for a major storm (Fig. 13-14). Consistency among consecutive 5-, 4-, and 3-day forecasts from the U.S. global model (indicated by “M” in Fig. 13-14) of a significant East Coast snowstorm led the NWS and U.S. media to issue early outlooks on possible severe impacts from this storm. The positive societal benefit in forecasting such a major weather event 3–5 days in advance was very unusual for 1993. In this case, forecasters and the media coordinated closely to issue watches and warnings for excessive snowfall and blizzard conditions well ahead of the storm, a major contrast from the U.S. blizzards of 1888, which were completely unanticipated. They communicated with the public effectively about the impending hazards. Such watches and warnings became increasingly accurate for severe weather events as this era progressed.

b. Observations used for forecasting

Huge investments in weather satellites occurred during this era to the benefit of global medium-range NWP and also for short-range NWP and human forecasting. A summary of weather-monitoring satellites is provided in this monograph by Ackerman et al. (2019). Rain-sensitive microwave imager data became available after 1992. Key milestones occurred in 1998, when ATOVS [consisting of the Advanced Microwave Sounding Unit/Microwave Humidity Sounder (AMSU/MHS) and the improved High Resolution Infrared Radiation Sounder (HIRS)] sounding data were introduced; in 2002, when hyperspectral infrared sounding data became available; and in 2006, when global navigation satellite system (GNSS) [global positioning system (GPS)] radio-occultation data were introduced (Ackerman et al. 2019). Very accurate precipitable water estimates derived from ground-based GPS timing delays also became available (e.g., Gutman et al. 2004). Satellite-based atmospheric motion vectors (AMVs; Nieman et al. 1997) derived from cloud-drift motion added to radiance-based information for forecast improvements. Buehner et al. (2018) recently studied the impact of various observation systems on the skill of 24-h global forecasts assessing impact with several recent approaches. They showed the most impact from AMSU radiances, scatterometer data, aircraft, AMVs, and rawinsondes, with lower but substantial contributions evident from many other instruments.

Rawinsondes continued to be a key component of the global and regional observation networks despite tremendous increases in global satellite data and commercial aircraft observations. Rawinsonde data were shown to be trailing only polar-orbiter radiances (AMSU) and AMVs in importance for global forecast impact (Gelaro et al. 2010) and trailing only aircraft, satellite cloud, and surface data for regional impact (James and Benjamin 2017). The number of global daily rawinsonde profiles was relatively constant during this era but with a slight decrease because of budgetary problems in some countries, where only one launch per day could be made instead of two.

The second critical remote observation technology—radar—dramatically improved short-range forecasts but affected forecasting at longer time scales as well, because radar broadened understanding of cloud and precipitation systems. The U.S. NOAA NWS procured a new radar, Weather Service Radar-1988 Doppler (WSR-88D) in 1988. The first production unit was installed in 1990. An upgrade to dual-polarization capability occurred from 2011 to 2013 (https://www.roc.noaa.gov/WSR88D/PublicDocs/NEXRAD.pdf, accessed 30 April 2018). The WSR-88D was a key component of the U.S. NWS modernization program. It introduced Doppler-sensed radial velocity and transformed severe weather and heavy precipitation forecasting. It
FIG. 13-14. Surface forecasts for the March 1993 U.S. East Coast snowstorm: the 5-, 4-, and 3-day manual surface fronts and isobar forecasts issued by the NCEP (NMC) Weather Forecast Branch on the basis of the 132-, 108-, and 84-h U.S. global and Met Office model forecasts and the 144-, 120-, and 96-h ECMWF forecasts. The forecasts are valid at 1200 UTC (left) 13 and (right) 14 Mar. The U, E, and M notations respectively refer to Met Office, ECMWF, and U.S. global (Medium-Range Forecast) forecast positions of the surface cyclone, with the last two digits representing central pressure (e.g., E94 = 994 hPa from the ECMWF). Solid, dashed, and dot–dashed lines in the right column refer to forecast tracks of the surface cyclone from their respective models. The figure is taken from Uccellini et al. (1995), their Fig. 3.
allowed forecasters to examine convective storms for potential rotation and was central to improving NWS warnings for tornadoes, severe winds, and hail. Collection and merging of all radar data in a central facility were accomplished through new technology (Kelleher et al. 2007). This was critical for subsequent national-scale radar DA; it ended reliance on local forecast offices as separate distribution points for high-resolution radar data. Assimilation of radial velocity data was found to be essential for very-short-range storm-scale prediction (e.g., Dowell et al. 2011). Dual-polarization radar data (e.g., Vivekanandan et al. 1999), allowing detection of different hydrometeor types and improved quantitative precipitation estimates, are only now being considered for advanced assimilation techniques. One such parameter is differential reflectivity, which is non-zero for rain drops, is near zero for ice (sleet, graupel or hail), and is variable for different types of snow crystals.

Near the start of this era, automated commercial aircraft observations became available for NWP and monitoring of local conditions (Moninger et al. 2003). In the United States, an FAA-led task force, along with the NWS, recommended key steps toward removing major deficiencies (especially in upper-level winds) in aviation weather forecast guidance (AWFTF 1986; McCarthy 2001). Air–ground communications, previously used for airline management and transmission of engine-performance data, were now to be used also for automated reporting of meteorological data. This was a vast improvement over the less frequent, sometimes inaccurate, manually transmitted pilot reports. Aircraft data had been previously recognized as the only in situ source of upper-air observations other than those from rawinsondes (recall special aircraft flights from Era 1 before the invention of rawinsondes). The accuracy of aircraft-based wind observations significantly improved with the introduction of a more accurate navigation system, starting in the mid-1980s (laser gyroscopes and, more significant, GPS). The number of global automated aircraft reports in the early 1990s, 7000 day⁻¹ (Benjamin et al. 1991), increased to 170000 day⁻¹ by 2002 (Moninger et al. 2003) and to over 800,000 by 2017 (https://www.wmo.int/pages/prog/www/GOS/ABO/data/statistics/aircraft_obs_cmc_monthly_ave_daily_reports_by_program.jpg). The number of automated aircraft observations has tripled in the 2010s alone, which fact is a major contribution to global and regional forecast skill (e.g., Buehner et al. 2018; James and Benjamin 2017).

Almost all observations needed for human forecasting and NWP up to the current time have been paid for by the government, the most notable exception being commercial aircraft data. Some governments pay for transmission of these data as well. Provision of aircraft reports to the government in return for more accurate aviation forecasts has been a unique quid pro quo, a private–public agreement that has had a strong positive effect on forecast accuracy. The global community has so far been able to meet NWP-center exchange requirements set forth in WMO Resolution 40 for commercial aircraft observations. Various national meteorological and hydrological services around the world work with hundreds of private vendors to obtain instruments for all of their observations.

Improved surface observations, many from state and local governments, gave more detailed information for defining small-scale features during this era. The state government of Oklahoma started a mesonetwork with over 100 stations in the early 1990s (Crawford et al. 1992). The University of Utah, working with the NWS, organized a collection of mesonet data from the western United States (Horel et al. 2002).

c. Science understanding enabling forecasting

The QG omega equation, derived by combining the QG vorticity equation and the thermodynamic equation (e.g., Carlson 1982), was used extensively by forecasters (Trenberth 1978), especially in the 1980s and 1990s. Forecasters in the United States used the QG omega equation to diagnose vertical motion from model output using 500-hPa 1) vorticity advection and 2) Laplacian of thermal advection. An alternative formulation in terms of so-called Q vectors was introduced by Hoskins et al. (1978) and also became popular, including diagnostics for rapidly developing oceanic extratropical cyclones (“bombs”) described by Sanders and Gyakum (1980). Merged use of satellite imagery and Q diagnostics from analyses and model-forecast grids [demonstrated by Astling (1976)] came into common use after the introduction of the Man Computer Interactive Data Access System (McIDAS: Suomi et al. 1983) and AWIPS workstations. The notion of potential vorticity conservation furthered understanding of frontal circulations (e.g., Hoskins 1971; Shapiro 1978; Hoskins et al. 1985); it became widely used for model diagnosis and forecasting. These studies clarified the importance of latent heating, arising from precipitation and frictional effects, on frontal circulation. Understanding of the three-dimensional cyclone/frontal structure was extended (e.g., frontal fracture, frontal T-bone, and bent-back front) into the Shapiro–Keyser cyclone model (Shapiro and Keyser 1990). More recent summaries of extratropical cyclone evolution have been provided by Shapiro and Grønås (1999) and Schultz et al. (2019, in this monograph).

Understanding of mesoscale convective systems (MCSs) continued to improve during this era (Houze 2019). More thorough conceptual models guided both
forecasts and model developers in this period, both for use of observations to assess current atmospheric conditions and for model evaluation. Scientists gained improved knowledge about midlatitude warm-season precipitating systems including diurnal cycles (e.g., Carbone and Tuttle 2008). Forecasters began to use understanding of tropical convective systems such as the Madden–Julian oscillation and tropical cyclones for assessing global forecast model evolution. Better understanding of convective storm structure (rotation, line structures, etc.; Moller et al. 1994; Brooks et al. 2019) became routine for forecasters. Doppler radar became critical for forecasting during this era, providing the ability to observe storm rotation. Improved satellite data, especially geostationary images, also became essential for understanding storms and their prediction.

Forecasting in regions where terrain strongly influences weather patterns also has benefited from theoretical advances during this era. There is a long history of research in mountain meteorology and terrain-induced flows, but much of this work through Era 2 and most of Era 3 was observational in nature and specialized to particular regions such as the Alps and the western United States, where terrain effects dominate the day-to-day weather. For example, in Era 3, observations from research aircraft and at the surface cast new light on the mechanics of downwind storms in the lee of the Rocky Mountains (Lilly and Zipser 1972). Forecasting of mountain weather depended on local experience and pattern recognition.

With advances in modeling technology that came with increased computing power, theoretical work expanded to take advantage of the ability to describe nonlinear dynamics, including flow blocking, mountain waves, and downwind and gap-flow windstorms, through numerical simulation. By early in Era 4, this work had led to a measure of consensus on the dynamics of these phenomena as a result of the work of Douglas Lilly, Terry Clark, Joseph Klemp, Dale Durran, Ronald Smith, and others (see Durran 1990 and references therein). Partly through education of NWS forecasters through the COMET program (the founding of which was mentioned earlier in the 1956–85 section) beginning in the early 1990s, this understanding has been translated into guidance for forecasters in interpreting operational model output to infer terrain influences on sensible weather. With the recent advent of operational high-resolution nonhydrostatic models such as the U.S. 3-km High-Resolution Rapid Refresh (HRRR; Benjamin et al. 2016a), accurate predictions of mountain waves, upstream blocking, and gap flows as well as realistic precipitation forecasts in mountainous areas are now being produced by operational NWP models.

d. Community of forecast providers

As the science of NWP advanced, the role of forecasters changed. During the previous 1956–85 era, forecasters relied on observations, upper-level guidance, and experience to issue worded forecasts. With the arrival of workstations (e.g., AFOS and AWIPS in the United States, and similar tools in Canada (Sills 2009)), forecasters had tools to work directly with gridded data (e.g., the Interactive Forecast Preparation System (IFPS); Ruth 2002). AWIPS was a centerpiece of the NWS modernization, allowing forecasters to view on a single system in situ observations, model data, and satellite and radar images, separately or in combination by means of overlays (MacDonald and Wakefield 1996). IFPS initially used digital data only at surface observing sites but later changed to a gridpoint-based version that could be modified by forecasters on AWIPS using a Gridded Forecast Editor. From the gridded forecast data, text, voice, and graphical products could be produced automatically. In recent years, the human–machine mix has shifted more to forecasters spending their time interpreting and communicating forecast information (R. McPherson 2018, personal communication).

