ENSEMBLE PREDICTION SYSTEMS

S. Potempski

CoE MANHAZ, Institute of Atomic Energy

CONTENTS

1. Chaotic behavior of atmosphere ... 362

2. Ensemble forecasting ... 363

2.1 Theory ... 363

2.2 Practical realization of ensemble forecasting ... 365

3. Ensemble graphics ... 366

4. Ensemble indicators ... 368

4.1 Arithmetic mean or average ... 368

4.2 Median ... 368

4.3 Mode ... 368

4.4 Measures of variability ... 369

4.5 Standard Deviation (SD) ... 369

4.6 Percentile ranking of data ... 369

4.7 Commonly used percentile rankings ... 369

4.8 Measures of shape ... 369

4.9 Skewness ... 369

4.10 Kurtosis ... 370

4.11 Multimodal PDFs ... 371

5. Using PDFs in the forecasting ... 371

5.1 Methods not based on NWP models ... 371

5.2 Methods using a single NWP forecast ... 371

5.3 Methods using ensemble forecasts ... 371

References ... 372

1. Chaotic behaviour of atmosphere

The atmosphere behaves in a chaotic way. This statement, although simple must be explained more deeply. According to E. Lorenz's papers (Lorenz, 1965) small perturbations of initial conditions can lead in time to completely different results. So the word "chaotic" is not understood as in colloquial language, but rather in order to describe some special behaviour. In fact the atmosphere is described by deterministic equations, which give a deterministic solution – however, the problem arises while evolution of this solution in time. Small initial errors can provide to big differences in results. So we should rather speak about "deterministic chaos". A typical picture of such a situation is shown on Fig. 1.1 (time 48 h), called "spaghetti plot" – results of all simulations for specific contours are presented (we can observe essentially different results at the Black Sea). It is generally agreed that the maximum reasonable simulation time is about two weeks. After a longer time results are so different that no practical conclusions can be drawn concerning prognosis.

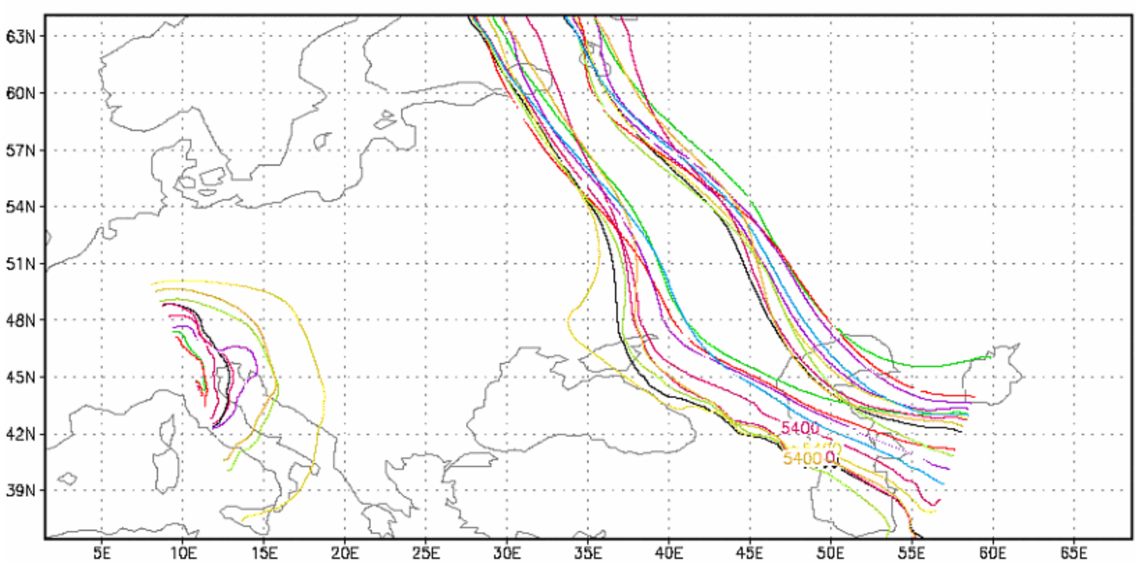

Fig.

Figure 1.1. Geopotential heights shown 5300-5400m – t48h.

Evidence of the chaotic nature of the atmosphere is also seen in different results obtained while running numerical weather prediction (NWP) models with identical initial conditions, but various model dynamics and parameterizations. Model forecasts can be sensitive to the design of the model as well as to the initial conditions. Each model configuration approximates the actual behaviour of the atmosphere differently, so this introduces another source of forecast uncertainty. We will never be able to construct an NWP model that includes the behaviour of the atmosphere in every detail at infinitely high resolution. But even if we could create such a "perfect" NWP model, its forecast would eventually break down because of errors in initial conditions, although such a breakdown might take longer to happen. The atmosphere's sensitive dependence on initial conditions means that model initial conditions would need to be "perfect" as well for there to be any hope of making a perfect forecast. Of course, the reality is that our observation and assimilation systems will never give us perfect initial conditions. We can, however, apply our knowledge that the atmosphere is chaotic, highly sensitive to initial conditions, to the forecast process.

