U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 301

PROCEDURE FOR THE STATISTICAL CORRECTION OF MEDIUM-RANGE SPECTRAL FORECASTS

EDWARD S. EPSTEIN
CLIMATE ANALYSIS CENTER

JANUARY 1985

PROCEDURE FOR THE STATISTICAL CORRECTION OF MEDIUM-RANGE SPECTRAL FORECASTS

by
Edward S. Epstein
CAC/NMC/NWS

1. Introduction.

It has long been recognized that statistical adjustments applied to the output of dynamically produced numerical progs can systematically reduce their errors. The success of MOS (Glahn and Lowry, 1972; Glahn, 1980) is partially due to this, although MOS has traditionally been applied to the specification of events not explicitly treated in the forecast model, and our interest here is in the height and temperature fields (or other predicted variables) that are specifically forecast by the model. Hughes (1982) has been using, successfully, a statistical correction procedure to improve 6-10 day forecasts of the height field. Our intent is to try to improve on Hughes' results by employing a somewhat more systematic developmental approach, and especially by applying the statistical corrections in the spectral domain of the model rather than in the spatial domain of the output.

The purpose of this note is to establish that a feasible procedure has been found and to justify and encourage more extensive testing and evaluation to demonstrate its operational value.

The likelihood that a statistical procedure applied in the spectral domain would be successful is strongly indicated by the work of Lorenz (1977). Of particular relevance to the present study was his finding that most of the variance of a particular spectral coefficient that could be explained by a linear combination of <u>all</u> the dynamically forecasted coefficients are accounted for by just the <u>two</u> real numbers that define the complex coefficient of the predictand wave number. If this proves valid in the context of the current problem (a medium range prediction based on a sophisticated multilevel spectral model (Sela, 1982), as opposed to a 24-hour barotropic forecast), then complex statistical screening can be avoided and very parsimonious statistical relations will be very effective. In fact, in terms of reduction of mean square error, the procedures we have developed on this basis appear to be very successful. No effort has been made to prove that they are optimal. Optimality in a statistical sense is, in any case, difficult to define. Certainly, in any developmental sample greater reductions in mean square error could be achieved by the addition of other predictors, but the question of the utility of and criteria for choosing such additional predictors has never been settled.

The predictands with which we shall deal are the global day-5 (120 hr) 500 mb and 1000 mb height forecasts expressed in terms of spectral coefficients. We assume at the outset that most of the predictibility remaining after 5 days resides in the longest waves. Our results will tend to substantiate and quantify this assumption. A 6x6 rhomboidal truncation includes 66 scalar coefficients (30 complex coefficients and 6 -- zonal wave no. 0 -- that are scalar). Each complex coefficient is defined by a scalar vector of dimension two, i.e. consists of two real scalars. We assume that the corrected forecast of each vector depends linearly on only the unmodified forecast of the two scalars that comprise that vector. In the case of zonal wave number zero, where each coefficient is a scalar, there is only one predictor; the raw forecast of the coefficient fully determines the corrected value.

The regression coefficients are determined by standard least squares techniques using the most recent m daily pairs of forecasts and verifying analyses (actually, the initialized fields for the verifying time). We have examined values of m in the range 20 to 60. On each day all coefficients are reevaluated, and corrected forecasts calculated. These are then subject to later verification and comparison with uncorrected forecasts. Note that all verifications are done on independent, i.e. subsequent, data. There is no *a priori* reason for achieving better verification results in this test than in subsequent applications of the procedure; that can certainly occur, but if so it will be a chance rather than a systematic result.

It was an early observation that the regression coefficients, especially for the lowest zonal wave numbers (0 and 1) varied rather slowly and systematically with the season of the year. The correction procedure was in large part attempting to introduce a correction for the seasonally varying climate drift of the model. This led us to investigate, within the constraints of a limited data set, the model's climate drift as it varies with season and from year to year. A companion note will be prepared to describe these results. It also led us to investigate a procedure of first correcting for the seasonally dependent model bias (climate drift) in each coefficient and then estimating additional corrections by regression involving recent forecasts and analyses. Results to be presented will indicate that this latter procedure results in somewht smaller rms prediction errors, but the advantage is relatively slight. Considering the disadvantages to a procedure that would have to rely on several years of accumulated experience with an unchanged model to estimate reliably its climate drift, a straightforward regression procedure seems preferable.