The availability of frequently updated radar and geostationary satellite information in this era, along with expanded surface data and introduction of forecaster workstations like AWIPS, allowed forecasters to focus on very-short-range extrapolation of current conditions (by definition, not NWP), now formally called “nowcasting” (Browning 1982). Nowcasting in more recent times is similar to that of the first 1919–38 era in that it is based on translation with growth or decay, but now with far more advanced and detailed observations. At the same time, very-short-range NWP skill also increased because of effective assimilation of these same observations, as noted by Mass (2012) in his more recent summary on nowcasting.

Early in this era, U.S. forecasters in offices did not trust centralized products (https://esrl.noaa.gov/gsd/eds/gfesuite/techreview/TechReview2003.pdf). This evolution of the role of forecasters, cited early by Snellman (1977), Bosart (2003), and others, has continued to the present time. The consolidation into 123 NWS offices in the 1980s under director Hallgren (reported under the previous 1956–85 era) took enormous persuasion, especially since the closed offices all had local Congressional districts whose representatives usually opposed the consolidation (R. McPherson 2018, personal communication). Collocation of the new full-service NWS offices with universities and establishment of the COMET
training program (Johnson et al. 2015) with its first director in 1989, William Bonner, underscored the increased interaction between NWS and the academic community in this era. Recent years have brought improved automated decision-making capability and improved NWP skill, and both public and private forecasters are spending less and less time editing gridded model output (e.g., Neilley et al. 2015).

In 1995, after its field office modernization and after Elbert “Joe” Friday replaced Hallgren as director in 1988 (Table 13-1), the U.S. NWS consolidated its National Meteorological Center where its NWP models were developed and run into a broader NCEP, including NCEP’s Environmental Modeling Center (EMC), the SPC, the Aviation Weather Center (AWC), and other centers for prediction of hurricanes, climate, marine conditions, space weather, and general weather/precipitation. The organizational consolidation was not accompanied by geographical collocation, and many of the centers remain distributed around the United States.

e. Forecasting applications

Aviation continued as a particularly important application requiring improved forecasts, especially in this era. In 1985, the Aviation Weather Forecasting Task Force (AWFTF 1986) articulated a need for more accurate enroute forecasts to 1) minimize fuel consumption, 2) minimize weather-related delays, and 3) improve safety in the air through avoidance of hazardous weather (thunderstorms, icing, and turbulence). Aviation’s need for improved forecasts for domestic flight planning (2–8 h) was the primary driver in the United States for establishment of an hourly updated model in the United States in 1998 [Rapid Update Cycle (RUC); Benjamin et al. 2004a] and, in 2014, the 3-km convection-allowing HRRR model with radar reflectivity assimilation (Benjamin et al. 2016a). Other transportation needs such as road weather also became acute for better forecast guidance (Haupt et al. 2019b) in recent years. Improved public response to severe weather outlooks (“watches” in the United States) and warnings has also driven implementations of convection-allowing models (as shown in Table 13-8, described below).

The energy industry has become increasingly tied to weather prediction, especially by high-resolution NWP models, both for best possible estimates of variable power generation from wind and solar sources and for anticipated loads (power needs). Additional needs for the best possible precipitation forecasts, beyond those just for the general public, have been forced by hydrological applications (Peters-Lidard et al. 2019, in this monograph). Construction, food-related (agriculture), and health-related (Rogers et al. 2010) decisions have become very closely tied to weather forecasting especially from NWP automated predictions and ensemble-probabilistic forecasts. Forecasts of aerosols, overall air quality, and smoke (Grell and Baklanov 2011) with off-line and online chemistry applications have been increasingly important applications during this era. Some of these applications have been made possible by coupling of two or more component models (e.g., atmospheric, chemistry, ocean, cryosphere, biosphere, and others) into more general Earth-system models as discussed at the end of this section and in the next looking to the future. Additional information on current applications for meteorology is provided in Haupt et al. (2019a,b,c).

f. Numerical weather prediction

The requirements for NWP models of this era extended far beyond emphasis on extratropical cyclones and cold-season phenomena as in the last era and now moved toward improved precipitation and near-surface forecasting, convective storm forecasts, transportation forecasts, and certainly extension of effective multiday global forecasts. With the maturation of digital computing and arrival of the Internet, the possibilities for enhanced digital environmental prediction data exploded in this period. Requirements for probabilistic prediction information motivated further development of ensemble prediction. Requirements for emergency response and even automated decision-making (e.g., air-traffic management) motivated more frequent updating of NWP models (also allowed by more frequent observations).

Operational NWP centers introduced numerous effective improvements to global and regional models that are dizzying when viewed over the 33-yr span of this period. Global operational models moved to higher horizontal and vertical resolution. Higher vertical resolution was accompanied by higher top levels to allow more effective satellite radiance assimilation and improved representation of stratospheric dynamics.

Forecast users wanted longer-period outlooks, achievable only with global models. They wanted skill at 5 days eventually to match then-current skill at 3 days. Global forecasts were extended to 10 days and longer, but it was found that accurate extended forecasts required accurate representation of tropical convection including Madden–Julian (Madden and Julian 1971, 1972) oscillations (Zhang 2005) and recurving tropical cyclones. Longer-range outlooks for seasonal to subseasonal forecasting became important. The U.S. Weather Bill of 2017 (https://www.congress.gov/bill/115th-congress/house-bill/353) expressed a need for improved subseasonal forecast accuracy. A goal to use the same global model for both weather and climate prediction began to be realized during this era, the earliest example being the Met

Improved forecasts of severe local weather and sudden weather changes gained a much higher priority during this era. Requirements for public safety and economic efficiency drove the need for high-resolution, high-frequency forecasts. Frequent, three-dimensional, in situ (e.g., aircraft) and remote sensing observations proliferated and thus supported development of systems for rapidly updating short-range forecasts, starting in the 1990s. Better understanding of high-impact weather events allowed incorporation of detailed, relevant physics into models. Regional, high-resolution NWP, including nonhydrostatic models, provided an early testing ground for forecasting such local events prior to this capability being realized in future operational global convection-allowing models.

By 2018, at least 14 NWP centers were running global weather prediction models (Table 13-5), generally with horizontal resolution of 20 km or finer. NCEP and ECMWF both used spectral models for global prediction applications through this entire period. These applications included deterministic, ensemble, and even seasonal predictions. In 1991, the United States increased resolution from R40 (~300 km) to T126 (~110 km) (common spectral model grid resolutions are provided here: https://climatedataguide.ucar.edu/climate-model-evaluation/common-spectral-model-grid-resolutions; see also section 4c) in its global spectral model with 18 vertical levels (Table 13-3). It also installed a new DA system (more below) using statistical interpolation directly with spherical harmonics. International NWP centers periodically increased computer power to improve their overall forecast skill by 1) capturing details through high spatial resolution (horizontal and vertical), 2) increasing the sophistication of parameterized physical processes (e.g., clouds and turbulence), 3) improving DA, 4) instituting ensemble forecasting, 5) issuing short-range predictions more frequently, 6) extending forecasts to subseasonal or seasonal (from 3 weeks to 9 months), and 7) reforecasting to allow more accurate statistical postprocessing. Highlights in the history of improvements to the U.S. NCEP deterministic global model are shown in Table 13-3, including key changes in resolution and DA.

g. Innovative computational grids

As pointed out by Randall et al. (2019), quasi-uniform icosahedral (Williamson 1968, Sadourny et al. 1968) and cubed-sphere grids (Sadourny 1972) were developed decades ago for general circulation models. But not until about 15 years ago were such innovative global horizontal grids considered for global weather forecasting. Their principal advantage is to better load-balance global model solutions geographically. Perhaps the best example of quasi-uniform horizontal grid structures used in operational forecasting is the icosahedral triangular grid now used by Germany in its global model (Table 13-5). Experimental icosahedral models were developed in Japan [Tomita and Satoh 2004; Satoh et al. 2008—Nonhydrostatic icosahedral atmospheric model (NICAM)] and the United States [Skamarock et al. 2012—Model for Prediction Across Scales (MPAS); Bleck et al. 2015—the finite-volume Flow-Following Icosahedral Model (FIM)]. Stretched global grids with smoothly varying mesh transitions, avoiding the need for lateral boundary conditions, were developed in Canada (Côté 1997; Côté et al. 1998) and France (Courtier and Geleyn 1988). More uniform Cartesian grids (yin-yang) were later applied by Canada (Qaddouri 2011; Qaddouri and Lee 2011). As of this writing, NOAA/NCEP is readying a global gridpoint, finite-volume cubed-sphere model (“FV3”; Lin and Rood 1996; Lin 2004) with nonhydrostatic capabilities for replacement of its global spectral model, planned for implementation in 2019 (Table 13-3). Miyamoto et al. (2013) first succeeded with a global, cloud-resolving, simulation at a uniform horizontal resolution of 870 m. Still more advanced experimental global model grids are now in testing, such as the spectral element model developed by the U.S. Navy (Miller et al. 2018), which appears to be better suited for future exascale computing environments with 10^6–10^7 processors.

The sigma vertical coordinate (Phillips 1957) was long employed to minimize complications in lower-boundary treatment of orography. A hybrid sigma–isobaric vertical coordinate introduced by Arakawa and Lamb (1977) became more popular in this era to reduce truncation error in calculation of the horizontal pressure-gradient force. Vertical resolution increased to better represent boundary layer processes, clouds, and frontal zones. A hybrid isentropic–sigma vertical coordinate that minimizes cross-coordinate vertical transport and related numerical diffusion was applied in the NCEP regional RUC model (Bleck and Benjamin 1993; Benjamin et al. 2004b) and the NOAA experimental global FIM (Bleck et al. 2015).

h. Parameterizations of subgrid-scale processes

Physical parameterizations to capture subgrid-scale features became much more sophisticated during this era and were applied even with far higher horizontal resolution down to 1–3 km (Table 13-8, described below). Back in the 1970s, atmospheric modelers began to introduce a new approach to subgrid-scale representation of turbulence using higher-order closure [e.g., Mellor and Yamada 1974; Randall et al. 2019, their section 5c(1)]. Higher-order closure used equations that described statistical moments of subgrid-scale variables,
### Table 13-5. Current status of deterministic global models run at international NWP centers (resolution and number of layers). Additional acronym definitions can be found online (https://www.ametsoc.org/PubsAcronymList). The table is adapted from the WGNE (http://wgne.meteoinfo.ru/nwp-systems-wgne-table/wgne-table/).