Through strategic use of imperfect initial conditions and/or imperfect NWP models in an *ensemble prediction system* (EPS), we try to:

- ascertain a range of *possible* forecast outcomes,
- estimate the *probability* for any individual forecast outcome,
- hopefully determine the most likely forecast outcome.

Some version of an EPS is in use at most forecast centers. These operational EPSs typically use initial condition uncertainty as a basis for their multiple forecasts, and in addition, some also use model uncertainty (imperfect structure and dynamics) and boundary condition uncertainty.

We present below the table summarizing the advantages of an ensemble system of deterministic forecasts over a single deterministic forecast (source: NCAR).

Table 1.1. Advantages of an ensemble system of deterministic forecasts over a single deterministic forecast.

Characteristic	Single Forecast	Ensemble Forecast System
Uncertainty in Initial Conditions	Data assimilation system is designed to minimize initial condition errors using various forms of data. Uncertainty is implicitly (but incompletely) accounted for through relative weighting of each piece of observational data and the model forecast first guess. Can assess initial condition error using satellite and other observations, but cannot explicitly assess subsequent impact on NWP forecast.	Initial condition uncertainty can be taken into account by determining the most important (i.e., rapidly growing) potential errors to the subsequent model forecast and scaling them down to a reasonable initial condition perturbation.
Atmospheric Predictability	Cannot be assessed from a single deterministic forecast. Can be incompletely inferred from the degree of consistency between consecutive model forecast runs.	Can be assessed by the rate of growth in spread of ensemble member forecasts. Ensemble size and adequate initial condition perturbation are important in obtaining an adequate ensemble spread and a measure of predictability.
Model Uncertainty: Dynamics	Only one numerical method can be used, for example, breaking down the flow into sines and cosines (spectral method).	Multiple numerical methods can be used, e.g., spectral, grid-point, grid-point with different variable configurations on the grid.
Model Uncertainty: Physics	Only one set of physical parameterizations can be used (e.g., one convective precipitation scheme).	Multiple combinations of physical parameterizations can be used (e.g., two types of convective precipitation schemes might be used to combine their individual strengths).

2. Ensemble forecasting

2.1 Theory

Since the actual state of the atmosphere at any time is known only approximately, a complete description of the weather-prediction problem should be formulated in terms of the time evolution of an appropriate probability density function (PDF) in the atmosphere's phase space. Although this problem can be formulated exactly through the continuity equation for probability, also known as the Liouville’s equation (Kalnay, 2001), its practical solution is impossible for nonlinear models with more than a few degrees of freedom. Even restricting attention to the evolution of the first- and second-order moments of the atmospheric PDF, one is still faced, for medium-range forecasting, with a system of nonlinear equations which have no well-defined closure and which cannot be solved for the large models currently used in numerical weather prediction.

Ensemble forecasting appears to be the only feasible method to predict the evolution of the atmospheric PDF beyond the range in which error growth can be described by linearized dynamics. In ensemble forecasting, the PDF at initial time is represented through a finite sample of possible initial conditions. A nonlinear model integration is carried out from each of these states, and the properties of the PDF at any forecast time are assumed to be described by the sample statistics computed from the ensemble.

The ensemble statistics will approximate the correct PDF if:

(i) the sample of initial states provides a realistic estimate of the probability distribution of analysis errors; and

(ii) the phase-space trajectories computed by the numerical model are good approximations of atmospheric trajectories.

Requirement (ii) is also necessary for deterministic NWP; hence, most of the recent research in ensemble forecasting has focused on point (i). However, systematic or regime-dependent model errors can severely affect the ability of the ensemble to forecast not only the first moment of the PDF, but also higher moments, such as the standard deviation.

Requirement (i) poses a problem of considerable theoretical and practical difficulty. Firstly, the PDF of analysis error is poorly known; secondly, the number of independent directions in phase space spanned by this PDF (essentially the dimension of the NWP model) exceeds by many orders of magnitude the maximum practicable ensemble size for a realistic NWP model. As demonstrated by early experiments in ensemble forecasting, a sparse random sampling of phase space (even taking into account geostrophic and hydrostatic constraints) will not produce a realistic distribution of forecast states. For any given initial flow, only certain directions in phase space are associated with dynamical instabilities which will determine the growth of small perturbations (or errors) in the forecast.

Forecasts started from successive data-assimilation cycles tend to diverge at a rate which is smaller than, but comparable with, the actual error growth (Lorenz, 1996). The difference between the analysis at a given initial time and a very-short-range forecast verifying at the same time can therefore be considered as a growing perturbation consistent with our uncertainty in the initial conditions. This idea is exploited in the lagged-average forecasting (LAF) method (Kalnay, 2001), in which ensembles are composed of forecasts started from consecutive analyses. In this method, the ensemble size is limited by the number of available analyses in a relatively short time interval (typically not more than two days), and the ensemble members cannot be considered as equally likely (at least in the medium range). These problems become less serious if one is mainly concerned with just the first moment of the sample PDF, namely the ensemble mean, or when longer forecast ranges are considered. In fact LAF has been used in experimental programmes on extended-range ensemble predictions in several NWP centres.