2. Basic Procedure.

Let the final analysis valid at time (day) t of any given parameter (say 500mb height) be expressed in terms of the complex spectral coefficients $A_{l,n}(t)$, where l and n are, respectively, the zonal and total wave numbers. Similarly, the forecast produced on the basis of that analysis and valid at time t+s, will be expressed in terms of $F_{l,n}(t,s)$. We will be introducing statistically modified forecasts given by another set of complex coefficients, $f_{l,n}(t,s)$. A superscript (r) or (i) will be used to indicate the real or imaginary parts of the coefficients. For l=0 the coefficients are real, so $A^{(i)}_{0,n}=0$, and similarly for F and f.

The statistically corrected forecasts are allowed to depend linearly on the dynamical forecasts for that wavenumber only:

$$
(1)\quad
\begin{aligned}
f^{(r)}_{l,n}(t,s) &= a_{l,n}(t,s) + b_{l,n}(t,s)F^{(r)}_{l,n}(t,s) + c_{l,n}F^{(i)}_{l,n}(t,s) \\
f^{(i)}_{l,n}(t,s) &= \alpha_{l,n}(t,s) + \beta_{l,n}(t,s)F^{(r)}_{l,n}(t,s) + \gamma_{l,n}F^{(i)}_{l,n}(t,s)
\end{aligned}
$$

The regression coefficients ($a,b,c;\alpha,\beta,\gamma$) are redetermined every day. They depend on the m immediately preceding forecasts and corresponding verifying analyses. In other words, to determine $a(t_0,s)$ one needs $F(t,s)$, $t_0-s\leq t\leq t_0-s-m+1$, and $A(t)$, $t_0\leq t\leq t_0-m+1$. The coefficients are determined by the ususal method of least squares.

To evaluate the statistically modified forecasts we have accumulated sums of

$$[f_{l,n}(t,s)-A_{l,n}(t+s)][\tilde{f}_{l,n}(t,s)-\tilde{A}_{l,n}(t+s)]$$

(The $\tilde{}$ indicates complex conjugate.) Summations are carried out over both time (t) and wavenumber. The sums over time have generally been over seasons (~90 days). Because of the orthonormality of the spectral representation, the sum of squares of the differences (or errors) in the coefficients is equal to the globally averaged mean square error in the parameter in question.

3. Modified procedure.

A large part of the error, especially at day 5 and longer range, is the climate drift of the model. The regression procedure described above has the effect, in large measure, of estimating the climate drift (i.e. the model bias) as it was manifested in the previous m days. To the extent there is a strong seasonal dependence to the climate drift, the fact of applying it m/2+s days later than the center of the period over which it was estimated must contribute to the residual errors. An alternative procedure was first to estimate the seasonally varying model climate drift and then to carry out the regression analysis in terms of departures from that drift.

For each of the calendar years 1982 and 1983 each of the coefficients [$A_{l,n}(t)$, $F_{l,n}(t,s)$, $l=0,\ldots,5$, $n=l,l+1,\ldots,l+5$) was fitted by an annual harmonic:

$$\hat{A}_{l,n}(t) = \underline{A}_{l,n} + C_{l,n}\cos(2\ t/365) + S_{l,n}\sin(2\ t/365)$$

$$\hat{F}_{l,n}(t) = \underline{F}_{l,n} + D_{l,n}\cos(2\ t/365) + T_{l,n}\sin(2\ t/365)$$

The coefficients $\underline{A}$, $\underline{F}$, C, S, D, T were again determined by least squares. $\hat{A}$ and F represent a highly smoothed harmonic describing the gross features of the behavior of the analyses and forecasts. $B_{l,n}(t)=F_{l,n}(t,s)-A_{l,n}(t+s)$ represents an annual march of the model climate bias. $A'_{l,n}(t)=A_{l,n}(t)-A_{l,n}(t)$, and $F'_{l,n}(t,s)=F_{l,n}(t,s)-F_{l,n}(t,s)$ are the departures of the analyses and forecasts from their respective "climatologies". The forecasts F' were used as the independent variables to produce statistically corrected forecasts, f', in a manner entirely analogous to (1), but where the data for the determination of the regression coefficients depended on the first m pairs of values of F' and A', rather than F and A.