<table>
<thead>
<tr>
<th>Forecast Center (Country)</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECMWF (Europe) spectral</td>
<td>Tco1279 L137 (~9 km)</td>
<td>Coupled ocean–atmosphere 0.25 Tco1279 L137 (~9 km)</td>
<td>Coupled ocean–atmosphere 0.25 Tco1279 L137 (~9 km)</td>
</tr>
<tr>
<td>Met Office (United Kingdom) grid point</td>
<td>UM 17 km L70</td>
<td>UM 10 km L70</td>
<td>UM 10 km L70</td>
</tr>
<tr>
<td>Météo France (France) stretched spectral</td>
<td>TL1198c2.2 L105 (8 km for western Europe; stretched elsewhere)</td>
<td>~TL1798c2.2 L105 (5 km for western Europe; stretched elsewhere)</td>
<td>~TL1798c2.2 L105 (5 km on western Europe)</td>
</tr>
<tr>
<td>DWD (Germany) icosahedral grid</td>
<td>13 km L90 (6.5 km for Europe)</td>
<td>13 km L90 (6.5 km for Europe)</td>
<td>13 km L90 (6.5 km for Europe)</td>
</tr>
<tr>
<td>Hydrometeorological Research Center of the Russian Federation (HMC; Russia) grid point</td>
<td>(0.16°–0.24°) × 0.225° L51 T339 L31</td>
<td>(0.16°–0.24°) × 0.225° L60 T339 L63</td>
<td>0.1° × 0.08° L90 T679 L63</td>
</tr>
<tr>
<td>NCEP (United States)</td>
<td>GSM T1534 L64 (13 km)</td>
<td>GSM T1534 L64 (13 km)</td>
<td>FV3-cubed-sphere grid C768 L64 (13 km) T1000 L100 (~12 km)</td>
</tr>
<tr>
<td>Navy/Fleet Numerical Meteorology and Oceanography Center (FNMO/NRL (United States) spectral</td>
<td>T425 L60</td>
<td>T681 L80</td>
<td></td>
</tr>
<tr>
<td>CMC (Canada) grid point</td>
<td>Yin-yang (0.23° × 0.23°) L80 coupled atmosphere-ocean–ice</td>
<td>Yin-yang (0.14° × 0.14°) L84 coupled atmosphere-ocean–ice</td>
<td>Yin-yang (0.14° × 0.14°) L84 coupled atmosphere-ocean–ice</td>
</tr>
<tr>
<td>Centro de Previsão de Tempo e Estudos Climáticos/Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE; Brazil) spectral</td>
<td>T666 L64 (~20 km)</td>
<td>9 km L64</td>
<td>9 km L64</td>
</tr>
<tr>
<td>JMA (Japan) spectral</td>
<td>Global/Regional Assimilation Prediction System (GRAPES) (0.25°, L60)</td>
<td>GRAPES (0.25, L60)</td>
<td>GRAPES (0.15, L90)</td>
</tr>
<tr>
<td>China Meteorological Administration (CMA; China) gridpoint</td>
<td>UM 17 km L70</td>
<td>UM 12 km L70</td>
<td>UM 12 km L120</td>
</tr>
<tr>
<td>Korea Meteorological Administration (KMA; Korea) grid point</td>
<td>17 km L70</td>
<td>10 km L70</td>
<td>10 km L70</td>
</tr>
<tr>
<td>National Centre for Medium Range Weather Forecasting (NCMRWF; India) grid point</td>
<td>UM 12 km L70</td>
<td>UM 12 km L70</td>
<td>UM 12 km L120</td>
</tr>
<tr>
<td>BoM (Australia) grid point</td>
<td>UM 12 km L70</td>
<td>UM 12 km L70</td>
<td>UM 12 km L120</td>
</tr>
</tbody>
</table>
including variances and covariances for wind components and temperature and explicit cloud variables. These schemes were critical for overall improved NWP accuracy since they controlled fluxes of heat, moisture, and momentum between Earth’s surface and atmosphere.

Even with the advent of nonhydrostatic models, efforts to improve the parameterization of deep convection continued, particularly in global models, where finer horizontal grid spacing still precluded explicit representation of convection. Forecasts of tropical convective clusters and large areas of convection associated with the Madden–Julian oscillation were poor. It became evident that effective representation of organized deep convection in the tropics and midlatitudes (e.g., Rodwell et al. 2013) was essential for improved 5–7-day (and longer) forecast skill. Expressing accurately the release of convective available potential energy (CAPE) by subgrid-scale plumes has arguably been the most important parameterization for forecasting (e.g., Stensrud 2007, his section 6).

In the United States, operational NWP models generally used a single variable to represent cloud condensate under stable conditions (e.g., Zhao and Carr 1997; Ferrier et al. 2002). Toward the end of this period, some models began to introduce stable cloud parameterizations with multiple cloud hydrometeor species. NCEP started this with the RUC regional model in 1998 with the Reisner et al. (1998) bulk, mixed-phase scheme, later replaced by versions of the Thompson et al. (2004, 2008) and Thompson and Eidhammer (2014) mixed-phase scheme. This sophistication in cloud-precipitation processes contributed to more specific prediction on winter weather hazards, including type of precipitation: snow, rain, freezing rain and sleet (e.g., Benjamin et al. 2016b), replacing nonexplicit snow–rain estimates such as using the 540-dam 1000–500-hPa thickness isoline.

Land surface/vegetation/snow treatments in operational NWP models went from crude flux estimation models at the beginning of this era to multilayer soil models with as many as nine layers and multilayer snow parameterizations. Boundary layer schemes using higher-order closure added complexity, going from simple specifications of surface drag and algebraic expressions for vertical eddy exchange coefficients to prediction of eddy kinetic energy and derivation of exchange coefficients based on second-order turbulence closure models such as formulated by Mellor and Yamada (1982). More recently, a more physically based treatment of the daytime convective boundary layer has been introduced, in which vertical transport of heat, momentum, and moisture by thermals also is represented (so-called eddy-diffusivity—mass-flux, or EDMF, schemes). Observations and large-eddy simulations with resolution down to 100–10 m indicate a need for gridscale-independent parameterizations of cloud and turbulence processes. Effective “scale aware” parameterizations remain a substantial challenge, limiting forecast skill. In this monograph, the reader will find more detailed discussion of these physical parameterizations in Randall et al. (2019) and, specifically of boundary layer parameterizations, LeMone et al. (2019).

### i. Data assimilation

Since NWP began in 1956 at the beginning of the previous era, research on DA has expanded rapidly—even more rapidly than that on atmospheric models. As late as 1980, DA was still considered a necessary but auxiliary component of NWP—that assessment changed quickly with far more sophisticated and far more effective DA. Indeed, over the past 10–15 years, the computer resources dedicated to DA have begun to significantly exceed those for deterministic NWP forecasts at most major NWP centers. This reflects the enormous importance of obtaining optimal initial conditions and the complexity of information embedded within the observations. In 2018, about 4 times as many computing resources were devoted to global DA as to global deterministic models at ECMWF (P. Bauer, ECMWF, 2018, personal communication). With the advent of statistical interpolation, analysis schemes moved from empirical to theoretically based techniques. Variational assimilation (Sasaki 1958; Lewis 1972; Lorenc 1980, 1986; Paillleux 1990; Derber et al. 1991; Parrish and Derber 1992) was built into operational models in the 1990s, including an option for direct analysis onto spectral coefficients. Three-dimensional variational (3D-var) analysis was implemented at NCEP in 1991 and later at ECMWF. This “best estimate” variational formulation requires an iterative process to minimize a cost function \[ J \] in Eq. (3) below with terms expressing the relative fit to current observations and background fields (“vectors,” i.e., three-dimensional grids of all prognostic variables) containing information from prior observations:

\[
J = J_b + J_o + J_c
\]

where

\[
J_b = (x - x_b)^T B^{-1} (x - x_b)
\]

\[
+ [H(x) - y_0]^T (E + F)^{-1} [H(x) - y_0] + J_c,
\]

\[ J = \text{Fit to background } (J_b) \text{ plus Fit to observations } (J_o) \text{ plus Constraints } (J_c), \]

all \(J\)s are scalars, \(x\) = analysis (a vector to be determined, valid at a given time), \(x_b\) = background (a vector,
supplied by a model forecast valid at the given time),
\( B \) = background error covariance (a matrix), \( H \) = forward model (a matrix operating on the analysis vector providing synthetic observations from the background using spatial interpolation and variable transformation as needed), \( y_o \) = observations (a vector, valid at the current time), and \( E + F = R \) = instrument error + representativeness error (covariance matrices). The above equation expresses mathematically the fundamental quantity to be minimized by modern DA schemes in solving for \( x \), the new best estimate of the model state at a new initial time. A key term in the variational cost function is the \( B \) term, or background error covariance. This term was estimated initially using long-term verification data, the raw data from which to estimate cross-variable, three-dimensional covariances in forecast error. This estimate of \( B \) was a static estimate, calculated initially by Parrish and Derber (1992), using a procedure later called “the NMC method,” after the National Meteorological Center, the prior name for NCEP. In minimizing this cost function, DA can be considered as estimating a probability density function of the 3D state of the atmosphere including the most likely values for all variables at all 3D points (e.g., Lorenc 1988).

In 1995, NCEP replaced assimilation of sounding profiles retrieved from satellite radiances with direct assimilation of satellite radiances themselves (Derber and Wu 1998). Retrieved profiles had previously been shown to have a neutral to slightly negative impact on forecast accuracy because of introduction of errors in the retrieval process. This change to satellite radiance assimilation was a transformative step in NWP overall. It has been critical for continued improvement of global modeling skill since its introduction. A forward radiation transfer model was applied to estimate radiances in measured channels (e.g., Eyre 1989). The \( H \) represents the generic forward model, an operator in the variational cost function above. In this instance, \( H \) calculates synthetic radiances on the basis of prognostic variables in the model background. It then interpolates these diagnosed radiances to the locations where the satellite radiances were observed. Introduction of radiance assimilation in the NCEP global model in 1995 (Table 13-3) ended a 10-yr period with little improvement in skill. Forecast skill steadily increased from that year onward. ECMWF introduced direct radiance assimilation in January 1996 (Andersson et al. 1994; Derber and Wu 1998). A key component of radiance assimilation was introduction of bias corrections for radiances, as discussed by Derber and Wu (1998).

