Improving Extended-Range Seasonal Predictions of Intense Atlantic Hurricane Activity

J. B. ELSNER

Department of Meteorology, The Florida State University, Tallahasee, Florida

C. P. SCHMERTMANN

Department of Economics, The Florida State University, Tallahassee, Florida

22 December 1992 and 31 March 1993

ABSTRACT

This study shows that hindcasts of seasonal numbers of intense Atlantic hurricanes made using a nonlinear statistical model are superior to those made by linear statistical models previously described in the literature. A fully cross-validated Poisson model achieves an increase of nearly 40% in hindcast skill when compared to a fully cross-validated linear model. Improvements are most evident for years with relatively large numbers of intense hurricanes. It is suggested that a significant improvement in forecast skill is possible with the Poisson model. A prediction for the 1993 season is made, and calls for two intense hurricanes to visit the Atlantic basin.

1. Introduction

The seasonal number of tropical storms and hurricanes over the Atlantic shows large interannual variability (Gray and Landsea 1992). Several factors have recently been associated with this variability, including stratospheric winds over the Atlantic and rainfall over western Africa. In general, as a result of strong stratospheric easterlies, increased vertical wind shear inhibits development of tropical storms and hurricanes (Gray 1984). Further, a drier than normal climate over western Africa is associated with weaker and less organized easterly waves originating in the region, and consequently, leads to reduced numbers of tropical storms and hurricanes (Landsea and Gray 1992). New evidence indicates that both the intensity of stratospheric easterlies and rainfall anomalies during a particular hurricane season are correlated with winds and rainfall anomalies during the previous season (Gray 1990; Gray et al. 1992; Gray and Landsea 1992). These relationships allow for long-term predictions of Atlantic tropical cyclone activity.

More specifically, Gray (1984, 1990) has shown that seasonal variability of Atlantic tropical cyclone activity can be explained to some extent by the phase of the stratospheric quasi-biennial oscillation (QBO), the presence or absence of a moderate to strong El Niño, the zonal component of the 200-mb wind, sea level pressure anomalies in the Caribbean basin, and the anticipated June–September rainfall in the western Sahel region of Africa. Of the many important variables influencing hurricane activity at the beginning of the Atlantic tropical cyclone season, Gray et al. (1992, hereafter G92) identifies two extended-range predictors available 6–11 months in advance. The two long-term predictors are 10-month forward-extrapolated QBO strength and Gulf of Guinea rainfall prior to 1 December of the previous year.

Physical explanations of the foregoing associations center on the strength of the easterly waves originating over western Africa and on the nature of the large-scale environment encountered by the waves as they move westward across the tropical Atlantic. For example, drier than normal conditions over the Sahel region of Africa tend to be associated with weaker and less organized easterly waves, resulting in fewer cyclones reaching tropical storm or hurricane strength over the Atlantic Basin. Extended-range predictions of this association are possible since Sahel rainfall during the Atlantic hurricane season is likely related to the strength of the western Africa monsoon (as indicated by rainfall in the Gulf of Guinea region) in the previous year through positive feedbacks of soil moisture and evapotranspiration (Gray and Landsea 1992).

The large-scale environment through which the easterly waves travel is also a factor. The greater the thickness of the stratospheric layer of westerly winds, the greater the amount of hurricane activity (Gray 1984). Fortunately, wind changes at 30 mb appear to be related to changing depths of the stratospheric westerlies and easterlies. When the stratospheric easterlies are considerably stronger than the westerlies, greater advection of convective elements from the storm center results, which restrains development and intensification. In addition, the absolute magnitude of vertical shear of zonal wind between 50 and 30 mb is relatively

Corresponding author address: James B. Elsner, Department of Meteorology, The Florida State University, B-161, Tallahassee, FL 32306.

large when easterlies are predominant, further contributing to a hostile stratospheric environment for deep cumulus and thus cyclone development. These explanations remain speculative.

Gray et al. (1992) have demonstrated skill with hindcasts of seasonal tropical cyclone activity in the Atlantic basin (including the Atlantic Ocean, Caribbean Sea, and Gulf of Mexico). Long-term (approximately 9-month lead time) hindcasts are made using linear regression models. Skill is derived from extrapolation of a relatively stable oscillation in tropical stratospheric zonal wind patterns (QBO) and from seasonally lagged regional correlations of rainfall anomalies over portions of western Africa. Using these factors, G92 estimates statistical forecast models for seasonal numbers of named storms, hurricanes, and intense hurricanes, as well as for seasonal numbers of named storm days, hurricane days, and intense hurricane days, and for hurricane destructive potential.

