Information Flow in Ensemble Weather Predictions

RICHAH KLEEMAN

Courant Institute of Mathematical Sciences, New York University, New York, New York

(Manuscript received 30 June 2005, in final form 12 June 2006)

ABSTRACT

In a weather prediction, information flows from the initial conditions to a later prediction. The uncertainty in the initial conditions implies that such a flow should be quantified with tools from probability theory. Using several recent developments in information theory, this flow is explored using a moderate-resolution primitive equation atmospheric model with simplified physics. Consistent with operational experience and other methodologies explored in the literature, such as singular vectors, it is found that the midlatitude flow is mainly in an easterly direction. At upper levels, the flow is primarily steered by advection of the jet stream; however, at low levels there is clear evidence that synoptic dynamics are important and this makes the direction of flow more complex. Horizontal rather than vertical flow is generally found to be more important, although there was evidence for propagation from the mid- to upper troposphere of zonal velocity.

As expected, as the length of the prediction increases, more remote areas become important to local predictions. To obtain reliable/stable results, rather large ensembles are used; however, it is found that the basic qualitative results can be obtained with ensembles within present practical reach. The present method has the advantage that it makes no assumptions concerning linearity or ensemble Gaussianicity.

1. Introduction

The uncertainty in the initial conditions of a weather prediction (e.g., Lorenz 1963) led many years ago to the realization that only probabilistic predictions were possible (e.g., Epstein 1969; Leith 1974). In general, for practical reasons, Monte Carlo techniques are used to explore the prediction probability distribution function and these now form part of the arsenal of the operational meteorologist (e.g., Toth and Kalnay 1993; Palmer 2000).

Recently, considerable interest has been shown in improving predictions by using improved observational platform design (e.g., Szunyogh et al. 2002). Thus, dynamical techniques such as singular and bred vectors are used to identify regions to which predictions are particularly sensitive. Given enough warning, the observational platform could be in principle made more complete in the sensitive regions identified using the dynamical tools, and hence final predictions made more accurate. Given the severe economic impact of certain weather events, such a strategy clearly has significant potential utility [see Shapiro and Thorpe (2004), their section 3 for further details].

In addition to this recent practical interest in improving predictions, there has been some development of the theoretical framework underlying ensemble predictions. In particular, a number of authors (e.g., Leung and North 1990; Schneider and Griffies 1999; Kleeman 2002; Roulston and Smith 2002; DelSole 2004) have advocated the use of information theoretic functionals to measure the information content of ensembles. This approach is particularly attractive since the meaning of such measures is intuitively clear and they have very general and appealing mathematical properties.

In the present context, we are interested in how this information flows from one particular region and physical variable to another and also in quantifying the likely reduction in uncertainty in a target region/variable. These ideas are naturally of general interest to physicists and so some interesting ideas have appeared in this literature over the past two decades (e.g., Kaneko 1986; Vastano and Swinney 1988; Schreiber 2000). In general, these approaches involve the introduction of information theoretical functionals related to those discussed in the previous paragraph. In particular, the two function-
als most studied have been the so-called time-delayed mutual information and the transfer entropy (TE). These enable one to quantify, respectively, how much uncertainty in a particular variable would be reduced by perfect knowledge of another variable at an earlier time and the amount and direction of information transferred from one location/variable to another at a later time. Clearly such measures are of interest to the practical problem outlined above. At present, these functionals have a clear interpretation in terms of uncertainty (see below); however, they are not part of a general rigorous flow formalism. The author and his coworker are presently developing just such a formalism (see Liang and Kleeman 2005) and will apply it to the practical problem discussed above in a future publication.

To test these ideas, we chose to utilize models able to simulate well the midlatitude mean flow and variability. In addition, since we are interested in exploring the possibilities offered by ensemble prediction, we chose to examine models with simplified physics that enable highly efficient integration. This enabled us to produce ensembles on the order of 10,000 within a day or so on current hardware.

This manuscript is organized as follows: section 2 contains a description of the information theoretic machinery used. Section 3 describes the atmospheric model used, briefly discusses its performance, and outlines the method used to initialize ensemble members. Section 4 contains the results. Finally, section 5 consists of a summary and discussion.

