

Flow-dependent predictability of the North Atlantic Jet

Article

Published Version

Frame, T.H.A. ORCID: <https://orcid.org/0000-0001-6542-2173>,
Methven, J. ORCID: <https://orcid.org/0000-0002-7636-6872>,
Gray, S.L. ORCID: <https://orcid.org/0000-0001-8658-362X> and
Ambaum, M.H.P. ORCID: <https://orcid.org/0000-0002-6824-8083> (2013) Flow-dependent predictability of the North Atlantic Jet. *Geophysical Research Letters*, 40 (10). pp. 2411-2416.
ISSN 1944-8007 doi: <https://doi.org/10.1002/grl.50454>
Available at <https://centaur.reading.ac.uk/32207/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/grl.50454>

Publisher: Wiley on behalf of the American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Flow-dependent predictability of the North Atlantic jet

T. H. A. Frame,^{1,2} J. Methven,² S. L. Gray,² and M. H. P. Ambaum²

Received 22 March 2013; revised 5 April 2013; accepted 5 April 2013; published 30 May 2013.

[1] The North Atlantic eddy-driven jet is a major component of the large-scale flow in the northern hemisphere. Here we present evidence from reanalysis and ensemble forecast data for systematic flow-dependent predictability of the jet during northern hemisphere winter (DJF). It is found that when the jet is weakened or split, it is both less persistent and less predictable. The lack of predictability manifests itself as the onset of an anomalously large instantaneous rate of spread of ensemble forecast members as the jet becomes weakened. This suggests that, as the jet weakens or splits, it enters into a state more sensitive to small differences between ensemble forecast members, rather like the sensitive region between the wings of the Lorenz attractor. **Citation:** Frame, T. H. A., J. Methven, S. L. Gray and M. H. P. Ambaum (2013), Flow-dependent predictability of the North Atlantic jet, *Geophys. Res. Lett.*, 40, 2411–2416, doi:10.1002/grl.50454.

1. Introduction

[2] It is well established [Whitaker and Loughe, 1998] that for a given lead time the spread of operational ensemble forecasts fluctuates from day to day. These fluctuations are frequently cited as evidence of the presence of atmospheric large-scale flow regimes; e.g., blocked and zonal regimes. They are also an indication that the predictability of the atmosphere is flow-dependent; i.e., ensemble spread increases more rapidly near some states than others because forecast trajectories diverge more rapidly near these states. A commonly cited analogue of this behavior is the model of Lorenz [1963]; e.g., Palmer [1993]. In the Lorenz model, the likelihood of rapid divergence of trajectories is much higher when they pass through the region between the two “wings” of the attractor than when they are on the “tips” of the wings. This means that the predictability of the Lorenz model can be said to be systematically flow-dependent; i.e., one could make a useful prediction of the rate of spread of an ensemble using prior experience and knowledge of the ensemble mean.

[3] A major difference between the Lorenz model and a numerical weather prediction (NWP) model is the huge disparity in the number of variables in the model state; the Lorenz model has three variables and a typical NWP model $\sim 10^8$. This means that the Lorenz model passes near most states on its attractor in a relatively short time. By contrast, when all variables are taken into account, the atmosphere

takes an extremely long time to pass near a previously visited state [Van den Dool, 1994]. A consequence of this is that systematic flow-dependent predictability in NWP can only feasibly be observed statistically by considering a small subset of variables or indices.

[4] Recent results [Woollings *et al.*, 2010; Franzke *et al.*, 2012; Hannachi *et al.*, 2012] suggest that statistics of North Atlantic eddy-driven jet indices possess significant inhomogeneities, indicating the presence of three regimes: a regime with the maximum wind-speed of the jet shifted south of its climatological mean latitude, one with it close to the mean latitude and one with it shifted north of the mean latitude. Further evidence has been found that the skill in forecasting the jet appears to vary with these regimes, with the skill being lowest when the forecast starts with the jet in the north regime [Frame *et al.*, 2011]. These differences in forecast skill could indicate the presence of flow-dependent predictability.

[5] In this letter the evidence for systematic flow-dependent predictability of the winter North Atlantic eddy-driven jet is examined. The aims are to identify properties from climatological data which link the statistics of the jet to its evolutionary behavior, and to determine whether the predictability of the jet shows flow dependence consistent with these properties.

