8 Predictability

8.1 Background

The term atmospheric predictability may be defined as the time required for solutions from two models that are initialized with slightly different initial conditions to diverge to the point where the objective (e.g., RMS) difference is the same as that between two randomly chosen observed states of the atmosphere. In the practical context of a forecast, the no-skill limit that defines the predictability may be the forecast lead time when the model-simulated state has no greater resemblance to the observed state of the atmosphere than does a reference forecast based on persistence or climatology. Many of the other chapters in this text address the various components of the modeling process that limit predictability, from data-assimilation systems to numerical methods to physical-process parameterizations, as well as metrics for quantifying it. This chapter will review the general concept of theoretical and practical limits to forecasting skill.

8.2 Model error and initial-condition error

As shown in the previous chapter, error that limits predictability originates in both the model and the initial conditions. Refer to Section 7.3 for more information, especially about the various sources of error associated with the model. Often the concept of predictability is discussed in the context of the system's response to infinitesimally small perturbations in the model initial conditions. This predictability is an inherent property of the fluid system and not of the model. Indeed, it is sometimes assumed in this hypothetical discussion that the model is perfect. In contrast, the term predictability is most-often used in the literature in a very practical sense to refer to the average length of useful forecasts that are obtainable from a particular operational modeling system, where all the sources of uncertainty in the modeling process contribute to error growth. For example, the impact of a particular new data source, or data-assimilation system, or parameterization will be evaluated in terms of its effect on the predictability of a particular variable.

Lorenz (1963a,b) describes simple modeling experiments that served as the foundation for later studies on the inherent predictability of the atmosphere. Using a form of identical-twin experiment (see Section 10.2), he initialized the same model with initial conditions that differed only very slightly, in the digit a few places to the left of the decimal point. He found that, after a few weeks of simulated time the two model solutions differed by as much as two random solutions. Many modelers have unintentionally replicated this experiment with contemporary models by changing computer compilers, or compiler optimizations, in the middle of a series of controlled, long-running simulations. Different compilers often perform arithmetic operations in an equation in a different order, which leads to different roundoff or truncation errors. Even if modeling-system configurations (initial conditions and physics) are identical, the subtle compiler-introduced differences in the model solutions can amplify to define the same type of predictability time limit that Lorenz observed. This growth of small perturbations in the atmosphere (or the model atmosphere), regardless of the source, led Lorenz to refer to the possibility that the flutter of a butterfly's wings could, after passage of significant time, influence the large-scale weather.

A more-contemporary identical-twin experiment is shown in Fig. 8.1, and illustrates error growth from small initial-condition perturbations. Here, the same atmospheric GCM is used for a control simulation, and for a perturbation simulation in which the initial