2. Mathematical tools

Suppose we have available the joint probability distribution function $p(x(t), y(t_0))$ of a particular prediction (random) variable $X(t)$ and another variable $Y(t_0)$ at the initial condition time $t_0$. 1 Now if we had perfect knowledge of $Y(t_0)$, then the (univariate) distribution for $X(t)$ would change to reflect this improved knowledge of our system. The resulting univariate distribution is the conditional distribution

$$p(x(t)|y(t)) = \frac{p(x(t), y(t))}{p(y(t))}.$$  

Entropic functionals on such distributions quantify the uncertainty associated with the particular variables of interest (see Cover and Thomas 1991). Of interest here are the absolute entropy associated with a random variable $X(t)$

$$H(X(t)) = -\int dx'(t)p(x'(t)) \ln[p(x'(t))]$$

and the conditional entropy

$$H(X(t)|Y(t_0)) = -\int dx'(t)dy'(t_0)p(x'(t), y'(t_0)) \ln[p(x'(t)|y'(t_0))]$$

$$= -\int dy'(t_0)p(y'(t_0)) \int dx'(t)p(x'(t)|y'(t_0)) \ln[p(x'(t)|y'(t_0))].$$

The second line here shows that the conditional entropy is the expected entropy of the (univariate) conditional distribution. Since this entropy reflects the expected uncertainty in the distribution given perfect knowledge of $Y(t_0)$, it will intuitively be less than $H(X(t))$, which measures the unconditioned uncertainty of $X(t)$. In fact, it is easily shown (see Cover and Thomas 1991; Leung and North 1990) that the difference of these two entropies is the so-called mutual information $I$ between $x(t)$ and $y(t_0)$:

$$I(X(t); Y(t_0)) = H(X(t)) - H(X(t)|Y(t_0)). \quad (1)$$

Intuitively then (see Leung and North 1990), such mutual information represents the expected reduction in uncertainty in $X(t)$ associated with perfect knowledge of the initial condition variable $Y(t_0)$.

The mutual information $I$ can be written as the relative entropy of the joint distribution and the direct product of the separate univariate distributions:

$$I(X(t); Y(t_0)) = D[p(x(t), y(t_0)), p(x(t))p(y(t_0))]$$

which shows (due to the properties of relative entropy) 3 that it is nonnegative and represents the “distance” between the joint distribution and the distribution that would hold if $X(t)$ and $Y(t_0)$ were completely independent. We shall refer to $I(X(t); Y(t_0))$; henceforth as the time-lagged mutual information (TLMI).

1 In a practical situation, such a distribution will only be available to us as a sample estimate at reasonably coarse resolution; however, for the present we shall ignore this technical difficulty. Later we shall discuss the practical implementation of these ideas.

2 We use lowercase to denote particular numerical choices for the random variables.

3 Relative entropy measures the difference of two distributions using a nonnegative functional that only vanishes when the distributions are identical. It is also invariant under general nonlinear transformations of state space. See Kleeman (2002) for more details.
The TLMI does not measure the information flow into \( X \) from other random variables. Consider a typical multivariate distribution for the initial time \( t_0 \). For spatial points that are close together, there is often a high correlation between random variables \( X(t_0) \) and \( Y(t_0) \) and hence, the TLMI between \( Y(t_0) \) and \( X(t) \) may simply measure the importance of persistence to the particular random variable \( X(t) \). Recently, Schreiber (2000) has suggested another functional that, in his view, better reflects information flow. Suppose there was, on the contrary, no information flowing into \( X \) from other random variables such as \( Y \). In such a case, one could model the sequence of random variables \( X(t) \) as a univariate process meaning that
\[
p(x(t)|x(t_0)) = p(x(t)|x(t_0), y(t_0)) \quad \forall x(t), x(t_0), y(t_0).
\]

(2)

If we now consider the spatially close \( x(t_0) \) and \( y(t_0) \), it is clear that if they are highly correlated, then \( p(x(t)|x(t_0)) \) will be close to \( p(x(t)|x(t_0), y(t_0)) \) since the addition of \( y(t_0) \) will not provide much further influence on \( x(t) \) beyond that already provided by \( x(t_0) \). We note that because of its definition in terms of the relative entropy functional \( D \), transfer entropy is nonnegative but also generally non-symmetric since usually
\[
T(Y \rightarrow X, t, t_0) \neq T(X \rightarrow Y, t, t_0),
\]
which allows a direction to be associated with information interchange between random variables.