[6] The rest of this letter is divided into four sections. Section 2 introduces a two-dimensional principal component space which describes the large-scale structure of the jet. In section 3, the frequency distribution and evolutionary behavior within this two-dimensional principal component space is examined using climatological data taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis data set (ERA-40) [Uppala *et al.*, 2005]. In section 4, the predictability of the jet is examined using ECMWF ensemble forecast data taken from the THORPEX Interactive Grand Global Ensemble (TIGGE) data set [Park *et al.*, 2008]. A summary of the main results and conclusions is presented in section 5.

2. Characterization of the Eddy-Driven Jet Using Principal Components

[7] Throughout this letter, the North Atlantic eddy-driven jet will be characterized using a two-dimensional coordinate system, in which the amplitude and latitude of the jet maximum are approximate radial and angular coordinates respectively. This coordinate system is derived from ERA-40 data as follows.

[8] Following Woollings *et al.* [2010] and Frame *et al.* [2011] the North Atlantic eddy-driven jet is defined from the zonal winds between 15°N and 75°N, zonally averaged between 60°W and 0°W, and vertically averaged between the 925 and 750 hPa pressure levels, to produce profiles of zonal wind as a function of latitude. Such “jet profiles” are

¹NCAS–Weather.

²Department of Meteorology, University of Reading, Reading, UK.

Corresponding author: T. H. A. Frame, Department of Meteorology, University of Reading, Earley Gate, Reading, RG6 6BB, UK. (t.h.a.frame@reading.ac.uk)

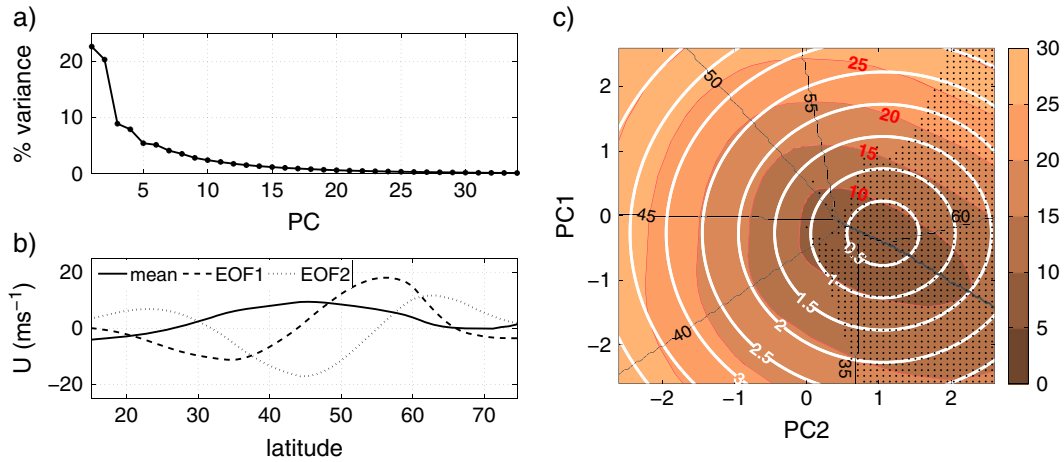


Figure 1. Summary of the EOF decomposition of the North Atlantic eddy-driven jet data from ERA-40. (a) Principal component spectrum. (b) Latitudinal structure of the time mean jet and first two EOFs. (c) Physical characteristics of the North Atlantic eddy-driven jet as a function of leading two principal components. Black contours: latitude of jet maximum. Colored contours: speed of jet maximum (ms^{-1}). Stippling indicates the region of PC-space in which the jet has multiple maxima. White contours: absolute amplitude coefficient of leading two EOFs (PCU).

calculated for each six hourly time-point in the ERA-40 data set. The resultant jet profile data is decomposed into empirical orthogonal functions (EOFs) [Jolliffe, 2002] as follows. It is weighted by the square root of cosine of latitude, so that the EOFs will be orthogonal with respect to area averaged energy. The time mean is subtracted and the data decomposed into EOFs. The EOFs form an orthogonal set of spatial structures similar to Fourier series, from which any jet profile can be constructed. Each EOF is paired with a principal component (PC). The PCs are a set of uncorrelated univariate time series describing the fluctuations of the amplitude coefficient, A , of the corresponding EOF about the time mean, $\langle A \rangle$; e.g., the amplitude coefficient of EOF1 is defined as $A_1(t) = \langle A_1 \rangle + \text{PC1}(t)$.

