# Multiple Indices of Northern Hemisphere Cyclone Activity, Winters 1949–99

CHRISTOPHER J. PACIOREK

*Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania*

JAMES S. RISBEY

*School of Mathematical Sciences, Monash University, Clayton, Victoria, Australia*

VALÉRIE VENTURA

*Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania*

RICHARD D. ROSEN

*Atmospheric and Environmental Research, Inc., Lexington, Massachusetts*

## ABSTRACT

The National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis is used to estimate time trends of, and analyze the relationships among, six indices of cyclone activity or forcing for the winters of 1949–99, over the region 20°–70°N. The indices are Eady growth rate and temperature variance, both at 500 hPa; surface meridional temperature gradient; the 95th percentile of near-surface wind speed; and counts of cyclones and intense cyclones. With multiple indices, one can examine different aspects of storm activity and forcing and assess the robustness of the results to various definitions of a cyclone index. Results are reported both as averages over broad spatial regions and at the resolution of the NCEP–NCAR reanalysis grid, for which the false discovery rate methodology is used to assess statistical significance.

The Eady growth rate, temperature variance, and extreme wind indices are reasonably well correlated over the two major storm track regions of the Northern Hemisphere as well as over northern North America and Eurasia, but weakly correlated elsewhere. These indices show moderately strong correlations with each of the two cyclone count indices over much of the storm tracks when the count indices are offset 7.5° to the north.

Regional averages over the Atlantic, the Pacific, and Eurasia show either no long-term change or a decrease in the total number of cyclones; however, all regions show an increase in intense cyclones. The Eady growth rate, temperature variance, and wind indices generally increase in these regions. On a finer spatial scale, these three indices increase significantly over the storm tracks and parts of Eurasia. The intense cyclone count index also increases locally, but insignificantly, over the storm tracks. The wind and intense cyclone indices suggest an increase in impacts from cyclones, primarily over the oceans.

## 1. Introduction

Long-term variability in extratropical cyclone activity is an aspect of climate of particular interest to society. The most recent Intergovernmental Panel on Climate Change (IPCC) assessment emphasizes the importance of extreme events (McCarthy et al. 2001) but finds that the numerous studies examining trends in extratropical cyclone activity have not come to consistent conclusions (Houghton et al. 2001). There is no standard measure of cyclone activity because cyclones are complicated phenomena that may exhibit changes in intensity, duration, location, and frequency. A given index of cyclone activity may be closely related to some, but not all, of these aspects of cyclone behavior. Certain indices may better reflect the impacts of cyclones on human society and ecosystems whereas others are better suited to understanding dynamics.

Cyclones have a variety of signatures on the atmosphere. Dynamicists are interested in the factors that control the life cycle of storms (growth, maintenance, decay) and will frequently use a parameter such as the Eady growth rate (Hoskins and Valdes 1990) to assess the potential for instability and cyclone growth on the basis of local temperature gradients and static stability. More broadly, the potential for cyclogenesis is set by the larger-scale meridional temperature gradient be-

*Corresponding author address:* Christopher J. Paciorek, Department of Statistics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213.
E-mail: paciorek@stat.cmu.edu

tween equator and pole (Gitelman et al. 1997). Cyclones are one of the main factors setting the variance of temperature, pressure, and moisture in the troposphere on timescales of 2.5–10 days. Accordingly, dynamicists use measures of variance, such as the bandpass variance of temperature at 500 hPa (Iskenderian and Rosen 2000), to characterize the intensity and locations of the major centers of synoptic-scale (cyclone and anticyclone) activity. Synopticians describe cyclones by the coherent closed low pressure systems that are identifiable on synoptic pressure maps. Various techniques have been developed for counting lows (Lambert 1996; Zhang and Wang 1997; Simmonds and Keay 2000). The counting methods all have their deficiencies, but they have an intuitive match to the phenomena being studied. The impacts of cyclones are often felt via the strong winds and extreme gusts that accompany them, so researchers in this area typically categorize storms by some measure of the wind speed (e.g., Hulme and Jones 1991).

Detection of trends in any cyclone index is difficult because of the short observational record, concerns about data inhomogeneity (WASA Group 1998), the shortage of good proxies in the historical record, and high year-to-year variability. Regional studies have focused on North America, Europe, and the northeastern Atlantic Ocean. Many North American studies have used historical storm track maps, and most find evidence for a decline in winter cyclone frequency during 1950–80 (Reitan 1979; Zishka and Smith 1980; Parker et al. 1989; Changnon et al. 1995), although Hirsch et al. (2001) find no trend in U.S. east coast winter storm frequency for 1951–97 with a marginal decrease in intense storms. Zhang et al. (1997) use tide gauge data on storm surges at two sites on the East Coast and find no long-term trend in cyclone activity since 1910. Studies of Europe and the northeastern Atlantic have focused more on wind and wave measurements and provide evidence for increases in activity since the 1960s but little prior trend (Schmidt and von Storch 1993; Alexandersson et al. 1998; Schmith et al. 1998; Grevemeyer et al. 2000). These studies use long-term station records to address the criticism that analyses based on pressure maps suffer from serious inhomogeneity (WASA Group 1998).

Other studies have analyzed storm activity on a hemisphere-wide basis, with results reported as averages over broad spatial regions. Lambert (1996) finds no trend before 1970 in intense cyclones but a positive trend after 1970 in the Atlantic and Pacific regions. Using reanalysis data for 1958–97, Key and Chan (1999) report a significant negative trend in cyclones at 1000 hPa over 30°–60°N and a significant positive trend over 60°–90°N at both 1000 and 500 hPa. Gitelman et al. (1997) find a decrease in the NH surface meridional temperature gradient (MTG) over the period 1880–1991, which suggests a decrease in midlatitude cyclone activity.

In recent years, the climatology of cyclones has been of increasing interest in the modeling community, with particular attention paid to how cyclone activity and forcing may change under enhanced $CO_2$. Models tend to agree best on large-scale predictions of cyclone activity, but the agreement deteriorates for regional changes. Among the large-scale changes predicted by GCMs and simpler models are increases in atmospheric moisture (Hall et al. 1994), high-latitude surface warming, and tropical upper-tropospheric warming, resulting in a decreased surface MTG and increased MTG in the upper troposphere (Held 1993). Although surface and upper-tropospheric changes have competing effects, one expectation is that a decreased surface temperature gradient would yield a reduction in cyclone frequency or intensity (Held 1993; Stephenson and Held 1993). Warming at high latitudes is expected to shift the region of greatest baroclinic instability and storm tracks poleward (Held 1993; Hall et al. 1994; Carnell et al. 1996). Increased moisture also has competing effects, and it is not clear whether it would cause more intense cyclones, or fewer or weaker cyclones (Held 1993). At finer spatial scales, there is evidence for enhanced downstream development in the GCM storm tracks (Hall et al. 1994; Carnell et al. 1996) and fewer but more intense cyclones (Carnell and Senior 1998; Knippertz et al. 2000). This is expected to increase medium and stronger intensity winds at the expense of weaker winds, with predictions of more extreme winds at the downstream end of the Atlantic storm track (Carnell et al. 1996; Knippertz et al. 2000).

This paper examines the activity of NH extratropical storms for the winters 1949–99 using six indices derived from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis. We analyze the indices first averaged over broad spatial regions and then on a finer spatial scale, as averaging can obscure important details. We compare the indices to assess how different aspects of cyclone activity and forcing are related. We estimate trends in the indices and assess how the trends compare across indices and to predictions from GCMs under enhanced $CO_2$.

## 2. Data

### a. NCEP–NCAR reanalysis

The NCEP–NCAR reanalysis uses a fixed global data assimilation system and model, along with a comprehensive database of land surface, ship, rawinsonde, pibal, aircraft, satellite, and other data, to produce analyses of the atmosphere every 6 h since 1 January 1948 on a 2.5° × 2.5° grid (Kalnay et al. 1996; Kistler et al. 2001). We use the fields of temperature, wind speed, geopotential height, and mean sea level pressure (MSLP) from 20° to 70°N to calculate six extratropical NH cyclone indices in the winter months (December–January–February) of 1948/49–1998/99 (hereafter 1949–99). The fixed data assimilation system and model

alleviate concerns about inhomogeneities caused by these sources, although changes in data availability, data quality, and the types of observations used may still cause spurious trends. We perform an independent check on the wind speed index to determine whether the data are likely to be corrupted in this way (appendix).

### b. Indices

The six indices, based on prior literature and described below, were chosen to capture important characteristics of cyclone activity and forcing. The spatial patterns of our index values closely match published values, although in some cases the magnitudes of the indices are scaled relative to those in the literature. For all indices, we exclude values for locations in the Tibetan Plateau with surface elevation exceeding 4000 m to prevent the indices from depending heavily on extrapolation of fields below the surface, in particular the fields used to calculate the two indices measured at 500 hPa.

#### 1) MERIDIONAL TEMPERATURE GRADIENT

The meridional temperature gradient (Gitelman et al. 1997) indexes the temperature contrast across midlatitudes that is related to extratropical eddy activity. The value of the MTG at longitude $l$ for a given day is

$$\text{MTG}_l = T_{32.5°N,l} - T_{52.5°N,l}, \qquad (1)$$

the difference between daily mean surface temperature, $T$, at latitudes 32.5° and 52.5°N. Our index is the average of the daily values in a winter. This index differs from the others that follow in that it is calculated for individual longitudes rather than on a latitude–longitude grid.

#### 2) EADY GROWTH RATE

The Eady growth rate assesses baroclinic instability through the vertical gradient in horizontal wind speed in the troposphere and a measure of static stability (Hoskins and Valdes 1990). We follow Hall et al. (1994) in calculating Eady growth at 500 hPa to reduce problems with 780 hPa, the height used in Hoskins and Valdes (1990), falling below the land surface in areas of high elevation. The daily Eady growth rate, $\sigma_{\text{BI}}$, at a grid point is calculated from daily mean values as

$$\sigma_{\text{BI}} = 0.31 \frac{f}{N} \left| \frac{\partial \mathbf{v}}{\partial z} \right|, \qquad (2)$$

where $f$ is the Coriolis parameter, $N$ is the Brunt–Väisällä frequency (Lee and Mak 1994), $z$ is vertical distance, and $\mathbf{v}$ is the horizontal wind vector at 400 and 600 hPa. The winter average of the daily Eady growth rates is the final index.

#### 3) TEMPERATURE VARIANCE

Iskenderian and Rosen (2000) propose an index of cyclone activity based on the temperature variability at 500 hPa. We modify their approach by applying a 29-point Lanczos moving average filter (Duchon 1979) to the daily mean temperature series at each grid point to exclude variability at periods other than 2.5–10 days, thereby removing any seasonal temperature trend. Our index is the sample variance of the filtered series during winter. This index reflects anticyclone as well as cyclone activity.

#### 4) CYCLONE COUNT

An actual count of cyclones is a widely used measure of cyclone activity. However, such counts are subject to numerous problems, including arbitrary cutoff values, directional biases (Taylor 1986), difficulties when using a uniform latitude–longitude grid (Hayden 1981), high counts in areas of persistent stationary lows (Lambert 1996), and the changing relationship between geopotential height and wind/vorticity with latitude (Wallace et al. 1988). Finally, a count index measures neither intensity nor duration. For these reasons, any cyclone count should be considered an index as well, rather than a complete measure of cyclone activity.