One can also write the transfer entropy in a form analogous to Eq. (1):
\[
T(Y \rightarrow X, t, t_0) = H(X(t)|X(t_0)) - H(X(t)|X(t_0), Y(t_0)),
\]
which shows that it has the interpretation of the expected reduction in uncertainty in the target random variable \( X(t) \) because of knowledge of the initial time variable \( Y(t_0) \), beyond that which knowledge of the target variable at the initial condition time would give.\(^4\)

In summary, TLMI allows one to identify which initial condition variables require better observation in order to reduce prediction error. The TE excludes the influence of persistence in such a calculation and might

\(^4\) This was pointed out to the author by T. DelSole.

In other words, knowledge of the other random variables \( Y \) at the initial conditions makes no difference to the future of the random variable \( X \) if \( x(t_0) \) is known. Schreiber suggests using the deviation from this property as a measure of the information flow from \( Y \) into \( X \). We can use the relative entropy of the two distributions as such a measure. For particular choices of the conditioning variables \( x(t_0) \) and \( y(t_0) \), this is
\[
D[p(x(t)|x(t_0), y(t_0)), p(x(t)|x(t_0))]
\]
\[
= \int dx(t)p(x(t)|x(t_0), y(t_0)) \ln \left( \frac{p(x(t)|x(t_0), y(t_0))}{p(x(t)|x(t_0))} \right).
\]

The TE is then the expected value of this deviation when all possible values of the conditioning variables are considered, that is,

\[
(3)
\]

3. Atmospheric model and initialization strategy

One of the requirements of the present theoretical study is the ability to produce large ensembles. This means that we shall focus our attention on computationally efficient models. As we shall see below, for many purposes, much smaller ensembles are sufficient; however, we chose here to mainly use rather large samples in order to demonstrate statistical convergence and also to explore the relationship between TLMI and TE.

Since “atmospheric physics” (radiation, convection, and boundary layer processes, primarily) often consumes considerable integration time in typical general circulation models, we chose to use a model in which such processes are replaced by relatively simple linear temperature, vorticity, and divergence relaxation terms. Such a modeling strategy was originally proposed by Held and Suarez (1994) in the context of understanding better extant and complex general circulation models, but it also offers a method for developing very efficient and reasonably realistic models (at least
of the midlatitudes). An additional attraction is that we have considerable confidence that the “dynamical core” represents realistically the midlatitude baroclinic turbulence processes.

In the present context, we utilized the Portable University Model of the Atmosphere (PUMA) developed recently at the University of Hamburg (Leslie and Fraedrich 1997; Frisius et al. 1998; Franzke et al. 2000). As mentioned, this uses Newtonian cooling, Rayleigh friction terms in the temperature, and vorticity/divergence equations, respectively, to replace radiation, convection, and surface frictional effects. The temperature was relaxed toward an idealized radiative/convective temperature profile that varied both meridionally and vertically but was zonally uniform. Such a profile can be used to simulate seasonal effects and we chose here to study the northern winter. The model is configured as a reasonably standard primitive equation spectral model and we chose to retain five vertical levels. As is usual, the model also uses hyperviscosity to control small-scale noise. Finally, realistic orography is included in the model, which has the effect of breaking zonal symmetry and locating storm tracks in reasonably realistic locations (the north Atlantic and North Pacific).

We found that a horizontal resolution of T42 was sufficient to give reasonably realistic predictability and circulation performance. Interestingly, a resolution of T21 was found to be noticeably deficient in this regard. In particular, the divergence of initially very close trajectories was found to take place much too slowly. The computational cost of T42 was almost an order of magnitude higher than T21; however, the quite significant increase in realism was thought worthwhile. Higher resolutions were also examined and no large qualitative change over T42 was noticed either in circulation or basic predictability properties.

