



Reflections on the Conception, Birth, and Childhood of Numerical Weather Prediction

Edward N. Lorenz

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

Annu. Rev. Earth Planet. Sci.
2006. 34:37–45

The *Annual Review of
Earth and Planetary Science*
is online at
earth.annualreviews.org

Copyright © 2006 by
Annual Reviews. All rights
reserved

0084-6597/06/0519-
0037\$20.00

Abstract

In recognition of the contributions of Norman Phillips and Joseph Smagorinsky to the field of numerical weather prediction (NWP), a symposium was held in 2003; this account is an amplification of a talk presented there. Ideas anticipating the advent of NWP, the first technically successful numerical weather forecast, and the subsequent progression of NWP to a mature discipline are described, with special emphasis on the work of Phillips and Smagorinsky and their mentor Jule Charney.

NWP: numerical weather prediction

IAS: Institute for Advanced Study

MIT: Massachusetts Institute of Technology

CONTENTS

A full century has passed since Vilhelm Bjerknes (1904) identified the problem of weather forecasting with that of solving the equations that express the laws governing the atmosphere's behavior. He maintained that the equations were already known, but noted that a practical means of finding the solutions was lacking. Not too many years later, Lewis Richardson (1922) proposed that the equations could be solved by numerical methods. He arranged a detailed computational procedure, but subsequently acknowledged that a team of 64,000 persons might be required to carry it out faster than the weather itself evolved. Nevertheless, numerical weather prediction (NWP) had been conceived. It is now a long half century since numerical procedures became practical, thanks to the advent of the computer. A moderately successful computer-produced weather forecast presently followed (Charney et al. 1950). NWP had been born.

Since that time, symposia devoted to NWP have been abundant, but in April 2003, one of them was held at Princeton University for the express purpose of honoring Norman Phillips and Joseph Smagorinsky, two of the great contributors to the NWP field. The present account is an amplification of the talk that I presented at that symposium.

As such, it does not aim to present a detailed history of the early days of NWP. It consists instead of a few reflections, and it is heavily slanted toward the contributions of Smagorinsky and Phillips and their mentor Jule Charney. Rather definitive histories have been written by some of the principal NWP players, including George Platzman (1979), Philip Thompson (1983), and George Cressman (1996), whereas I have simply had the pleasure of watching the action from a grandstand seat.

It was in 1952 that Jule invited Victor Starr and me to visit his group at the Institute for Advanced Study (IAS) in Princeton. Victor was the leading dynamic meteorologist at the Massachusetts Institute of Technology (MIT), and I had a research appointment there. Victor was my mentor. The group at IAS had been assembled by the great mathematician John von Neumann, who was championing the extensive use of computers, and had identified the weather forecasting problem as one where numerical computation might prove particularly fruitful. He was either very fortunate or very far-seeing in having chosen Charney as the principal meteorologist.

Two members of the group whose acquaintance I immediately made were Joe and Norm. In the years since then, I have seen both of them rather frequently and have had many occasions to cite their works—Norm joined our department at MIT a few years later—but I have known them best, along with Martha Phillips and Margaret Smagorinsky, as close friends.

To place the status of NWP at that time in a proper context, let me drop back ten years from our visit to IAS. The United States had just entered World War II, and I had just halted my graduate studies in mathematics to join what was then the Army Air Corps as an aviation cadet and begin my training as a weather forecaster. The regular two-year graduate course in meteorology at MIT, where I was assigned, was crowded into eight months. Lectures in such subjects as synoptic meteorology and dynamic meteorology filled the mornings, and the afternoons were devoted to map

analysis and forecasting. With my mathematical background, I naturally found dynamic meteorology to my liking. Our teacher in that subject was Bernhard Haurwitz, who a year earlier had also written our textbook (Haurwitz 1941). We learned the equations of motion and some of their applications, and I waited eagerly for the time when we would learn how to use the equations to forecast the weather. The course finally ended, and the moment never came.

Why did the Army have us devote so much effort to dynamic meteorology and other subjects when its purpose was to produce forecasters for the field? One can think of a number of possible reasons, but it is likely that the correct answer can be given in one word, or rather one name—Rossby. Carl-Gustaf Rossby, who seemed to be involved in everything that was new and important in meteorology, was active in the planning stages of the Army's weather training program, and with his persuasive powers he could easily have convinced the authorities that we would be better weather officers with an all-around meteorological education.