Some cyclone counting approaches attempt to track individual cyclones using sophisticated algorithms (Sinclair 1994; Simmonds and Keay 2000; Gulev et al. 2001). For ease of computation, our cyclone count definition follows the nontracking approach of Zhang and Wang (1997) with minor modifications to fit the reanalysis grid. Our index counts the number of cyclones per winter at each grid point, thereby measuring cyclones in a way that relates directly to their impacts. We check for a cyclone at each location every 12 h. We take a cyclone to exist if the location is a minimum in the 1000-hPa geopotential height field relative to eight surrounding grid points 5° away (N, S, E, W, NE, SE, NW, SW). We use grid points 5° distant rather than 2.5° distant to better match the characteristic space and time scales of cyclone movement. If a minimum exists, the algorithm checks that the vorticity (based on the grid points 5° to the N, S, E, and W) exceeds $2 \times 10^{-5}$ s$^{-1}$ and that it remains positive 12 and 24 h later. Any detected cyclone is also considered to occur at the surrounding eight grid points 2.5° distant, which reduces undercounting due to fast-moving cyclones. To avoid double counting, only one cyclone is allowed within any 48 h at any location. To better assess impacts by capturing cyclones that remain at a location for several days, we calculated a variation on the index in which a cyclone is allowed every 24 h, but we found little empirical difference between the two definitions.

Some authors area-normalize the cyclone counts to account for the change in area of grid boxes with latitude, but others argue that this introduces a bias toward

higher counts at high latitudes [see Changnon et al. (1995) for an overview]. We do not area-normalize because we consider the cyclone index to count the number of cyclones at each grid point rather than cyclone activity per unit area. Our cyclone count index is not a completely consistent measure of cyclones across latitudes, because the distances involved depend on latitude. Since our analysis focuses on trends and correlations with other indices at individual locations, latitudinal differences are relatively unimportant provided that the method at each location is consistent through time.

### 5) COUNT OF INTENSE CYCLONES

An index that may be more closely related to impacts is the number of intense cyclones per winter, obtained as above but including only the cyclones with an MSLP of 980 hPa or less. Previous authors have used a 970-hPa cutoff (Carnell et al. 1996; Lambert 1996), but the small number of such events at individual locations makes a 970-hPa index difficult to analyze. For those locations with sufficient storms for analysis, the results for the 980- and 970-hPa indices were similar.

An error in the NCEP–NCAR data assimilation process caused some surface pressure observations of less than 1000 hPa to be omitted from the reanalysis during 1948–67, primarily in Europe (see http://www.cdc.noaa.gov/cdc/reanalysis/problems.shtml). The relatively small geographical extent of the problem suggests that it does not have a major impact on our analysis, but to the extent that other types of observations cannot replace the pressure data, the reanalysis may overestimate central pressures, thereby decreasing the number of intense cyclones, in this area.

### 6) EXTREME WINDS INDEX

A second index related to impacts uses the 95th percentile of four times daily wind speeds during winter (Alexandersson et al. 1998) at the $\sigma_{0.995}$ level (~40 m above the surface) of the reanalysis model. The four times daily data better capture the temporal variability at small spatial scales than daily values would. A wind index based on the 99th percentile yielded similar results. Either percentile should be regarded as an index rather than as a realistic value, because the assimilation by the reanalysis model and the 6-h sampling result in systematically lower values than percentiles based on instantaneous observations at point locations (Smith et al. 2001).

Observational wind data are subject to many biases. The reanalysis does not use land surface anemometer data but does use ship-based wind measurements. Possible biases in ocean wind measurements have caused concern as a source of spurious long-term trends in wind field estimates (Wright 1988; Cardone et al. 1990; Ward 1992; Gulev 1999). Wright (1988) and Ward (1992)

suggest using geostrophic winds derived from sea level pressure differentials to reduce bias. For comparison, we also calculate a version of the wind index based on the geostrophic wind field from the reanalysis. Finally, a more homogeneous source of geostrophic wind data for the northeastern Atlantic and northwestern Europe allows us to investigate data inhomogeneity in the reanalysis there (appendix).

## 3. Methods

The goal of the analysis is to compare the indices, in terms of both year-to-year covariation and long-term behavior. We focus on analyses at two spatial resolutions: broad regional averages and the full grid of 3024 points. Regional averages show the complexity of the temporal behavior but with a loss of spatial detail, such as activity specific to the storm tracks. Full grid analyses allow greater understanding of the spatial heterogeneity in the behavior of the indices over time but lose temporal detail because low-frequency behavior is difficult to display simultaneously at individual grid points other than as linear trends. Plots of the climatological mean fields are shown for reference in Fig. 1. The spatial patterns of the intense storm count mean field (not shown) are similar to those for the storm count index but without the local maxima west of Greenland and over the Mediterranean and Hudson Bay. In all analyses we assume independence between winters, since the autocorrelation and partial autocorrelation functions suggest no or very weak temporal autocorrelation even at lags of 1 yr.

### a. Regional analyses

We consider four regions (Fig. 1a), all restricted to 20°–70°N: 1) the entire NH belt; 2) a Eurasian region, 10°–127.5°E; 3) a Pacific region, 130°–247.5°E, which includes the Pacific storm track; and 4) an Atlantic region, 250°–7.5°E, which includes the Atlantic storm track. For MTG, the regional average is taken across the appropriate longitudes. For the other indices, we area-weight to account properly for the area represented by each grid point.

We fit the low-frequency variability in the indices using the Bayesian spline method of DiMatteo et al. (2001). We use a spline model for its flexibility and work within a Bayesian paradigm to estimate not only low-frequency variability but also confidence intervals for this behavior based on the posterior distribution. Cubic splines are piecewise cubic polynomials connected at points called knots, whose locations are often chosen arbitrarily. In contrast, this Bayesian framework allows the number of knots and their locations to be influenced by the data, but with a prior distribution that summarizes our belief regarding the number of knots. Our prior distribution is Poisson with mean two, reflecting a desire to estimate low-frequency variability
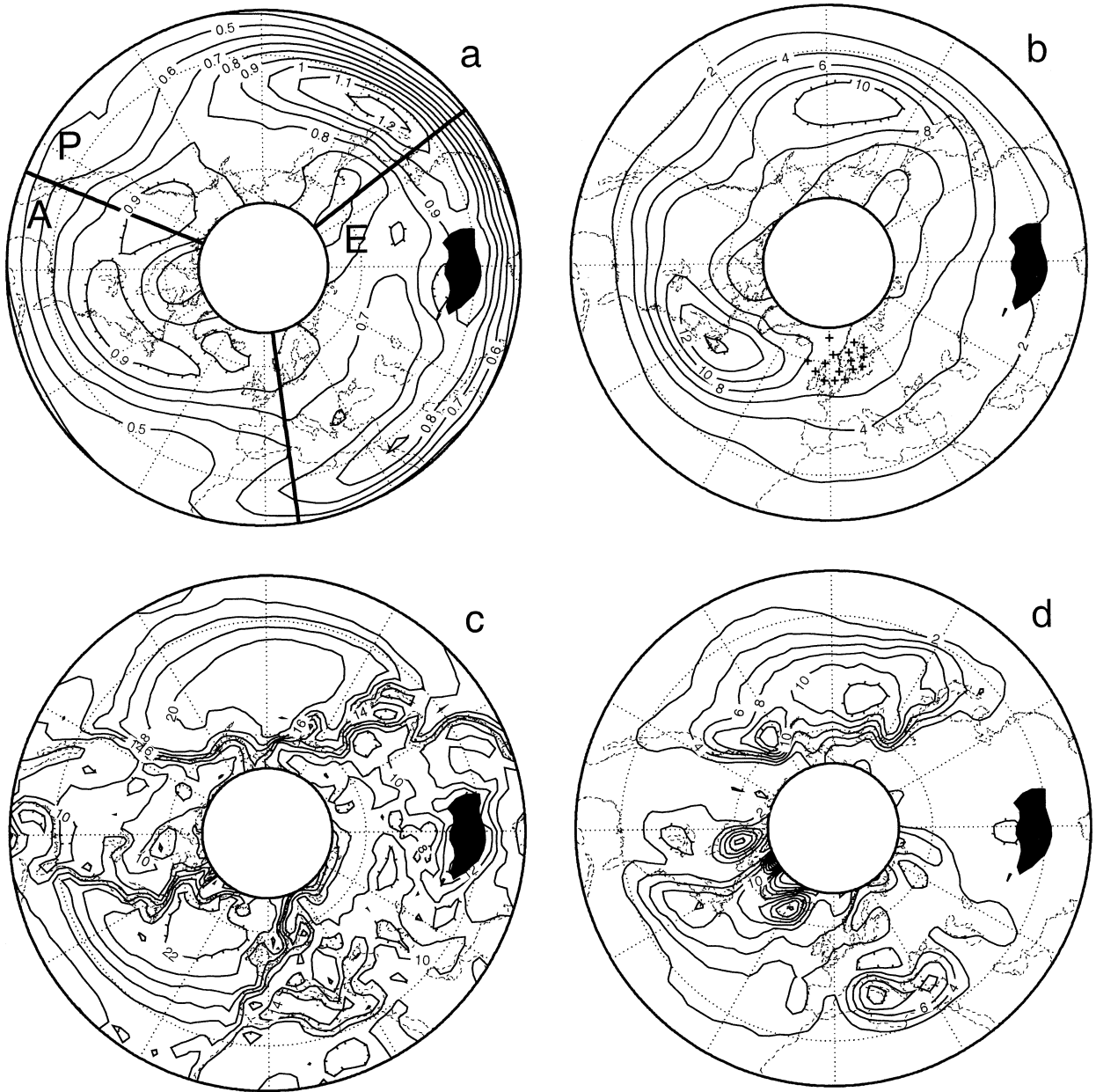
FIG. 1. Climatological mean fields for (a) Eady growth rate (day$^{-1}$), (b) temperature variance (°C$^2$), (c) wind index (m s$^{-1}$), and (d) cyclone count index (no. of events). Fields are calculated as the average of the winter values for 1949–99. Areas of high topography (at least 4000 m) are blacked out. Eurasian (E), Pacific (P), and Atlantic (A) regions are indicated in (a). The crosses in northwestern Europe and the northeastern Atlantic Ocean in (b) indicate the 16 NCEP–NCAR grid locations for which WASA triangles of stations were available for comparison with NCEP–NCAR extreme winds (see appendix and Fig. A1).

(whether natural or forced) rather than high-frequency variability.

Long-term behavior may be related to the behavior of atmosphere–ocean phenomena with characteristic teleconnection patterns. We focus on the Southern Oscillation (Trenberth 1984), North Pacific Oscillation (NP; Trenberth and Hurrell 1994), and North Atlantic Oscillation (NAO; Hurrell 1995) indices and fit linear regression models of the form

$$Y_l(t) = \alpha_l + \beta_l P(t) + \epsilon_{l,t}, \tag{3}$$

where $P(t)$ is the value of the teleconnection pattern averaged over December, January, and February of year $t$, $Y_l(t)$ is a cyclone index for region $l$, and $\epsilon_{l,t}$ are independent Gaussian errors with constant variance. The time series of residuals, $\epsilon_{l,t}$, represents the variability in the cyclone index after removing the linear influence of the teleconnection pattern. For each cyclone index-

teleconnection pair for a region, we compare the time series of residuals with the original time series to see if the teleconnection pattern influences the low-frequency variability. In addition, short-term variability in the cyclone indices may be related to short-term variability in the teleconnection patterns. To check this, we calculate the Spearman rank correlation (see below) between each of the cyclone indices in each region and each of three teleconnection patterns. To account for multiple significance tests, we use the false discovery rate (FDR) method described in section 3d.

We quantify the relationships between storm indices at the regional level with correlations between all pairs of indices in each region, using the 51 winters, 1949–99, as replicates. We use the nonparametric Spearman correlation, rather than the more common Pearson correlation, because the former accommodates non-normal data, such as our cyclone count indices. For consistency, we also use the Spearman correlation for continuous indices.