More recently, however, techniques to generate initial perturbation have been based on strategies (commonly used in dynamical systems theory) to identify those directions in phase space where dynamical instabilities are strongest. One possibility is to assume that errors in the initial conditions are dominated by those instabilities of the flow which have developed over a series of previous assimilation cycles. This assumption is the basis of the "breeding" method proposed by Toth and Kalnay (Toth and Kalnay, 1993), which corresponds to the computation of the vectors associated with the largest Lyapunov exponents of the NWP model. The NMC ensemble prediction scheme is based on a combination of the breeding and LAF techniques, and at least partially satisfies conditions (i) and (ii) above.

However, even assuming an isotropic PDF in phase space for the error at the initial time, the different amplification rates of perturbations along different axes would soon stretch the PDF along the directions of maximum instability during the early stages of the forecast period. In this way a particular phase-space direction, which perhaps was not necessarily associated with exceptional analysis error, may turn out to dominate the forecast error after a day or two. It would appear to be important to ensure that this direction was properly sampled by the ensemble of initial states.

As first shown by Lorenz (Lorenz, 1965) in a meteorological context, for any finite time interval in which the dynamics of perturbations is assumed to be linear, the axes of maximum instability can be computed as the eigenvectors of a symmetric operator defined as the product of the linear propagator by its adjoint. In dynamical systems theory, this operator is sometimes referred to as the Oseledec operator (Alligood et al., 1996). In linear algebra notation, these eigenvectors are the singular vectors (SVs) of the linear propagation itself. SV growth can be much faster than either normal-mode growth (for stationary flows) or Lyapunov exponent growth (for time-evolving flows see, for example,(Buizza et al., 1993), (Farrell, 1988), (Molteni and Palmer, 1993)).

Ensemble forecasting experiments in which unstable SVs computed from a 3-level quasi-geostrophic model were used to construct initial perturbations for a multilevel primitive-equation model were carried out at the ECMWF in the past years, and have been reported by Mureau *et al.* (1993), and Palmer *et al.* (1993). This approach proved to be more successful than alternative ensemble techniques tested at the ECMWF. However, the inconsistency between the vertical coordinates of the quasi-geostrophic and primitive-equation model created difficulties in the vertical interpolation over high topography. Although this problem did not affect the strongly unstable SVs localized on the western side of the oceans, continental features often had a smaller growth rate in the primitive-equation than in the quasi-geostrophic model (Mureau *et al.* 1993). Efforts were, therefore, directed towards the computation of SVs in a simplified primitive-equation environment, using an iterative Lanczos algorithm for the solution of the eigenvector problem (Buizza, 1993).

SV structures are dependent on the choice of inner product. In relating these structures to analysis error we give arguments to suggest that a suitable inner product can be defined in terms of perturbation energy. Determining analysis-error statistics from conventional data-assimilation techniques is difficult, and relatively little quantitative knowledge about flow-dependent 3-dimensional structure of analysis-error statistics is available. Recently, however, the development of adjoint models allows investigation of the component of analysis error which on any given day has the greatest impact on short-range-forecast error. This technique provides forecast “sensitivity” fields which can be directly compared with the SVs at initial time.

2.2 Practical realization of ensemble forecasting

Historically in 1992 in NCEP and ECMWF introduced the first ensemble approach in the operational NWP Systems. Concerns the perturbation of initial conditions one may take into account the following possibilities:
- Monte Carlo perturbations of controls,
- dynamical perturbations connected with the „growing errors of the day” or dynamic perturbations.

But even the best realistic Monte Carlo perturbations, compatible with the average estimated analysis errors, do not contain the finite size „growing errors of the day”, which are present in the analysis, and which are the fastest growing errors of the day.

Therefore there have been elaborated two practical methods of dynamic perturbations:
- „breeding” of „bred vectors” (BV) or „growing modes”(BGM), NCEP,
- singular vectors (SV), ECMWF.

At NCEP was introduced a method of finding in phase space vectors indicating the magnitude and direction of fastest growing modes or perturbations. It is achieved by introducing to the first basic analysis (often called control analysis) perturbations and running the model and then subtracting from the results of the perturbed run the control run results. The obtained difference is normalized and again used in the next run. After simulations of few days period one gets the perturbed forecast (the Bred Vector) consisting of a superposition of the fastest growing modes of the atmosphere during the analysis cycle. Those vectors are closely related to Lyapunov’s vectors. They may be considered as a rough (finite) but nonlinear approximations to leading Lyapunov’s vectors during the assimilation cycle.

In contrast, in ECMWF the method based on Singular Vectors is used. They compute the fastest growing modes but in linear way using tangent linear model propagator and its adjoint. Those modes are also very close to Lyapunov’s vectors but rather belonging to the linear counterpart of the model propagator. Unfortunately this method is very expensive as concerns necessary computer resources. In contrast to BV the SV strongly depends on the selected norm.