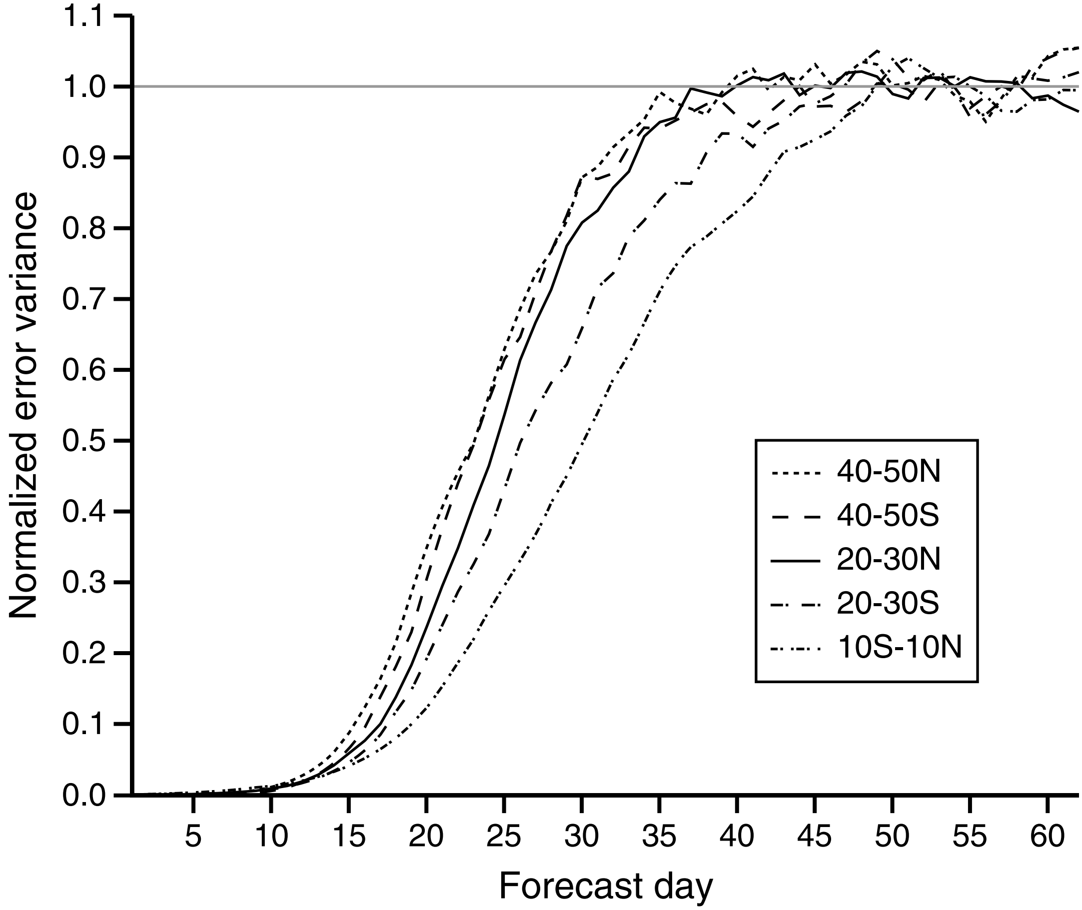

Fig. 8.1 Error variance of the 850-hPa zonal wind component plotted as an average for different latitude bands. The variance is normalized in terms of the saturation value (maximum) for each curve. The error is the difference between a control and a perturbation simulation of an identical-twin experiment that employed an atmospheric GCM. Adapted from Straus and Paolino (2009).

conditions were slightly different. For the perturbation simulation, the initial conditions were perturbed according to the following method,

$$\delta A = 0.001 rA,$$

where A is a model dependent variable, and r is a random variable that ranges from −1 to 1. This perturbation was applied independently at each grid point to temperature (K), both horizontal wind components, specific humidity, and surface pressure. The area-average difference in a variable between the control and perturbation experiments defines an *error*, and the square of this error is the error variance that is plotted in the figure. Specifically, the error variance in the u velocity component is illustrated as a function of integration time. There is an induction period of 10–15 days during which the error grows very slowly, for the next 20 days it grows rapidly and approximately linearly, and then it reaches a saturation level where the two simulated fields are uncorrelated.

The relative contributions of model error and initial-condition error to the total forecast error, and therefore to the predictability, are not well understood nor are they easy to individually quantify. One approach is to define the growth in the total error through a comparison of the forecast with observations, or analyses of observations. Then, initial-condition-related error growth is estimated by calculating the difference between two forecasts that are initialized from slightly different initial times, for the same integration period. This is illustrated in Fig. 8.2 for the ECMWF ensemble-prediction system. The