### b. Relationships between indices

As with the regional averages, we calculate Spearman rank correlations between pairs of indices at each grid point to assess how indices vary jointly over the NH. Because we want to capture both the high-frequency variability and the long-term trends, we report correlations calculated without detrending the data; note that correlations calculated after detrending the data are very similar. The $p$ value at each location is based on a two-sided $z$ test that uses the standard error of the correlation coefficient (Conover 1999). We map the correlations and assess their significance as described in section 3d.

For pairs that include the MTG, we calculate the correlation between the MTG at a given longitude and both 1) the other index at 42.5°N (halfway between the two latitudes used to calculate the MTG) and 2) the average of the other index over 20°–70°N.

We find that some pairs of indices are more highly correlated when one index is offset spatially with respect to the other. The best offset is defined to maximize the average correlation, $\bar{r}$, across all locations in the region 20°–70°N with respect to $(y, x)$:

$$\arg\max_{(y,x)} \left[ \bar{r} = \sum_{\phi=20°}^{70°} \sum_{\lambda=0°}^{360°} \frac{r(Y_{i,\phi,\lambda}, Y_{j,\phi+y,\lambda+x})}{n} \right], \quad (4)$$

where $r(\cdot)$ denotes the correlation between index $i$ at lat–lon location $(\phi, \lambda)$ and index $j$ offset spatially from that location by $(y, x)$ and $n$ is the number of locations. This maximization attempts to find the shift in one index that results in the largest positive point-by-point correlations with another index. We search for the best offset over $y \in (-15°, 15°)$, $x \in (-45°, 45°)$, namely, the region bounded 15° south, 15° north, 45° west, and 45° east of the focal location.

Pairwise scatterplots of indices at a sample of locations showed no substantial nonlinearity in the relationships, which suggests that correlations are adequate statistics for relating the indices. Correlation maps based separately on the periods 1948–73 and 1974–99 were similar, suggesting relatively constant relationships through time.

### c. Linear trends

To examine spatial patterns in the long-term linear trends for Eady growth rate, temperature variance, and the wind index, we fit simple linear regressions for each index individually at each grid point $l$:

$$Y_l(t) = \alpha_l + \beta_l t + \epsilon_{l,t}, \quad (5)$$

where $\epsilon_{l,t}$ are independent Gaussian errors with constant variance, and $t$ is the year. For count indices, we use Poisson regression (McCullagh and Nelder 1989), where $Y_l$ is now taken to be a Poisson random variable with mean

$$E[Y_l(t)] = \exp(\alpha_l + \beta_l t). \quad (6)$$

The Poisson distribution is the most commonly used distribution for discrete data, as the Gaussian distribution is for continuous data, and the exponential relationship is the standard way to prevent prediction of negative counts. The exponential relationship may appear unrealistic, but for the time period considered here, examination of the model at a sample of locations shows that the exponential fits are close to linear. We then map the maximum likelihood estimates of the slope coefficients, $\hat{\beta}_l$ from (5)–(6), and perform simultaneous testing as described in section 3d. The $p$ values are calculated from two-sided tests based on the usual $t$ statistic for the linear regressions and on a likelihood ratio test for the Poisson regressions.

We performed standard regression diagnostics for a sample of locations and found that for the continuous indices the assumption of Gaussian errors with equal variance is reasonable. The count data have higher variance than expected under the Poisson assumption, but quasi-likelihood models (McCullagh and Nelder 1989) that allow for higher variance gave reasonably similar results. To diagnose possible nonlinear trends, we fit nonparametric models by replacing $\beta$ in (5)–(6) with a cubic smoothing spline (section 3a) with four degrees of freedom (Hastie and Tibshirani 1990), using the generalized additive model (gam) function in S-PLUS (Venables and Ripley 1997). The improvement in fit is assessed via the proportion reduction in the sum of squared errors (residual deviance for the Poisson models).

### d. Simultaneous inference

Analyses of the full grid require that inference be made at many locations simultaneously, but use of the usual 0.05 significance level may falsely reject many null hypotheses (on average 151 = 0.05 × 3024, if there

is no real trend at any location). Various corrections have been suggested, none of which is entirely satisfactory. Here we use an approach popular in the climatological literature, that of Livezey and Chen (1983), and a promising new procedure for controlling the false discovery rate, defined as the proportion of true null hypotheses among rejected hypotheses (Benjamini and Hochberg 1995).

For the Livezey and Chen (1983) analysis, the entire field is deemed significant if the number of significant locations, $n_{\text{sig}}$, exceeds the 95th percentile of its distribution under the null hypothesis, which is obtained by simulating data under the full null hypothesis of no real trends at any location. For location $l$, a simulated sample is $Y_l(t^*)$, where $t^*$ is a vector of 51 years sampled at random and with replacement from the years (1949, 1950, . . . , 1999). This destroys any temporal trends that may exist, yet preserves the spatial structure of the data when the same $t^*$ is used at every location. Note that different $t^*$'s are used for different indices when analyzing correlations. For each analysis, the null distribution of $n_{\text{sig}}$ is based on 1000 such simulated samples. While this field significance approach is useful, it does not specify which individual locations show significant trends, suggesting that tests at every location may be helpful.

Traditional methods of simultaneous testing specify the individual tests to reject so that the probability of at least one falsely rejected null hypothesis (familywise error control), is $\alpha$, usually 0.05. This typically gives too little power, failing to reject many null hypotheses for which the alternative hypothesis actually holds. A more powerful approach controls the false discovery rate (FDR), defined as the proportion $q$ of falsely rejected null hypotheses relative to the total number of rejected hypotheses (Benjamini and Hochberg 1995). To apply the procedure, $p$ values for each location are calculated in the usual manner. A threshold is determined based on the number of tests, the observed $p$ values, and $q$; locations with $p$ values smaller than this threshold are deemed significant. This guarantees that on average the proportion of falsely rejected null hypotheses out of all rejected null hypotheses is no greater than $q$. As with significance levels, the choice of $q$ is arbitrary, but 0.05 seems sufficiently protective against falsely rejected null hypotheses. The original Benjamini and Hochberg (1995) procedure assumes that the $p$ values of the true null hypotheses are independent of the $p$ values of the true alternative hypotheses and that the $p$ values for the true null hypotheses are all independent; these assumptions are certainly violated for climatological data with high spatial correlation. Another FDR procedure (henceforth called FDR-A) assumes only independence of the two groups of $p$ values [Yekutieli and Benjamini 1999, Eq. (9)]. While this assumption is also likely to be violated, the violation is restricted to the boundary between the groups. An alternative (FDR-B) makes no such independence assumption (Benjamini and Yeku-

tieli 2001, theorem 1.3), but it is very conservative, rejecting few null hypotheses, as it protects against extreme dependence between test statistics. In the next section, our maps make use of FDR-A, but we also report the results of FDR-B in the captions.

## 4. Results

### a. Regional analyses

Figures 2–5 show the regional averages of the six indices plotted against time for the four regions defined in section 3a. Eady growth rate shows an increase until 1980 in all but the Atlantic region. The MTG shows little change over the period, except possibly a long-term decline in Eurasia. The year-to-year variation in our MTG is similar to that of Gitelman et al. (1997), who show a century-long decline not seen in the shorter time period available through the reanalysis. Temperature variance shows a clear increase in the NH and Eurasia and mild evidence for an increase in the Pacific and an increase in the 1970s in the Atlantic. The cyclone count index has declined since 1970 in the NH and Eurasia, with less evidence for a decline in the Atlantic and little clear trend in the Pacific. In contrast, the intense cyclone count shows increases in all regions, particularly in the Pacific, with the increase in the Atlantic possibly leveling off in recent years. The distinct behaviors of our two cyclone count indices are of particular interest in light of some model predictions of increases in intense cyclones but no trend or a decline in the total cyclone count (Carnell and Senior 1998; Knippertz et al. 2000).

Extreme winds increase sharply in all regions in the 1950s, followed by a smaller increase in the Pacific, a decrease in Eurasia, and a slight increase in the Atlantic. One possible explanation for the jumps in the 1950s involves the changes in the techniques used for ocean wind measurements, with particular concerns about changes during the 1940s and 1950s (Cardone et al. 1990). We cannot directly assess this possibility; however, two pieces of circumstantial evidence fail to support the hypothesis. Regional averages of the wind index based on the geostrophic wind field (not shown) show very similar behavior to the $\sigma_{0.995}$-based index, including the jumps in the 1950s. The geostrophic field is based on pressure differences, which should reduce the effect of any trends in pressure measurements (Wright 1988; Ward 1992). Second, we decomposed the behavior at the grid points into EOFs of time. One of these EOFs primarily reflects the jump in the 1950s, and for this EOF, large positive coefficients are not confined to ocean grid points.

Comparison of the low-frequency behavior in Figs. 2–5 with that based on residuals from the storm index–teleconnection regressions [Eq. (3)] suggests that for most indices, the low-frequency behavior is not strongly influenced by the teleconnection patterns. The largest
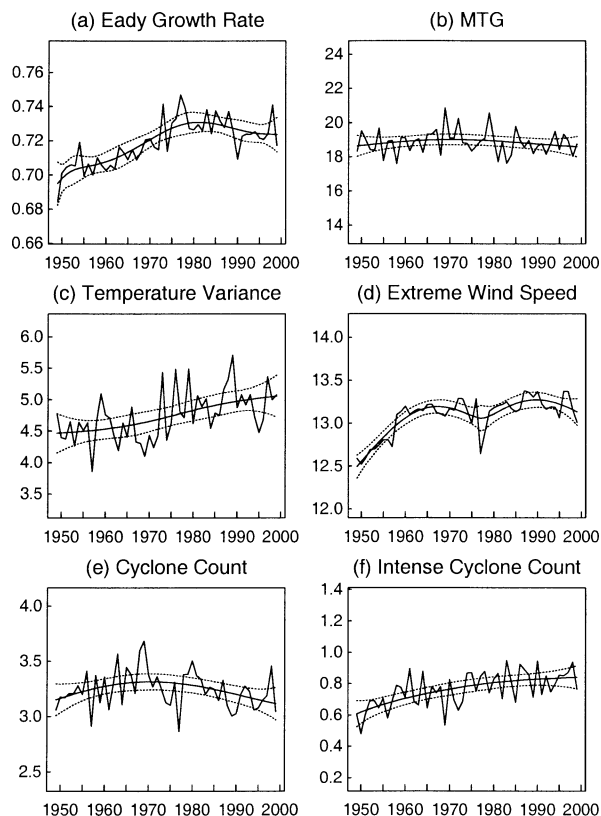
FIG. 2. Time series plots of the area-weighted averages over the NH, 20°–70°N, for (a) Eady growth rate (day$^{-1}$), (b) MTG (°C), (c) temperature variance (°C$^2$), (d) extreme wind speed (m s$^{-1}$), (e) cyclone count (no. of events), and (f) intense cyclone count (no. of events). The jagged solid line connects the winter means. The smooth solid line is the trend from the Bayesian spline model, with pointwise 95% confidence intervals (dashed lines). For a given index, the difference between the minimum and maximum values on the $y$ axis is constant for the regions in Figs. 2–5. We also calculated linear trends ($p$ values) for each of the indices: (a) 0.00061 (2.9 × 10$^{-8}$), (b) −0.0025 (0.71), (c) 0.013 (0.00013), (d) 0.0098 (2.7 × 10$^{-7}$), (e) −0.0014 (0.40), and (f) 0.0045 (3.4 × 10$^{-6}$).