[9] The basic properties of the principal component decomposition are summarized in Figure 1. Figure 1a shows the percentage of the total variance explained by the leading 34 PCs. The first two PCs explain more than 40% of the variance. The stepped structure of Figure 1a indicates that the leading PCs should be grouped into twos, each group accounting for roughly equal variance.

[10] The zonal wind structures associated with the first two EOFs and time mean are plotted in Figure 1b. For ease of viewing, the plotted EOFs have been multiplied by the standard deviation of their respective principal component time-series. The first two EOFs are similar to those found for simulated data by Monahan and Fyfe [2006, 2009], and may be considered “typical” zonal jet EOFs. In combination with the mean, they are sufficient to describe the large-scale variability of jet, allowing for both latitudinal shifts and amplification of the jet. Henceforth, we shall use only PC1 and PC2 to characterize the variability of the jet.

[11] Figure 1c summarizes the relationship between the values of PC1 and PC2 and the jet structure obtained using only the time mean jet profile, EOF1 and EOF2. The coordinate axes have been normalized by the pooled standard deviation of PC1 and PC2. The pooled standard deviation is used as it preserves the relative scaling of the two axes. We shall refer to the unit of distance in PC-space as a PC unit (PCU). The white contours show the absolute amplitude,

$\sqrt{A_1^2 + A_2^2}$, of the first two EOFs normalized by the pooled standard deviation. They are offset from zero due to the contribution from the time mean. Since the EOFs were defined to be orthogonal with respect to area averaged energy and the relative scaling of the axes has been preserved, the square of absolute amplitude is proportional to the contribution of the first two EOFs to the total energy of the jet, hence, we shall use these contours as an indicator of energy.

[12] The black radial and colored contours show the latitude and wind-speed of the jet maximum [Woollings *et al.*, 2010], respectively. They are constructed systematically by calculating the latitude and amplitude of the jet maximum for different additive combinations of EOF1, EOF2 and the time mean. The stippling indicates the region of PC-space in which the jet has two maxima. The jet can have two maxima for two reasons. First, when the amplitudes of EOF1 and EOF2 are zero (the zero point of the white contours), the jet structure must be constructed from only EOFs 3, 4, etc. (not shown), which are wavelike with wave number greater than 1 and hence must produce multiple maxima (cf. Fourier series). Second, EOF2 (Figure 1b) has two maxima (at approximately 25°N and 62°N), so that when its amplitude is large and positive, the jet will tend to have two maxima. The thicker black contour between 35°N and 60°N in Figure 1c demarks the point at which the two separate maxima have equal magnitude. Crossing this boundary in a clockwise direction implies the decay of the jet in the north and growth of the jet in the south, the converse being true for anti-clockwise crossing. Except near this boundary, where the jet maxima can jump rapidly between latitudes, the plotted values are representative of those obtained using the full data.

3. Frequency and Persistence of the Jet

[13] We have described the properties of a two-dimensional PC-space which encapsulates the large-scale structure of the North Atlantic eddy-driven jet. We shall now consider the frequency distribution of the ERA-40 data