On NCEP servers one can now find every day controls plus positive and negative perturbations from Global Ensemble Forecast (GSF) for consecutive assimilation and forecasting hours. They are intensively used by meteorological community for performing their specific tasks. They are known as global controls and global bred (perturbation) vectors.

Another way of performing ensemble forecasting is based on introducing ensemble of randomly perturbed observations. Such a method is being used by Canadian Meteorological Center and it can be considered as very specific example of ensemble based data assimilation system or an Ensemble Kalman Filter.

Very important and interesting is an approach connected with introducing errors in the model equations (due for example to sub-grid processes, truncation and other approximations in numerical methods and physical description of atmospheric processes). One can perturb the physical parameters of the model used and have an ensemble of models or simply take into consideration many models (Multi-Model Systems).

Still another way of constructing Ensemble Kalman Filter is to take into account simultaneously analyses from many meteorological centres and to take them as a basis to create an ensemble of forecasts (Multianalysis Systems). Such an approach is frequently called as a Poor Man's Ensemble Forecasting System.

The presented ideas can be certainly introduced also to the Regional Mesoscale Models (RMM). Usually those models are in some way embedded into Global Forecasting Systems (GFS) by importing from them the necessary lateral boundary conditions. In majority of cases also the Bred Vectors or Singular Vectors are taken from GFS to form the perturbed Initial Conditions.

Summarizing the following approaches can be implemented practically:
- singular vectors, ECMWF,
- bred vector, NCEP.
- perturbation of measurement data (used in analysis), Canada,
- multi analysis (different analyses, one model),
- multi model (different models or parameterizations),
- any combination of above.

3. Ensemble graphics

To help with the interpretation of the distribution of forecasts, ensemble output is frequently displayed in various graphical forms, particularly "postage stamp"' maps, "spaghetti"' diagrams and "plume" diagrams. Postage stamp maps consist of forecast maps for all of the ensemble members, presented on one page, covering a limited domain of interest, and at a particular level. They most often depict the selected contour or surface fields, for a particular forecast projection. These present enormous amounts of information, and may be hard to read and interpret, especially if the ensemble has many members. To help with the interpretation of spatial fields from the ensemble, the forecasts may be processed to help group "similar"' patterns or, alternatively, identify patterns that are markedly different from the others. Two such processing methods in operational use are clustering and tubing. Clustering involves calculating the differences among all pairs of ensemble members, grouping together those for which differences are small, and identifying as separate clusters those members with larger differences. Tubing involves identifying a central cluster which contains the higher density part of the ensemble distribution, then identifying "outliers" - extreme departures from the centroid of the main cluster. The line joining the centroid of the main cluster and each outlier then defines the axis of a tube, which extends in the direction of the outlier from the main cluster. The tubes may be interpreted as an indication of the directions in which the forecast may differ from the ensemble mode. Whatever the similarities and differences between these two methods of categorizing or grouping the members of the ensemble distribution, the aim is the same: to organize the massive information content of the ensemble to become more easily interpretable by forecasters.

Spaghetti plots represent a sampling of the full ensemble output in a different way (Fig. 1.1). In a spaghetti plot, a single contour is plotted for all the ensemble members. Thus the main ridge and trough features of the contour can be seen along with the way in which they are forecast by each ensemble member. Areas where there is greater uncertainty immediately stand out as large scatter in the position of the contour. One must be cautious in the interpretation of the chart, however, because large spatial scatter may not be significant when it happens in areas of flat gradient. For this reason, a measure of the ensemble spread such as the standard deviation is usually plotted on the map (Fig. 3.1). Then, areas of large apparent uncertainty in the contour forecast can be checked to verify that they also coincide with larger ensemble spread.

Plume diagrams are equivalent to spaghetti plots. Instead of plotting the spatial distribution of the ensemble member forecasts at a particular forecast projection, the distribution of the ensemble forecasts is plotted for a particular location as a function of projection time. Plume diagrams are most often prepared for elements such as concentration, for specific stations. For ease of interpretation, the probability density of the ensemble distribution might be contoured on a plume chart.

Another format in which single-location ensemble output is presented are "box-and-whisker" plots of direct model output element forecasts (Fig. 3.2). The boxes shown at each 6 h indicate the range of values forecast by the ensemble between the 25th and 75th centile (the middle 50% of the distribution), while the median of the ensemble is indicated by the horizontal line in the box. The ends of the lines extending out of both ends of the box (the "whiskers") indicate the maximum and minimum values forecast by the ensemble. These kinds of plots are an effective graphical way of depicting the essential characteristics of a distribution. Asymmetries and the spread of the distribution can be immediately seen, and variations over the period of the forecast are also immediately apparent.