FIG. 3. Time series plots of the area-weighted averages of each index for the Eurasian region, 20°–70°N, 10°–127.5°E. Details are as in Fig. 2. We also calculated linear trends ($p$ values) for each of the indices: (a) 0.0011 (5.0 × 10$^{-11}$), (b) −0.023 (0.20), (c) 0.016 (1.4 × 10$^{-5}$), (d) 0.0016 (0.44), (e) −0.0043 (0.047), and (f) 0.0017 (0.016).

effect is that the observed increase in intense cyclones in the Pacific region is associated with the long-term decline in the NP index (Trenberth and Hurrell 1994). This relationship between the NP index and the intense cyclone index in the Pacific is also seen in their strong negative correlation, −0.94, showing that the high-frequency variability is also strongly related between the two. Other significant correlations between cyclone indices and teleconnection patterns, indicating correspondence in their high-frequency variability, are seen for MTG (Eurasia)–NAO, Eady (Pacific)–NP, cyclone count (Atlantic)–NAO, and extreme winds (Pacific)–NP (Table 1). With these exceptions, nearly all involving a storm index and a teleconnection pattern in the same area, the results suggest that the regional averages of the storm indices are not strongly influenced by short-term variability in the teleconnection patterns.

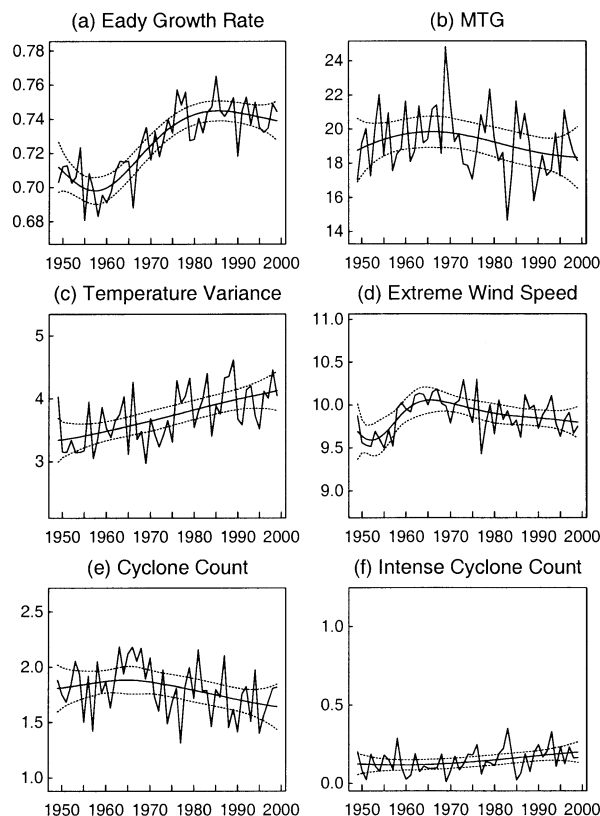The correlations between pairs of storm indices at the

regional level show no consistent patterns of large positive or negative values. This occurs because high correlations are confined to areas smaller than the regions, as we show below in the full grid correlation analysis.

### b. Relationship between indices

Maps of correlations between indices highlight grid points at which pairs of indices are closely related. Figure 6 shows the Spearman rank correlation map between Eady growth rate and temperature variance, the pair of indices (apart from the two count indices) with the highest correlations across the hemisphere as a whole. The highest values are over the two major storm tracks of the NH, with correlations higher in the downstream regions than in the upstream regions. Eady growth rate and extreme winds are also most strongly correlated over the oceans (Fig. 7), with significant correlations largely confined to the oceans and northern Eurasia. Similar results are observed for the correlations between temperature variance and extreme winds (not shown). The correlations between any of the Eady, temperature variance, and wind indices, and either of the two count
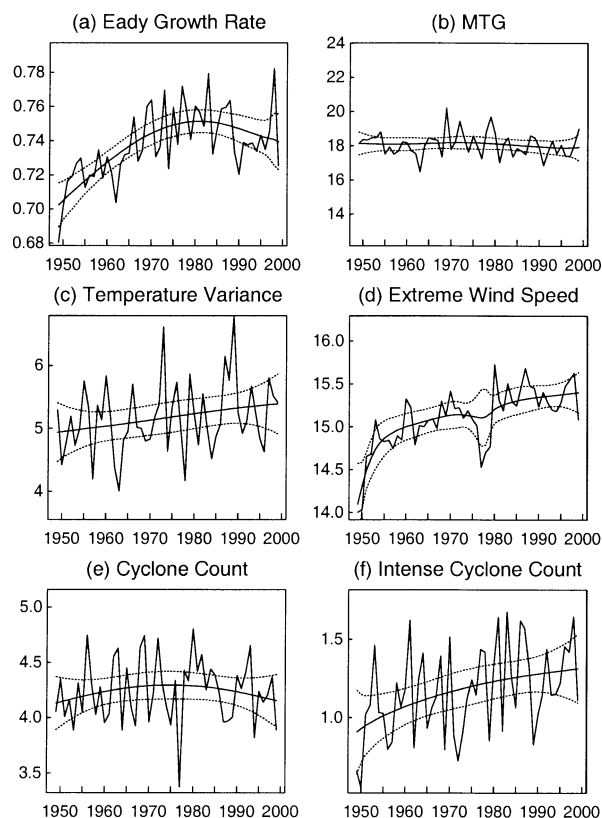
FIG. 4. Time series plots of the area-weighted averages of each index for the Pacific region, 20°–70°N, 130°–247.5°E. Details are as in Fig. 2. We also calculated linear trends ($p$ values) for each of the indices: (a) 0.00075 ($3.9 \times 10^{-5}$), (b) $-0.0045$ (0.50), (c) 0.0094 (0.078), (d) 0.016 ($5.3 \times 10^{-8}$), (e) 0.00047 (0.87), and (f) 0.0075 (0.0040).
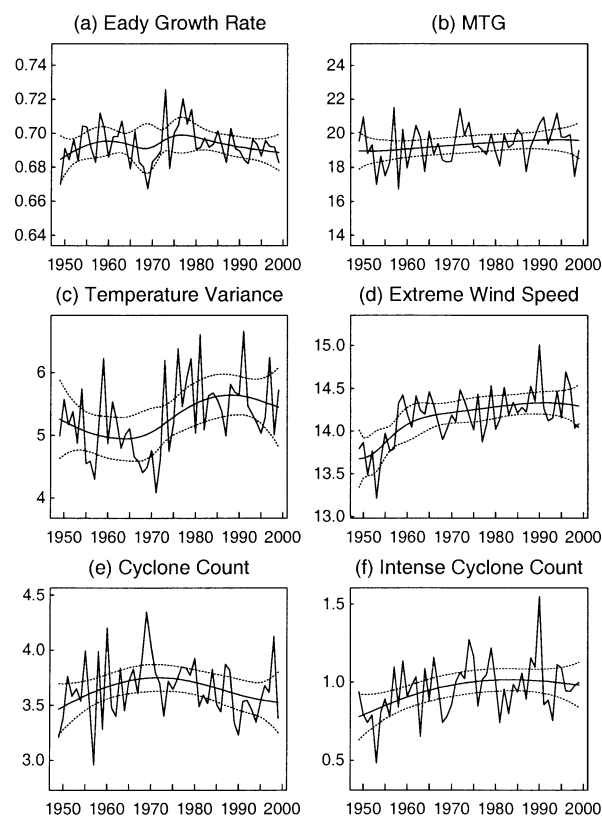


FIG. 5. Time series plots of the area-weighted averages of each index for the Atlantic region, 20°–70°N, 250°–7.5°E. Details are as in Fig. 2. We also calculated linear trends ($p$ values) for each of the indices: (a) $8.6 \times 10^{-6}$ (0.94), (b) 0.016 (0.14), (c) 0.014 (0.011), (d) 0.011 ($2.7 \times 10^{-5}$), (e) $-0.00035$ (0.89), and (f) 0.0040 (0.016).

indices are weak and generally significant at few locations (for both FDR-A and FDR-B); they do not show the tendency toward correlations over the oceans (not shown). The exception is significant correlations of 0.3–0.5 between the wind and intense cyclone count indices over portions of the North Atlantic and Pacific Oceans. The correlations between the two cyclone count indices are greater than 0.6 over the storm tracks but cannot be determined over much of the continental areas because of the small number of intense events in those locations (not shown). All of these correlation fields are field-significant according to the Livezey and Chen (1983) approach with $p \leq 0.001$.

Some of these results change substantially when a spatial offset is introduced into the correlation analyses. For pairs of indices in which either cyclone count index was involved, the best offset is 5°–10°S, and in some cases approximately 5° to the west, as shown in Fig. 8. We also obtain the same offset when searching within individual regions, although the southerly offsets in the land areas tend to be somewhat less pronounced (c. 5°) than in the oceanic areas (c. 10°). The best offsets are determined based on the average of 3024 correlation

TABLE 1. Spearman rank correlations between cyclone index regional averages and teleconnection patterns. Significant correlations based on the FDR-B method are indicated by italics.

| Cyclone index | NH | Eurasia | Pacific | Atlantic |
|---|---|---|---|---|
| Southern Oscillation Index | | | | |
| Eady | −0.29 | −0.10 | −0.45 | −0.06 |
| MTG | 0.19 | 0.12 | −0.22 | 0.29 |
| Temperature variance | −0.30 | −0.21 | −0.06 | −0.19 |
| Wind | −0.07 | 0.01 | −0.16 | −0.01 |
| Cyclone count | −0.11 | −0.22 | 0.01 | −0.03 |
| Intense cyclone count | −0.38 | −0.12 | −0.39 | 0.00 |
| North Pacific Oscillation | | | | |
| Eady | *−0.50* | −0.39 | *−0.53* | −0.08 |
| MTG | 0.31 | −0.05 | 0.37 | 0.37 |
| Temperature variance | −0.14 | −0.30 | 0.23 | −0.19 |
| Wind | −0.19 | 0.04 | *−0.49* | 0.00 |
| Cyclone count | −0.06 | −0.08 | 0.19 | −0.28 |
| Intense cyclone count | *−0.71* | −0.02 | *−0.94* | 0.13 |
| North Atlantic Oscillation | | | | |
| Eady | 0.06 | 0.21 | −0.11 | −0.12 |
| MTG | −0.37 | *−0.69* | 0.03 | 0.33 |
| Temperature variance | 0.13 | −0.12 | 0.02 | 0.31 |
| Wind | 0.07 | −0.15 | 0.15 | −0.07 |
| Cyclone count | −0.46 | −0.37 | −0.04 | *−0.52* |
| Intense cyclone count | 0.14 | 0.36 | −0.02 | 0.13 |
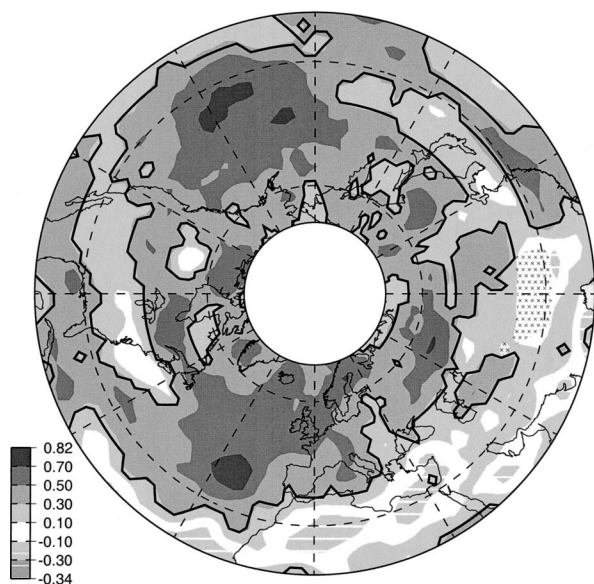
FIG. 6. Spearman rank correlations between Eady growth rate and temperature variance. Gray shading indicates value of correlation coefficient with darker colors indicating values larger in magnitude. Negative values are indicated by horizontal hatching. The areas inside the thick lines have correlations significantly different from zero based on the false discovery rate procedure (FDR-A). Here 66% of grid points are significant based on FDR-A (36% based on FDR-B). Areas with high topography (at least 4000 m) are marked with ×s on a white background.