In 1997, ECMWF was first to introduce four-dimensional variational (4D-var) assimilation (Rabier et al. 2000), using a full model adjoint previously investigated by Thépaut and Courtier (1991), Courtier et al. (1994), and others. The computational expense of 4D-var was very high, thus limiting the resolution at which it could be run. Nevertheless, solving the minimization using a simplified model in incremental space in an inner loop, followed by a second, less-frequent, outer loop with full model adjoint, reduced the computational burden enough to allow implementation at ECMWF (Rabier et al. 2000). Use of a full model adjoint ensures that the final analysis is very close to a realizable model state, one that is in balance, and thus obviating the need for a separate initialization step.

In 2007, Honda et al. (2005) implemented a non-hydrostatic regional 4D-var DA system operationally in Japan to supply initial conditions to a 5-km mesoscale model. As described by Bauer et al. (2015), the 4D extension to variational assimilation has been found to be an essential component of today’s global NWP systems. Other NWP-center implementations of 4D-var assimilation include Météo-France in 2000, Met Office in 2004, Japan and Canada in 2005, and the U.S. Naval Research Laboratory in 2009. NCEP implemented in 2016 a 4D “ensemble-var” using ensemble-based 4D forecast error covariances instead of a model adjoint. Overall, 4D-var with a model adjoint is used by most major NWP centers, at least as a component of DA, but it is not used at NCEP. Out of its resources for DA and deterministic model prediction, ECMWF currently uses an overwhelming proportion for DA: 80% (P. Bauer, ECMWF, 2018, personal communication). At NCEP, by contrast, about 42% of that resource is currently used for DA (M. Kane, NCEP, 2018, personal communication).

Ensemble-based DA (EDA) methods subsequently emerged as a much less computationally demanding alternative to 4D-var but also as an effective addition to better describe 4D evolution of background error covariance \( B \). Evensen (2003) and Hamill (2006) provide comprehensive descriptions of EDA methods. Ensemble-based \( B \) can be estimated and applied in EDA with a much more realistic, weather-dependent representation than with the static estimate. Houtekamer et al. (2005) implemented EDA in the Canada global model with performance similar to that of their 3D-var assimilation. Whitaker et al. (2008) described initial testing of EDA with the NCEP Global Data Assimilation System with positive results, outperforming 3D-var assimilation. EDA required multiple short-range forecasts, so Whitaker et al. applied EDA at coarser resolution. To ensure adequate long-wave error covariance representation, which is not a certainty with a limited number of ensemble members, a hybrid ensemble-variational DA method was developed with fractional contributions to \( B \) from both the static \( B \) (e.g., the NMC method) and...
the time-varying ensemble-based B. This hybrid method was implemented in the NCEP global model in 2012 (Wang et al. 2013), resulting in the greatest single improvement in Global Forecast System (GFS) skill in this decade (shown below in Fig. 13-15). As of this writing, combined EDA/4D-var schemes are applied effectively at both ECMWF and the Met Office with their global models. Houtekamer and Zhang (2016) provide an excellent summary of ensemble-Kalman-filter (EnKF) DA methods, including the application of localization to avoid spurious effects at long distances when using a limited number of members (Anderson 2007).

j. Ensemble prediction systems

Ensemble forecasting increasingly became the backbone for NWP, especially at the global scale. In 1991, introduction of initial ensemble prediction systems (EPSs) occurred during the same year at both NCEP (Toth and Kalnay 1993) and ECMWF (Molteni et al. 1996). Ensemble prediction provides multiple solutions to longer-range predictions. Early in the twentieth century, Poincaré (1914) had noted that small perturbations in the initial conditions of nonlinear systems of equations (including the meteorology application) could grow to the point that solutions diverged markedly, casting doubt on their validity. Lorenz (e.g., Lorenz 1963, 1993) explored this finding in much greater depth, as noted in the previous 1956–85 section. In his review paper on the roots of ensemble forecasting, John Lewis (Lewis 2005, p. 1865) began,

> Eric Eady was the first meteorologist to voice concern over strict determinism—that is, a future determined by the initial state without account for uncertainties in that state. By the end of the decade, Philip Thompson and Edward Lorenz explored the predictability limits of deterministic forecasting and set the stage for an alternate view—a stochastic–dynamic view that was [fully] enunciated by Edward Epstein.

Lewis explained why a marriage between probability and strict dynamical prediction is both justified and highly useful. He discussed the bridge between theory and practice, which is neither easy nor straightforward.

Global ensemble prediction systems are now run by all major NWP centers as shown from a recent Working Group on Numerical Experimentation (WGNE) survey and presented in Table 13-6. Increasing proportions of NWP computing resources are allocated for global ensembles, up to 40% at ECMWF (P. Bauer, ECMWF, 2018, personal communication) as compared with 10% for NCEP (M. Kane, NCEP, 2018, personal communication), another notable difference between these centers in computer prioritization along with the DA difference.

For 50 years, computing speed increased exponentially with time within an individual chip, following “Moore’s Law” (Gordon Moore), but this is no longer true. Instead, the trend is toward cheaper and more numerous chips, which has made ensemble NWP practicable. This is a fortunate coincidence because the demand for probabilistic guidance has spurred development of ensemble prediction methods.

k. Increasing forecast skill

Over this most recent 1985–2018 era, NCEP, ECMWF, and other global models increased forecast skill steadily at 500 hPa [Figs. 13-10 and 13-16a,b—anomaly correlation coefficient (ACC) for 500-hPa height]. A steep increase started in 1994 with ECMWF global skill (Fig. 13-15). A steeper increase in global 500-hPa skill started in 1998 for the United States and United Kingdom (Fig. 13-15). The key improvements in this decade were related more to DA than to model improvements. Variational DA started in the early to mid-1990s at most of these centers (see Table 13-3 for the NCEP global model), with direct assimilation of satellite radiances in 1995. Radiance assimilation and bias correction, and the introduction of TOVS data in 1998, coincided with an upward surge in GFS skill (Fig. 13-16b).

As noted earlier, 4D-var was implemented at ECMWF in late 1997, followed by France, the United Kingdom, Japan, and Canada in 2004 [Lorenc (2017), from the ECMWF Symposium on the 20-yr anniversary of 4D-var; https://www.ecmwf.int/en/learning/seminars/symposium-20-years-4dvar]. The ECMWF shows nearly equal skill between the Northern and Southern Hemispheres by 2003 via effective satellite radiance assimilation (Fig. 13-16a) despite many fewer rawinsonde and aircraft data in the Southern Hemisphere. Relatively flat U.S. global skill for 500-hPa height ACC was apparent over the 2005–11 period. This flat period ended in 2012 when the United States introduced hybrid ensemble-variational assimilation instead of using fixed background error covariances, after improved results were found (Wang et al. 2013). All major NWP centers adopted ensemble-variational DA in the last 10 years [e.g., Buehner et al. 2013 (Canada); Kleist and Ide 2015 (United States); Bonavita et al. 2012, 2015 (ECMWF)]. The key breakthrough was that the current weather-dependent 3D (and 4D) background error covariance could be represented through an adequately sized short-range ensemble. Long-term statistics provided a longwave component (wavenumbers 1–6) to the ensemble-based background error covariances. It was discovered that straight 4D-var could
be incorporated within the 4D-ens-var assimilation method.

1. Reanalyses

Reanalyses using state-of-the-art global DA and modeling systems [suggested first by Bengtsson et al. (1982)] are extremely important for assessing possible climate change (e.g., Dee et al. 2011; Compo et al. 2011) and to capture a wide range of atmospheric conditions over many decades in a gridded dataset. Reanalyses take extraordinary computing resources using all observations, assimilated usually through a 6-h cycle, resulting in almost 44 000 DA cycles to cover a 30-yr period. ECMWF produced a 40-yr reanalysis in 2005 (Uppala et al. 2005).

Reanalyses use the same DA and model configuration through a multidecade period (Bengtsson and Shukla 1988). Full twentieth-century reanalyses have been developed by NOAA (Compo et al. 2011) and others. These reanalyses have become the fundamental datasets for atmospheric diagnostics, leading to improved understanding. NCEP has performed a reanalysis [the Climate Forecast System Reanalysis (CFSR)] three times, most recently with its CFSv2 coupled model (Saha et al. 2014). In Fig. 13-16b, the relative skill of the real-time GFS global model and

![Fig. 13-15. Annual mean 500-hPa anomaly correlation for GFS, ECMWF, CMC (Canada), "UKM" (Met Office), and FNMOC (U.S. Navy) from 1985 to 2017 for the (top) Northern and (bottom) Southern Hemispheres. The figure is provided through the courtesy of F. Yang of NOAA/NCEP.](image-url)
TABLE 13-6. Operational global ensemble prediction systems: resolution (spectral or grid point), number of layers (L), number of members (M), forecast (FC) range in days (D), and number of ensemble runs per day. The table is adapted from the WGNE (http://wgne.meteoinfo.ru/nwp-systems-wgne-table/wgne-table/).