As students, we were quite familiar with Rossby's name because of the then-popular Rossby diagram, used to analyze the stability of columns of moist air. Today's students are probably much more familiar with the Rossby number and Rossby waves. Rossby's ideas, and particularly his conviction that the key to understanding atmospheric behavior lay in the wind field rather than the more commonly analyzed pressure field, influenced Charney's thoughts about the atmosphere—probably more so than the ideas of any other meteorologist.

If I had read some of the unassigned pages of Haurwitz's book, I would have seen that he did mention the possibility of forecasting by integrating the dynamic equations, and he briefly described Richardson's attempt to do so. Nevertheless, he disposed of the topic in less than one page and concluded that we needed to know more about what is and what is not important before we could exploit the method. In this respect, he was far ahead of some of his contemporaries, who maintained that dynamic forecasting was not possible at all.

To understand this attitude, let us drop back another 20 years, to Richardson (1922). A characteristic of his procedure was his inclusion of almost every process in his equations. He did, however, adopt one important simplification. Strictly speaking, if we neglect water and other variable constituents, there should be five prognostic equations governing the three velocity components and temperature and density; pressure is related to the latter two by the diagnostic equation of state. Richardson, like others before and after him, realized the impracticality of using the vertical equation of motion, and he replaced it with the diagnostic hydrostatic equation, which equates the dominating terms in the vertical equation. When one does this, a perhaps unexpected thing happens. Because the hydrostatic equation is assumed to hold at all times—not just momentarily—the time derivatives of the two sides of the equation may be equated, yielding a second diagnostic equation, solvable for the vertical velocity, whereupon the number of prognostic equations drops from five to three.

What is happening is that the horizontal motion alters both the density and pressure fields in such a way as to disrupt hydrostatic equilibrium. The new system assumes that the vertical motion field, which also alters density and pressure, is the one

field whose effect exactly cancels the disruptive effect of the horizontal motion, thus preserving hydrostatic equilibrium.

Richardson next concluded that the geostrophic approximation was inadequate, and then went on to present a complete procedure for solving what we now call the primitive equations. His effort has been reviewed many times, perhaps most thoroughly by Platzman (1967). As everyone seems to know, his one forecast was disastrous. He correctly attributed the extremely large pressure tendencies to unrepresentative observed winds, but he apparently did not recognize that no wind observations available in his day could have led to reasonable pressure changes. Had he fully realized this he would almost surely have written his book anyway. His message was clearly addressed to a future generation.

With hindsight one can maintain that those who concluded from Richardson's work that NWP was unfeasible should have recognized that geostrophic equilibrium could be introduced analogously to hydrostatic equilibrium. The procedure is less straightforward because the wind is a vector, and introducing both components of the geostrophic equation would wipe out both equations of motion. What is needed is some readily measured scalar function of the wind, and a second scalar function whose time derivative depends largely on the difference between the first function of the wind and the same function of the geostrophic wind. When this difference and its time derivative are equated to zero, two more diagnostic equations result, and the number of prognostic equations drops from three to one.

What Charney (1948) succeeded in doing was equivalent to identifying the two scalars as vorticity and divergence and then carrying out the procedure, although his line of reasoning was entirely different and his choice of what to include and what to omit depended critically on a consideration of scales. Concurrently, Arnt Eliassen (1949) and Eric Eady (1950) were proceeding along similar lines, and their papers appeared soon after Charney's.

What happens here is that the rotational part of the wind—the part that is readily observed—alters both the wind and the pressure fields in such a way as to disrupt geostrophic equilibrium. The new system assumes that the less readily observed divergent part of the wind field, which also alters wind and pressure, is the one field whose effect exactly cancels the disruptive effect of the rotational part, so that geostrophic equilibrium is preserved.

I personally regard the successful reduction of the dynamic equations to a single prognostic equation by means of the geostrophic relationship, entirely apart from any applicability to NWP, as the greatest single achievement of twentieth-century dynamic meteorology. Consideration of the processes described by the new equation enabled me to see why cyclones and anticyclones and other weather systems move as they do—an understanding that the primitive equations never conveyed.

Like Richardson, Charney looked ahead to a day when weather forecasts would be numerically produced, but his approach to NWP was rather different. He felt that one should begin by simplifying the equations as much as permissible, thus minimizing the number of technical problems to be faced all at once; after these had been overcome, some omitted and supposedly secondary features could be reintroduced, preferably not too many at a time.