coefficients, so we have a high degree of confidence in the location of the maxima by virtue of the central limit theorem. Less notably, the pairs of indices involving Eady growth, temperature variance, and the wind index
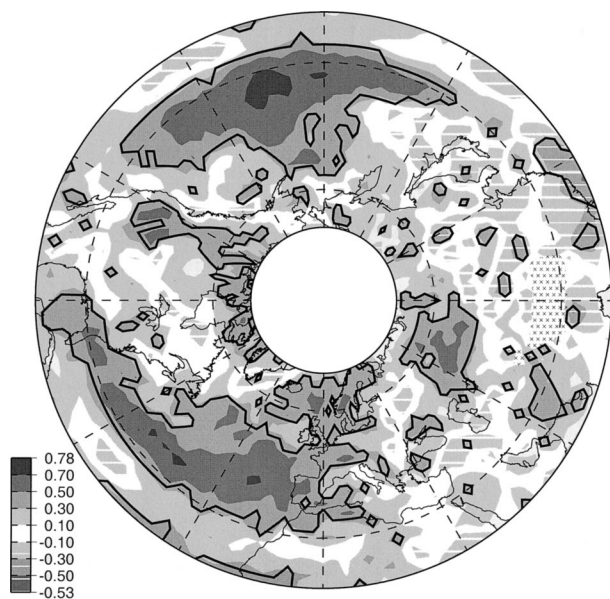


FIG. 7. Spearman rank correlations between the Eady growth rate and wind indices. Here 30% of locations are significant based on FDR-A (11% based on FDR-B). Details are as in Fig. 6.
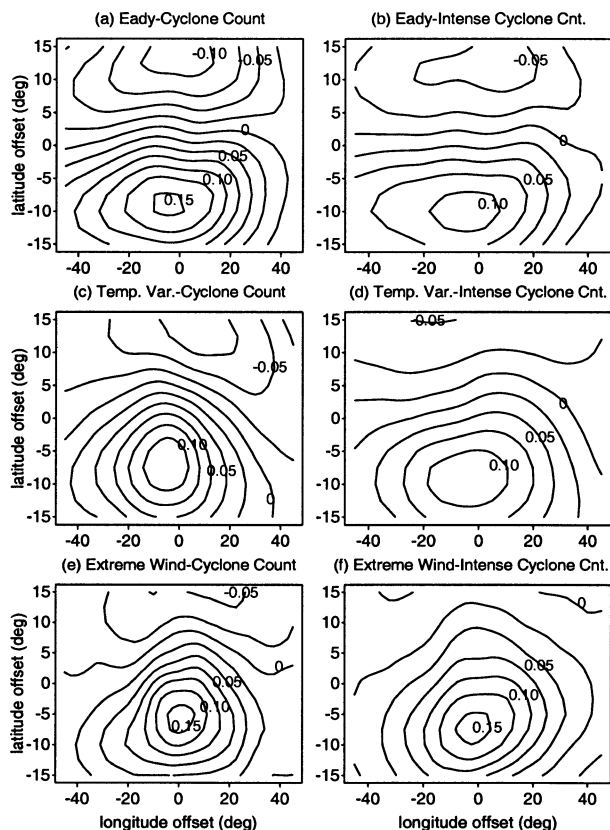


FIG. 8. Contours of average correlation, $\bar{r}$ in (4), as a function of longitude and latitude offsets, $x$ and $y$, respectively, for (a) Eady-cyclone count, (b) Eady-intense cyclone count, (c) temperature variance-cyclone count, (d) temperature variance-intense cyclone count, (e) extreme wind-cyclone count, and (f) extreme wind-intense cyclone count. For a given plot, any fixed point indicates the average, $\bar{r}$, over all grid points, of the correlation between the count index at a given grid point and the noncount index at the grid point offset in space by the indicated amount on the axes of the plot. The individual correlations are calculated based on the 51 years of data.

are slightly more highly correlated with Eady growth offset 5°W of either temperature variance or extreme winds (not shown).

Figure 9 shows the correlation between Eady growth and the cyclone count index with Eady offset 7.5°S and 5°W. Significant positive correlations of the order of 0.5 are seen over the northeastern Atlantic and Pacific Oceans, with weak correlations elsewhere, compared to correlations no larger in magnitude than 0.3 without the offset. The results are similar for the correlations between Eady growth rate (with the above offset) and intense cyclone counts (not shown). The same can be observed when Eady is replaced by temperature variance, although the correlations over the oceans are somewhat weaker. With the offset, there are significant positive correlations between extreme winds and the cyclone count index over much of the North Atlantic and North Pacific Oceans (Fig. 10). The correlations with offset between extreme winds and intense cyclones (not shown) are markedly stronger, suggesting that high
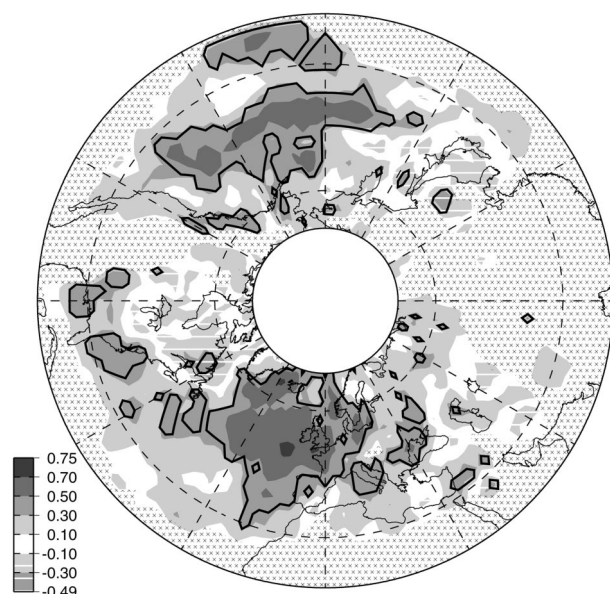
FIG. 9. Spearman rank correlations between the Eady growth rate and cyclone count indices with the former offset 7.5°S and 5°W. Here 25% of locations are significant based on FDR-A (11% based on FDR-B). Areas with high topography (at least 4000 m) or fewer than one storm per year on average are marked with ×s on a white background. Other details are as in Fig. 6.
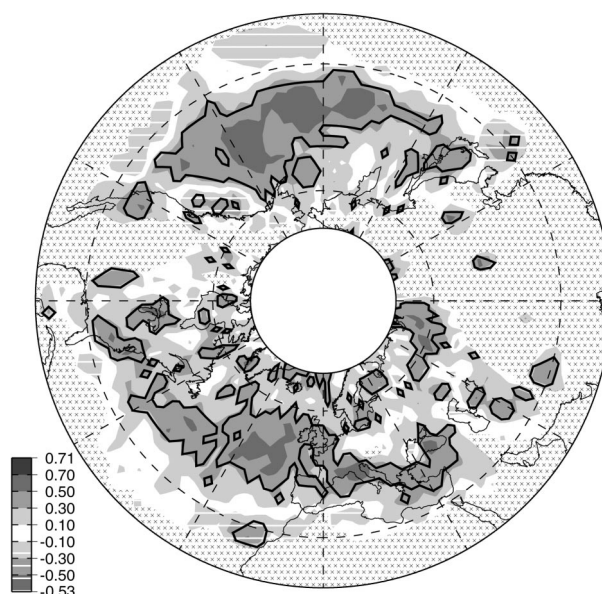


FIG. 10. Spearman rank correlations between the wind index (with offset) and cyclone count index. Here 28% of locations are significant based on FDR-A (6% based on FDR-B). Map details and offset are as in Fig. 9.

wind speeds are more closely related to the intense cyclone count index than the cyclone count index, as we would expect. In general, the strongest positive correlations between count and noncount indices occur predominantly over the middle and downstream portions of the storm tracks. For all pairs of a count index and a noncount index offset in space, the correlations are field significant with $p < 0.001$.

MTG computed by longitude is generally uncorrelated with other indices, regardless of whether these are considered at 42.5°N or as averages over latitudes 20°–70°N. The highest correlations, on the order of 0.4, are between MTG and Eady, both indices of cyclone forcing, at 42.5°N. Correlations between MTG and other indices offset in longitude are not appreciably higher.

### c. Linear trend models

Figures 11–13 display the estimated linear trends, $\hat{\beta}_l$ in (5), for the Eady growth rate, temperature variance, and wind indices. There are positive trends in both the North Atlantic and North Pacific Oceans, with large areas of significance in the Pacific and smaller areas in the Atlantic. In the Pacific, the area of largest positive trends in Eady growth rate is upstream relative to the areas of maximum trends in temperature variance and wind. This is consistent with the notion of Eady growth rate as an index of cyclone forcing, with the strongest responses in cyclone activity expected downstream from the locus of change in Eady growth rates (Chang and Orlanski 1993). In the Atlantic, Eady growth has a sig-
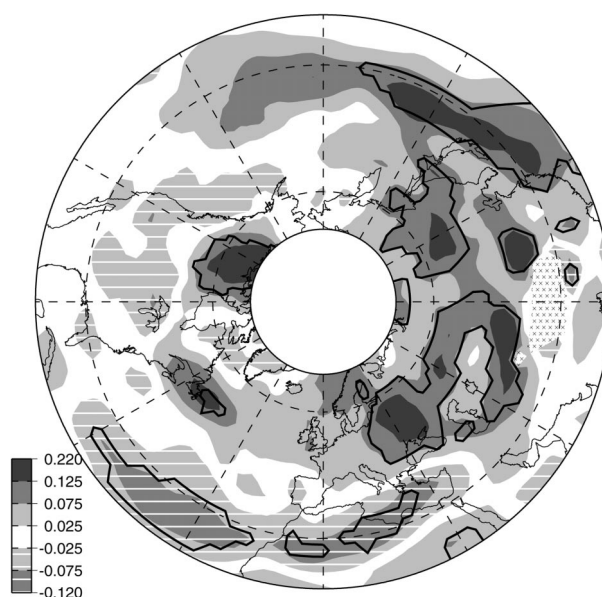