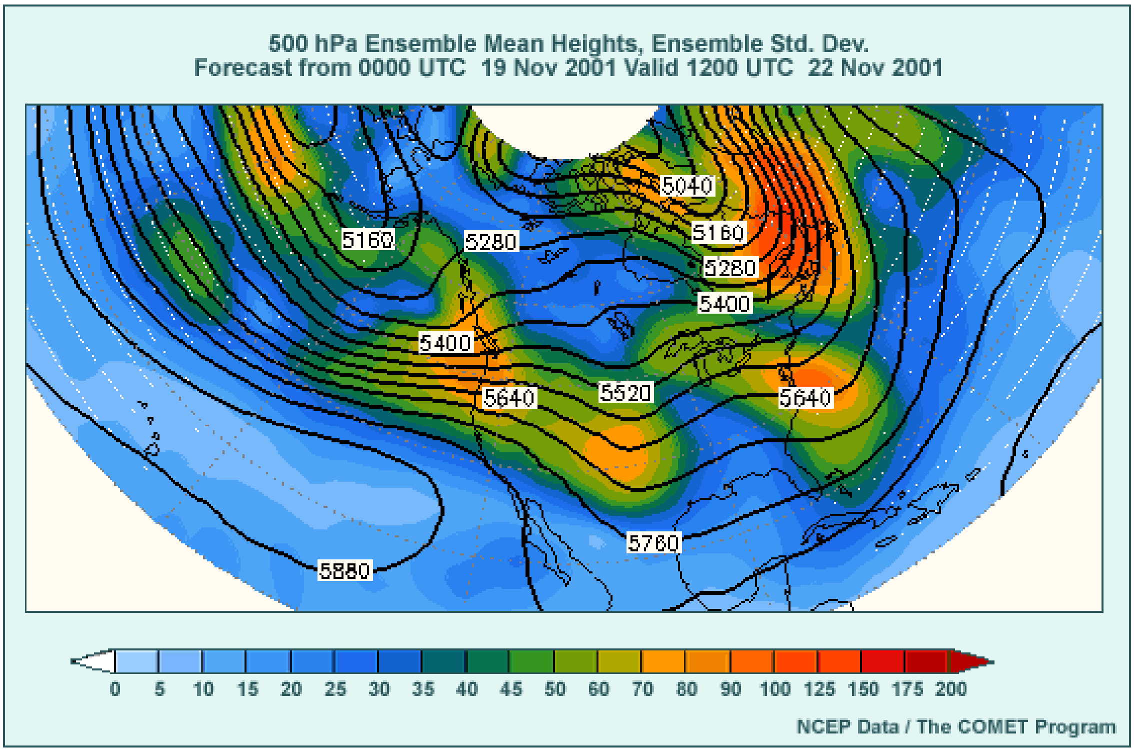

Figure 3.1. Ensemble mean heights and standard deviation (NCEP).

Other types of pictures are also often related to statistical analysis. Among them are:

- histograms for specific location,
- variable thresholds for probability of exceedance,
- plume diagrams (time analysis) for specific location,
- ensemble soundings (profiles), which corresponds to “vertical spaghetti” plots.

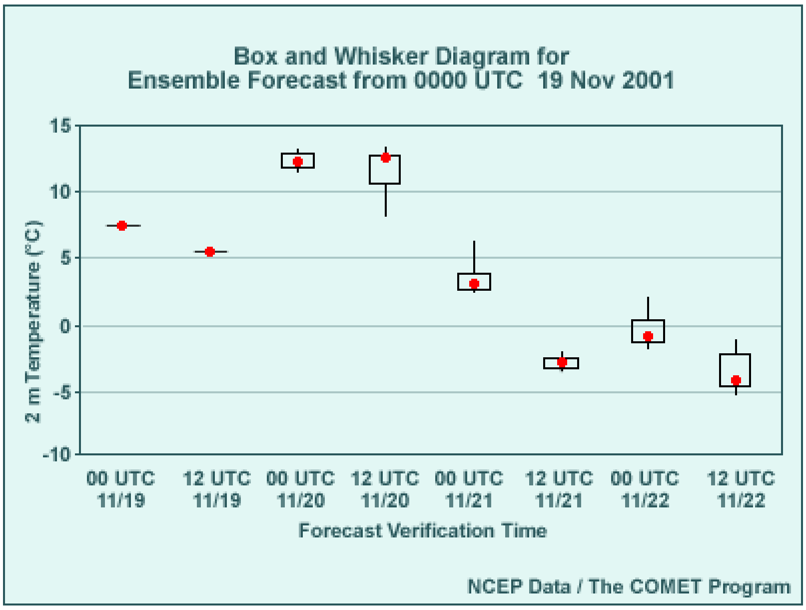

Figure 3.2. Box and whisker diagram (NCEP).

4. Ensemble indicators

Several statistical measures are used to describe the distribution of a theoretical data distribution in terms of the location of its "middle" or central tendency, its variability or spread, and its general shape. Where is the middle of the data? A number of statistical measures give us a notion of where the middle of a data sample is, but since they use different concepts of middleness, they can give very different answers when the data does not fit a bell-shaped normal distribution. The following sections describe three different concepts of middleness.