The performance of the model was assessed using a 50-yr control integration, and results can be seen in Fig. 1. The zonally averaged mean upper-tropospheric ($\sigma = 0.3$) zonal winds are shown in Fig. 1a and these show a jet stream of approximately realistic strength. Figure 1b shows the variance of lower-tropospheric ($\sigma = 0.9$) winds, and regions of intensification may be noted in the mid North Atlantic and Pacific corresponding reasonably well with the observed storm tracks. Incidentally one would not expect better quantitative agreement with the observations since the model surface temperature is zonally uniform. Figures 2a–c show a (random) 5-day sequence of $\sigma = 0.9$ streamfunction in the North America–Europe sector. This appears broadly consistent with observed weather variations in this region both in the scale of disturbances as well as the general direction of propagation (northwestward).

The approach we shall follow here is one of statistical predictability. In this scenario, initial conditions for dynamical forecasts are assumed (realistically) to be uncertain due to limitations in the observing network. A “Monte Carlo” approach is used in this situation in that (fully defined) initial conditions are taken as random samples from the assumed distribution for the initial conditions. These sample initial conditions are then all integrated forward in time to produce a so-called ensemble prediction. In this study, we used rather large samples/ensembles by conventional standards.

Defining the appropriate multivariate distribution for the initial conditions is well recognized as a difficult issue and forms much of the focus of data assimilation theory and practice. Since we are interested here primarily in prediction issues, we choose to use a particularly simple distribution that nevertheless is similar in concept to those often produced in statistical data as-

---

5 The vorticity and divergence were relaxed toward rest.
The standard deviations of the above variables are spatially uniform and take the values of 0.75°C for temperature, $7.3 \times 10^{-7}$ s$^{-1}$ for both divergence and vorticity, and 1.0 mb for surface pressure. These represent values somewhat less than an order of magnitude smaller than typical climatological fluctuations. In a more realistic scenario, there should be some nonuniformity to reflect the nature of the observing network. We defer this interesting case to a later publication when we examine in more detail (and more rigorously) the relationship between such networks and initial condition distributions. The means for the distribution were chosen to be a random initial condition drawn from the 50-yr integration discussed above. The purely random nature of the perturbations about the mean initial condition resulted in some degree of gravity "shock," which was most evident at large scales. Large scales are associated with low-frequency gravity waves, and the numerical scheme used quite effectively filtered the high-frequency/small-scale waves often seen in other studies. The waves were removed with a standard operational gravity wave filter (Lynch and Huang 1994).

The selection of an appropriate ensemble size is dependent on a consideration of what is required for a robust evaluation of the distributions in Eqs. (1) and (3). These require, respectively, two- and three-dimensional spaces [the variables $x(t)$ and $y(t_0)$ for TLMI and the further variable $x(t_0)$ for transfer entropy]. As was mentioned in section 2, an ensemble represents a sample evaluation of the required multivariate distributions and the size of the ensemble determines the degree of coarseness of the resolution of this distribution. More specifically, in order to use an ensemble as an estimate of a probability distribution requires that the (multidimensional) target space be divided into a large number of bins or "partitions." The estimate is then obtained by counting the number of ensemble members that fall within each partition element. A detailed and rigorous discussion of this issue may be found in Kleeman and Majda (2005). To avoid the loss of information due to sampling "error," one requires that there be sufficient ensemble members in each partition element we choose to use in our two- or three-dimensional space. For the present study, we decided to use 12 partitions in each dimension and chose these partitions so that there were equal numbers of ensemble members in each univariate direction. This gives a total of 1728 partitions for the transfer entropy calculation. In order that there be a sufficient sample size in each partition to avoid significant information loss [see Kleeman and Majda (2005) for more details on this point], we chose an ensemble of 9600 members. For

---

This is defined as $a$, where the covariance drops by a factor $\exp(-r/a^2)$ at a distance $r$ from the central variable.

\[\text{Fig. 2.} \ (a) \text{ Day 0 of a 5-day random sequence of } \sigma = 0.9 \text{ streamfunction.} \ (b) \text{ Same as in (a), but for Day 2.} \ (c) \text{ Same as in (a), but for day 4.}\]
TLMI, we retained the 12 partitions per dimension for consistency with the transfer entropy. This of course meant that there was a considerably larger number of ensemble members per partition for TLMI. For the particular ensemble size chosen here, we found very little sensitivity of our results to moderate changes in the number of partitions per dimension. Later we shall test the sensitivity of our results to using much smaller numbers of ensemble members and consequently much coarser resolutions for our partitioning.