FIG. 11. Estimated change in Eady growth rate based on linear regression of Eady growth rate against time in units of day$^{-1}$ (50 yr)$^{-1}$. Gray shading indicates estimated change over 50 years with darker shading indicating values larger in magnitude. Negative values are indicated by horizontal hatching. The areas inside the thick lines have trends significantly different from zero based on the false discovery rate procedure (FDR-A). Here 18% of locations are significant based on FDR-A (4% based on FDR-B). Areas with high topography (at least 4000 m) are marked with ×s on a white background.
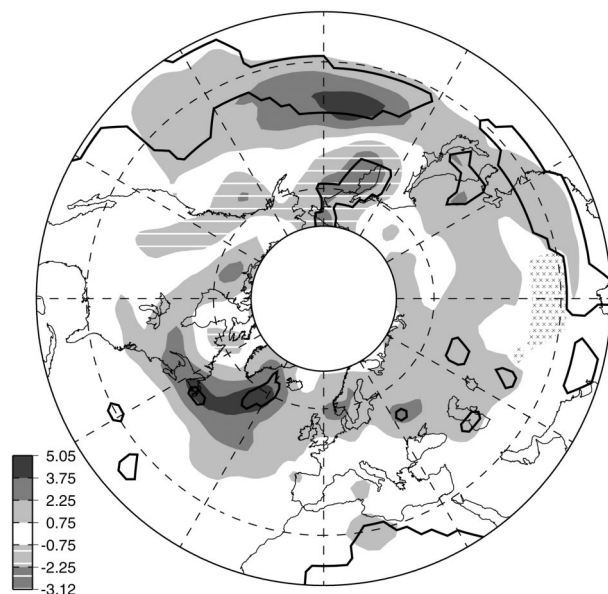
FIG. 12. Estimated change in temperature variance based on linear regression of temperature variance against time in units of °C² (50 yr)⁻¹. Here 12% of locations are significant based on FDR-A (1% based on FDR-B). Map details are as in Fig. 11.
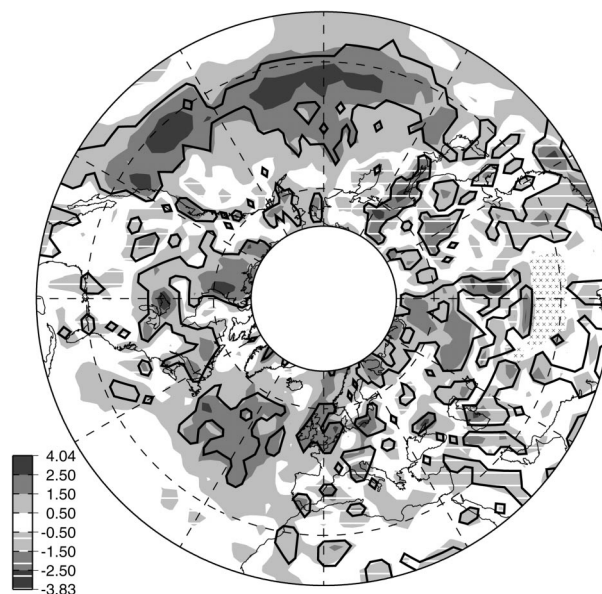


FIG. 13. Estimated change in extreme winds from linear regression of the wind index against time in units of m s⁻¹ (50 yr)⁻¹. Here 33% of locations are significant based on FDR-A (11% based on FDR-B). Map details are as in Fig. 11.

nificant positive trend only in a small area east of New-foundland. The maximum positive trend for temperature variance stretches from Newfoundland northeast to Iceland, with a significant increase in the downstream portion of this area, very similar to the result in Iskenderian and Rosen (2000) based on unfiltered reanalysis temperature data for 1958–96. The wind index shows significant positive trends in a large area east of Newfoundland, as well as over the United Kingdom and the North Sea. All three indices suggest increasing activity west of Hudson Bay, although this is not significant for temperature variance. Eady growth and temperature variance both show positive trends over much of Eurasia, which are significant for large areas only for Eady growth. Trends in the wind index over Eurasia are heterogeneous. The results are field significant for all the indices ($p < 0.001$).

We also fit linear trends in extreme winds excluding the period of sharp increase, 1949–59, seen in Figs. 2–5. These trends (not shown) are similar to those in the full time series, except in the southwestern North Atlantic; however, the areas of significance are much smaller, which could be caused in part by fewer data points, because statistical power generally increases with sample size. This suggests that the increase in extreme winds is not solely caused by the jumps in the 1950s, although these certainly contribute to the positive trends seen in Fig. 13. The field of trends based on the geostrophic wind is similar to, but smoother than, that based on the $\sigma_{0.995}$ wind, although with broader areas of significant declines over eastern Eurasia (not shown).

Trends in the count indices are difficult to analyze

outside the storm tracks, where few cyclones are recorded; we do not fit models at locations for which fewer than one storm or intense storm per year on average was recorded. The count indices show positive trends over parts of the North Pacific and North Atlantic that are similar to the other indices, although they are not significant based on the FDR procedure (Figs. 14–15), and there are decreases in some areas for the cyclone count index. Field significance is achieved with $p = 0.003$ ($p = 0.018$) for the cyclone count index (intense cyclone count index). This difference between the field significance and FDR results demonstrates the difficulty in interpreting field significance when few individually significant locations can be identified.

Comparison of the trend plots with plots of the climatological mean fields (Fig. 1) suggests a possible downstream shift in the Atlantic storm track based on Eady growth rate and temperature variance; there is much less evidence for such a shift in the Pacific storm track. No index shows a significant trend that could indicate a poleward shift, although the cyclone count index does show a possible shift to the north in the Atlantic. For the wind index the areas of significant increase are somewhat equatorward of the maxima in the mean field.

Linear trends may not be appropriate everywhere. Maps of the improvement in fit achieved by using smoothing spline models in place of the linear term in the regression models [Eqs. (5)–(6)] suggest some significant nonlinearity for some locations for most indices (not shown). The linear model appears to fit best in the case of temperature variance, although the spline model
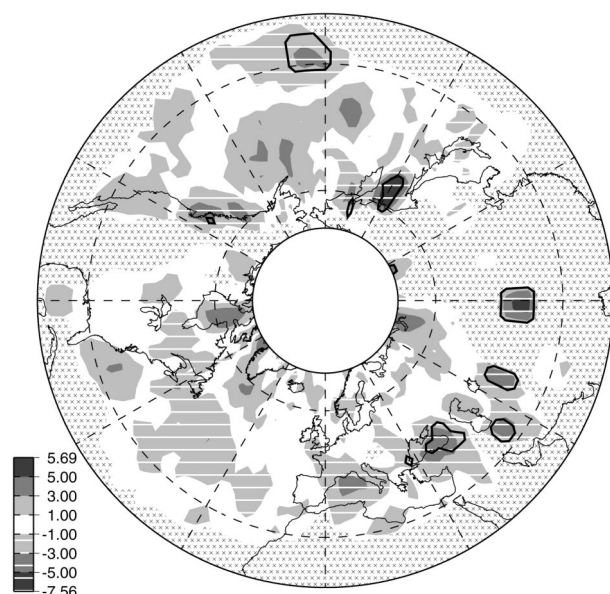
FIG. 14. Estimated change in number of storms in units of $(50 \text{ yr})^{-1}$ based on a Poisson regression model of the cyclone count index against time. Gray shading indicates estimated change over 50 years (difference between model-estimated number of storms in 1999 and 1949) with darker shading indicating values larger in magnitude. Negative values are indicated by horizontal hatching. Here 3% of locations are significant based on FDR-A (3% based on FDR-B). Areas with high topography (at least 4000 m) or fewer than one storm per year on average are marked with ×s on a white background.
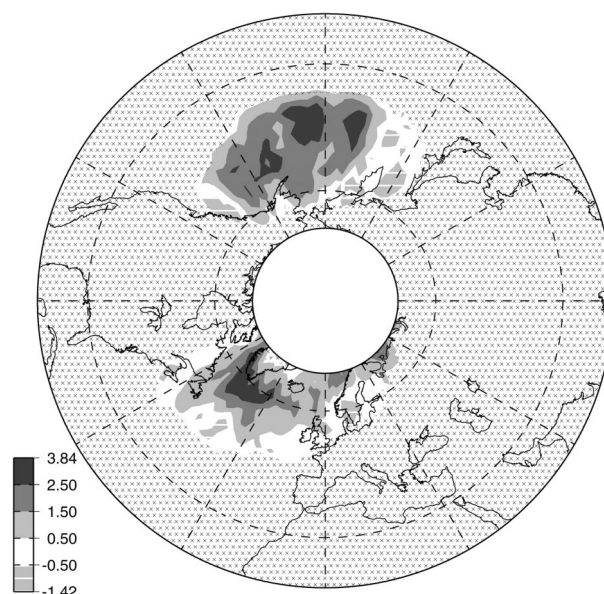
FIG. 15. Estimated change in the number of intense storms in units of $(50 \text{ yr})^{-1}$ based on a Poisson regression model of the intense cyclone count index against time. No locations are significant based on FDR-A (0.2% based on FDR-B). Map details are as in Fig. 14.

is a significant improvement across much of Canada and the Atlantic at 45°–65°N. For Eady, areas of significant improvement are concentrated in the Atlantic and Eurasia. For extreme winds, there are large areas of significant improvement over Eurasia and North America, but the linear model appears to fit reasonably well over the oceans. Areas of improvement in the count index models are scattered.

## 5. Discussion

We report correlation and trend analyses for six cyclone indices over the NH on the full reanalysis grid and as regional averages for the winters 1949–99. The reanalysis allows comparison of indices derived from the same dataset over a large spatial area, at the cost of changing data sources over time and the heavy influence of the assimilation model in areas of sparse observations. The regional averages allow examination of low-frequency variability and show the temporal complexity of the indices (Figs. 2–5). Comparison of regional average and full grid results shows that large-scale averaging obscures many of the small-scale correlation and trend patterns, particularly the localized behavior in the storm tracks. However, simple linear trend analyses at grid points miss some nonlinear behavior.

Correlations among indices are on the order of 0.4 and higher in the storm tracks, but much smaller in most

other locations, particularly over the continents, which suggests that the indices provide reasonably consistent measures of cyclone activity in the storm tracks. The weaker correlations elsewhere are not surprising given the smaller number of storms in these areas, but this points out the difficulty in reliably measuring storm activity away from the tracks.

The cyclone count indices do not correlate well with other indices unless the latter are spatially offset 5°–10° to the south, after which, correlations over the storm tracks become much larger. The offset between cyclone counts and temperature variance agrees with other reports of offsets between cyclone counts and eddy activity as measured by the variance of the MSLP or geopotential height fields. Wallace et al. (1988) report that positive and negative anomalies in the high-pass-filtered geopotential height field have similar tracks; but that with the addition of the mean climatology, the negative anomalies tend to move toward areas of low mean pressure, while the positive anomalies move toward the areas of high mean pressure in the subtropics. For this reason, Wallace et al. (1988) recommend that the term ''storm track'' be replaced by ''baroclinic wave guide,'' as the storm tracks are traditionally based on eddy statistics. Hoskins and Valdes (1990) note that cyclonic circulation tends to occur on the poleward side of the storm tracks as defined by eddy statistics. GCM simulations have shown that the maxima in cyclone counts tend to be poleward of the maxima in eddy activity (König et al. 1993; Carnell et al. 1996). Finally, the concordance between rotational and translational components of cyclone movement in the area south of the

center of the cyclone would contribute to the offset between extreme winds and the count indices. There is a much less pronounced offset in the east–west orientation; temperature variance and Eady growth rate are displaced to the west of the cyclone centers, and Eady growth is displaced to the west of temperature variance and extreme winds. The westward Eady growth offsets are consistent with Chang and Orlanski (1993), who show that Eady growth maxima tend to lie on the upstream portions of the storm tracks. In addition, the characteristic east–west vertical tilt of storm systems would typically yield surface storm centers slightly downstream of these midtropospheric (500 hPa) storm indicators.

There are significant positive trends over portions of the storm tracks for the Eady growth rate, temperature variance, and wind indices. The cyclone count indices show positive trends over the storm tracks, but these are not significant based on FDR, although field significance is achieved. Increases are seen in most regional averages for Eady, temperature variance, extreme winds, and intense storms, whereas the cyclone count index shows some evidence of decline. Our results are generally consistent with other empirical studies of storm activity. Lambert (1996) reports a post-1970 increase in intense cyclones, and McCabe et al. (2001) find an increase in storm intensity at both midlatitudes and high latitudes, consistent with the increase in our regional averages. The significant declines in the number of cyclones in the $30°–60°N$ band seen in Key and Chan (1999) and McCabe et al. (2001) are consistent with the NH-wide $(20°–70°)$ decline in our regional average cyclone counts since 1970 (Fig. 2). Using a tracking approach, Gulev et al. (2001) find similar location-specific trends in cyclones and intense cyclones, but their regional results differ from ours, perhaps due to differences in definitions of the regions and the northern winter season. Geng and Sugi (2001) use a principal component analysis of tracked storms to find an increase in North Atlantic storm track cyclone activity, while Graham and Diaz (2001) find intensification in North Pacific cyclones. Studies of North American trends in cyclone counts based on historical cyclone track maps have found evidence for a decline in winter cyclones during 1950–80 (Reitan 1979; Zishka and Smith 1980; Parker et al. 1989; Changnon et al. 1995). Our results show little trend in cyclone counts during that period in our Atlantic region, which covers North America, and mixed trends over the continent in the full grid analyses for the various indices, with a tendency for increases over Canada and decreases over the United States. The EOF analysis of Chang and Fu (2002) shows evidence for an increase in NH storm track intensity based on the filtered 300-hPa meridional velocity variance, supporting the increases over the storm tracks seen in our temperature variance index. Several studies have reported increases in the last few decades in waves, wind, or pressure-based indices over northwestern Europe and the northeastern Atlantic (Schinke 1993; Alexandersson et al. 1998; Grevemeyer et al. 2000), although other results are mixed or show no trend (Schmidt and von Storch 1993; Schmith et al. 1998). Our results show significant increases in the wind index (both the $\sigma_{0.995}$ and geostrophic versions) over much of the area, with positive but insignificant trends for other indices.