4. Results.

Table 1 presents some summary statistics for day-5 forecasts of the 500-mb height field. The quantities given are the globally averaged root mean square errors contributed by waves 0 through 5 ($0\leq l\leq 5$, $l\leq n\leq l+5$). During Spring 1982 (March - May) the root mean square error, i.e. the root mean square difference between the forecast field made up of waves 0 to 5 and the analysis made up of those waves, was 63.4 meters. For statistically corrected forecasts, where the regression coefficients were based on the most recent 40 days, the rms error was 51.8 meters. When each of the coefficients of the forecasts and the analyses, separately, were approximated by an annual harmonic, using only the data from 1982, and these were subtracted from the forecasts and analyses, the differences between the forecast and observed anomalies for Spring 1982 was 50.3 meters. This is a smaller error than obtained by regression alone. However adding a regression step to the anomalies, using 40 days of most recent history, further reduces the rms error to 49.7 m.

Of course one will not know theseasonal pattern of analysis and forecast as well as this is general -- the same data were used to determine the seasonal pattern as were used to determine the errors. A second calculation was made, in which the seasonal patterns of forecast and analysis were based on 1983 data. The rms anomaly errors for Spring 1982 were then 52.6 meters, further reduced by regression to 50.5 m (using 40-day periods to deterine the coefficients), and to 50.4 m when the record period was increased to 50 days.

The pattern repeats with some regularity. Regression alone provides a reduction in rms error of mostly more than 10 m, and sometimes more than 20 m. A 50-day record length does slightly, but rather consistently better than 40-day periods. Using 30-day or 20-day training periods produces consistently much poorer results. Removing the annual cycles of forecast and analysis -- i.e. removing the seasonally varying model bias -- does a little better than regression alone when the annual cycle is based on the year in which the data are embedded. When the seasonal cycles are more crudely and realistically estimated -- as in using data from a preceding or subsequent year -- then the raw anomalies have slightly larger rms errors than regression alone, but with regression added, the combination of approximate seasonal adjustment plus regression improves upon regression alone, or seasonal adjustment alone, especially at the longer record length.

This pattern is of course not invariant. Notice, for example, that in Fall 1982 the regression step, when applied after removing the 1982 harmonics, results in a larger mean square error. However even then the resulting error is less than that of regression alone. Also regression still improves upon the unregressed forecast based on removing the 1983 annual cycle. In general, however, the pattern is consistent and carries over to the more extensive 1000 mb results shown in Table 2. Also, at least at 1000 mb, 60-day training periods (not tested at 500 mb) result in small but persistent further improvement.

The results shown in these first two tables are illustative of the gross results of the statistical correction procedure, as measured by root mean square errors. It is of considerable interest to diagnose the contributions to the error reduction that are attributable to various steps in the procedure. For this purpose it is better to quantify results in terms of mean square errors (rather than root mean square errors) because the mean square errors are additive.

Tables 3 and 4 exemplify and characterize some of these results. For each of the seasons shown, the total mean square error in zonal wave numbers 0 through 5 (including meridianal waves 0 through 5) of the uncorrected forecasts is given, plus the apportionment of that error between the zonal wave no. 0 and wave nos. 1-5. In general the contributions to the errors in the day 5 forecasts from these two sources are about equal. These contributions are shown both in terms of total mean square error (expressed as the square of the root mean square error) and as a percent of the total mean square error for waves 0-5. All other results are expressed as percentages, using the same base.

Regression alone generally results in a 30% to a 50% reduction in variance. Most of this is attributable, usually, to improved forecasts of wave no. 0. In some cases, however, such as Summer 1982, almost as much error reduction comes from the waves 1-5 as from wave no. 0. The consistently positive reductions of variance for wave nos. 1-5 as a whole, and the occasional occurrence of a substantial error reduction from this source, argues for the value of including this correction in any operational procedure. It also suggests that if new models are introduced having substantially less climate drift, then the total improvement of which adjustment by regression will be capable will be less, but it will still be a useful procedure.