In the 1950s an additional reason for simplification was computer capacity. Charney's quasigeostrophic equation was three-dimensional, and would have overburdened the Electronic Numerical Integrator and Computer (ENIAC) at the Aberdeen Proving Grounds—the computer available to the group at IAS. Thus it was that he chose for the first test the two-dimensional barotropic vorticity equation—effectively what the three-dimensional equation reduces to if one disregards vertical variations of the flow. The forecast (Charney et al. 1950) was considered a success, and a small celebration followed (Thompson 1990).

ENIAC: Electronic Numerical Integrator and Computer

GCM: general circulation model

It is sometimes assumed that the idea of dynamical forecasting became dormant after Richardson, until Charney's breakthrough. One person who knows that this is not the case is Norman Phillips. Norm has critically reviewed a long sequence of papers extending through the period of supposed dormancy, each dealing with some aspect of the geostrophic relationship, and he has produced a veritable history of geostrophic ideas (Phillips 1990). Some of the papers that he examined are difficult to follow, and a lesser reader would likely have neither discovered their true significance nor identified their failings. In particular, it appears that Jeffreys (1919), in Richardson's time, and Kibel' (1940), somewhat before Charney, had derived closed systems of prognostic equations, and might have reached the goal ahead of Charney, but for some untenable assumptions.

After the first dynamical forecast, what was next? A milestone was Phillips's simplification of the atmosphere's vertical dimension to two discrete layers (Phillips 1952). This brought baroclinic forecasting to within the reach of the computers of that day. Soon afterward came a virtual flood of refined two-layer models, led by papers by Eady (1952), Eliassen (1952), Sawyer & Bushby (1953), and Thompson (1953).

Using the two-layer model, Phillips introduced a new field of endeavor—numerical modeling of the general circulation. The mechanics were much like those of NWP, but the aim was to reproduce typical circulation patterns rather than those occurring at a specified moment. The initial state of a general circulation model (GCM) could be made intentionally like nothing seen in nature, to see whether the model equations would guide the state toward something more realistic. Dissipation and external forcing, which in the earliest short-range forecasts had been placed among the “secondary” effects, had to be included now, being essential elements of the guiding process.

It has often been noted that a piece of pure research can lead, sometimes much later, to a practical application very likely not anticipated by the scientist performing the pure research. This is, in fact, one of the arguments often advanced by scientists when they are seeking funding for some esoteric piece of work. A classic example is the non-Euclidean geometry of Lobachevsky and Riemann, which much later led to the geometry of Einstein's four-dimensional space-time, and still later to useful applications of nuclear energy, not to mention some less friendly contraptions. This process can evidently take place on a two-way street; practical endeavors can sometimes lead to pure research that was originally unanticipated.

I doubt, for example, that the developers of the two-layer model, which has long since outgrown its intended usefulness as an operational forecasting tool, imagined

that it would, half a century later, continue to be one of the widely used tools in pure research. I would have trouble counting the number of students, or even just the MIT students, who have used a form of the two-layer model in their doctoral research, and to these we must add the more established scientists who merely find its simplicity attractive. The two-layer model is alive and well.

Let me cite an esoteric example. It is common practice to use the sum and difference of the stream functions in the two layers of the model as prognostic variables. These may be identified geostrophically with the pressure and temperature at some upper level. Several years after the two-layer model had become established, I derived from it a system of six ordinary differential equations by approximating the pressure and temperature fields each by three terms in a double Fourier series. The retained trigonometric functions formed an interacting triad, with one function representing the zonal flow and two representing the phases of a traveling disturbance.

Some 25 years later, I realized that the six equations could be reduced to three without sacrificing the baroclinic behavior (Lorenz 1983), in a manner suggested by the procedures for incorporating the hydrostatic and geostrophic approximations. The temperature field looks somewhat like the pressure field, displaced westward. We simply replace pressure and temperature by two new quantities, one of which is a “temperature anomaly”—the departure of the temperature field from the displaced and suitably scaled pressure field. Then, noting that the temperature anomaly is often small, we discard it. There remain three equations, governing the coefficients in the series for the other new quantity.