Sinclair and Watterson (1999) argue that positive trends in intense cyclones seen in some GCM experiments are an artifact of a decrease in global sea level pressure over time. However, in the reanalysis, the hemisphere-wide $(20°–70°N)$ average sea level pressure declines by only 0.56 hPa over 50 years, which would not seem to markedly impact our count of intense cyclones. Sinclair and Watterson (1999) suggest that cyclone centers be located based on maxima in vorticity rather than minima in pressure to avoid artifacts caused by changing sea level pressure climatology. Accordingly, we calculated new cyclone count indices in which a grid point is required to be the local maximum of vorticity rather than the local minimum of pressure and in which intense cyclones are required to have vorticity exceeding $5 \times 10^{-5}$ $s^{-1}$ rather than MSLP of 980 hPa or less. The climatological mean fields for the vorticity-based count indices are very similar to those for the pressure-based indices, as are the trend plots. Further there is no evidence for an offset between the vorticity- and pressure-based indices, as implied by Sinclair and Watterson (1999). The low-frequency behavior in the regional averages for the vorticity-based indices are similar to, but weaker in magnitude in some regions than, that for the pressure-based indices. These results suggest that the increases in intense cyclones reported here are not a spurious effect of changes in sea level pressure.

Easterling et al. (2000) have suggested comparing GCM predictions with changes in the observational record as a way of attributing empirical changes to anthropogenic influence. Our results show some historical changes consistent with differences between typical GCM control and enhanced $CO_2$ scenarios. The regional averages show increases in intense cyclones with little change in total cyclones, consistent with predictions from some GCM studies (Carnell and Senior 1998; Knippertz et al. 2000). The increases seen in extreme winds are consistent with changes seen in intense cyclones, the events that support high wind speeds. Model predictions of a decrease in surface MTG match the long-term observed trends in Gitelman et al. (1997). We find some decline in MTG, at least since 1970, in most regions except the Atlantic. Full grid results for Eady growth rate and temperature variance indicate a possible downstream shift or extension in the Atlantic storm track. Such shifts are consistent with predictions in Hall et al. (1994), Carnell et al. (1996), and Knippertz et al. (2000). We see little evidence for a poleward shift in storm activity, suggested by some GCMs and theoretical models as a result of high-latitude warming (Hall et al. 1994; Carnell et al. 1996), and found empirically in the

reanalysis for hemisphere-wide latitudinal bands (Key and Chan 1999; McCabe et al. 2001); in fact, the wind index provides evidence to the contrary. Given the modest current levels of $CO_2$ and warming relative to GCM-enhanced $CO_2$ scenarios (Carnell et al. 1996), the similarities we do find between our results and the GCM predictions are more pronounced than might be expected.

The results leave open the possibility that trends in cyclone activity are a consequence of increased data availability and correspondingly less model-based influence over time. The more data are collected, the more likely are extreme values in the temperature, wind, and pressure fields, possibly causing spurious trends. This problem is not limited to reanalysis data. The Waves and Storms in the North Atlantic (WASA) project geostrophic wind analysis of Alexandersson et al. (1998) is one source of reasonably homogeneous data. For northwestern Europe and the northeastern Atlantic, there is no evidence for an upward bias in the $\sigma_{0.995}$-based wind index compared to the WASA-based index (see Fig. A1). Other studies report general agreement between reanalysis results and more homogeneous data sources. Graham and Diaz (2001) find that the limited observational data support the intensification seen in the North Pacific based on the reanalysis. Radiosonde observations support reanalysis results of an increase in storm track intensity based on the meridional velocity variance, but there is evidence that this index is biased low before the 1970s in the reanalysis (Chang and Fu 2002). Iskenderian and Rosen (2000) compare temperature variance at 500 hPa calculated from the Oort and Liu (1993) radiosonde dataset with the index calculated from the NCEP–NCAR reanalysis for 1958–96. There are large areas of agreement over the United States and the North Atlantic. However, over Asia there are major differences with mainly negative trends for all seasons from the Oort dataset and a mixture of positive and negative trends of lesser absolute value for the reanalysis. This provides some support for the Atlantic storm track trends reported here but also raises some doubts for trends over Asia.

The societal impacts of storm activity drive much of the interest in measuring cyclone activity. The indices investigated here behave similarly in several respects, especially over the storm tracks, but they do not fully mimic each other. Even the cyclone and intense cyclone count indices show important differences, best illustrated in the contrasting low-frequency behavior of their regional (including hemisphere wide) averages. These differences amongst indices are not unexpected, given their focus on different aspects of cyclone activity, but do cause difficulty in choosing indices that are best suited for estimating changes in impacts over time. Despite their intuitive match to the idea of cyclone activity, the count indices do have drawbacks. First, they do not lend themselves to robust statistical analysis because they are discrete indices with small counts in many locations,

particularly the intense cyclone counts over land. Second, their definition involves several arbitrary choices. The noncount indices suffer less from these shortcomings. The correspondence between extreme winds, one of our measures of impacts, and both temperature variance and Eady growth rate suggests that these latter two indices may be useful measures of cyclone activity, despite being measured at 500 hPa rather than at the surface. Their fields are smoother, and they raise fewer concerns about data quality, than the wind index. The peak values for all the indices, but particularly the count indices, are concentrated over the storm tracks, as are the highest correlations between the indices. As most impacts occur over land, future work on cyclone activity may benefit from consideration of indices that more directly measure impacts or activity there (e.g., Hirsch et al. 2001).

Use of the false discovery rate procedure permits more spatially focused statistical analysis than the standard field significance approach of Livezey and Chen (1983), but the approach has not been widely used in the climatological literature. It remains an open question which version of FDR control should be employed, as different assumptions about the relationships of the test statistics result in different levels of conservativeness in testing. There has been increasing attention to the topic in the statistical literature, and we expect that future research will provide more guidance.

## 6. Conclusions

In this study, NCEP–NCAR reanalysis data for the winters 1949–99 are used to calculate six cyclone indices at the 3024 grid points in the extratropical NH from 20° to 70°N and as broad regional averages. The key findings are as follows.

- Cyclone indices are reasonably well correlated over the North Atlantic and North Pacific storm tracks, but often weakly correlated elsewhere.
- Regional averages for large sectors of the hemisphere provide some evidence for increases in storm activity and forcing, but results vary by region and decade. The number of cyclones does not appear to be increasing, but there is evidence for an increase in intense cyclones.
- Location-specific results reveal details hidden in the regional averages; significant increases in Eady growth, temperature variance, and extreme wind speeds are most noted in the storm track areas. Increases in the number of cyclones and intense cyclones occur in the same areas but are not significant based on the false discovery rate approach.
- Control of the false discovery rate is a new technique in the statistical literature that offers higher power to detect significant effects than previous approaches and allows one to determine significance at individual locations.

- Positive trends in cyclone activity in the storm tracks are reasonably consistent across indices based on the reanalysis, but trends caused by data inhomogeneity cannot be dismissed despite our efforts to compare results with more homogeneous sources.

## APPENDIX

### Wind Index Comparison

We compare trends at a set of locations in northwestern Europe and the northeastern Atlantic Ocean using three variants of the wind index, each based on a separate wind field: 1) the WASA-based geostrophic winds (WASA Group 1998), 2) the NCEP–NCAR-based $\sigma_{0.995}$ winds, and 3) NCEP–NCAR-based geostrophic winds. If the reanalysis trends are positively biased because of data inhomogeneity (WASA Group 1998), this should be reflected in more positive slopes and smaller $p$ values for the NCEP–NCAR-based $\sigma_{0.995}$ wind index than the WASA-based index. The WASA-based geostrophic wind field is calculated at the centers of triangles of WASA stations as in Alexandersson et al. (1998). For many stations, WASA data are available for the winters 1949–95, although for some the period was restricted to the winters 1959–95. We selected 11
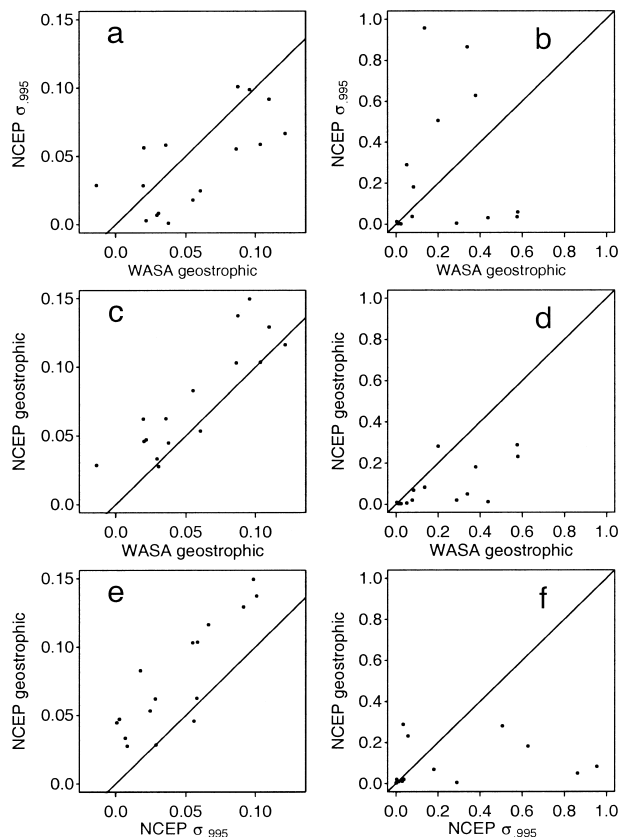


FIG. A1. (left) Slopes from the regression of the extreme wind index against time based on one wind field plotted against slopes based on a different wind field for the 16 locations (marked in Fig. 1b) in northwestern Europe and the northeastern Atlantic: (a) WASA geostrophic winds and NCEP–NCAR $\sigma_{0.995}$ winds, (c) WASA geostrophic winds and NCEP–NCAR geostrophic winds, and (e) NCEP–NCAR $\sigma_{0.995}$ and NCEP–NCAR geostrophic winds. (right) Scatterplots for the $p$ values associated with (a), (c), and (e) [(b), (d), and (f), respectively]. In each subplot, the solid line is the 1:1 line.

NCEP–NCAR grid locations at the approximate center of WASA triangles and for which WASA interstation distances are approximately the same as distances between NCEP–NCAR grid points. These 11 locations are restricted to southern Scandinavia and the Baltic Sea, so we also relaxed the interstation distance restriction and selected a second set of 16 locations (Fig. 1b) for comparison of the indices.