Ensemble members were initialized using the multivariate Gaussian distribution discussed above and each member was run for at least 10 days. Results discussed below were generally drawn from the first few days of these integrations.

4. Results

We chose to examine the information flow of three variables (potential temperature, zonal and meridional winds) within a domain encompassing much of North America and the North Atlantic (20° through 65°N and 90°W through 0°). Figure 2 shows the particular region studied. We also primarily considered two (sigma) vertical levels of 0.9 and 0.3, but occasionally 0.7 and 0.5. For most of the results shown, we chose to consider flow to a point at the center of the domain of consideration (this is marked in figures by a filled box) and at the lower level σ = 0.9. In terms of the notation adopted in section 2 above, the plots shown shall be of the functionals at the location y(t₀) while the filled box shall denote x(t₀) and x(t). It is to be emphasized that all information theoretic plots have identical units (bits) and that one may directly compare quantitative information between plots. There is considerable variation in the magnitude of the functionals from situation to situation, and to convey this, different choices in contouring are often made between figures.

Since we are considering flow from a particular set of initial conditions and with respect to a particular location (the central North Atlantic in most cases), it is of value to consider the synoptic flow present during the predictions. As may be noted from Figs. 2a–c, at the lower levels there is a significant cyclone present just off the Carolinas’ coast, and as prediction time increases this feature progresses eastward, amplifies, and becomes more complex. At upper levels (not shown), there is considerably less (meridional) structure, although some mild wavelike disturbances may be seen in the dominant jet stream.

a. One-day near-surface predictions

We first examined the information flow for short-range predictions. Displayed in Figs. 3a,b are the TLMI and TE for flow from initial condition temperature data to the temperature at the point marked with a filled square at prediction time one day. Results are for σ = 0.9, that is, near the surface. We shall refer to this particular point as the “target.” Note that this terminology differs from other papers in this area that refer to “target regions” for improving the observational network. Regions where we find elevated values of TLMI and TE would be where our results are suggesting an improved network.

For this particular case, the importance of persistence to near-surface temperature prediction is quite clearly seen: the TLMI shows a strong peak around the target point while the TE that has the persistence effect removed (see section 2 above) shows that independent information from the west (and hence basically upstream) has some importance to surface temperature. The independent information revealed by the TE is clearly considerably less than that due to persistence as revealed by TLMI.

It should be noted that the values of TLMI and particularly TE do not go to zero as one approaches the boundary of the domain. This is not a consequence of real information flow but results from the sampling error involved in the estimation of our information functionals [for a more detailed and rigorous analysis of this phenomenon, see Kleeman and Majda (2005)]. This “background level” can be used as a rough reference for estimating the importance of the more elevated values of the functionals close to the target regions.

The importance of persistence noted for temperature is not present for zonal velocity. Displayed in Figs. 3c,d are the analogous results to Figs. 3a,b but for zonal velocity. A vector plot of the ensemble mean flow field at the initial conditions is superimposed for reference. The structures for TLMI and TE are basically similar but are very different from the temperature results. The two important regions are located west of the target and the southern node is closely associated with aspects of the cyclone present in the initial conditions. This structure suggests a strong role for dynamics as opposed to straightforward advection. Note that the southern node is relatively stronger in the TE, suggesting that some part of the influence of the northern node (which is close to the target point and hence, by the construction of the initial condition, correlated with the target) is due to persistence.