<table>
<thead>
<tr>
<th>Forecast Center (Country)</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECMWF (Europe)</td>
<td>Tco639 L91 to D + 15 days</td>
<td>Tco639 L91 to D + 15 Tco319 L91 to D + 45 coupled ocean–atmosphere 0.25°</td>
<td>Tco639 L91 to D + 15 Tco319 L91 to D + 45 coupled ocean–atmosphere 0.25°</td>
</tr>
<tr>
<td>Met Office (UK)</td>
<td>33 km L70 to D + 7; 12 M and 24 M lagged products</td>
<td>20 km L70 to D + 7; 18 M and 36 M lagged products</td>
<td>20 km L70 D + 7; 18 M and 36 M lagged products</td>
</tr>
<tr>
<td>Météo France (France)</td>
<td>TL798c2.4L90 M35; 4 days 4×/day</td>
<td>~TL1198c2.2L90; M35; 4 days, 4×/day</td>
<td>~TL1198c2.2L90; M35; 4 days, 4×/day</td>
</tr>
<tr>
<td>DWD (Germany)</td>
<td>40 km (20 km for Europe) L90; M40 FC range 7.5 days 2×/day, 5 days 2×/day</td>
<td>40 km (20 km for Europe) L90; M40 FC range 7.5 days 2×/day, 5 days 2×/day</td>
<td>40 km (20 km for Europe) L90; M40 FC range 7.5 days 2×/day, 5 days 2×/day</td>
</tr>
<tr>
<td>HMC (Russia)</td>
<td>T168 L31 + 0.72° × 0.9° L28; M14; 10 days</td>
<td>T339 L63 + 0.18° × 0.225° L60; M22; 10 days</td>
<td>T339 L63 + 0.18° × 0.225° L60; M22; 10 days</td>
</tr>
<tr>
<td>NCEP (United States)</td>
<td>T574 L64, M21 to D + 8 T382 L64, M21 for D + 8–16</td>
<td>GSM—T574 L64, M21 to D + 8 T382 L64, M21 for D + 8–16</td>
<td>GSM—T574 L64, M21 to D + 8 T382 L64, M21 for D + 8–16</td>
</tr>
<tr>
<td>Navy/FNMOC/NRL</td>
<td>T359 L60; M20; 16</td>
<td>T359 L60; M20; 16</td>
<td>T359 L60; M20; 16</td>
</tr>
<tr>
<td>CMC (Canada)</td>
<td>Yin-yang (0.36° × 0.36°) L84; M21; 16 – 32</td>
<td>Yin-yang (0.36° × 0.36°) L84; M21; 16 – 32 coupled atmosphere–ocean–ice</td>
<td>Yin-yang (0.27° × 0.27°) L84; M21; 16 – 32 coupled atmosphere–ocean–ice</td>
</tr>
<tr>
<td>CPTEC/INPE (Brazil)</td>
<td>T126 (~100 km) L28; M15; 15; 2/day</td>
<td>T213 L42; M20; 15; 2/day</td>
<td>T213 L42; M20; 15; 2/day</td>
</tr>
<tr>
<td>JMA (Japan)</td>
<td>TL479 L60; M27; 0–11 days</td>
<td>TL479 L100; M27; 0–5.5/11 days; TL479 L100; M13; 11–18 days; TL319 L100; M13; 18–34 days</td>
<td>TL479 L100; M27; 0–5.5/11 days; TL479 L100; M13; 11–18 days; TL319 L100; M13; 18–34 days</td>
</tr>
<tr>
<td>CMA (China)</td>
<td>T639 (0.28125, L60); M15; 2/day</td>
<td>GRAPES (0.5°, L60); M40; 2/day</td>
<td>GRAPES (0.5, L60); M40; 2/day</td>
</tr>
<tr>
<td>KMA (Korea)</td>
<td>Met Office Global and Regional Ensemble Prediction System (MOGREPS), 32 km L70; M49; D + 12</td>
<td>MOGREPS, 32 km L70; M49; 12d</td>
<td>MOGREPS, 25 km L110; M101; 12d</td>
</tr>
<tr>
<td>NCMRWF (India)</td>
<td>~33 km L70; 44; 10</td>
<td>17 km L70; 22/44; 10</td>
<td>12 km L70; 22/44; 10</td>
</tr>
<tr>
<td>BoM (Australia)</td>
<td>MOGREPS, 60 km L70; M24; 10</td>
<td>MOGREPS, 30 km L70; M24; 10</td>
<td>MOGREPS, 30 km L70; M24; 10</td>
</tr>
</tbody>
</table>
CFSR reanalysis/reforecast skill for 5-day forecasts are both shown. Skill from the CFSR 5-day forecasts increased the 500-hPa ACC by 0.06 for the Northern Hemisphere and by 0.08 for the Southern Hemisphere for the 30-yr period from 1987 to 2017, indicating the contribution by the improvement to the global observing system over this period. By contrast, over the same period, real-time forecast skill for the GFS also for 500-hPa ACC increased by 0.24 in the Northern Hemisphere and 0.28 in the Southern Hemisphere. This suggests that the increasingly effective observational network over that same 30-yr period contributed about 25% of the overall improvement in the Northern Hemisphere and 28% in the Southern Hemisphere.

Fig. 13-16. 500-hPa height anomaly correlation for the Northern Hemisphere vs the Southern Hemisphere, from 1984 through 2017 for (a) ECMWF at 3-, 5-, 7-, and 10-day projections (the figure is from https://www.ecmwf.int/en/forecasts/charts/catalogue/plwww_m_hr_ceaf_adrian_ts?facets=undefined&time=2018041100) and (b) NOAA/NCEP at 5-day projection only for deterministic GFS model and for CFSR (the figure is provided through the courtesy of F. Yang of NOAA/NCEP).
Hemisphere, leaving 72%–75% of the 30-yr improvement attributable to the gradual improvement in real-time assimilation and modeling. We call this combined assimilation/modeling improvement the \textit{NWP contribution} to skill, a contribution largely separate from the contribution from observations.

The skill of reanalysis-based forecasts relative to that from real-time global models provides a metric through which NWP progress can be isolated and assessed (a new metric that is introduced by this chapter). Reanalysis forecast skill can be subtracted from the skill produced in real time to examine a trend for this \textit{NWP (DA + model) contribution to forecast skill}. Years with easier or more difficult predictability are removed from this trend. Also removed are the effects of changes in observational density and quality. The remainder isolates the effect of combined improved DA and modeling methods. The overall NWP contribution to forecast skill for the U.S. GFS system was relatively constant between 2004 and 2011 as shown in Fig. 13-17—more constant than the overall skill in Fig. 13-16b. A surge in NWP contribution to forecast skill started again in 2012, apparently from introduction of hybrid ensemble-variational DA that year. In contrast, the trend in reanalysis forecast skill (Fig. 13-16b) represents a trend in observational information extractable by the DA–modeling system.

\textit{m. Regional models}

As in the previous era, many NWP centers (ECMWF excepted) also dedicated some resources to shorter-range \textit{regional} models with higher resolution and more sophisticated parameterizations of subgrid-scale processes. These regional models provided improved precipitation fields, orographic and coastal effects, and representation of clouds and near-surface details. These were essential steps toward direct NWP of sensible (surface) weather including more accurate precipitation for winter storms.

The United States introduced the NGM (Hoke et al. 1989) in 1985 (Table 13-4) at the beginning of this era. The NGM featured the then-novel sigma vertical coordinate and then-impressive 80-km resolution with regional statistical interpretation (OI) used for DA (DiMego 1988). In 1993, the United States introduced the Eta model in which a vertical step-mountain coordinate (Mesinger et al. 1988) replaced the sigma coordinate. Horizontal resolution was initially 80 km, was reduced to 29 km in 1995 in the mesoEta model, and was reduced to less than 10 km in 2011 (Table 13-4) for a nonhydrostatic application with a sigma coordinate developed by Janjić (2003).

In 1998, the United States introduced hourly updated NWP, the RUC (Benjamin et al. 2004a), using observations from the latest hour. Rapid updating with the RUC and its successors, the 13-km Rapid Refresh and
the 3-km HRRR (Benjamin et al. 2016a; Table 13-4) provided “situational awareness” forecast guidance. It exploited very recent observations to meet short-range forecast requirements for aviation, severe weather, weather-dependent power generation, and hydrology. James and Benjamin (2017) showed that assimilation of cloud data (supplied by ceilometers and cloud-top retrievals) and radar reflectivity were effective in rapid updating. Application of digital filter initialization (Lynch and Huang 1992) was essential to add sufficient mesoscale balance for 1-h intermittent assimilation (Benjamin et al. 2004a; Peckham et al. 2016).

To improve forecasts of convective storms, precipitation, and mountain waves, it was necessary to add vertical acceleration to models and leave behind the hydrostatic assumption, which had been suitable for grid lengths of 10 km and coarser. Tapp and White (1976) developed the first fully compressible nonhydrostatic model for forecasting, although it was used initially for horizontal resolutions for which the hydrostatic approximation is still acceptable. Development of other nonhydrostatic models for forecasting started early in this era, and implementation at operational NWP centers commenced in the 2000s (Saito et al. 2007). These CAM forecasts (Gustafsson et al. 2018) correctly predict clusters of convective storms and sometimes even individual convective cells. They result in better forecasts of severe weather and local wind storms. Recent studies (Davis et al. 2003; Wilson and Roberts 2006) show that numerical models that cannot resolve the convective scale often have trouble in representing even the diurnal cycle of convection and are unable to propagate convection systems accurately, thus limiting predictive skill. Forecast improvements with CAMs also result from better representation of orography, coastlines, and other variations in surface conditions. The higher-resolution Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5) nonhydrostatic community model (Dudhia 1993; Grell et al. 1994). This led to a new community nonhydrostatic model developed by NCAR: the Weather Research and Forecasting (WRF) Model (Skamarock et al. 2005; Skamarock and Klemp 2008). WRF was designed for both research and operational NWP and has been developed through a partnership among NCAR, NOAA, and universities. It has been extensively used to train university students in nonhydrostatic model development and application. It also provides the dynamic core for the NOAA hourly updated 3-km HRRR and 13-km Rapid Refresh models. The United Kingdom modified their UM to a nonhydrostatic version as described in Davies et al. (2005). As of 2018, many NWP centers have started running these nonhydrostatic models over regional domains at Δx < 5 km, at which resolution they are defined as convection-allowing models (CAMs). Many of these operational CAM runs are listed in Table 13-8.

### Table 13-7. History of initial operational implementations of nonhydrostatic models at a resolution of 5 km or finer. Reproduced from Saito et al. (2007).

<table>
<thead>
<tr>
<th>Name</th>
<th>Resolution/layers</th>
<th>Year</th>
<th>Frequency</th>
<th>Domain</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany (DWD) Lokal Modell (LM)</td>
<td>2.8 km L40</td>
<td>2007</td>
<td>Germany and surroundings</td>
<td>Doms and Schättler (1997) Steppeler et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>JMA Non-Hydrostatic Model (NHM)</td>
<td>5 km L50</td>
<td>2004</td>
<td>Japan and surroundings</td>
<td>Ikawa and Saito (1991), Saito et al. (2006)</td>
<td></td>
</tr>
<tr>
<td>U.S. NCEP HRRR (WRF-ARW)</td>
<td>3 km L50</td>
<td>2014</td>
<td>1 h</td>
<td>CONUS</td>
<td>Benjamin et al. (2016a), Skamarock et al. (2005), Skamarock and Klemp (2008)</td>
</tr>
<tr>
<td>Météo-France “AROME”</td>
<td>2.5 km L60</td>
<td>2008</td>
<td>3 h</td>
<td>France and surroundings</td>
<td>Seity et al. (2011)</td>
</tr>
</tbody>
</table>
space and time-scale prediction has also opened a wide variety of new and/or improved services, for example, local infrastructure forecasts such as for large public events or airports and other transportation hubs, precipitation forecasts for small basins including flash flood predictions, winds and wind-driven waves in coastal regions, input to local storm surge and hydrological models, and temperature and fog forecasts under stable conditions.

Toward the end of this era as operational non-hydrostatic models reached the range of 1–3 km in resolution, it became increasingly evident that subgrid-scale parameterizations would be needed down to 100-m resolution [confirming the findings/recommendation of Bryan et al. (2003)].

Higher-resolution and coupled data assimilation

The same variational and ensemble-variational DA applies on the regional scale as on the global scale, but usually at higher spatial resolution. Japan applied 4DVAR in their regional model before their global model (Koizumi et al. 2005). Regional models give more emphasis to moisture-related variables: accurate initial moisture fields are less essential for longer-range global models but are critical for short-range regional models, especially those using convective-allowing resolution (<5 km; Tables 13-7 and 13-8). Operational NWP systems have added assimilation of radar reflectivity and cloud ceilometer and satellite-based retrievals of cloud layers (e.g., Koizumi et al. 2005; Benjamin et al. 2016a). Storm-scale EDA has been demonstrated in experimental configurations. Dowell et al. (2011) described an early application of EDA for assimilation of radar-based radial wind and reflectivity for rapidly developing precipitating systems with complex dynamics. Coupled DA linking atmosphere and ocean is now being initially explored, at least for a coupled reanalysis (e.g., Laloyaux et al. 2016; ECMWF: Coupled ERA reanalysis). Snow and soil moisture DA are being increasingly addressed by NWP centers (De Rosnay et al. 2014; Benjamin et al. 2016a) and found to be important for near-surface, cloud, and convection forecasts.