The present work addresses the question of whether improvements can be made in extended-range forecasts of seasonal numbers of intense hurricanes. The paper is divided as follows. In section 2 the data and model used are described. A comparison of forecast model hindcast skill is given in section 3, and a forecast for the 1993 hurricane season is presented in section 4. Conclusions are outlined in section 5.

2. Methodology

To perform the strictest possible comparison between the proposed forecast model and the forecast model of G92, the same data and the same method of evaluating skill used by G92 are employed here. The predictor variables are taken from Tables 1 and 5 of G92 and consist of 1) previous year August–November Gulf of Guinea precipitation anomalies (Rg), 2) previous year August–September western Sahel precipitation anomalies (Rs), 3) November–September 10-month extrapolated QBO zonal winds at 50 mb (U50), 4) November–September extrapolated QBO zonal winds at 30 mb (U30), and 5) absolute difference between 50- and 30-mb zonal winds (Ud). The predicted variable is the seasonal number of *intense hurricanes* (I), taken from Table 1 of G92. The intense hurricane season runs from August through October (Gray and Landsea 1992).

Hurricane strength is measured by wind and storm surge intensity on a scale from 1 to 5 (weakest to strongest)—called the Saffir–Simpson (S–S) scale (Simpson 1974). Intense hurricanes are defined as those reaching a 3, 4, or 5 on the S–S scale, with a category 3 storm having maximum sustained winds exceeding 50 m s^{-1} and a category 5 storm having maximum sustained winds in excess of 69 m s^{-1}. These powerful storms account for three-quarters of all U.S. tropical cyclone damage (Landsea and Gray 1992), despite an average of less than one making U.S. landfall per year. Both the predictor and predicted variables span the 41-year period from 1950 to 1990 inclusive. The complete dataset for intense Atlantic hurricanes used in this study is reproduced in Table 1.

The number of intense Atlantic hurricanes over the past four decades is shown in Fig. 1. Two characteristics are clear—large interannual variability and a trend toward fewer intense hurricanes in the most recent decades. Of particular interest, however, is the relatively low number of intense hurricanes in any year. Indeed, there are 14 years in which less than two intense hur-

TABLE 1. Data used in the model comparison study from G92. The columns show 1) seasonal numbers of intense Atlantic hurricanes (I), 2) precipitation anomalies expressed as a standard deviation for previous year August–November Gulf of Guinea (Rg), 3) August–September precipitation anomalies for previous year in the western Sahel (Rs), 4) 10-month extrapolated QBO zonal winds in meter per second for September at 50 mb (U50), 5) at 30 mb (U30), and 6) magnitude of vertical shear (|U50 − U30|) of extrapolated zonal winds (Ud) in meters per second. See Gray et al. (1992) for details.

Year	I 1	Rg 2	Rs 3	U50 4	U30 5	Ud 6
1950	8	1.07	−0.14	−3	−3	0
1951	5	−0.66	1.68	−4	−13	9
1952	3	0.65	0.49	−23	−26	3
1953	4	0.41	0.93	0	−18	18
1954	2	−0.16	0.20	−23	−32	9
1955	6	0.64	0.60	0	−4	4
1956	2	0.41	1.00	−19	−33	14
1957	2	−0.36	0.47	−2	−3	1
1958	5	1.03	0.58	−12	−28	16
1959	2	−0.74	1.45	−9	−5	4
1960	2	0.12	0.25	−6	−21	15
1961	7	1.05	0.23	−3	−3	0
1962	1	−0.74	0.48	−12	−32	20
1963	2	0.73	0.28	−17	−3	14
1964	6	1.18	−0.12	−4	−18	14
1965	1	−0.68	0.59	−23	−32	9
1966	3	−0.17	0.75	−9	−2	7
1967	1	−0.14	0.34	−8	−25	17
1968	0	−0.51	0.72	−23	−14	9
1969	5	1.28	−0.82	0	−9	9
1970	2	−0.31	0.38	−12	−30	18
1971	1	−0.23	−0.45	−3	−2	1
1972	0	−0.40	−0.19	−8	−31	23
1973	1	−0.88	−1.10	−5	−5	0
1974	2	0.43	−0.72	−14	−32	18
1975	3	−0.08	−0.04	−5	−2	3
1976	2	−0.55	0.06	−6	−21	15
1977	1	−0.59	−0.50	−21	−24	3
1978	2	−0.50	−0.75	−2	−9	7
1979	2	−0.73	−0.36	−19	−31	12
1980	2	0.55	−0.92	−5	−1	4
1981	3	0.36	−0.34	−8	−30	22
1982	1	−0.93	−0.44	−21	−5	16
1983	1	−0.61	−0.90	−2	−13	11
1984	1	−1.32	−1.24	−23	−14	9
1985	3	0.04	−1.23	−2	−9	7
1986	0	0.13	−0.51	−12	−32	20
1987	1	−0.48	−0.01	−21	−9	12
1988	3	1.37	−0.63	−3	−13	10
1989	2	0.35	0.29	−23	−32	9
1990	1	0.19	0.10	−3	−1	2