4.1 Arithmetic mean or average

The *arithmetic mean* or *average* of a data sample is simply the sum of the values divided by the total number of values, or

$$\overline{x} = \sum_{i=1}^{n} x_i / n$$

where x is the variable of interest, the over bar indicates a mean quantity, and n is the number of values.

4.2 Median

If we rank a data set from lowest to highest, we can locate the position where one-half of the data is lower and one-half is higher. The value at this position is the *median* of the data sample. Note that taking the median reduces the influence of extremely high and/or low values, or outliers, in the data sample, which can make the mean less representative of the true middle. Also, if there is an even number of data, the median is defined as the average of the *N/2-th* and the *(N/2 + 1)-th* ranked data values.

4.3 Mode

The statistic of central tendency given by the most frequently observed value or interval is the *mode* of the data sample. In the case of the mode, *no* values other than those in the most frequently observed category affect the statistic.

The following table lists advantages and disadvantages of the measures of central tendency (source: NCAR).

Table 4.1. Advantages and disadvantages of the measures of central tendency.

Statistic	**Advantages**	**Disadvantages**
Mean	Takes entire data sample into account For normal distributions, is the most stable measure of central tendency when using sample data to infer the central tendency of the larger population	Not representative of the central tendency when the data sample is skewed (see the Shape section) Can be strongly affected by extreme values, particularly for small samples
Median	Is not affected by extreme values Good for skewed (non-symmetric) distributions	Must sort data Does not use all data values
Mode	Is not affected by extreme values Can be used for non-numerical data (e.g., precipitation type) Can uncover multiple maxima (if they have the same count)	Does not use all the data Sensitive to how data intervals are assigned

4.4 Measures of variability

Now that we've covered the central tendency estimates, the middleness, for our ensemble data, let's take a look at measures of variability to get a better notion of the spread in the data. Good measures of variability use all the data and increase as the population or the sample spread in the data increases.

4.5 Standard Deviation (SD)

The first of these measures, the SD, assumes that the data have a normal distribution. The SD is the square root of the *variance*, which is the average squared difference between each datum and the mean of the data sample. The formula for the standard deviation is:

$$S = \sqrt{\sum_{i=1}^{N} \frac{(x_j - \bar{x})^2}{N-1}}$$

where *N* is the size of the data sample, *x* is the variable of interest, and *s* is the sample standard deviation. In the formula, *N* is reduced to *N-1* because it can be shown that using *N* in the denominator underestimates the true (population) variance. The SD is used to measure distance from the mean.

4.6 Percentile ranking of data

Another way to describe ensemble data is by ranking it so that we can describe the relative position of a particular member within the full ensemble. A measure that expresses this position in terms of a percentage is called **a** percentile. The percentile of a value gives the percentage of the total set of data that falls below the value. The median, by definition, is the 50th percentile.

4.7 Commonly used percentile rankings

Quartiles are used to describe data broken into 4 equal parts and constitute the 25th percentile, the median, and 75th percentiles. Consequently, the 25th percentile, the median, and the 75th percentile are the breaks between the lowest and second quartiles, the second and third quartiles, and the third and upper quartiles, respectively, for a group of data. Deciles break the data into 10 equal parts of 10% each, and have breaks at the 10th percentile, 20th percentile, and so on through the 90th percentile.

If a percentile falls between two ranked items, the percentile break point is determined by interpolation between the ranked items.

4.8 Measures of shape

Experienced forecasters certainly understand that many atmospheric processes are *not* normal (or normally distributed, for that matter)! Rather, the chaotic behaviour of the atmosphere and the physical limits placed on quantities often result in asymmetrical PDFs. Processes and quantities that result in asymmetric distributions include individual rainfall events, cloudiness, and relative humidity, for example. How can we measure this asymmetry ?

4.9 Skewness

We may want to measure how far to one side data are "skewed" from the symmetrical normal distribution. Skewed distributions are found when quantities we are concerned with have physical characteristics that lead to limits on their values, such as daily precipitation amount (lower boundary at 0.00", physical limitation as to maximum possible precipitation) and wind speeds (lower boundary at zero, physical limitation on maximum values based on pressure gradient forcing, viscosity, friction).

The formula for the skew ness parameter is:

$$\sum_{i=1}^{N} \frac{(x_i - \bar{x})^3}{N\sigma^3}$$

Skewness measures the placement of the mean relative to the total distribution. A normal distribution will have a skewness of 0.0, as will other perfectly symmetric PDFs. A PDF with positive skewness—that is, "skewed to the right"—will have its maximum frequency (its mode) to the left of the median, and the arithmetic mean further to the right in the long tail. One with negative skewness—"skewed to the left"—will have its maximum frequency to the right of the median and arithmetic mean, and a long tail to the left. Hypothetical examples of both are shown in the graphic below (Fig. 4.1).