I managed to get a few papers out of the equations, but it was another ten years before I noticed that the reconstituted pressure field would always contain a single high and a single low, whence, because the system had only three variables, the longitude and latitude and central pressure of the low could be used as prognostic variables in place of the Fourier coefficients. Here, in miniature, was the synoptic meteorologist’s dream—a system where forecasting could be based entirely on the positions and intensities of the highs and lows. I offer this example not as a significant part of NWP history, which it is not, but as an illustration of the extent to which a piece of research can wander from the practical application that spawned it.

Eventually, it became evident that quasigeostrophic models would never produce forecasts of the quality that had been hoped for, even if one stayed away from the tropics, and thoughts were turning to the long-neglected primitive equations. Joe Smagorinsky, who had already succeeded in incorporating water and some other supposedly secondary effects into the quasigeostrophic equations (Smagorinsky 1957), now, together with some of his coworkers, notably Syukuru Manabe and Kikuro Miyakoda, produced a two-level primitive equation model (Smagorinsky 1958). Later, he enlarged it to become a two-level GCM (Smagorinsky 1963), which was followed in due time by a nine-level GCM (Smagorinsky et al. 1965). In discussing the latter model in his Wexler Memorial Lecture (Smagorinsky 1969), he emphasized its potentiality for extended-range forecasting.

I do not mean to imply that the bulk of the significant work in NWP and GCMs was performed by those who had once been part of the group at IAS, even though

a great deal was. Once computers had become generally available, models of all sorts, and modelers to manage them, proliferated, and much work of a similar nature and often of comparable quality appeared. As computers became ever larger and the temporal range of the forecasts became ever longer, there was more need to extend the spatial range to the whole globe, and ultimately the distinction between operational forecasting models and GCMs seemed to dissolve. NWP had come of age.

I recently reread Joe's Wexler Memorial Lecture because I partly remembered something that he had quoted from a newspaper question-and-answer column, and I thought that this was the place. The question proved to be, "Besides the appendix, what other parts of the body can man live without?" The answer given was, "Man can afford to lose his thymus, thyroid, three of his parathyroids, gonads and other internal and external reproductive organs, spleen, esophagus, ureters and urinary bladder, most of his liver, his bowel, one lung, one kidney, and part of his brain. Not all at the same time, of course."

The last sentence was Joe's vivid and long-remembered way of illustrating a type of redundancy, in this case a redundancy in weather observations that he had suspected and then verified with the nine-level model. In one of his integrations he replaced the observed moisture field with a field that did not vary at all with longitude. Of course, the north and south winds, and the upward and downward motions, immediately set in to produce longitudinal variations, but what Joe discovered was that after a day or so they were producing essentially the correct variations. He also found that he could do without observations of surface pressure and numerous other quantities, but, of course, not all at the same time.

I also ran into something in the lecture that I had not expected. About four years after Joe had given his talk, I presented a talk in Washington, entitled "Does the flap of a butterfly's wings in Brazil stir up a tornado in Texas?" (Lorenz 1993). I was careful not to answer the question, but apparently some listeners or readers assumed that I had, in the affirmative. In any event, I somehow became associated with a butterfly. I received several letters inquiring as to the origin of the butterfly metaphor, including one that mentioned Ray Bradbury's short story "A Sound of Thunder," where the death of a prehistoric butterfly changes the outcome of a present-day presidential election (Bradbury 1980).

I have mentioned on several occasions that I did not compose my title. This was done by the session convenor, Philip Merilees, who was unable to locate me when he had to submit tentative titles. I managed to get Phil's chosen locations changed to "Brazil" and "Texas," but this was solely for alliterative reasons. Previously, I had always referred to a seagull.

In reading Joe's paper I came across the questions "If we [satisfy certain conditions], then could we predict the atmospheric evolutions from the initial time with infinite precision infinitely distant into the future? Or would the flutter of a butterfly's wings ultimately amplify to the point where the numerical simulation departs from reality...?" In the same paragraph, he mentions the butterfly again, and in a later paragraph he talks of "the flap of the butterfly's wings"—words just like those that were to appear in my title. Let me submit the proposition that Joe is the original butterfly man.

ACKNOWLEDGMENTS

This work has been supported by the Large-Scale Dynamic Meteorology Program of the Lower Atmospheric Research Section of the Division of Atmospheric Science, National Science Foundation under Grant ATM-0216866.