 Ensemble prediction from convection-allowing models (generally ≤ 5-km resolution) has been increasingly applied since the turn of the century. The United States started a two-member ensemble using different versions of the WRF Model (Tables 13-7 and 13-8—ARW and NMM dynamic cores; Powers et al. 2017). These mixed-model ensembles, generally not yet using CAM-scale EDA, have been organized and described in the United States in an ongoing NOAA spring forecasting experiment (Clark et al. 2018). Schwartz et al. (2017) describe an experimental EPS at 1-km horizontal resolution. Again, Houtekamer and Zhang (2016) summarize applications of CAM-scale regional ensemble predictions, sometimes using CAM-scale EDA.

Not only has accuracy of forecasts improved but also their reliability. In particular, there is a great reduction in poor global forecasts. Figure 13-18 shows the percentage of anomaly correlations (500-hPa heights) for 5-day GFS forecasts that are smaller than 0.7 for each
year from 1996 through 2017. As shown, the number of “busts” in each hemisphere has dropped sharply in the past 20 years. Figure 13-19 shows the frequency distribution of the same 20-yr sample of anomaly correlations (NH only). The peak in the distribution shifts to the right in succeeding years; in the most recent three years, more than 40% of anomaly correlations exceed 0.90. Ensemble forecasting has reduced the threat of completely unpredicted events. Especially poor global forecasts seldom occur any more. At the same time, the bar has been raised on the quality of forecasts needed because human lives and safety are increasingly endangered by extreme weather events (e.g., excessive precipitation).

Short-range regional forecasts have improved apace with global forecasts, contributing increasingly in recent decades to better skill in precipitation forecasts issued by NCEP’s Weather Prediction Center from 1961 through 2017 (Fig. 13-13). Figure 13-20, like Fig. 13-18, shows that major forecast errors have become much less likely in recent years. It compares 250-hPa RMS vector wind errors in 6-h RUC forecasts for 1999–2004 (left panel) with those in 6-h Rapid Refresh forecasts for 2010–18 (right panel). (The Rapid Refresh was the successor to the RUC model, with both operating on a 1-h assimilation/prediction cycle.) Errors greater than 6 m s\(^{-1}\) have become much less frequent in recent years, and those over 8 m s\(^{-1}\) are nonexistent. Progress in eliminating forecasts of particularly poor skill, at least in some categories, gives new confidence to decision-makers.

**o. Relation to climate models**

The development of climate models in different levels of complexity is described by Randall et al. (2019), especially at GFDL, NCAR, and UCLA. These models ranged from early nongrided radiative-equilibrium models (e.g., Manabe and Wetherald 1967) to current Earth-system models with coupled atmospheric and ocean model components using time-stepping and grids similar to NWP models but at coarser resolution. A true coupled global model by Manabe and Wetherald (1975) gave an early indication of anthropogenic global warming. The climate modeling community generally first considered detailed treatment of Earth-modeling extensions such as clouds, land surface, and radiation before NWP application. These treatments were necessary to consider the overall energy budget of the Earth system and later were introduced into weather prediction models at far higher spatial resolution to add accuracy to forecasts of only days or even hours in duration. Later treatments included the oceans and cryosphere and, more recently, the carbon cycle. The evolution of Earth-system models over the last 100 years is treated in more detail by Randall et al. (2019) in this monograph.

**p. Coupled modeling: NWP toward NEWP**

International NWP centers during this era increasingly accounted for Earth-system components and worked toward a numerical environmental and weather prediction
Anticipating this change in scope, the United States renamed its NMC to the broader National Centers for Environmental Prediction in 1995.

Weather events account for approximately two-thirds of natural disasters in the form of floods, wind storms, slides (e.g., primarily mud slides and avalanches induced by precipitation), extreme temperatures (e.g., urban heat wave), and wildfires (typically associated with high winds, low humidity, and high temperatures). Floods and wind storms (primarily tropical cyclones) are by far the most common disasters affecting our planet, accounting for over one-half of the disasters. The character, intensity, and impacts of these disasters vary on regional and even local scales.

![Diagram](image1.png)

**FIG. 13-20.** The 250-hPa RMS vector error for 6-h wind forecasts against rawinsonde data over the contiguous United States for NCEP hourly updated models: (left) RUC for 1999–2004 and (right) RAP for 2010–18. Blue shows all forecast valid times (0600 and 1800 UTC initial times, verified every 12 h), and red shows 30-day mean values.
The human influences on the climate system have also likely changed the probability of occurrence of some types of high-impact weather, particularly extreme temperature events, over the past decades. Furthermore, many of the observed changes in extremes are by and large consistent with what the scientific climate community anticipates of human-induced climate change. In the long term we can expect more-frequent warm extremes, less-frequent cold extremes, and more high-impact hydrometeorological events as a result of climate change. A list of climate models (https://pcmdi.llnl.gov/mips/cmip5/availability.html) participating in the Climate Model Intercomparison Project is maintained by the Lawrence Livermore National Laboratory in the United States.

The improved quality of modern NWP systems run operationally at NMHSs has formed the basis for the development of new specialized environmental prediction tools to support policies and operational objectives of their respective government and industrial sectors. Improved numerical environmental forecasts bring significant health (Rogers et al. 2010), safety, and economic benefits. This trend toward increasing complexity as models move from solely predicting weather toward broader environmental conditions includes forecasting capability for atmospheric constituents, soil–vegetation–ecosystems, the cryosphere, hydrology, and the coupling of atmospheric–ocean and even freshwater systems (Belcher et al. 2015; Roy et al. 2015; Valcke et al. 2012). NCEP has included the effects of sudden solar storms (e.g., coronal mass ejections) on the upper atmosphere (especially, the ionosphere) with its Space Weather Prediction Center, which provides critical guidance for aviation, microwave communications, and GPS operations. This overall trend toward Earth-system prediction increases the utility of the atmospheric forecasting systems but also improves the quality of the coupled systems when feedbacks of these diverse processes on atmospheric flows are properly represented. As examples of new applications, Earth-system predictions can be used to determine the factors that affect water levels and overflows in hydrographic basins and lake systems, to develop and test potential new regulation plans, and to assess the impacts of these potential plans on the ecosystem, people’s lives, and economic interests.

Observations of air chemistry and the ocean and coastal environments are facilitating development of NEWP systems. New satellite instruments [e.g., high-spectral-resolution sounders such as AIRS and the infrared atmospheric sounding interferometer (IASI); also, Earth-observing system instruments] have provided a wealth of information on the chemical state of the atmosphere as it relates to regional and urban air quality. Starting earlier in this most recent era, 3D trajectory transport tools such as NOAA’s HYSPLIT (Draxler and Hess 1998, Stein et al. 2015) were used to provide information on emergencies such as from radiological emissions or volcanic eruptions by importing wind information from NWP models. Increasingly in the last 5–10 years, coupled atmosphere–chemistry prediction models extending into the middle atmosphere, with assimilation of space- and ground-based observations of chemical species, are now the basis for air-quality forecasts, including ozone monitoring. Not surprisingly, coupled atmospheric–chemistry modeling entails substantial additional computational costs. Options range from a single smoke tracer to schemes containing some few tens of species, each requiring integration of a conservation equation. Global emission inventories, required for effective atmospheric–chemistry coupled NWP, have been insufficient over some geographical areas, resulting in increased international efforts for gathering this information (e.g., Frost et al. 2013).

For oceans and coastal areas, the international Argo program is providing routine soundings of the ocean equivalent to (or better than) those that the global ra-winsonde network provides for the atmosphere. Together with new satellite measurements of the ocean surface, Argo has revolutionized observation of the global ocean. The DA techniques used for atmospheric models has been extended to assimilate these data into coupled ocean–atmosphere models with positive effect on even medium-range NWP forecasts (e.g., Smith et al. 2018). This has important applications for understanding and forecasting ocean regional circulation, ice cover, and wave activity. Regional coupling efforts have been extended to large freshwater bodies and coastal areas (Durnford et al. 2018; Joe et al. 2018).

q. Computing and technology

Tremendous advances in computing technology since the mid-1950s enabled the spectacular growth of numerical weather prediction (Lynch 2006). For instance, NOAA/NCEP has made 13 major computer upgrades (Fig. 13-10) to support its continued model improvements over this period. Today’s top NWP computers rank within the top 100 of the 500 most powerful computers (https://www.top500.org/list/2018/06), with at least six international NWP centers now using computers with at least one petaflop (10^{15} floating-point operations per second) (https://www.top500.org/news/noaa-to-add-28-petaflops-to-supercomputing-capacity/). Over the last 20 years, NWP centers shifted from single vectorized computers with shared memory to parallel computers with rapid exchange of information between
nodes. Bauer et al. (2015) describe the progress in NWP with its aptly named article in *Nature*, “The Quiet Revolution of Numerical Weather Prediction.” They describe the current progress in solving the colossal computational problem of NWP and NEWP and challenges in the future, including limits in power consumption. Redesign of NWP/NEWP systems is required, given the needs for data input/output, storage, and archiving data, especially of ensemble forecast information for full Earth-system prediction systems.

Rapid, reliable communication of observations based on international cooperation has been essential for forecasting, even for subjective analyses during the first era and yet much more so for automated DA and NWP in the last era. Equally essential is efficient transfer of exponentially increasing volumes of gridded data ("big data") to consumers and value-adding members of the community of forecast providers.

During this era, display systems backed by powerful computers revolutionized the ability of field forecasters to merge atmospheric information, apply conceptual models, and examine the consistency and accuracy of models. AWIPS, developed in the 1990s, is currently used in NWS forecast offices throughout the United States, with emphasis on local forecast areas, although it is fully capable of displaying national, continental, and global domains. AWIPS displays detailed radar and satellite images, in situ observations, and gridded analyses and forecasts, either separately or with overlays. Compared with the facsimile machine in the 1970s, AWIPS workstations now supply orders-of-magnitude more information to forecasters and in far more accessible formats, with the contrast evident in Fig. 13-21. A separate display capability called National AWIPS (NAWIPS) was developed to assist forecasters in viewing larger areas at the NCEP Centers (for precipitation, aviation, and tropical cyclones). In 2003, the U.S. NWS introduced a digital merged National Digital Forecast Database (NDFD). Forecasts (graphics, text, and recorded messages) were generated automatically from the NDFD gridded data. NOAA continues to modernize its gridded data in the 2010s for NDFD with a merged gridded dataset called the National Blend of Models (NOAA Meteorological Development Laboratory 2018) using grids from different models at different durations from the Global Forecast System down to the 3-km HRRR model during the first 36 h.