In addition to the influence of field variables on themselves, one can also analyze the influence of one variable in the initial conditions on another at prediction time. Displayed in Figs. 3e,f is the influence of meridional velocity in initial conditions on predicted zonal velocity. The regions of influence are quite different...
Fig. 3. Plots of TLMI and TE at various prediction lags. Units for both functionals are bits and so are directly comparable between figures. Note that differing contour intervals are used for clarity. (a) TLMI of temperature at day 1 in the center of the domain (the solid square) with the same variable at the initial time at all domain locations and at $\sigma = 0.9$. (b) Same as in (a), but for the TE. (c) Same as in (a), but for zonal velocity. The ensemble mean zonal velocity at the initial time is superimposed as vectors for reference. (d) Same as in (c), but for TE. (e) The TLMI of Day 1 zonal velocity at domain center with meridional velocity at the initial time. (f) Same as in (e), but for TE. (g) The TLMI of Day 1 zonal velocity at domain center and $\sigma = 0.9$ with the same variable at the initial time but for $\sigma = 0.3$. 
from the regions already seen but again appear influenced by the synoptic events occurring during the prediction since they are located to the west and have a very similar spatial scale to synoptic variability in this region.

All of the results above concern horizontal influence, that is, the source regions of information are on the same horizontal level. We next tested the importance of vertical information propagation. Shown in Fig. 3g is the TLMI from the upper-level initial conditions to the lower level for zonal velocity (TE is very similar). Comparing this to Fig. 3c, we see that the degree of influence is considerably reduced, suggesting that horizontal information flow is more important than vertical. This reduced amplitude of vertical influence was noted for all variables examined.

b. Upper-level predictions

The information flow for the near-surface level showed some complexity (see Figs. 3c,e), suggesting a significant role for synoptic dynamics. This did not appear to be the case for upper levels. The TLMI for upper-level temperature is shown in Fig. 4a (results for TE were similar) together with a vector plot of the upper-level mean winds. The region of information origin appears directly to the west and quite significantly farther upstream than in the lower-level case. In fact, the advection by the model jet stream at this level seems to explain these results quite well. A jet stream velocity of around 35 m s\(^{-1}\) (as seen here) results in a displacement of around 3000 km, consistent with the results reported. Note there is some southward displacement of the source region, suggesting some minor role for synoptic dynamics here as well. Note also that the source region is weaker in magnitude than at lower levels perhaps because the jet stream strength is dispersive of information. Similar results were seen for other variables.

Unlike the low-level target results above, we noted for this case evidence for the importance of the vertical propagation of information. This is depicted in Fig. 4b, which shows a vertical–longitude section of TLMI at the target latitude of 40°N. Notice the strong region of influence at \(\sigma = 0.5\), which is comparable in magnitude with the region at the target vertical level of \(\sigma = 0.3\). This vertical propagation was not evident for temperature where the \(\sigma = 0.5\) region was much reduced compared to the \(\sigma = 0.3\) target level. These results seem consistent with linear analyses (e.g., Hakim 2005), which have argued for the importance of vertical propagation from the midtroposphere to the tropopause.

c. Longer predictions

Predictions for longer ranges were examined next. The lower levels were considered here, as the upper-level results appeared a straightforward extension of the previous subsection. The TLMI results for temperature and zonal velocity at 3 and 6 days are shown in Figs. 4c–f (TE results are similar although noisier). As perhaps unexpected, the region of influence spreads farther upstream and weakens as time goes on. It is rather surprising that at 6 days, the region of maximum influence for the central Atlantic is only located in the eastern United States and not farther upstream (at upper levels, this region is located over Asia).

The behavior of temperature on temperature as time increased was a little surprising, as it shows that the source region stays fixed at the target location for around 3 days and then moves fairly uniformly upstream between this time and 6 days (not shown). The quasi-stationary TLMI for temperature during the first 3 days shows the importance of persistence but can also be seen if small perturbations are introduced at the target and their evolution tracked (not shown). The low-level temperature perturbations remain at the target for 3 days before moving fairly uniformly downstream later.

d. Target region sensitivity

The lower-level results above suggest the importance of local synoptic dynamics to information flow. To test this, we shifted our region of analysis 48° of longitude in a westerly direction. The target (still in the domain center) was now located approximately over Philadelphia, Pennsylvania. The prediction initial conditions and time were identical to previous experiments. Displayed in Figs. 5a,b is the TE for 1-day predictions of temperature and zonal velocity. Comparing these with the analogous figures for the North Atlantic target region (Figs. 3b,d) shows considerable differences, and in particular for zonal velocity, broader and farther afield regions of importance for information. Such a result is consistent with the hypothesized importance of synoptic dynamics for our results.