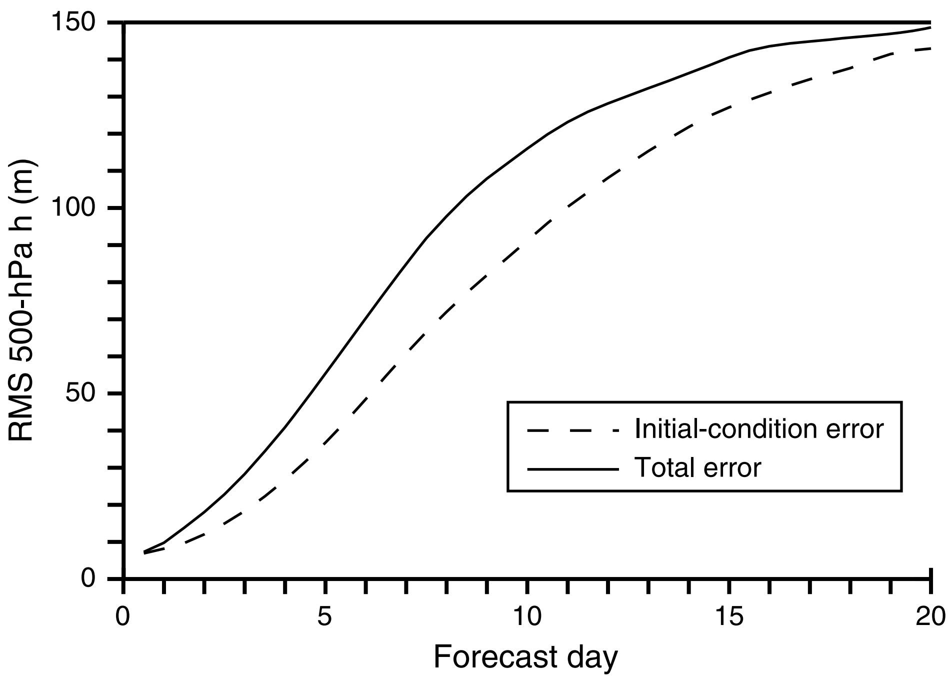

Fig. 8.2 Growth of 500-hPa error associated with initial-condition error only (dashed line) and with both model and initial-condition error (solid line) in an ECMWF global model. The total error is based on a comparison of the model solution with observations and the initial-condition error results from the growth of differences between parallel simulations from the same model initialized 12 h apart. Adapted from Leutbecher and Palmer (2008).

upper curve corresponds to the total model error, and the lower one is based on the growth of differences in the initial conditions of forecasts whose initialization times differed by 12 h. In this example, the initial-condition error is a large fraction of the total error.

8.3 Land-surface forcing's impact on predictability

The existence of the diurnally and seasonally varying solar forcing of Earth's surface and atmosphere causes processes that are tied to this cycle. Thus, correctly defining this forcing in the model will produce circulations and structures that are seasonally and diurnally dominant, without the processes necessarily being well observed in model initial conditions. Examples of diurnally varying phenomena that require thermal forcing are some low-level jets, sea breezes, mountain-valley breezes, urban heat-island circulations, and a variety of moist-convective processes. Figure 8.3 illustrates the near-surface *u* wind component observed over three diurnal periods near the slope of a north–south-oriented mountain range. If a model resolves the orography and represents the diurnal heating and cooling cycle reasonably well, these reoccurring winds, which are a dominant local feature, should be predictable, especially for weak synoptic-scale regimes. On seasonal time scales there are monsoons, the migration of the Hadley circulation and associated precipitation and trade winds, and the subtropical high-pressure systems whose seasonal migration with the Sun produces the Mediterranean-type climates described later. Getting the

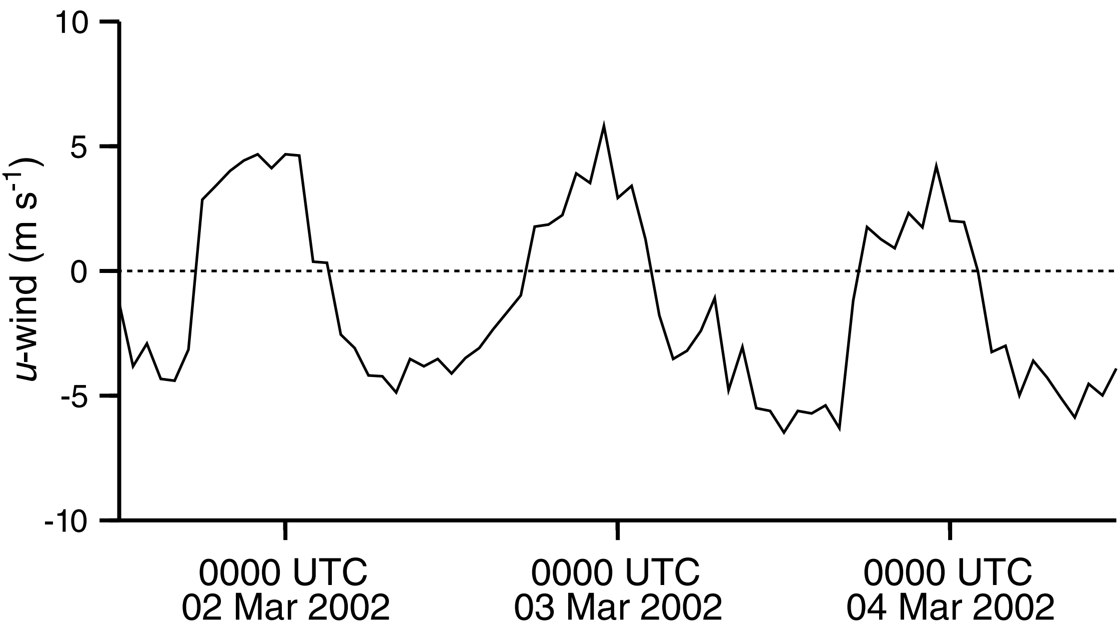

Fig. 8.3 The 10-m AGL *u* wind component observed over three diurnal periods near the slope of a north–south-oriented mountain range. The weak 3-day-average wind for the period has been subtracted from each point in the series. Adapted from Rife *et al.* (2004).