Communication and forecast comprehension by public

Communication of weather information to the public has grown exponentially during this era with new media, specifically the introduction of cable television, the Internet, and cellular phones (Lazo et al. 2009; Henson 2010). The availability of dedicated weather channels on cable such as The Weather Channel (United States) and the Weather Network (Canada) provided much more detailed forecast and storm information. Cable and satellite-based television became widely available, even in rural and lower-income areas. The Internet allowed access to a great number of weather forecast websites including those from government agencies and private weather companies. The introduction of smart cellular phones facilitated access to these Internet sites and to cellular-phone-specific applications. Cellular phones have provided access to weather information for large populations without previous access, especially in developing countries, since they have
become relatively commonplace in all countries as of 2018. Such widely available weather forecasts influence public decision-making more than ever before (e.g., Morss et al. 2017).

6. The future of forecasting

The skill of NWP models over the past few decades has increased because of close ties between research and development efforts; theoretical advances in predictability, fluid dynamics, and numerical methods; subgrid-scale physics parameterizations (cloud, mountain, etc.); assimilation of diverse observations from the surface and space; and high-performance computing (HPC) systems. During the upcoming period of 2020–50, HPC capability is expected to grow further, moving beyond the petascale to the exascale and thus driving NWP advances and innovative applications. NWP systems of the future will be significantly larger than today’s systems, because computational tasks will require more power than is available with current HPC technology. To meet this challenge, the current technology [e.g., Graphics Processing Units (GPUs) and quantum computing] and HPC architecture (e.g., massively parallel) must change drastically, accompanied by a shift in thinking on how to solve the NWP problem (Bauer et al. 2015). However, nobody expects this to be a showstopper.

There will be an increase in predictions and observations of all components of the Earth system. Even now, numerical prediction is no longer limited to weather; it includes all of the weather-sensitive Earth-system components (Shapiro et al. 2010; Nobre et al. 2010). Over the next three decades, work will center on combining all the Earth-system components to establish what can be called a NEWP system. Since all NEWP forecasts will be probabilistic in nature, EPS techniques will be used for all Earth-system components. Consumers will thus be able to estimate the range of likely future weather and environmental conditions central to decision-making. Earth-system models including chemistry, hydrology, and other components are expected to be applied in NEWP systems down to subhourly updating.

The long-range roadmap for achieving this progress in numerical weather, climate, and environmental prediction science was discussed at the 2014 World Weather Open Science Conference (WWOSC) in Montreal. The outcomes of the 2014 WWOSC are outlined in the book Seamless Prediction of the Earth System: From Minutes to Months (Brunet et al. 2015). The science presented at the conference ranged from basic research, expanding our knowledge of processes and methods, to the applied research required to assemble the prediction system and assess the impacts of weather and climate events.

The NEWP research challenges of today are windows into the future of forecasting. Four fields of research have been identified that will lead to significant changes in the NEWP landscape over the next 30 years. These will be discussed in the next four sections for this future era. The fifth and last section focuses on how the GWE will deliver services based on this science. Initial versions of many of these visions for the next 30 years are already being developed and tested in different laboratories and centers.

a. Data assimilation and observations

Research on observing systems and DA will improve forecasts of high-impact weather. For example, this research supports international efforts to optimize the current WMO Integrated Global Observing System, design regional observing networks, and develop well-founded strategies for enhancing observations. The next three decades should see better integration of observing systems (including disparate regional radar networks) to support optimal global and regional NEWP systems. An integrated observing network will meet the demands of weather and climate forecasting spanning time scales from minutes to decades in a cost-effective manner. This will include climate monitoring with NEWP-generated analyses that meet the exacting requirements of climate science: not only accurate quantities, but accurate fluxes (for example) as well.

The assimilation of a wide range of observations, from satellite radiances to in situ observations, is now done at the main NEWP centers. Modern global NEWP systems have horizontal resolutions on the order of 10 km and are in hydrostatic balance; they include subgrid-scale parameterizations for unresolved dynamical and physical processes. In the next 30 years, we will see global nonhydrostatic convective-scale NEWP models with resolutions on the order of 1 km (eventually down to subkilometer scale) that will increasingly model nonlinear and turbulent processes explicitly. Subgrid-scale representation of cloud, turbulence, and land surface processes will be required until a resolution of 100 m (Bryan et al. 2003) or even 25–50 m is attained. Innovative and efficient DA systems will need to account for these characteristics properly, including different modeling and observational scales, in the context of evolving global and regional observation networks.

NEWP centers will continue to use in situ surface observations over land and aviation reports of water vapor, pressure, temperature and clouds aloft. Such measurements are made in most countries. In the future we should see a significant increase in these observations,
with an expanded role for drones and mobile radars (Bluestein et al. 2014), and improved international dissemination of the data to all NEWP centers. New variables, such as greenhouse gas concentrations, will be measured by aircraft, including commercial aircraft. NEWP centers will make use of new sources of data, especially for convective-scale models, such as lidar (light detection and ranging; e.g., Raman lidar for vertical profiles of water vapor), crowd-sourcing sites (e.g., cellular-phone pressure sensors; Mass 2012), and mobile-phone technology.

Scheduled deployment of drones over wide networks to improve boundary layer observations appears to be very possible. Cost-effective, adaptive or targeted, drone-based observations will be routinely obtained from certain locations with the aim of improving NEWP forecasts for high-impact weather events, as has already been demonstrated for tropical cyclones. However, these should not be seen as a replacement for the conventional observations, which anchor the observing network, but rather as a means to improve coverage in the future.

Satellite-based observations of the atmosphere and surface provide good global coverage. There will be a dramatic increase in these measurements in the future. The difficulties of using data acquired over land, clouds, and sea ice and the need to thin the data to reduce horizontal error correlations between measurements will be overcome by directly accounting for these observation error correlations. This will permit substantially greater use of observations in assimilation cycles when compared with the current 20%. Data from active space-based microwave radars like those used in the Tropical Rainfall Measuring Mission and the Global Precipitation Measurement mission will be routinely assimilated, providing valuable information for the prediction of precipitation and moist processes.

In situ surface observations of the ocean (pressure, wind, and sea surface temperature) come mainly from the drifting buoy network introduced in the late 1990s, supplemented by moored buoys and ships. We should see an increase in the use of floats to provide profiles of temperature and salinity to ocean depths beyond the Argo 2000-m limit. Marine mammals will be increasingly instrumented to provide a new source of three-dimensional ocean data in rich ecosystems where needed. Underwater vehicles will routinely collect samples from the ocean and from coastal regions at various depths over periods ranging from weeks to months. Notably, satellite ocean-color observations will be routinely assimilated and predicted for use in predicting chlorophyll concentration and phytoplankton characteristics.

Data assimilation will be the cornerstone of NEWP, which will connect models to atmosphere–ocean–land–cryosphere–biosphere observations. These coupled DA systems will begin to run on a continuous basis, never leaving computer memory, using hybrid ensemble and variational DA schemes. Data-assimilation ensembles will have far more members to obtain better error covariances. Even on a global scale, there will be direct assimilation of far more satellite and radar data. Satellite radiances being assimilated will include channels sensitive to atmospheric constituents (aerosols and gases). Most important, ensemble-based background error covariances will be fully flow dependent, and their use will circumvent the need for error models that involve poorly understood balance relations and complex physical processes. While still computationally intensive, especially for increasing ensemble sizes, these algorithms will be suitable for implementation on the next generation of massively parallel HPCs. As horizontal resolution increases, these assimilation techniques will allow practitioners to focus more on observation types with non-Gaussian error distributions, such as cloud and precipitation. The stochastic perturbation approaches for the physics will have a more rigorous basis.

b. Predictability and dynamical/physical/chemical processes

In the next few decades, research in dynamical, physical, and chemical processes will generate the knowledge needed to advance our understanding of the limits of predictability for seamless forecasting of the Earth system. Considerable efforts will be required to quantify and reduce the uncertainties in the representation of physical processes in numerical models. This will include evaluation and improvement of parameterizations and explicit representations like those used to model cloud microphysics and aerosol interaction, convection, orographic and turbulent drag, surface interactions, radiative transfer, global teleconnections, and nonlinear dynamics (e.g., intensification of tropical cyclones). Making significant progress may require challenging traditional parameterization approaches. For example, future convection schemes may include multiple columns that have memory and an inherent representation of uncertainty. These parameterizations will be scale aware and able to cope with the problem of processes, including convection and turbulence, being partially resolved. Moreover, unified physical parameterizations treating together convection, turbulence, cloud, radiation, and even aerosol processes will be developed and implemented in atmospheric models rather than the separate treatment currently used.
The comprehensive calculations performed with very-high-resolution models on massively parallel HPC will require extensive development of new numerical techniques. Finite-volume solvers on unstructured and adaptive meshes will slowly gain more acceptance as the best tool for construction of meteorological models that resolve the global convection-allowing scale. Various innovative methods in NEWP will be developed to address the numerical stability, accuracy, computational speed, and flexibility required to handle a larger number of prognostic variables and the interaction between resolved and unresolved dynamic and physical processes. In the future, an HPC system used by interdisciplinary teams must be capable of routinely handling many petabytes of data and attaining a peak performance of tens of petaflops. This will no doubt entail significant changes to hardware and code design.

Air-quality forecasting systems will achieve much greater accuracy with improvements in 1) transport and dispersion of chemical species, 2) cloud microphysics, especially interaction with aerosols, 3) radiative transfer, and 4) turbulent fluxes in the boundary layer. More accurate prediction of transport and dispersion of chemical species will be the most important achievement of the next generations of numerical schemes for global atmospheric composition models. The numerical schemes will have better conservation, shape preservation, and prevention of numerical chemical mixing. The latter enhancement is important since inaccurate mixing is equivalent to introducing erroneous chemical reactions.

We will have increased understanding of weather systems in terms of regional phenomena (e.g., polar lows and tropical cyclones), their multiscale interaction with other systems (e.g., atmosphere–ocean–waves), dynamical processes (e.g., moist and dry air interactions) and physical processes (e.g., interactions between aerosol and cloud microphysics). The potential vorticity framework (Hoskins et al. 1985; Brennan et al. 2008) will serve as the dynamical meteorological foundation for the analysis of synoptic weather systems: cyclones (Chen et al. 2003; Martinez et al. 2011; Wernli and Sprenger 2007; Joos and Wernli 2012), anticyclones (Brunet et al. 1995), large-scale Rossby waves (Zadra et al. 2002), waveguides, and teleconnections. Global kilometer-scale NWP models will improve our fundamental knowledge of mesoscale weather systems and their interaction with large-scale circulation and topographic features.