e. Importance of ensemble size

All the results reported above involve rather large ensembles, so it is of interest from a practical perspective to see how robust the results are when considerably smaller (and possibly operationally feasible in the future) ensembles are used. To examine this, we subsampled 200 members of our ensembles and recomputed results. In general, the TE results at this level were no longer usable since the calculation of this functional requires a trivariate probability density function (pdf) implying very coarse resolution to obtain adequate sampling (see Kleeman and Majda 2005). On the other hand, the TLMI that uses a bivariate pdf
Fig. 4. (a) Same as in Fig. 3c, but for $\sigma = 0.3$ for both random variables. (b) A vertical-longitude plot of TLMI along the target latitude of 40°N. The target is shown with a solid box at the upper target level. (c) Same as in Fig. 3a, but for prediction time of 3 days for the first random variable. (d) Same as in (c), but for zonal velocity. (e) Same as in (b), but for prediction time of 6 days. (f) Same as in (d), but for prediction time of 6 days.
apparently produced adequate results with the calculations performed above. In the original calculation with the large ensemble, we used 12 partitions per dimension. In the reduced sample we dropped to four, which represents quite a significant coarse graining of our data [again see Kleeman and Majda (2005) for more background on the technical issues here].

As an illustration, Fig. 6 shows results for the zonal velocity of 1-day predictions. A comparison with Fig. 3c shows quantitative agreement, although the noisiness associated with the small sample may be clearly discerned.

The convergence of TE results was examined by considering the intermediate ensemble sizes of 400, 800, 1600, and 3200 members. It was found that rather rapid convergence occurred for the TE between wind components (both zonal and meridional) with an ensemble of around 400 being adequate to reproduce the large ensemble results. Convergence for temperature was, however, considerably slower with an ensemble size of at least 3200 being required to reproduce the full ensemble results (see, however, the Gaussian results below).

f. Dynamical interpretation of results

The dynamical cause of results where the target region is in the upper midlatitude atmosphere seems quite clear at least for the horizontal case: advection downstream by the jet stream. This was confirmed in several perturbation experiments (not shown) where a small patch of zonal and meridional velocities and temperature with the horizontal structure of TE were added to the standard model run and the standard run subtracted. The patch propagated downstream to the target region in the expected time period.

The lower-level target results were more difficult to interpret; however, it is reasonably clear that they are either strongly influenced by persistence in the case of temperature or linked to upstream synoptic events in the case of wind. A careful comparison of the results of Figs. 3e–f shows for velocity a close correspondence between sensitive regions and the mean synoptic conditions at the initial time. In particular, note that the zonal velocity sensitivity patches correspond with regions where the mean zonal velocity due to the low-level cyclone are strong. Similarly, the meridional velocity sensitivity regions correspond approximately with regions where the mean meridional velocity is large. In general, there is some bias in this correspondence toward regions closer to the target, which is what one would expect intuitively.

The general impression is that greater knowledge of all aspects of the important low-level cyclonic circulation improves zonal velocity predictions at later times. A viewing of the synoptic evolution during the prediction period shows a straightforward propagation of this cyclone in an easterly direction at about 1000 km day$^{-1}$.

---

Fig. 5. (a) Same as in Fig. 3b, but for a target region based over Philadelphia rather than the central Atlantic. (b) Same as in (a), but for zonal velocity rather than temperature.

Fig. 6. Same as in Fig. 3c, but with an ensemble size of 200 rather than 9600.
This is sufficient to significantly influence velocity fluctuations in the target region. We tested this idea by adding a perturbation in zonal and meridional velocity with the structure of the TE plots in Figs. 3d,f. We changed the sign of the perturbation to reflect the local sign of the mean velocities and also subtracted the mean far-field level (before the sign adjustment). The resulting perturbation field then strongly resembled the initial condition mean cyclone. During the prediction period, this perturbation evolved (not shown) so as to influence primarily the northeast sector of the cyclone, that is, the sector nearest the target at the prediction time of 1 day. It retained approximately the same horizontal scale as the cyclone and also showed some mild amplitude damping.