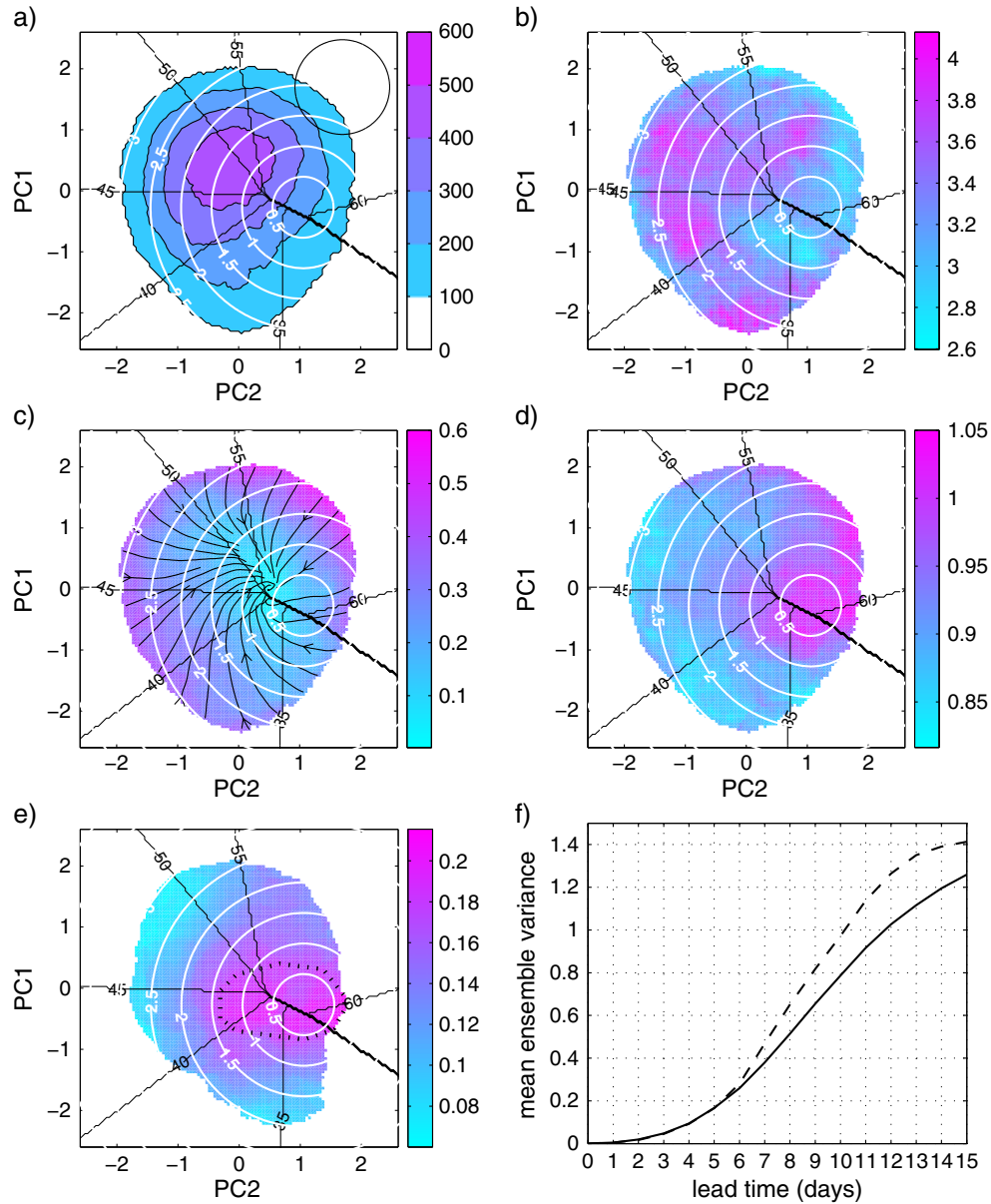


Figure 2. Statistics of winter (DJF) North Atlantic eddy-driven jet data projected into a two component PC-space. From the 45 winters of the ERA-40 data set: (a) Kernel-smoothing frequency estimate: The number of data points (per 2025) within radius 0.78 PCU of each location in PC-space. The black circle illustrates a radius of 0.78 PCU. (b) Mean residence time (days): The mean length of time trajectories remain within 0.78 PCU of each location. (c) Magnitude (colors) and direction (streamlines) of the mean PC-space increments: The mean net change in position over 2 days of trajectories originating within 0.78 PCU of each location. Interpretable as mean velocity (PCU per 2 days). (d) RMS PC-space increments: The RMS net distance (PCU) trajectories originating within 0.78 PCU of each location travel in 2 days. Interpretable as RMS speed (PCU per 2 days). From six-winters of ECMWF ensemble forecast data: (e) Mean rate of change in ensemble variance (PCU²/day) versus ensemble mean. Calculated at 6 day lead time. Dashed black line bounds the “high spread region”. (f) Mean ensemble variance (PCU²) versus lead time for two strata of the forecast data. Solid line: Ensemble mean trajectory has not entered high spread region by day 6. Dashed line: Ensemble mean trajectory first enters the high spread region on day 6. Black and white contours in Figures 2a–2e as in Figure 1c.

within that PC-space, and how this relates to the way in which trajectories move through it.