We have already seen, in the discussion of Tables 1 and 2, that corrections for the seasonally varying climate bias of the forecast model account for most of the forecast improvements. It can be seen in Tables 3 and 4 that this is a characteristic due mostly to wave number 0. Indeed, for zonal wave no. 0 only meager residual forecast improvement can be achieved by regression once the seasonally varying model bias is accounted for. However, for wave nos. 1-5, the limited improvement that is possible is due more to the regression step in the procedure than to the seasonal adjustment.

In this statistical correction procedure the criterion for selection of coefficients is the minimization of mean square error. There are at least two factors of meteorological concern that are not dealt with explicitly in this procedure but warrent attention. These factors are the wave amplitudes and phases in the resulting forecasts. The pairwise regressions involving the coupled real and imaginary (or sine and cosine) coefficients will introduce systematic phase shifts to the various waves (1-5). In the calculations to date we have no measures of these phase shifts. In future tests the phases of the individual waves can be verified, or, better, anomaly correlation coefficients can be calculated. Cursory examination of the regression coefficients indicates substantial contributions from the c and β terms of equation (1) indicating that there are indeed significant phase adjustments implied in the statistical corrections.

There are some numerical results that pertain to the amplitudes of the waves. Some of these are shown in Table 5. Regression, by its very nature, is expected to give rise to smaller amplitudes. Part of the effectiveness of regression is its tendency to 'regress' toward the mean when the excursions from the mean cannot be statistically validated. The systematic decrease in the ratio of root mean square amplitudes of the adjusted forecasts to the analyses (or of its square, the ratio of kinetic energies) implies a decrease in linear predictibility with increasing wave number. The fact that these ratios are substantially greater than 0 at wave number 5 implies that, for day 5 forecasts, predictibility extends well beyond wave no. 5 and that higher wave numbers should be included in future tests and calculations. Wave no. 0 is not shown in Table 5 because it is dominated, especially at 500 mb, by large mean values, and rms ratios would have very little meaning. When the amplitudes of both analysis and forecast are calculated after removal of the seasonally varying mean values, the rms ratios (regression forecast to analysis) range from .8 to .95 at 500 mb, and are near .7 or .8 at 1000 mb. Although maps have not been drawn of the daily forecasts, it is anticipated that maps of wave nos. 0-5 will have comparable meteorological content and "feel" to similar forecast maps now being prepared.

5. Conclusions.

Regression over recent experience in the spectral domain is an economical procedure that can routinely and reliably improve the numerical prediction of the standard fields that are now the basis of mid-range forecasts. Operational test are warranted for day-5 predictions of height and temperature fields. A 60-day training period, without prior corection for model dependent climate drift, is recommended. Test should also be carried out to ascertain the value of applying the method to longer-range predictions (say day-8 to start), and to develp related procedures for prediction of averages over days 6 to 10.

REFERENCES

Glahn, H.R., 1980: Methods and accuracy of statistical weather forecasting in the United States. Collection of papers presented at the WMO Symposium on Probabilistic and Statistical Methods in Weather Forecasting, Nice, 387-396.

Glahn, H.R., and Dale A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11, 1203-1211.

Hughes, F.D., 1982: An objective method to modify NMC/NWP model mean sea level progs. NMC Office Note No. 259.

Lorenz, E.N., 1977: An experiment in non-linear statistical weather forecasting. Mon. Wea. Rev., 105, 590-602.

Sela, Joseph G., 1982: The NMC Spectral model. NOAA Technical Report NWS 30.