5. Discussion and summary

Two different measures of information flow in dynamical systems proposed in the physics literature are examined in the context of statistical atmospheric prediction. The first is the time-lagged mutual information (TLMI), which measures the reduction in uncertainty or entropy of a particular prediction variable as a result of the perfect knowledge of another variable at an earlier time. The particular case where this earlier time is the initial condition time is obviously of practical interest to those considering targeted observations for improved predictions, such as in weather forecasting.

While TLMI has obvious practical application, it is of less value as a measure of information transfer. This is because the perfect knowledge of a particular random forecast variable at the initial time can obviously reduce uncertainty in the same variable at the prediction time. If a particular random variable did not receive information from other variables, then its probability at prediction time would only depend on the same variable at the initial time and not on the other variables. Motivated by this, the second measure of information flow, the transfer entropy (TE), measures the influence of external variables at earlier times in altering the probability of the prediction variable conditioned on the same variable in the initial conditions. If this conditional probability was not influenced by external variables, then one could consider the prediction variable as an isolated system and hence, not subject to information flow. The degree to which this is not true is a measure of the influence of external variables and hence, information flow. Obviously, from its definition, TE excludes the importance of persistence to prediction.

We calculated these two flow measures for the case of midlatitude atmospheric statistical prediction from a particular set of initial conditions. Consistent with the above discussion, we found that the two measures (TLMI and TE) differed sharply when persistence was important to prediction, as is the case (in this model) for temperature. Interestingly, when persistence was not important, the two measures were not qualitatively different. Further important properties deduced were as follows.

1) The horizontal scale of important source regions of information flow was generally on the order of 500 km and restricted to very specific areas. The low-level regions of importance appeared organized by synoptic features of the mean flow. This all suggests that this technique could have practical value for targeted observation strategies.

2) Horizontal as opposed to vertical flow of information was generally more important. There was, however, evidence of important vertical propagation of zonal velocity information from the midtroposphere to the upper levels of the jet stream.

3) Upper-level flow in the midlatitudes was generally explicable in terms of zonal advection by the jet stream. Lower-level flow, which might be considered more important practically, was more complex and apparently strongly influenced by local synoptic evolution.
4) As prediction time increases, source regions move farther “upstream” (i.e., westward and become more dispersed). This upstream movement was considerably reduced for the low levels of the atmosphere compared to upper levels.

5) Ensembles on the order of 200 were sufficient to describe TLMI (but not TE).

To date, basically two different sensitivity techniques have been proposed for targeted observations: the first uses linearization techniques such as singular vectors (see, e.g., Palmer et al. 1998) while the second (see, e.g., Bishop and Toth 1999) uses ensemble filter methods.

The present method is less restrictive in that it is fully nonlinear and makes no assumptions concerning the Gaussianity of the prediction ensemble. The present method does, of course, require rather large ensemble sizes. It would clearly be of significant interest to intercompare the three proposed methods and find out when the various restrictions of the methods are important. A linearization analysis using singular vectors would also be of value in better understanding the importance of synoptic dynamics reported here. These analyses are presently underway and will be reported elsewhere.

From the viewpoint of practical operational weather prediction, the essential content of the results described here can be obtained with ensembles of size around 200. We noted no qualitative difference between TLMI and TE beyond the persistence effect described above. Although operational ensemble predictions presently are typically of maximum size 50, it seems reasonable to conclude that increases in computational power in the near future may allow the operational calculation of TLMI.

Acknowledgments. It is a pleasure to acknowledge useful discussions with G. Eyink of the Johns Hopkins University at the IPAM Data Assimilation meeting at UCLA in February 2005 where this work was first presented. The author also wishes to acknowledge the work of the PUMA group at the Meteorologisches Institut, Universitaet Hamburg, and in particular, Klaus Fraedrich for help in running this atmospheric model. The comments of three reviewers are gratefully acknowledged, as they resulted in a much improved manuscript. This work was supported by NSF Grants CMG 0417728 and ATM 0430889.

REFERENCES