solar forcing prescribed correctly in a model, and using a good LSM to translate this energy source into appropriate sensible-heating patterns at the lower boundary of the atmosphere, will positively impact predictability.

8.4 Causes of predictability variations

There are numerous dependencies of the predictability of models (and the skill of human forecasters) on the geographic region, the local climatology, the season, the weather-regime, the time scale of a phenomenon, and the dependent variable. The following sections discuss a few examples.

8.4.1 Regional and climatological variability

Because of the planetary-scale circulation of the atmosphere, there are some regions for which the weather patterns on both large scales and mesoscales vary little from day to day, so the weather is relatively easy to anticipate using even simple tools like diurnal persistence. For example, when the trade winds prevail in a tropical region the wind speed and direction tend to be very regular when they are not interrupted by convective events. Again for the tropics, if a location is near a coastline, the diurnal variation of the winds associated with a sea-breeze circulation, as they are superimposed on the trade winds, is very predictable. And, the development of cloud and precipitation associated with the inland penetration of the sea-breeze front are regular parts of the local climatology and easy to anticipate as well. Lastly, there are some climates that are so dominated by the planetary circulation that their weather is virtually identical every day of the year. For example, there is a large area of northeastern Africa that experiences cloudiness less than 2% of the time.

8.4.2 Seasonal variability

For some regions that are seasonally influenced by subtropical high-pressure centers, half the year is dominated by subsidence from the Hadley circulation. During those months, days are generally cloud free, with no precipitation or other disturbances. Again, simple forecasting methods such as diurnal persistence are difficult to improve upon. During the rest of the year, when the latitudinal positions of the subtropical high-pressure centers and the storm track are different, the region can be dominated by synoptic-scale storms that are challenging to predict. Regions having seasonal variability in the predictability of their weather, with cloud-free warm-season months and cyclones in the cold season, include the Mediterranean Sea and the west coast of North America.

8.4.3 Weather-regime dependence

Predictability varies by weather regime, and on longer time scales, for reasons that are sometimes understood, and sometimes not well understood. The existence of the longer time-scale trends suggests that the predictability variation is not random, and may be associated with low-frequency variability or oscillations. Of course, one of the benefits of ensemble prediction is that we are provided with information about predictability on a day-to-day, or regime-to-regime, basis. It is simply pointed out here that this predictability can vary significantly in an organized way over relatively long time-scale shifts in weather regimes. As an illustration, Fig. 8.4 is of the Northern Hemisphere anomaly correlation (see Chapter 9 for a discussion of this verification metric) between 108, 15-day global-model forecasts and corresponding analyses of observations, for approximately a 3-month period. Clearly the predictability varies on time scales of a few weeks to a month. The plots in Fig. 7.3, showing the great difference in the ensemble dispersion of two forecasts one year apart, could be illustrating a regime dependence, or some variation on shorter time scales. Various studies have formally documented the relationship between predictability and the existence of various types of flow patterns and low-frequency variability. For example, Tracton (1990) states that a strong association exists between atmospheric predictability and the existence of blocking events in midlatitudes.[1] The onset of a blocking pattern results in a dramatic drop in predictability, and the collapse of the blocking