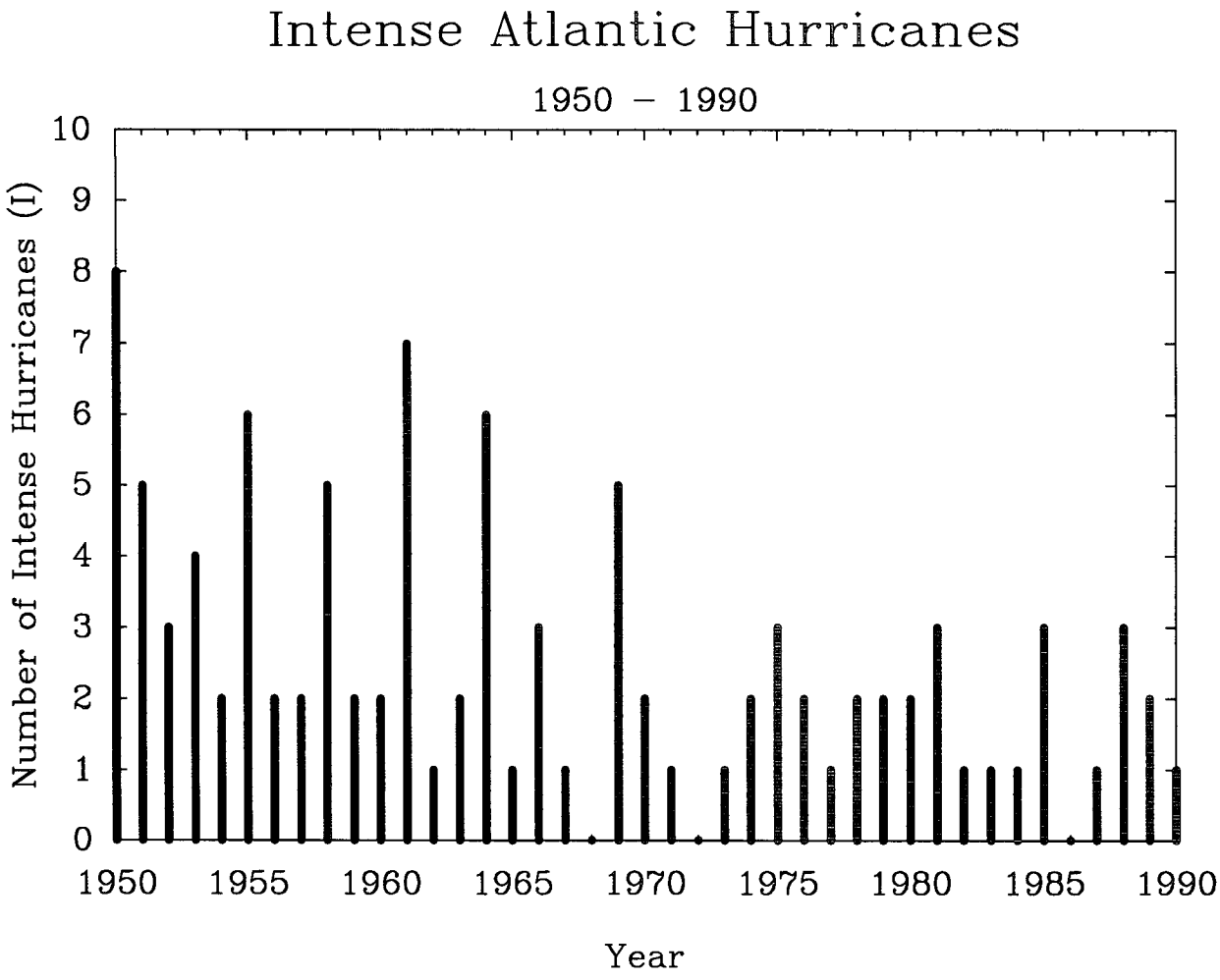

FIG. 1. Seasonal number of intense Atlantic hurricanes over the period 1950–1990.