[14] We shall use 45-winters of six hourly ERA-40 data sampled every 2 days, providing a total of 2025 data points. The 2 day sampling frequency is chosen because 2 days was found to be the shortest time-scale for which a linear constant coefficient first-order auto-regressive model (hereafter AR1 model) fit the auto-correlation function of the

data. The AR1 model is sufficient to explain the mean, variance, auto-correlation, and distribution of increment lengths of the data; it therefore provides a useful null hypothesis for testing whether variations in the statistical properties of the data with location in PC-space require “interesting” [Christiansen, 2009], inhomogeneous statistics which depend nonlinearly on PC1 and PC2 to explain them, or can easily be explained by sampling error from a homogeneous

random walk [Stephenson *et al.*, 2000]. The 2 day timescale was identified, and the AR1 model fitted using software and methods described by Schneider and Neumaier [2001] and Neumaier and Schneider [2001]. It is important to note that trajectories do not travel far on timescales shorter than 2 days, so we can still think of 2 day increments as representing behavior in local regions of PC-space.

[15] The first thing we shall consider is the frequency distribution of the data in PC-space. Figure 2a shows a kernel smoothing estimate of the frequency distribution (colored contours). The black and white contours show the jet latitude and absolute amplitude of the first two EOFs (Figure 1c), and will also be shown in subsequent figures. To produce the frequency distribution, a flat circular kernel of radius 0.78 PCU was used. This radius is the median distance traveled along trajectories in 2 days. The black circle in the top right illustrates the size of the kernel. Since the kernel is flat, the frequency distribution is the number of data points (out of the total 2025) within a distance 0.78 PCU of each location in PC-space. The same kernel was used in the production of Figures 2b–2e.

[16] Several features of the frequency distribution indicate non-Gaussianity. First, the mode is displaced away from the mean towards the top left, and the separation between contours is greater toward the bottom right than the top left. Second, a “bulge” is apparent in the distribution in the region around $PC1 = -1.5$, $PC2 = -0.5$ associated with strong southward-shifted jets. One quantitative measure of the deviation of the data from Gaussianity is the multivariate skewness of Mardia and Zemroch [1975]. The multivariate skewness of the ERA-40 data sample is 0.198. A value this large would be unusual given the AR1 model, with a probability of, occurring by chance, only 0.0002, estimated from 50,000 simulated data sets.

[17] To provide a further test of the robustness of these non-Gaussian features, we reproduced the distribution using a six-winter sample of day 15 ECMWF 50 member ensemble forecasts from the TIGGE archive. All the features which indicate non-Gaussianity were found to be present in this second distribution, and the multivariate skewness of the TIGGE data was found to be even higher than the ERA-40 data, with a value of 0.228.

[18] To investigate the possible explanation for these non-Gaussian features, Figure 2b shows the mean residence time within local regions of PC-space. This is defined as the mean length of time that trajectories entering a circle of radius 0.78 PCU centered on a point in PC-space remain within that circle. Since the sampling frequency of the data is 2 days and the data consists of 90 day segments, the shortest observable residence time is 2 days and the longest 86 days. We exclude trajectories which are within the circle at the start or end of each 90 day segment, since the total residence time of such trajectories cannot be known. Only points with frequency (see Figure 2a) greater than or equal to 100 (out of 2025) are shown.

[19] Two things are of note in Figure 2b. First, there is a region of relatively low mean residence time (~ 3 days or less) centered on $PC1 = -0.5$, $PC2 = 1$ coincident with the region where the jet has lowest energy. Second, the longest mean residence times (~ 4 days) occur in regions associated with higher energy jets. Notably, the longest residence time occurs where $PC1 < -2$ and $PC2 \simeq 0$ coincident with the bulge in the frequency distribution seen in Figure 2a. This

second feature implies that the existence of bulge in the frequency distribution is a reflection of the long residence times of southward shifted jets.