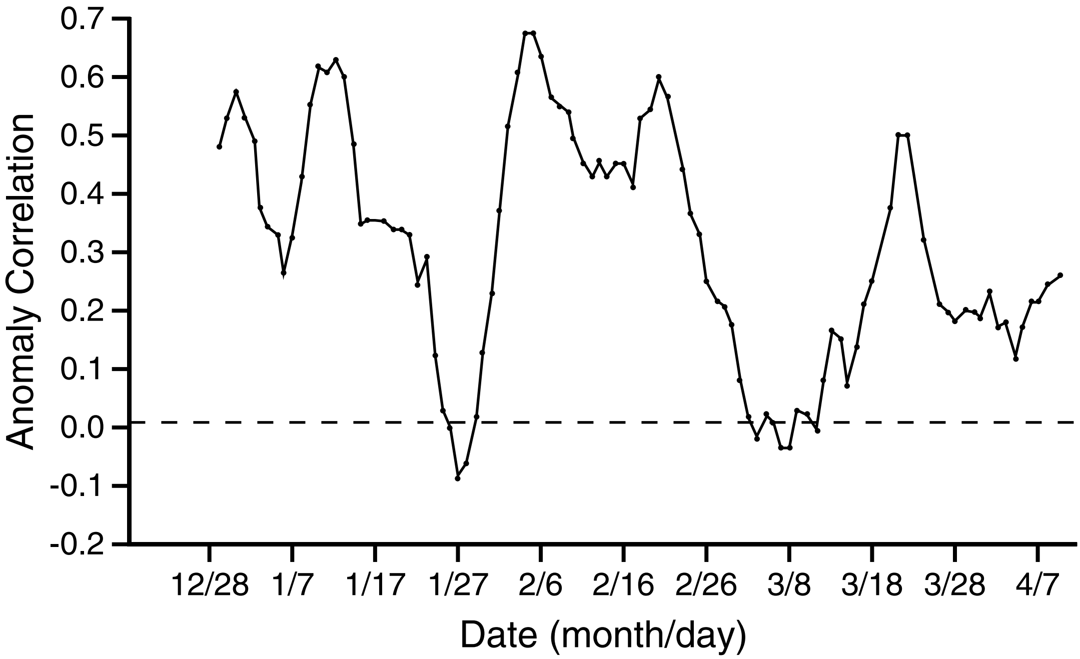

Fig. 8.4 The Northern Hemisphere anomaly correlation between 108, 15-day global-model forecasts and analyses of observations, for approximately a 3-month period. Adapted from Tracton (1990) and Tracton *et al.* (1989).

[1] Blocking refers to a situation where there is an obstruction to the west-to-east progress of migratory cyclones and anticyclones in midlatitudes. This situation is normally associated with upper-tropospheric closed anticyclonic circulations at high latitudes and cyclonic circulations at low latitudes. It may be viewed as an extreme-amplitude pattern of the ridges and troughs that normally prevail in the westerlies. This anomalous circulation remains stationary or moves slightly to the west, and can persist for weeks.

situation leads to a recovery in the predictability. In general, when and where there exist states of the atmosphere that are close to an instability threshold, whether it be baroclinic instability or some type of convective instability, the predictability is less because small perturbations that are poorly observed or poorly resolved by the model can cause the system to take alternative trajectories in phase space.

8.4.4 Phenomenon time-scale dependence

As described in Chapter 9, it is illustrative to consider model predictability in the context of three time-scale regimes: periods longer than diurnal (super-diurnal), periods that are approximately diurnal, and periods that are shorter than diurnal (sub-diurnal). Meteorological features with longer-than-diurnal periods are generally associated with synoptic-scale or planetary-scale processes, and are therefore reasonably predictable by global or regional models. Diurnal time-scale motions of course are related in some way to the heating cycle. In the wind field, a diurnal signal could be related to stability-related vertical momentum mixing, mountain-valley circulations, coastal circulations, etc. Provided that the model reasonably represents the land-surface and boundary-layer processes, features with these time scales should be reasonably predictable. Motions with sub-diurnal time scales include mesoscale features or circulations that are not thermally forced by the diurnal heating cycle. They can result from orographic or other landscape forcing, perhaps far upstream, or from nonlinear interactions. Given the sparse nature of the radiosonde network, these mesoscale features are not represented well, or at all, in three dimensions by the observation network, and therefore are not in the model initial conditions. Unless they are locally generated through nondiurnal forcing, they are not deterministically predictable by any model, no matter how good the resolution and physics. Time-scale dependences of predictability for longer-period variations in the atmosphere are discussed by van den Dool and Saha (1990).