ricanes were reported, and only two years (1950 and 1961) in which seven or more storms reached the intense hurricane category. A Poisson distribution, which restricts possible outcomes to the nonnegative integers, is thus an obvious choice for modeling the seasonal number of intense hurricanes. It should be noted that for the other Atlantic tropical storm variables considered in G92 the choice of a Poisson model is not as apparent, and the variables are not considered here.

In a Poisson model the probability of exactly n intense hurricanes is

$$\text{Prob}(\mathrm{I} = n \mid \lambda) = \exp(-\lambda)\lambda^n/n! \quad \text{for} \quad n = 0, 1, 2, \cdots, \quad (1)$$

where $\lambda > 0$ is the distributional parameter. A Poisson regression allows the distribution of I to vary with the vector $\mathbf{x} = (1, \mathrm{U50}, \mathrm{U30}, \mathrm{Ud}, \mathrm{Rs}, \mathrm{Rg})'$ by assuming λ is some positive function of $\mathbf{x}$. This is modeled as

$$\lambda(\mathbf{x}) = \exp(\gamma'\mathbf{x}), \quad (2)$$

where $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_5)'$ is the unknown parameter vector to be estimated.

Given γ and $\mathbf{x}$, the expected number of intense hurricanes in this model, $E(\mathrm{I} \mid \mathbf{x})$, is

$$E(\mathrm{I} \mid \mathbf{x}) = \exp[\gamma_0 + \gamma_1\mathrm{U50} + \gamma_2\mathrm{U30} + \gamma_3\mathrm{Ud} + \gamma_4\mathrm{Rs} + \gamma_5\mathrm{Rg}]. \quad (3)$$

Note that in Eq. (3) the expected number of intense hurricanes is a nonlinear function of the explanatory variables and that the expected value is nonnegative.

A maximum likelihood procedure is used to estimate γ from Eqs. (1) and (2). Specifically, for a given γ, first λ is calculated for each year's $\mathbf{x}$ from Eq. (2) and then the likelihood (probability) of the observed number of intense hurricanes is estimated from Eq. (1). The γ used for the forecast (or hindcast) is the one that maximizes the product of the probabilities in Eq. (1) over all years. A forecast (or hindcast) of the number of intense hurricanes for a season is made using

the selected γ in Eq. (3). This provides an estimate of I's probability distribution, including not only the mean in Eq. (3) but also estimated probabilities for each possible number of intense hurricanes ($n = 0, 1, 2, \cdots$) from Eq. (1). A more detailed examination of a particular forecast is therefore possible.

3. Comparisons of hindcast skill

Hindcast skill is assessed by computing the agreement coefficient (ρ) between the observed I and the cross-validated predictions of I. The agreement coefficient is given by

$$\rho = 1 - \delta_0/\mu_\delta, \tag{4}$$

where δ_0 is

$$\delta_0 = 1/n \sum_i |\mathrm{I}_i - I_i| \tag{5}$$

and μ_δ is the expected value of δ under the null hypothesis of equally likely permutations of predicted values (Mielke 1991). Instead of computing μ_δ by taking all $n!$ permutations, it is calculated directly using

$$\mu_\delta = 1/n \sum_i [1/n \sum_j |\mathrm{I}_i - I_j|]. \tag{6}$$

As indicated by G92, it is essential to cross validate a model in order to assess its ability to predict future events. Cross validation in this context means predicting, for each of the 41 years in the sample, each year's intense hurricane number (I) from the *other* 40 years' data on I and **x**. For the Poisson model, this requires that the calculation of γ in Eq. (2) simply not include own-year observations. The result of doing this is that the maximum likelihood estimate of γ is independent of the year for which predictions are required.