[20] In tests using 1000 simulated data sets generated with the AR1 model, the longest mean residence time obtained was shorter than that found for the ERA-40 data, being only 3.05 days. More importantly, the longest residence times were found to always lie in a patch near the center of PC-space, 95% of the simulated data sets being within 0.63 PCU of the center. This second point, makes the existence of the high residence times of the southward shifted jet and low residence times near the center of the distribution unusual without a non-homogeneous statistical model, i.e., one with parameters that vary with location in PC-space either because of low frequency variations in external boundary conditions (non-stationarity) or simply due to the structure of the dynamical equations themselves.

[21] To examine what such a model might look like, we shall consider the mean PC-space increments. These are the mean vector change in location in PC-space over a finite time interval, and can be viewed as an estimator of the mean velocity of trajectories originating in a local region of PC-space. Mathematically, they may be represented as $\langle \Delta \mathbf{x} \rangle = \langle \mathbf{x}_{t+\Delta t} - \mathbf{x}_t \rangle$, where \mathbf{x}_t is a two element vector containing the values of PC1 and PC2 at time t , Δ indicates an incremental change, and the angle brackets indicate a mean. For the purposes of this work, the time increment Δt is 2 days. The mean PC-space increments are referred to as mean phase-space tendencies by Franzke *et al.* [2007] and are viewed as an estimator of the drift velocity in the context of the Fokker-Planck equation by Sura *et al.* [2005].

[22] The mean PC-space increments are shown in Figure 2c. The streamlines indicate the direction of the mean PC-space increments, and the colored shading shows their magnitude. These were calculated by taking the subset of all data points within a radius 0.78 PCU of each location in PC-space, calculating the mean increment and attributing it to that location. Several points can be made about them. First, they form a single “swirl” spiraling inward. Second, the direction rotation indicates that the jet migrates northwards on average. Third, the magnitudes of the mean PC-space increments are largest at the extremities of the distribution and decrease to zero towards the center. This third point is not surprising. From a statistical point of view, it is a requirement for the distribution to be approximately stationary. From a physical point of view, it is a requirement for the energy of the jet to remain bounded. All three of these properties are consistent with the AR1 model, however, the mean PC-space increments lack the symmetry that would be expected from a homogeneous statistical model; e.g., the inward flow in the region around $PC1=1$, $PC2=1.5$ is significantly stronger than that in the region around $PC1 = -2$, $PC2 = -0.5$.

[23] A second means of examining the way in which trajectories move around PC-space is to quantify how rapidly they move. We shall do this by considering the root-mean-squared (RMS) PC-space increments. These are the mean distance traveled from a given starting point in 2 days. They can be represented mathematically as $\langle \|\Delta \mathbf{x}\|^2 \rangle^{\frac{1}{2}} = \langle \|\mathbf{x}_{t+\Delta t} - \mathbf{x}_t\|^2 \rangle^{\frac{1}{2}}$. The RMS PC-space increments are shown in Figure 2d. These were calculated using the same subsetting technique as Figure 2c; however, whereas the colors in

Figure 2c show the magnitude of the mean increment, those in Figure 2d show the mean magnitude of increments. The most striking thing about Figure 2d is the inverse relationship it has with residence time (Figure 2b). Regions which have larger RMS increments, and by implication move more rapidly through PC-space, have shorter residence time. Quantitative tests using simulated data generated using an AR1 model with parameters which vary in PC-space consistent with Figures 2c and 2d could explain the 4 day residence times in Figure 2b.

4. Flow-Dependent Predictability of the Jet

[24] We have so far provided evidence that the behavior of trajectories in PC-space varies with location. In this section we shall determine whether there is evidence that predictability varies with location in a manner consistent with these variations in trajectory behavior. We shall use ECMWF ensemble forecast data projected into PC-space. This data is taken from the TIGGE archive [Park *et al.*, 2008], and consists of daily 15 day, 50 member ensemble forecasts spanning six winters from December 2006 to January 2012. First, however, we shall say briefly what we mean by predictability and how we shall determine its variation with location in PC-space.