8.5 Special predictability considerations for limited-area and mesoscale models

An aspect of LAMs that affects predictability is the existence of LBCs. It was explained in Section 3.5 that information from upstream LBCs will sweep across a model grid during an integration, as information from the initial conditions exits the grid through the outflow boundary. This means that, unlike the situation with global models, the predictability of phenomena will depend less on initial-condition error for longer forecast lead times and more on LBC error. The importance of the LBCs to predictability with mesoscale LAMs is easily understood when one considers the number of important mesoscale phenomena that occur only when the large-scale atmospheric characteristics produce a conducive environment. Examples include frontal squall lines, mesoscale convective complexes, coastal fronts, and freezing rain. Also, land–atmosphere interaction has especially important controls on the solution of mesoscale LAMs (see Chapter 5). Thus, as noted in Section 8.3

above, the ability of the LAM to properly represent these processes will be an important factor in defining the predictability.

There is also perhaps a practical difference in the criteria used to define the predictive skill in forecasts by limited-area, mesoscale models, relative to synoptic- and global-scale models. For forecasts on the large-scale, predictive skill may arguably be defined with a continuous scale in terms of the phase error of waves or accumulated-precipitation amounts. On the mesoscale, in contrast, predictability may be in the context of whether individual high-impact events were forecast or not. The prediction of the existence of a severe-weather event, even in an incorrect location, may be viewed as a very successful forecast. In contrast, forecasters may consider the model solution to have zero utility beyond the point in the simulation when a major precipitation event was not forecast by the model.

Some of the smaller-space-scale processes that motivate the use of mesoscale models, such as moist convection, also have short time scales. And, it is sometimes stated that a practical limit to predictability is one life cycle of a physical process. This leads to a predictability limit of less than a few hours for individual convective events (not long-lived convective complexes).

Predictability based on mesoscale dynamic models has historically been considered to be somewhat complementary with the predictability based on algorithms that perform some sort of extrapolation from the current state. This is shown in the schematic of Fig. 8.5.

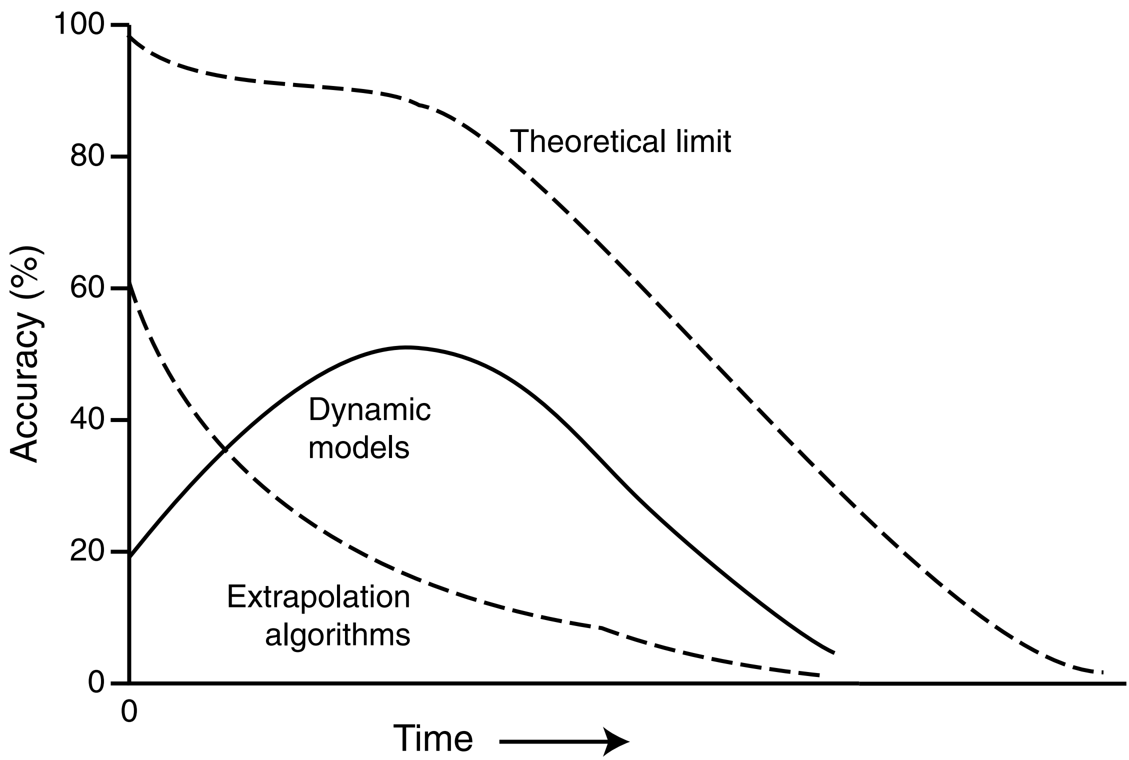

Fig. 8.5 Schematic of the accuracy of forecasts from dynamic models and extrapolation algorithms, relative to the theoretical limit to predictability.