Actual (I_i) versus cross-validated predicted (I_i) scatterplots of the G92 linear model and the present Poisson model are shown in Fig. 2 along with the corresponding agreement coefficients. The scatter is considerably closer to the I = I line for the present model compared to the G92 model, and the agreement coefficient of 0.530 is 38% larger, indicating improved hindcast skill with the current model. This improvement in hindcast skill is for seasonal numbers of intense hurricanes only and does not extend to the other variables considered in G92.

It should be noted here that the scatterplot and agreement coefficient for the G92 linear model shown in Fig. 2a are not as they appear in Fig. 11e of G92. The reason is that a strict comparison of the present model with the linear model requires that both models be fully cross validated. To understand this point it is necessary to examine the G92 model estimation procedure in more detail. Following the notation in G92 [Eqs. (1)–(3)], and using the identifying restrictions

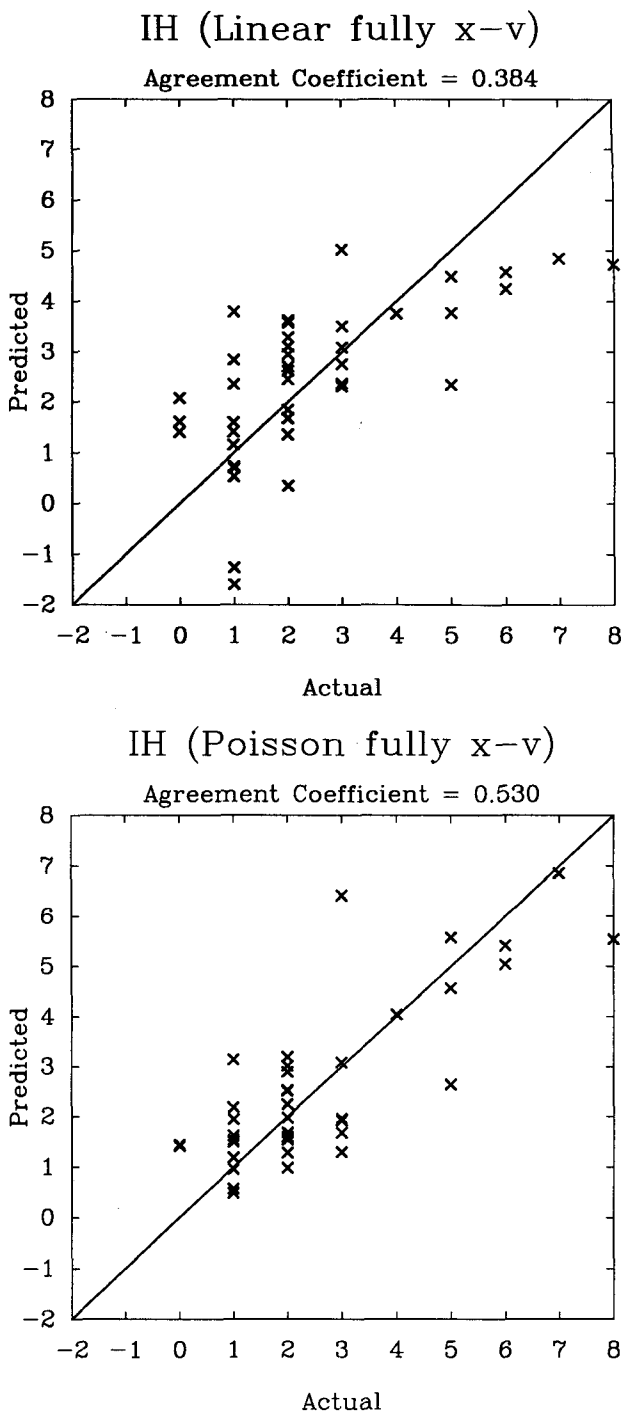

FIG. 2. Scatterplots of actual seasonal number of intense Atlantic hurricanes for the period 1950–1990 versus fully cross-validated hindcasts using (a) a linear model and (b) a Poisson model.

implicit in their Table 8 ($a_1 = a_4 = 1$), the G92 model is

$$E(\mathrm{I}) = \beta_0 + \beta_1[\mathrm{U50} + a_2\mathrm{U30} + a_3\mathrm{Ud}] + \beta_2[\mathrm{Rs} + a_5\mathrm{Rg}]. \tag{7}$$

The G92 procedure cross validates hindcasts of I by estimating $\beta = (\beta_0, \beta_1, \beta_2)'$ separately for each year via least absolute deviation (LAD) regression, based on data from the other 40 years. However, jackknifed estimates are computed *only after setting* $a = (a_2, a_3, a_5)$ *at the values that maximize the sample-wide (41 years) agreement coefficient for jackknifed* β *estimates.* At best, then, only three of six coefficient estimates in the G92 model are "jackknifed"; the other three are estimated from information that includes own-year data. Once own-year information is incorporated into the model, any jackknifed procedure will not represent a true cross-validated test of model skill. Consequently, it can be said that the predictions for the linear model in G92 are not fully cross validated.