LITERATURE CITED

- Bjerknes V. 1904. Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik. *Meteorol. Zeit.* 21:1–7
- Bradbury RD. 1980. A sound of thunder. In *The Stories of Ray Bradbury*, pp. 231–41. New York: Alfred A. Knopf
- Charney JG. 1948. On the scale of atmospheric motions. *Geophys. Publ.*, 17(2):1–17
- Charney JG, Fjörtoft R, von Neumann J. 1950. Numerical integration of the barotropic vorticity equation. *Tellus* 2:237–54
- Cressman GP. 1996. The origin and rise of numerical weather prediction. In *Historical Essays on Meteorology, 1919–1995*, ed. JR Fleming, pp. 21–39. Boston, MA: Am. Meteorol. Soc.
- Eady ET. 1949. Long waves and cyclone waves. *Tellus* 1(3):33–52
- Eady ET. 1952. Note on weather computing and the so-called 2-1/2-dimensional model. *Tellus* 4:157–67
- Eliassen A. 1949. The quasi-static equations of motion with pressure as independent variable. *Geophys. Publ.* 17(3):1–44
- Eliassen A. 1952. Simplified models of the atmosphere, designed for the purpose of numerical weather prediction. *Tellus* 4:145–56
- Haurwitz B. 1941. *Dynamic Meteorology*. New York: McGraw Hill. 365 pp.
- Jeffreys H. 1919. On travelling atmospheric disturbances. *Phil. Mag.* 37(6):1–8
- Kibel' IA. 1940. Prilozhenie k meteorologi uravenii mekhaniki baroklinnoi zhidkosti. *Izv. Akad. Nauk SSSR Ser. Geogr. Geofiz.* 5:627–37
- Lorenz EN. 1983. Irregularity: a fundamental property of the atmosphere. *Tellus* 36A:98–110
- Lorenz EN. 1993. Appendix I. The butterfly effect. In *The Essence of Chaos*, pp. 181–84. Seattle: Univ. Wash. Press
- Platzman GW. 1967. A retrospective view of Richardson's book on weather prediction. *Bull. Am. Meteorol. Soc.* 48:514–50
- Platzman GW. 1979. The ENIAC computations of 1950—gateway to numerical weather prediction. *Bull. Am. Meteorol. Soc.* 60:302–12
- Phillips NA. 1951. A simple three-dimensional model for the study of large-scale extratropical flow patterns. *J. Meteorol.* 8:381–94
- Phillips NA. 1956. The general circulation of the atmosphere: a numerical experiment. *Quart. J. R. Meteor. Soc.* 82:123–64
- Phillips NA. 1990. The emergence of quasi-geostrophic theory. In *The Atmosphere—A Challenge*, pp. 177–206. Boston, MA: Am. Meteorol. Soc.
- Richardson LF. 1922. *Weather Prediction by Numerical Process*. Cambridge, UK: Cambridge Univ. Press. 226 pp.

- Sawyer JS, Bushby FH. 1953. A baroclinic model suitable for numerical integration. *J. Meteorol.* 10:54–59
- Smagorinsky J. 1957. On the inclusion of moist adiabatic processes in numerical prediction models. *Ber. Dtsch. Wetterd.* 5(38):82–90
- Smagorinsky J. 1958. On the numerical integration of the primitive equations of motion. *Mon. Wea. Rev.* 86:457–66
- Smagorinsky J. 1963. General circulation experiments with the primitive equations. I. The basic experiment. *Mon. Wea. Rev.* 91:99–164
- Smagorinsky J, Manabe S, Holloway JL. 1965. Numerical results from a nine-level general circulation model of the atmosphere. *Mon. Wea. Rev.* 93:727–68
- Smagorinsky J. 1969. Problems and promises of deterministic extended range forecasting. *Bull. Am. Meteorol. Soc.* 50:286–311
- Thompson PD. 1953. On the theory of large-scale disturbances in a two-dimensional baroclinic equivalent of the atmosphere. *Quart. J. R. Meteor. Soc.* 79:51–69
- Thompson PD. 1983. A history of numerical weather prediction in the United States. *Bull. Am. Meteorol. Soc.* 64:755–69
- Thompson PD. 1990. Charney and the revival of numerical weather prediction. In *The Atmosphere—A Challenge*, pp. 93–119. Boston, MA: Am. Meteorol. Soc.