[25] Ensemble variance acts as an estimator of forecast uncertainty, larger variance indicating larger uncertainty. The rate at which variance grows can be used as a measure of predictability. When variance grows rapidly, the estimated forecast uncertainty increases rapidly with lead time, and the predictability is lower. Conversely, when variance grows slowly, the estimated forecast uncertainty increases slowly with lead time, and the predictability is higher. Predictability does not quantify forecast skill; it does, however, provide an a priori estimator of the rate at which forecast skill is lost under the assumption that the forecast model is consistent with the physics of the atmosphere and the ensemble is a well-constructed sample of the initial uncertainty.

[26] The two questions we shall answer are: does the predictability of the jet vary systematically with location in PC-space, and can this variation be related to the climatological behavior? To answer these questions we shall use the rate of change of ensemble variance as a measure of predictability and the ensemble mean as a measure of location. We shall use the terms ensemble-mean and ensemble-variance to refer specifically to values calculated within the two-dimensional PC-space.

[27] Figure 2e shows the mean rate of change of ensemble variance (PCU² per day) at forecast day 6 at different points in PC-space. For each point in PC-space, the mean rate of change of ensemble variance is calculated by averaging the change in ensemble variance in 1 day over all forecasts with ensemble mean lying within a radius ~ 0.78 PCU of that point. We have chosen to show results from forecast day 6, since at this lead time, the early quasi-exponential phase [Lorenz, 2006] of growth should have largely ended, but the variance is still sufficiently small for the ensemble to be contained within a local region. However, the results and their interpretation are not particularly sensitive to the chosen lead time.

[28] Two points can be taken from Figure 2e. First, the mean rate of change of ensemble variance depends on the location of the ensemble mean, implying that the trajectories

of ensemble members diverge at different rates in different regions of PC-space. Second, the region of PC-space associated with the largest rate of change of ensemble variance corresponds to the region associated with weak or split jets (Figure 1c) and of largest RMS PC-space increments (Figure 2d) and shortest residence times (Figure 2b). The dotted line encloses a region containing rates of change of ensemble variance larger than 80% of the range. We shall call this region the “high spread region”.

[29] Figure 2f shows the mean ensemble variance as a function of lead time for two mutually exclusive subsamples of the forecast data. The dashed line shows the mean ensemble variance of all forecasts with ensemble mean first entering the high spread region *on* forecast day 6. The solid line shows the mean ensemble variance of all forecasts with ensemble mean first entering the high spread region *after* day 6. Forecasts with ensemble mean entering the high spread region before day 6 are neglected. Notably, there is no discernible difference between the mean ensemble variance of the two subsets of forecasts prior to day 6. After day 6, forecasts with ensemble mean first entering the high spread region on day 6 have larger mean ensemble variance than those with ensemble mean first entering after forecast day 6. The fact that differences between the mean ensemble variances of the two subsamples of the forecast data only become apparent after entry into the high spread region supports the interpretation of it as a local source of ensemble variance leading to increased forecast uncertainty and lower predictability. An analogy for this would be the central region between the wings of the Lorenz attractor [Lorenz, 1963] associated with rapid separation of trajectories.

5. Summary

[30] In this letter we have used a two-dimensional principal component space to examine the evidence for systematic flow-dependent predictability of the North Atlantic eddy-driven jet. We have shown that the frequency distribution of atmospheric data projected into this PC-space is non-Gaussian, and that deviations from Gaussianity can be linked to variation in persistence (residence time) with location in PC-space. It has been shown that persistence is anomalously low in the region of PC-space associated with weak or split jets. The key result of this letter is that the predictability of the jet decreases systematically, when its trajectory passes through this region. Franzke *et al.* [2012] and Hannachi *et al.* [2012] suggest that when the jet is shifted to the north, it tends to transition south via wave breaking. This transition implies temporary disruption of the zonal flow and passage of trajectories through the less predictable region of PC-space associated with weak or split jets. The increased rate of spread associated with passage through this region may be the explanation for the observation of Frame *et al.* [2011] that ensemble predictions of the location of the jet are less skillful when the initial conditions have the jet shifted to the north.