Due to the incorporation of own-year information in jackknifed estimates, the hindcast skill of the linear model is overstated in G92. To demonstrate this, the G92 model [their Eqs. (1)–(3)] is recalculated, substituting the identifying restrictions ($\beta_1 = \beta_2 = 1$) for their ($a_1 = a_4 = 1$). This equally valid pair of restrictions yields a simpler version of the G92 model,

$$E(\mathrm{I}) = \beta_0 + a_1\mathrm{U50} + a_2\mathrm{U30} + a_3\mathrm{Ud} + a_4\mathrm{Rs} + a_5\mathrm{Rg}, \quad (8)$$

which incorporates exactly the same linear effects and should, therefore, yield the same hindcast skill. Fully cross-validated LAD predictions from Eq. (8) yield an agreement coefficient of only 0.384, compared to the 0.498 reported in G92. Purging own-year information from predictions therefore results in a considerable decrease in hindcast skill of the linear model.

Hindcast skill over the 41-year period is compared in Fig. 3 and Table 2. Overall, the Poisson model does better in nearly 75% of the hindcasts. Also evident is the fact that the Poisson model outperforms the linear model for years with relatively few intense hurricanes and years with a relatively large number of such storms. The improvement in skill over the linear model is

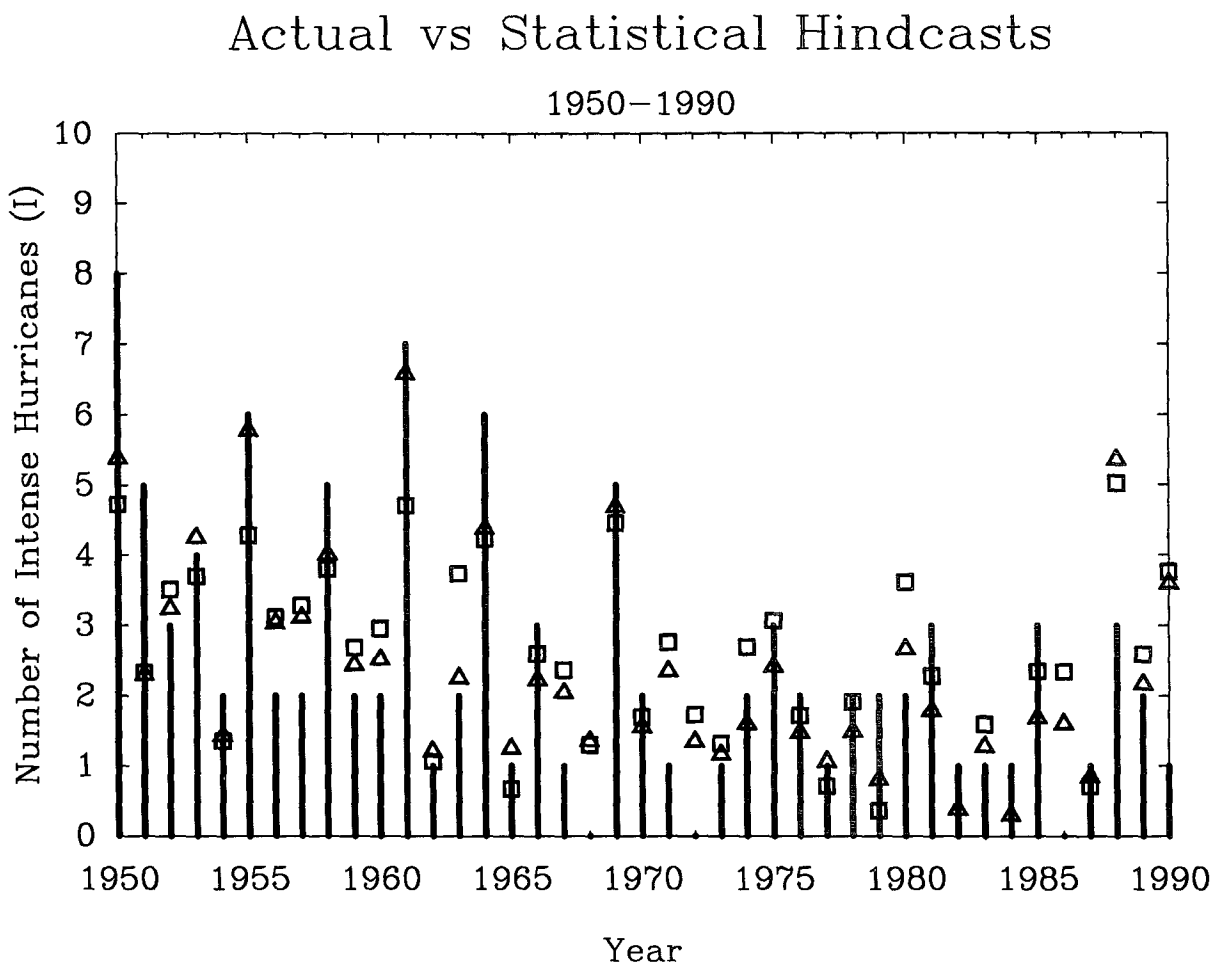

FIG. 3. Actual (impulses) versus fully cross-validated hindcasts of seasonal number of intense Atlantic hurricanes over the 41-year period 1950–1990. Poisson model hindcasts are the triangles, and linear model hindcasts are the squares.

TABLE 2. Comparison of fully cross-validated hindcast skill of the Poisson model with the linear model. Skill is compared as a function of observed number of intense hurricanes (I). Column 2 contains the number of years in which there were I number of intense hurricanes, columns 3 and 4 indicate how many of those years the Poisson model outperformed the linear model, and vice versa. The next two columns are the root-mean-squared errors as a function of number of intense hurricanes for the Poisson and linear models, respectively, and the last column is the difference in rmse.

I	No. of years with I intense hurricanes	No. of years $\lvert E(I) - I \rvert$ is smaller for		rmse		