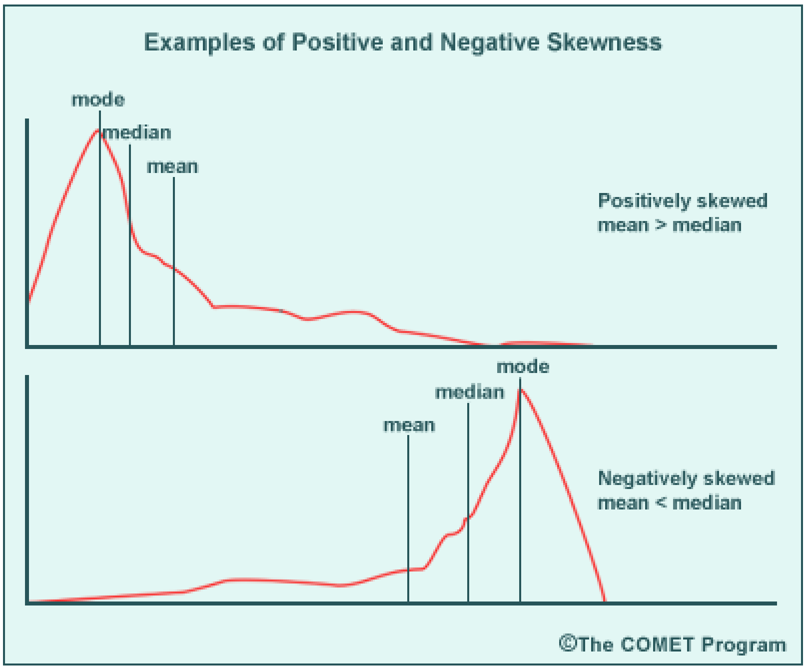

Figure 4.1. Examples of positive and negative skewness (NCEP).

4.10 Kurtosis

Kurtosis measures the size of the tails in a sample PDF compared to a theoretical normal distribution. Positive kurtosis indicates a "peaked" distribution and negative kurtosis indicates a "flat" distribution (i.e., tails smaller and larger than the normal distribution, respectively).

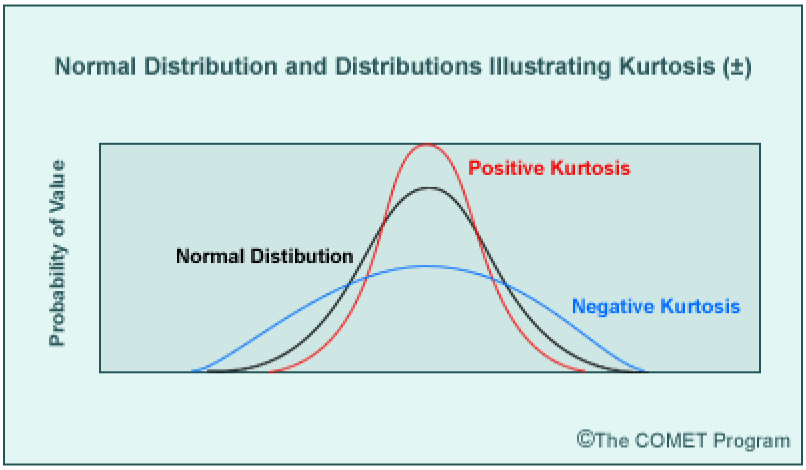

Figure 4.2. Distribution with Kurtosis (NCEP).

As the normal distribution has a kurtosis of 3.0, the formula for *excess kurtosis* subtracts 3 from the parameter:

$$\sum_{i=1}^{N} \frac{(x_i - \bar{x})^4}{N\sigma^4} - 3$$

4.11 Multimodal PDFs

EPS forecast probability distributions sometimes do not fit the theoretical distributions we've discussed up to this point. One common distribution is where there is more than one obvious peak. These are also known as multimodal distributions. Even though there is only one *true* mode for the distribution, the presence of similar, but separate, peak values makes this and other statistics of "middleness" less useful. For example, how representative of the middle are the *mean* and *median* for the distribution below, especially since their values fall between two bins with only 2 ensemble members, surrounded by bins with much higher probability?

A PDF with multiple probability maxima can indicate that the ensemble forecast doesn't have enough members to give an accurate reading on the probabilities of forecast outcomes, or, more frequently, there are multiple flow regimes that have a significant chance to verify.

5. Using PDFs in the forecasting

5.1 Methods not based on NWP models

Long before we had NWP models, weather forecasters made use of probability distributions obtained from observations. These methods include considering:

- Local climatology to assess likely future value of forecast variables.
- Current values for meteorological variables in making the forecast (persistence).
- Past evolution of the atmosphere in similar situations (forecast analogues).

Some of these methods still have their place today in the forecast process, if for nothing else than as a "sanity check" against what NWP models are telling us.

5.2 Methods using a single NWP forecast

Forecasters informally use probability distributions in the forecast process every day, based on their past experience in similar forecast situations. For example, a forecaster might assign the likelihood of reaching advisory-criteria heat index levels based in part on NWP forecasts for 850-hPa temperature and PBL relative humidity. By doing so, the forecaster essentially places these model variables on a subjective probability distribution for reaching the advisory criteria.