Using both the group of 11 and the group of 16 locations, we calculate pairwise correlations between the three variants at each location based on the replications across winters, finding that average correlations between the two NCEP–NCAR-based indices are higher than between the WASA index and either NCEP–NCAR index, but in all cases the variants of the wind index are well correlated (Table A1). The high correlations seen here between the two reanalysis variants are mirrored by high correlations in all areas of the NH (generally greater than 0.8 over the oceans and 0.6 over land) except the most southerly latitudes and in areas of high

TABLE A1. Average correlations across locations in northwestern Europe and the northeastern Atlantic between variants of the wind index using different data sources. Eleven locations were chosen so that WASA interstation distances were similar to distances between NCEP–NCAR grid points. This restriction was relaxed for the set of 16 locations. Standard deviations of the correlations are indicated in parentheses.

| | 11 locations | 16 locations |
|---|---|---|
| NCEP-geostrophic, NCEP–$\sigma_{0.995}$ | 0.90 (0.03) | 0.92 (0.03) |
| NCEP-geostrophic, WASA-geostrophic | 0.82 (0.04) | 0.83 (0.07) |
| NCEP–$\sigma_{0.995}$, WASA-geostrophic | 0.75 (0.06) | 0.79 (0.08) |

topography (not shown). We next compare the location-specific trends and their significance levels among variants. For the group of 16 locations, Fig. A1 shows scatterplots of both the slopes and $p$ values for these slopes for one variant plotted against those at the same locations for a second variant. The slopes from the NCEP–NCAR geostrophic index tend to be larger than those from either other variant, and the $p$ values are smaller. The WASA-based index tends to have larger slopes and smaller $p$ values than the NCEP–NCAR $\sigma_{0.995}$-based index. The results for the slopes and $p$ values for the group of 11 locations (not shown) are similar to those for the 16 locations, but the slopes are less tightly grouped around the $x = y$ line. For this region, the results do not suggest that data inhomogeneity causes systematic upward biases in the $\sigma_{0.995}$ wind index, but there do appear to be unexplained systematic differences between the wind index based on the NCEP–NCAR geostrophic wind field and the two other fields.

## REFERENCES

Alexandersson, H., T. Schmith, K. Iden, and H. Tuomenvirta, 1998: Long-term variations of the storm climate over NW Europe. *Global Atmos. Ocean Syst.,* **6,** 97–120.

Benjamini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.,* **57B,** 289–300.

——, and D. Yekutieli, 2001: The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.,* **29,** 1165–1188.

Cardone, V., J. Greenwood, and M. Cane, 1990: On trends in historical marine wind data. *J. Climate,* **3,** 113–127.

Carnell, R., and C. Senior, 1998: Changes in mid-latitude variability due to increasing greenhouse gases and sulphate aerosols. *Climate Dyn.,* **14,** 369–383.

——, ——, and J. Mitchell, 1996: An assessment of measures of storminess: Simulated changes in Northern Hemisphere winter due to increasing $CO_2$. *Climate Dyn.,* **12,** 467–476.

Chang, E., and I. Orlanski, 1993: On the dynamics of a storm track. *J. Atmos. Sci.,* **50,** 999–1015.

——, and Y. Fu, 2002: Interdecadal variations in Northern Hemisphere winter storm track intensity. *J. Climate,* **15,** 642–658.

Changnon, D., J. Noel, and L. Maze, 1995: Determining cyclone frequencies using equal-area circles. *Mon. Wea. Rev.,* **123,** 2285–2294.

Conover, W., 1999: *Practical Nonparametric Statistics.* Wiley, 584 pp.

DiMatteo, I., C. Genovese, and R. Kass, 2001: Bayesian curve fitting with free-knot splines. *Biometrika,* **88,** 1055–1071.

Duchon, C., 1979: Lanczos filtering in one and two dimensions. *J. Appl. Meteor.,* **18,** 1016–1022.

Easterling, D., G. Meehl, C. Parmesan, S. Changnon, T. Karl, and L. Mearns, 2000: Climate extremes: Observations, modeling, and impacts. *Science,* **289,** 2068–2074.

Geng, Q., and M. Sugi, 2001: Variability of the North Atlantic cyclone activity in winter analyzed from NCEP–NCAR reanalysis data. *J. Climate,* **14,** 3863–3873.

Gitelman, A., J. Risbey, R. Kass, and R. Rosen, 1997: Trends in the surface meridional temperature gradient. *Geophys. Res. Lett.,* **24,** 1243–1246.

Graham, N., and H. Diaz, 2001: Evidence for intensification of North Pacific winter cyclones since 1948. *Bull. Amer. Meteor. Soc.,* **82,** 1869–1893.

Grevemeyer, I., R. Herber, and H.-H. Essen, 2000: Microseismological evidence for a changing wave climate in the northeast Atlantic Ocean. *Nature,* **408,** 349–352.

Gulev, S., 1999: Comparison of COADS release 1a winds with instrumental measurements in the northwest Atlantic. *J. Atmos. Oceanic Technol.,* **16,** 133–145.

——, O. Zolina, and S. Grigoriev, 2001: Extratropical cyclone variability in the Northern Hemisphere winter from the NCEP/NCAR reanalysis data. *Climate Dyn.,* **17,** 795–809.

Hall, N., B. Hoskins, P. Valdes, and C. Senior, 1994: Storm tracks in a high-resolution GCM with doubled carbon dioxide. *Quart. J. Roy. Meteor. Soc.,* **120,** 1209–1230.

Hastie, T., and R. Tibshirani, 1990: *Generalized Additive Models.* Chapman and Hall, 335 pp.

Hayden, B., 1981: Cyclone occurrence mapping: Equal area or raw frequencies? *Mon. Wea. Rev.,* **109,** 168–172.

Held, I., 1993: Large-scale dynamics and global warming. *Bull. Amer. Meteor. Soc.,* **74,** 228–241.

Hirsch, M., A. DeGaetano, and S. Colucci, 2001: An East Coast winter storm climatology. *J. Climate,* **14,** 882–899.

Hoskins, B., and P. Valdes, 1990: On the existence of storm tracks. *J. Atmos. Sci.,* **47,** 1854–1864.

Houghton, J. T., Y. Ding, D. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. Johnson, Eds., 2001: *Climate Change 2001: The Scientific Basis.* Cambridge University Press, 881 pp.

Hulme, M., and P. Jones, 1991: Temperatures and windiness over the United Kingdom during the winters of 1988/89 and 1989/90 compared with previous years. *Weather,* **45,** 126–135.

Hurrell, J., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperature and precipitation. *Science,* **269,** 676–679.

Iskenderian, H., and R. Rosen, 2000: Low-frequency signals in mid-tropospheric submonthly temperature variance. *J. Climate,* **13,** 2323–2333.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Key, J., and A. Chan, 1999: Multidecadal global and regional trends in 1000 mb and 500 mb cyclone frequencies. *Geophys. Res. Lett.,* **26,** 2053–2056.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Knippertz, P., U. Ulbrich, and P. Speth, 2000: Changing cyclones and surface wind speeds over the North Atlantic and Europe in a transient GHG experiment. *Climate Res.,* **15,** 109–122.

König, W., R. Sausen, and F. Sielmann, 1993: Objective identification of cyclones in GCM simulations. *J. Climate,* **6,** 2217–2231.

Lambert, S., 1996: Intense extratropical Northern Hemisphere winter cyclone events: 1899–1991. *J. Geophys. Res.,* **101D,** 21 319–21 325.

Lee, W.-J., and M. Mak, 1994: Observed variability in the large-scale static stability. *J. Atmos. Sci.,* **51,** 2137–2144.

Livezey, R., and W. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.,* **111,** 46–59.

McCabe, G., M. Clark, and M. Serreze, 2001: Trends in Northern Hemisphere surface cyclone frequency and intensity. *J. Climate,* **14,** 2763–2768.

McCarthy, J., O. Canziani, N. Leary, D. Dokken, and K. White, Eds., 2001: *Climate Change 2001: Impacts, Adaptation, and Vulnerability.* Cambridge University Press, 1032 pp.

McCullagh, P., and J. Nelder, 1989: *Generalized Linear Models.* 2d ed. Chapman and Hall, 511 pp.

Oort, A., and H. Liu, 1993: Upper-air temperature trends over the globe, 1958–1989. *J. Climate,* **6,** 292–307.

Parker, S., J. Hawes, S. Colucci, and B. Hayden, 1989: Climatology of 500 mb cyclones and anticyclones, 1950–1985. *Mon. Wea. Rev.,* **117,** 558–570.

Reitan, C., 1979: Trends in the frequencies of cyclone activity over North America. *Mon. Wea. Rev.,* **107,** 1684–1688.

Schinke, H., 1993: On the occurrence of deep cyclones over Europe and the North Atlantic in the period 1930–1991. *Beitr. Phys. Atmos.,* **66,** 223–237.

Schmidt, H., and H. von Storch, 1993: German Bight storms analysed. *Nature,* **365,** 791.

Schmith, T, E. Kaas, and T.-S. Li, 1998: Northeast Atlantic winter storminess 1875–1995 reanalysed. *Climate Dyn.,* **14,** 529–536.

Simmonds, I., and K. Keay, 2000: Variability of Southern Hemisphere extratropical cyclone behavior, 1958–1997. *J. Climate,* **13,** 550–561.

Sinclair, M., 1994: An objective cyclone climatology for the Southern Hemisphere. *Mon. Wea. Rev.,* **122,** 2239–2256.

——, and I. Watterson, 1999: Objective assessment of extratropical weather systems in simulated climates. *J. Climate,* **12,** 3467–3485.

Smith, S., D. Legler, and K. Verzone, 2001: Quantifying uncertainties in NCEP reanalyses using high-quality research vessel observations. *J. Climate,* **14,** 4062–4072.

Stephenson, D., and I. Held, 1993: GCM response of northern winter stationary waves and storm tracks to increasing amounts of carbon dioxide. *J. Climate,* **6,** 1859–1870.

Taylor, K., 1986: An analysis of the biases in traditional cyclone frequency maps. *Mon. Wea. Rev.,* **114,** 1481–1490.

Trenberth, K., 1984: Signal versus noise in the Southern Oscillation. *Mon. Wea. Rev.,* **112,** 326–332.

——, and J. Hurrell, 1994: Decadal atmospheric–ocean variations in the Pacific. *Climate Dyn.,* **9,** 303–319.

Venables, W., and B. Ripley, 1997: *Modern Applied Statistics with S-Plus.* 2d ed. Springer-Verlag, 548 pp.

Wallace, J., G.-H. Lim, and M. Blackmon, 1988: Relationship between cyclone tracks, anticyclone tracks and baroclinic waveguides. *J. Atmos. Sci.,* **45,** 439–462.

Ward, M., 1992: Provisionally corrected surface wind data, worldwide ocean–atmosphere surface fields, and Sahelian rainfall variability. *J. Climate,* **5,** 454–475.

WASA Group, 1998: Changing waves and storms in the northeast Atlantic? *Bull. Amer. Meteor. Soc.,* **79,** 741–760.

Wright, P., 1988: On the reality of climate changes in wind over the Pacific. *J. Climatol.,* **8,** 521–527.

Yekutieli, D., and Y. Benjamini, 1999: Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *J. Stat. Plann. Inference,* **82,** 171–196.

Zhang, K., B. Douglas, and S. Leatherman, 1997: East coast storm surges provide unique climate record. *Eos, Trans. Amer. Geophys. Union,* **78,** 389–400.

Zhang, Y., and W.-C. Wang, 1997: Model-simulated northern winter cyclone and anticyclone activity under a greenhouse warming scenario. *J. Climate,* **10,** 1616–1634.

Zishka, K., and P. Smith, 1980: The climatology of cyclones and anticyclones over North America and surrounding ocean environs for January and July, 1950–1977. *Mon. Wea. Rev.,* **108,** 387–401.