		Poisson	Linear	Poisson	Linear	Linear–Poisson
0	3	2	1	1.439	1.841	0.402
1	11	10	1	0.990	1.506	0.517
2	13	10	3	0.677	1.039	0.361
3	6	1	5	1.274	0.955	−0.320
4	1	1	0	0.258	0.298	0.040
5	3	2	1	1.666	1.715	0.049
6	2	2	0	1.152	1.740	0.588
7	1	1	0	0.416	2.285	1.869
8	1	1	0	2.618	3.262	0.644
Total	41	30	11	1.113	1.444	0.331

greatest for the relatively few years with many intense storms (especially 1955 and 1961). However, improvements are also made with the Poisson model for the many years with only one or two intense hurricanes. These results are a consequence of the asymmetric distribution around the mean in the Poisson model, compared to the symmetric distribution implicit in LAD estimates of the linear model. Additionally, some improvement in hindcast skill with the nonlinear model can be attributed to the fact that the Poisson model is incapable of predicting a negative number. These limitations of the linear model are noted in G92.

Both the linear and Poisson model hindcasts are poor for 1950, 1951, and 1988. It should be noted that the five model predictors are related much more closely to the activity of easterly waves originating in western Africa than to tropical waves having extratropical origin. An attempt to separate seasonal numbers of intense hurricanes depending on their genesis region reveals that for 1950 only six of the eight intense storms were mentioned as having an African origin (Norton 1951) and that for 1951 one of the five intense hurricanes did not originate from an African wave (Norton 1952). This explains part of the poor performance during these years.

Also, recent work by Gray et al. (1993) indicates that in many cases Gulf of Guinea rainfall during the previous year is a good indicator of western Sahel rainfall during the current wet season. However, for some years like 1988, this direct relationship is not apparent. Indeed, for August–November of 1987, Gulf of Guinea rainfall was well above average but was followed by below-average western Sahel rainfall during June and July of 1988. This poor prediction of rainfall conditions over western Africa during the hurricane season of 1988 partly explains the overpredictions made by the models during this year.