However, the data from which the statistics were derived include many different forecasts under numerous different regimes. Perhaps the 850-hPa temperature error is regime-dependent, for example, because under wet conditions the forecast temperature tends to be too high due to excess sensible heat flux from the model surface, while in dry conditions the forecast temperature is too low for analogous reasons. This kind of information is not available in the sample.

Additionally, we need information from a "frozen" model so we know the model error statistics are stable. (Changes to a model may and often do change the characteristic errors in that model.) If the warm bias noted above is eliminated, reduced, or even replaced by a negative temperature bias because of model changes, our range and expected value for the 850-hPa temperature will be based on incorrect data.

5.3 Methods using ensemble forecasts

Use of PDFs developed from relationships between observations and model variables, as shown in the previous examples, can be helpful in the forecast process. However, there are a number of disadvantages. For example, relationships between model forecasts and the subsequent verification are often flow regime dependent, which means that application of these relationships in at least some cases will not be valid. We also do not get a quantitative sense of how predictable the flow regime might be.

A logical method to avoid these problems is to use ensemble forecasts to get a PDF of possible forecast outcomes. Ensemble forecasts have distinct advantages over single deterministic forecasts because they take into account the following:

- Current initial condition uncertainty and atmospheric predictability
- Current flow regime effect on NWP model predictability and bias
- Current model configuration (Previous versions of the NWP model may have different characteristic errors and biases.)

Additionally, when appropriately calibrated, ensembles can adjust for imperfections in NWP models and recent model performance in a systematic, objective way.

The number of forecasts we can have in an ensemble run is obviously constrained by available computing power to create the different ensemble members, and this constraint makes it important to carefully construct our ensemble prediction system (EPS).

References

Alligood K. T., T. D. Sauer and J. A. Yorke, 1996: Chaos: an introduction to dynamical systems. Springer-Verlag, New York.

Buizza R., J. Tribbia, F. Molteni and T. Palmer, 1993: Computation of optimal unstable structures for numerical weather prediction models. Tellus, 45A, 388-407.

Cai, M., E. Kalnay and Z. Toth, 2001: Potential impact of bred vectors on ensemble forecasting and data assimilation in the Zebiak-Cane model. Submitted to J of Climate.

Errico, R. and R. Langland, 1999: Notes on the appropriateness of bred modes for generating perturbations used in ensemble forecasting. Tellus, 51A, 442-449.

Farrell, B., 1988: Small error dynamics and the predictability opf atmospheric flow, J. Atmos. Sciences, 45, 163-172.

Kalnay, E 2001: Atmospheric modeling, data assimilation and predictability. Chapter 6. Cambridge University Press, UK. In press.

Kalnay E., Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003, ISBN 0 521 79629

Lorenz, E.N., 1965: A study of the predictability of a 28-variable atmospheric model. Tellus, 21, 289-307.

Lorenz, E.N., 1996: Predictability- A problem partly solved. Proceedings of the ECMWF Seminar on Predictability, Reading, England, Vol. 1 1-18.

Molteni F. and TN Palmer, 1993: Predictability and finite-time instability of the northern winter circulation. Q. J. Roy. Meteorol. Soc. 119, 269-298.

Morss, R.E.: 1999: Adaptive observations: Idealized sampling strategies for improving numerical weather prediction. Ph.D. Thesis, Massachussetts Institute of Technology, 225pp.

Ott, E., 1993: Chaos in Dynamical Systems. Cambridge University Press. New York.

Palmer, TN, R. Gelaro, J. Barkmeijer and R. Buizza, 1998: Singular vectors, metrics and adaptive observations. J. Atmos Sciences, 55, 633-653.

Patil, DJ, BR Hunt, E Kalnay, J. Yorke, and E. Ott, 2001: Local low dimensionality of atmospheric dynamics. Phys. Rev. Lett., 86, 5878.

Patil, DJ, I. Szunyogh, BR Hunt, E Kalnay, E Ott, and J. Yorke, 2001: Using large member ensembles to isolate local low dimensionality of atmospheric dynamics. This volume.

Snyder, C. and A. Joly, 1998: Development of perturbations within growing baroclinic waves. Q. J. Roy. Meteor. Soc., 124, pp 1961.

Szunyogh, I, E. Kalnay and Z. Toth, 1997: A comparison of Lyapunov and Singular vectors in a low resolution GCM. Tellus, 49A, 200-227.

Toth, Z and E Kalnay 1993: Ensemble forecasting at NMC – the generation of perturbations. Bull. Amer. Meteorolo. Soc., 74, 2317-2330.

Toth, Z and E Kalnay 1997: Ensemble forecasting at NCEP and the breeding method. Mon Wea Rev, 125, 3297-3319.

Toth, Z., I Szunyogh and E Kalnay, 1999: Response to Notes on the appropriateness of bred modes for generating perturbations used in ensemble forecasting. Tellus, 51A, 442-449.