One of the greatest future challenges is identifying sources of predictability at the subseasonal-to-seasonal (S2S) time scales and exploiting them in NEWP systems. It is expected that the number of degrees of freedom that are predictable at this scale is significantly greater than the number beyond seasonal time scales but far less than the number of degrees of freedom for the complete weather prediction problem (Brunet 1994). Improved S2S prediction has the potential to open up a large number of truly novel applications with important socioeconomic benefits. S2S forecasting depends on the initial conditions and slowly varying boundary forcing (e.g., oceans). It is fundamentally a combination of the already difficult NEWP and climate prediction problems, and it will need careful attention to the interaction of the different components of the Earth system (Brunet et al. 2010; Kirtman et al. 2013).

c. Interactions among Earth-system components

The trend toward including additional environmental couplings with atmospheric models will continue. These coupled models will be applied to address environmental emergencies and to ensure optimal watershed and ecosystem management. Already, increasingly accurate prediction of surface variables such as high winds, low humidity, and dry soils is being applied to reduce risk from forest fires (limiting human activity in some conditions) and to minimize the use of pesticides in agriculture. For the first time, we will see forecast models that balance the water and energy budgets at different scales. One of the major sources of error in hydrological predictions relates to uncertainties in the predicted rainfall. This problem will be solved largely through the future development of global and regional convection-allowing NEWP models, including ensemble models.

The land surface and the atmosphere interact through the global water cycle (exchanges of precipitation and evaporation) and control high-impact environmental events (e.g., floods and droughts) at regional to continental scales. Better representation of the water cycle will pave the way for more accurate prediction of floods and will lead to better management of water levels in large catchments such as the combined Great Lakes–Saint Lawrence River drainage basin. This will also improve operational prediction of floods, energy production from hydroelectric power plants, evaluation of the potential for bulk water exports, mapping of floodplains, planning of river diversions, and construction of new dams. These socioeconomic benefits will be even more substantial with the increasing realism of the submodels representing chemical processes, hydrology, the biosphere, and ocean circulation.

Overall, regional demand for many environmental prediction systems at subkilometer scale in complex terrains will grow. For urban areas and megacities, there is a need for air-quality forecasts at this very fine resolution. Better water management in urban areas has become more important as shortages of potable water
d. Putting it all together: Application of numerical environmental and weather prediction

The skill of global numerical weather predictions will continue to improve significantly. From past progress, we can expect an improvement of 3 days in deterministic global NEWP skill in the next three decades. We will approach the theoretical limit of predictability of the chaotic atmosphere a century after the first NWP forecasts by Charney, Fjortoft, and von Neumann in 1950. This will be accompanied by significant improvement in global quantitative precipitation estimation and forecasting as a result of convection-allowing models and DA advances, as discussed above.

As the skill of NEWP systems increases over the coming decades, more and more components of the Earth system must be taken into account to meet emerging economic as well as safety and health requirements. NEWP and climate models will continue to converge. We are approaching a new era of environmental prediction in which these geophysical subsystems coupled to the atmosphere need to be better simulated not only to advance weather prediction but also to provide new forecast variables (e.g., river flow and sea ice) with undeniable socioeconomic benefits.

Improved coupling is especially needed in polar regions (Nordeng et al. 2007). In support of this coupling, many international scientific efforts have been promoted recently such as the Year of Polar Prediction (YOPP), a major initiative of WMO, World Weather Research Program (WWRP). YOPP is a period of intensive observing, modeling, verification, user engagement, and education activities for 2017–19 (Jung et al. 2016). For example, the central YOPP contribution in Environment and Climate Change Canada (ECCC) is the development of a new high-resolution coupled environmental prediction system that integrates atmosphere–land surface–sea ice–ocean–waves elements, which will enhance understanding of the complex interactions of the Arctic environment and improve the models that are the basis for forecast services.

Overall, global NEWP systems with ensemble modeling will become ubiquitous and will be the cornerstone of automated decision-making. The future will see more emphasis on further automation of the forecasting process, allowing for more effective manual intervention in critical weather and environmental situations and improved techniques for forecasting the impacts of weather-related hazards. The information from these forecasts will feed applications tailored to the needs of end users. In light of the complexity and the huge size of the dataset generated by next-generation NEWP systems, it is evident that artificial intelligence (e.g., deep learning) and machine-learning techniques will be used routinely to integrate forecasts into the decision-making process.

John von Neumann, one of the contributors to the first successful NWP integration, was also the first scientist who suggested the concept of geoengineering by...
covering high-albedo polar regions with colorants to enhance the absorption of radiation to raise global temperatures. Today, the use of geoengineering is discussed from the perspective of reducing global temperatures by Paul Crutzen (Crutzen 2006). He suggested that the global temperature could be controlled by selective injection of sulfur dioxide into the stratosphere. This would boost the persistent stratospheric sulfate aerosol layer and could lead to a reduction in the global temperature similar to that caused by major volcanic eruptions. Despite the progress in NEWP and observations, there are still very significant unknowns that preclude the large-scale utilization of such optimal-control approaches. However, the current state of debate about these matters suggests that, in the future, these geoengineering solutions might be given serious consideration, and future state-of-the-art NEWP modeling tools will be needed to quantify the impact of such approaches and attempt to avoid drastic unanticipated consequences.

The complexity of NEWP systems will increase so much over the coming decades that a unified modeling approach for seamless climate and weather applications will be essential. A unified approach creates challenges but is greatly beneficial in terms of scientific research and cost-effectiveness (Brown et al. 2012). To streamline, simplify, and open up research, development, and operational processes to broader participation from the scientific community, many NMHSs are seeking to advance strategies for unified systems at all space and time scales. In addition to this unification trend, we will witness an increase in international collaborations devoted to developing and maintaining different components of the NEWP systems. An example is the Nucleus for European Modeling of the Ocean (NEMO), a general model of ocean circulation developed by a European consortium and used in many European countries and Canada. The socioeconomic benefits of NEWP are so diverse that this approach promises to be a game changer for all aspects of the GWE (monitoring, research, development, operations, and services) and has potential to attract resource and infrastructure investments.

e. The larger forecasting community: The global weather enterprise

The GWE is under constant pressure from users and operational centers, both public and private, to improve predictions and their dissemination in order to save lives, protect infrastructure, improve public safety and the quality of life, protect the environment, safeguard economic sectors, and increase socioeconomic benefits. This interweaving of the community to accomplish forecasting over the last 100 years and into the future is reflected in Table 13-1 at the beginning of this chapter. Forecasting of the weather and environment has become increasingly complex across all time and space scales with more and more credible environmental applications (e.g., forecasting sea ice, water quantity and quality, air quality, health, and food). Coordinating weather and environmental research and development internationally and between academic and operational sectors will continue to be an important goal of the GWE. Forecasters must become even better communicators so that consumers can evaluate the full range of options and make the right decisions.

The WMO will continue to be an important contributor to the GWE in many aspects. It will help to tackle international research challenges related to NEWP. The WMO sponsors WWRP, WGNE, and the atmospheric composition research program of the Global Atmosphere Watch, in addition to cosponsoring the World Climate Research Program (WCRP). It will promote a unified approach to multidisciplinary research and also encourage investments by the NMHSs in HPC to coordinate and accelerate NEWP development, validation, and use. Some of these activities will continue to engage the International Council of Scientific Unions (ICSU) through its cosponsorship of the Global Climate Observing System and WCRP and its academic constituency. In particular, strong links exist with the International Association for Meteorology and Atmospheric Science (IAMAS) and the International Association for Hydrological Science (IAHS). The ICSU’s Integrated Research on Disaster Risk program and the WWRP cohost a joint working group on the societal and economic research and applications of weather forecast products and services.

Another major issue for the GWE is a commitment to full and open access to environmental data. The United States has taken a strong position in this area with NOAA, NASA, NSF, and EPA all having policies in place to ensure that data on which scientific studies are based are preserved and made available at little or no cost. The AMS policy on open environmental data access (AMS 2013) developed with private, academic, and government partners sets an appropriate direction for the GWE.

The WMO will play a crucial role in coordinating, disseminating, and implementing NEWP innovations and products developed by NMHSs through its Global Data Processing and Forecasting System (GDPFS) program, which will benefit all nations. Seamless and integrated NEWP will expand the number of potential applications for weather and climate modeling systems by orders of magnitude. This will revolutionize cooperation and partnerships among the public, private, and academic sectors. Collaboration is crucial for advancing the required science and technology through high-risk and upstream research endeavors and training the next generation of highly
qualified personnel. Already there is evidence that the provision of capital-based infrastructure (e.g., observing technology hardware) for a service-based approach (e.g., provision of data services) will reshape the GWE landscape. This is nothing new since the private sector has been providing telecommunications and HPC infrastructure for the GWE for many decades. We will see private companies operating their own NEWP systems for targeted applications and customer-specific requests as is already happening with IBM. These NEWP private entities will, in one way or another, depend on the public and academic sectors for their innovations. No nation can advance NEWP science in isolation, let alone a single company. The added value of the private sector will continue to be demonstrated as it provides innovative and cost-effective contributions (Thorpe and Rogers 2018).

Success in following this future roadmap of GWE challenges will depend on the collaboration, strength, commitment, and excellence of the public, private, and academic sectors involved (e.g., Thorpe and Rogers 2018). Hooke and Pielke (2000) use a musical metaphor to describe this broader weather forecast process including forecast production but also communication and incorporation of forecast information into user decisions: the collaboration of these sectors on an international basis requires harmony among these components, like a symphony orchestra, with effective leadership and direction. This challenge will be thought provoking for all of the GWE communities, but past achievements and partnerships provide a solid base for confidence.

Acknowledgments. We very gratefully acknowledge background interviews with Alexander E. MacDonald, Ronald McPherson, Louis Uccellini, James Hoke, and Thomas Fahey (Delta and Northwest Airlines) in early 2018 and with Melvin Shapiro in early 2017. We also thank Rainer Bleck (NOAA/ESRL/GSD and NASA/GISS) and Thomas Lefebvre (NOAA/ESRL/GSD) for informative background discussions and Susan Cobb and Eric James (NOAA/ESRL/GSD) for excellent reviews. We gratefully acknowledge NWP table information from WGNE members [Ayrton Zadra—ECCC (Canada), Günther Zängl—DWD (Germany), François Bouyssou—Météo France, Keith Williams—Met Office, and Jun-ichi Ishida—JMA (Japan)]. We also found helpful an unpublished report by H. Robert Glahn, The United States Weather Service, First 100 Years, with detailed references to U.S. Weather Bureau annual reports. We thank the NOAA Library staff in Boulder (Sue Visser and Mike Robinson) for.unearthing many obscure references. Last, we are grateful for the opportunity to peer back a bit of the history of the larger forecasting community and to have been a tiny part of it.

REFERENCES


Buehner, M., J. Morneau, and C. Charette, 2013: Four-dimensional ensemble-variation data assimilation for global deterministic


——, 1954: Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Videnskaps-Akademiets Institut for Vaer-Og Klimaforsking Rep. 5 (appendix), 12 pp. [Available from Norwegian Meteorological Institute, P.O. Box 43, Blindern, N-0313 Oslo, Norway.]


Ferrier, B. S., Y. Jin, Y. Lin, T. Black, E. Rogers, and G. DiMego, 2002: Implementation of a new grid-scale cloud and precipitation scheme in the NCEP ETA model. 19th Conf. on