4. Forecast for 1993

Relevant wind and rainfall data for the September–November period of 1992, kindly provided by C. Landsea (1992, personal communication), allow for a forecast of the 1993 tropical cyclone season. Here, coefficients are estimated over the 43-year (1950–1992) sample. Shown in Table 3 are predicted probabilities for numbers of intense hurricanes for the 1993 hurricane season. A forecast, representing a rounded mean of this distribution, calls for two intense hurricanes to develop over the Atlantic basin. Note that this forecast was made in December of 1992. This forecast compares to a forecast of three intense hurricanes made based on a linear model (Gray 1992).

A few notes are important here. First, the Gray (1992) linear model also predicts two intense hurricanes for the 1993 season, but his actual forecast is subjectively adjusted upward to three intense hurricanes to account for the possibility of cold-phase El Niño–Southern Oscillation conditions developing during 1993. In addition, these forecasts (both the Poisson and the linear) do not specify 1) when during the hurricane season an intense storm is most likely, and 2) whether one will strike the United States.

5. Conclusions

It is demonstrated that a better extended-range (approximately 9-month lead) forecast of seasonal numbers of intense Atlantic hurricanes is probable with a Poisson model. The conclusion is based on significant improvements in hindcast skill over a linear model

TABLE 3. Forecast (out of sample) probabilities of number of intense Atlantic hurricanes for the 1993 hurricane season from the Poisson model. The coefficients estimated from data in the period 1950–1992 inclusive are $\gamma_0 = 1.2262$, $\gamma_1 = 0.0362$, $\gamma_2 = -0.0108$, $\gamma_3 = -0.0301$, $\gamma_4 = 0.3238$, $\gamma_5 = 0.6140$.

I	1993 Probability (I)
0	0.168
1	0.299
2	0.267
3	0.159
4	0.071
5	0.025
6	0.008
Mean	1.786
Poisson model forecast	2
Gray's forecast	3

when both models are evaluated using a common dataset and an identical measure of skill. Skill is derived from extrapolation of a relatively stable oscillation in tropical stratospheric zonal winds and from correlations among rainfall anomalies over portions of western Africa.

A Poisson model is suggested based on the fact that the number of intense hurricanes in any particular season is a small positive integer. Skill over the linear model is most noticeable for years with very few intense hurricanes and for years with relatively many intense storms. This latter observation is especially important considering recent speculation that the overall trend toward fewer strong storms may be reversing and considering a projection of a future more favorable for the development of intense hurricanes (Gray 1990; Gray and Landsea 1992).

Acknowledgments. Thanks are extended to Jason Hess for assistance with an earlier nonlinear model, to Gregg Rohaly for help in locating some of the data used in forecasting, and to Renato Assuncao for advice on statistical software.

REFERENCES

Gray, W. M., 1984: Atlantic seasonal hurricane frequency. Part II: Forecasting its variability. *Mon. Wea. Rev.*, **112**, 1669–1683.

——, 1990: Strong association between west African rainfall and U.S. landfall of intense hurricanes. *Science*, **249**, 1251–1256.

——, 1992: LAD multiple linear regression forecasts of Atlantic tropical storm activity for 1993. *Exp. Long-Lead Forecast Bull.*, **1**, 12–14.

——, 1993: Predicting Atlantic basin seasonal tropical cyclone activity by 1 August. *Wea. Forecasting*, **8**, 73–86.

——, and C. W. Landsea, 1992: African rainfall as a precursor of hurricane-related destruction on the U.S. east coast. *Bull. Amer. Meteor. Soc.*, **73**, 1352–1364.

——, C. W. Landsea, P. W. Mielke, Jr., and K. J. Berry, 1992: Predicting Atlantic seasonal hurricane activity 6–11 months in advance. *Wea. Forecasting*, **7**, 440–455.

Landsea, C. W., and W. M. Gray, 1992: The strong association between western Sahel monsoon rainfall and intense Atlantic hurricanes. *J. Climate*, **5**, 435–453.

Mielke, P. W., 1991: The application of multivariate permutation methods based on distance functions in the earth sciences. *Earth-Sci. Rev.*, **31**, 55–71.

Norton, G., 1951: Hurricanes of the 1950 season. *Mon. Wea. Rev.*, **79**, 8–15.

——, 1952: Hurricanes of 1951. *Mon. Wea. Rev.*, **80**, 1–4.

Simpson, R. H., 1974: The hurricane disaster potential scale. *Weatherwise*, **27**, 169–186.