Extrapolation methods, of which persistence is perhaps the simplest, begin with a best estimate of the current state of the atmosphere, and evolve it using statistical or other algorithmic approaches. Because there is no atmospheric dynamics involved, nonlinear processes cause the accuracy to deteriorate quickly, and thus the predictability limit for this approach is short. On the other hand, mesoscale models sometimes require a period of time early in the forecast during which local, thermally driven circulations spin up. And, models can suffer from a dynamic adjustment period after initialization, which causes the solution during at least the first 6–12 h to be problematic. Both the algorithmic and dynamic approaches begin from an imperfect state because of the coarseness of the observation network. Even though the details of the curves will be model and data-assimilation-system and weather-situation dependent, there is likely a period of time during which algorithmic methods have greater predictability than do dynamic models.

8.6 Predictability and model improvements

New data-assimilation methods, numerical algorithms, and physical-process parameterizations are routinely being developed with the goal of increasing the predictability of models for use in operational prediction or research. This section provides advice to consider when methods that are thought to represent improvement are implemented, but there is no apparent benefit to the accuracy of the model products. One issue is that a new parameterization, for example, may be very sophisticated and include more interactions among components of the physical system than did the previous version. However, new required inputs for the more-complex parameterization may be so poorly known that the new method performs worse than the older and simpler approach – more complicated methods do not necessarily lead to better performance. An example is that new multi-level LSMs of one or two decades ago had difficulty producing more-accurate predictions than were obtainable from single-substrate-layer slab models. Or, cloud-microphysics models with more microphysical particle types will not perform better than simpler methods if the conversion among types is not parameterized realistically.

Another consideration is related to the forecast verification metric that is used to assess the impact of a model change on predictability. It will be shown in Chapter 9 that standard metrics such as RMSE and MAE often produce better verification scores for smoother model solutions (e.g., Fig. 9.4). Thus, if a change to the model increases its ability to represent small-scale features, the predictability may appear to decrease. Model changes that could lead to this misleading response of the predictability metric are decreases in the horizontal grid increment; improvements in the filtering properties of the numerical algorithms such that finer-scale features are retained; or the use of higher-resolution estimates of the variability of landscape properties, which could lead to finer-scale structures in the boundary layer.

Lastly, there is the weak-link concept. The modeling system consists of many interacting components, and weaknesses in one of them can prevent improvements in another from leading to a better model prediction. There are numerous possible examples of this, but an obvious one is that a new observing system will not improve predictability if the data-assimilation scheme is not able to adequately use the observations.

8.7 The impact of post processing on predictability

The post processing of model output, described in Chapter 13, should be viewed as an integral part of the modeling system, at least for operational applications. This processing can take many forms, but one involves the use of methods for reducing the systematic error in the forecast products. The resulting better correspondence with observations will result in improved predictability for the entire system.

SUGGESTED GENERAL REFERENCES FOR FURTHER READING

Holloway, G., and B. J. West (eds.) (1983). *Predictability of Fluid Motions*. New York, USA: American Institute of Physics.

Kalnay, E. (2003). *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge, UK: Cambridge University Press.

Leutbecher, M., and T. N. Palmer (2007). Ensemble forecasting. *J. Comp. Phys.*, **227**, 3515–3539.

PROBLEMS AND EXERCISES

1. Read about Observing System Simulation Experiments (OSSEs) in Section 10.2, and explain how that type of experiment might be used to estimate the relative contribution of model and initial-condition error.
2. Refer to the weak-link concept in Section 8.6, which is related to how model changes impact predictability, and provide additional examples of why anticipated benefits may not be realized.
3. What types of mesoscale processes that can be predicted by LAMs have predictability limits of at least one day?
4. Referring to the accuracy curves for mesoscale models and algorithmic nowcasting methods, shown in Fig. 8.5, discuss what modeling-system, weather-regime, and scale factors will influence the shapes of the curves and their relationship with each other.
5. The predictability of clouds is quite low. Speculate on the reasons.