[31] **Acknowledgments.** This work was supported via the National Centre for Atmospheric Science—Weather directorate, a collaborative center of the Natural Environment Research Council. The authors gratefully acknowledge the help of the European Centre for Medium-Range Weather Forecasts for providing access to the TIGGE data set. The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

- Christiansen, B. (2009), Is the atmosphere interesting? A projection pursuit study of the circulation in the Northern Hemisphere winter, *J. Climate*, *22*, 1239–1254, doi:10.1175/2008JCLI2633.1.
- Frame, T. H. A., M. H. P. Ambaum, S. L. Gray, and J. Methven (2011), Ensemble prediction of transitions of the North Atlantic eddy-driven jet, *Q. J. R. Meteorol. Soc.*, *137*, 1288–1297, doi:10.1002/qj.829.
- Franzke, C., A. J. Majda, and G. Branstator (2007), The origin of nonlinear signatures of planetary wave dynamics: Mean phase space tendencies and contributions from non-Gaussianity, *J. Atmos. Sci.*, *64*, 3987–4003, doi:10.1175/2006JAS2221.1.
- Franzke, C., T. Woollings, and O. Martius (2012), Persistent circulation regimes and preferred regime transitions in the North Atlantic, *J. Atmos. Sci.*, *68*, 2809–2825, doi:10.1175/JAS-D-11-046.1.
- Hannachi, A., T. Woollings, and K. Fraedrich (2012), The North Atlantic jet stream: A look at preferred positions, paths and transitions, *Q. J. R. Meteorol. Soc.*, *138*, 862–877, doi:10.1002/qj.959.
- Jolliffe, I. T. (2002), *Principal Component Analysis, Second Edition*, Springer-Verlag, New York.
- Lorenz, E. N. (1963), Deterministic non-periodic flow, *J. Atmos. Sci.*, *20*, 130–141.
- Lorenz, E. N. (2006), *Predictability a Problem Partly Solved*, in *Predictability of Weather and Climate* Palmer, T., and R. Hagedorn (eds), pp. 40–58, Cambridge University Press, Cambridge.
- Mardia, K. V., and P. J. Zemroch (1975), Algorithm AS 84: Measures of multivariate skewness and kurtosis, *J. Roy. Stat. Soc. C-App.*, *24*, 262–265.
- Monahan, A. H., and J. C. Fyfe (2006), On the nature of zonal jet EOFs, *J. Climate*, *19*, 6409–6424.
- Monahan, A. H., and J. C. Fyfe (2009), How generic are dipolar jet EOFs, *J. Climate*, *19*, 6409–6424.
- Neumaier, A., and T. Schneider (2001), Estimation of parameters and eigenmodes of multivariate autoregressive models, *ACM Trans. Math. Softw.*, *27*, 27–57.
- Schneider, T., and A. Neumaier (2001), Algorithm 808: ARfit—A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models, *ACM Trans. Math. Softw.*, *27*, 58–65.
- Palmer, T. N. (1993), Extended-range atmospheric prediction and the Lorenz model, *Bull. Amer. Meteor. Soc.*, *74*, 49–65.
- Park, Y. Y., R. Buizza, and M. Leutbecher (2008), TIGGE: Preliminary results on comparing and combining ensembles, *Q. J. R. Meteorol. Soc.*, *134*, 2029–2050, doi:10.1002/qj.334.
- Stephenson, D. B., V. Pavan, and R. Bojariu (2000), Is the North Atlantic oscillation a random walk? *Int. J. Climatol.*, *20*, 118.
- Sura, P., M. Newman, C. Penland, and P. Sardeshmukh (2005), Multiplicative noise and non-Gaussianity: A paradigm for atmospheric regimes? *J. Atmos. Sci.*, *62*, 1391–1409.
- Uppala, S. M., and P. W. Källberg, et al. (2005), The ERA-40 re-analysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012, doi:10.1256/qj.04.176
- Van den Dool, H. M. (1994), Searching for analogues, how long must we wait? *Tellus A*, *46*, 314–324, doi:10.1034/j.1600-0870.1994.t01-2-00006.X.
- Whitaker, J. S., and A. F. Lough (1998), The relationship between ensemble spread and ensemble mean skill, *Mon. Wea. Rev.*, *126*, 3292–3302.
- Woollings, T., A. Hannachi, and B. Hoskins (2010), Variability of the North Atlantic eddy-driven jet stream, *Q. J. R. Meteorol. Soc.*, *136*, 856–868, doi:10.1002/qj.625