Table 1

Root Mean Square Prediction Errors (meters)
(day 5, 500 mb)

Season	Training period	Unadjusted	82 harmonic removed	83 harmonic removed
	none	63.4	50.3	52.6
	40 days	51.8	49.7	50.5
	30 days	52.7	50.9	
Spring 1982	20 days	57.1	54.6	
	none	63.3	50.1	52.6
	50 days	51.7	49.7	50.4
	none	63.8	47.4	48.8
	40 days	46.7	46.3	46.2
	30 days	49.6	48.9	
Summer 1982	20 days	51.1	50.7	
	none	64.2	47.4	48.8
	50 days	46.0	45.6	45.6
	none	64.9	50.6	53.2
	40 days	53.5	51.9	52.8
	30 days	55.6	54.1	
Fall 1982	20 days	58.8	57.4	
	none	64.4	50.6	53.2
	50 days	53.0	51.1	52.9
	none	62.2	50.8	49.2
	40 days	49.6	48.8	48.9
	30 days	50.8	50.1	
Winter 1982/3	20 days	54.4	53.7	
	none	63.7	50.8	49.2
	50 days	49.7	48.1	47.4
	none	72.6	57.9	53.7
	40 days	54.2	53.5	53.4
Winter 1983/4	none	72.3	57.9	53.7
	50 days	53.2	52.7	52.3

Note: For a given season the uncorrected forecasts are the same whether 40 or 50 day regression training periods are used and the rows opposite "none" should be the same. However at times the season was inadvertently allowed to begin one day too early or too late. In those cases the "none" values correspond properly to the adjacent training period length.

Table 2

Root Mean Square Prediction Errors (meters)
(day 5, 1000 mb)

Season	Training period	Unadjusted	82 harmonic removed	83 harmonic removed
	none	49.0	38.0	40.5
Spring	40 days	38.6	37.4	37.9
1982	50 days	38.6	37.0	37.6
	60 days	38.6	36.5	37.3
	none	49.9	37.6	39.7
Summer	40 days	36.6	35.1	35.3
1982	50 days	36.0	34.4	34.8
	60 days	36.2	34.2	34.6
	none		39.4	42.2
Fall	40 days		38.1	38.4
1982	50 days		37.7	38.3
	60 days		37.3	38.0
	none	51.8	40.4	39.6
Winter	40 days	37.0	37.3	37.1
1982/83	50 days	36.6	36.7	40.7
	60 days	36.3	36.6	36.2
	none	48.8	42.1	39.6
Spring	40 days	37.5	37.1	36.9
1983	50 days	37.3	36.9	36.6
	60 days	37.3	36.5	36.2
	none	51.4	41.4	39.0
Summer	40 days	41.2	40.0	39.6
1983	50 days	40.1	38.9	38.1
	60 days	39.8	38.9	37.9
	none	60.8		44.1
Winter	40 days	41.2		41.0
1983/84	50 days	40.7		40.3
	60 days	40.8		39.9

Table 3
Contributions to the Reduction of Mean Square Error
(1000mb -- day 5 forecasts)

Season	Wave Nos.	Total Mean Square Error	%	Mean Square Error Reduction as Percent of Total Forecast Mean Square Error: Regression Only	Removal of Seasonal Bias*	Regression (60 d)	Net
Spring 1982	0 - 5	$(48.96m)^2$	100.0	37.9	31.7	8.3	40.0
	0	$(34.35m)^2$	49.2	29.1	28.0	0.8	28.8
	1 - 5	$(34.89m)^2$	50.8	8.8	3.7	7.5	11.2
Summer 1982	0 - 5	$(34.35m)^2$	100.0	47.4	36.6	15.5	52.1
	0	$(25.36m)^2$	42.6	25.8	27.1	2.9	30.0
	1 - 5	$(23.17m)^2$	57.4	21.6	9.5	12.6	22.1
Winter 1982/83	0 - 5	$(51.83m)^2$	100.0	51.0	39.1	11.1	50.2
	1	$(36.96m)^2$	50.9	37.8	35.9	1.4	37.3
	1 - 5	$(36.33m)^2$	49.1	13.2	3.2	9.6	12.8

* 1983 seasonal bias harmonics used to make seasonal adjustment for Spring and Summer, 1982; 1982 harmonics used to adjust Winter 1982/83 forecasts and analyses.

Table 4
Contributions to the Reduction of Mean Square Error
(500mb — day 5 forecasts)

Season	Wave Nos.	Total Mean Square Error	%	Regression Only	Removal of Seasonal Bias*	Regression	Net
				Mean Square Error Reduction as Percent of Total Forecast Mean Square Error			
Spring 1982	0 - 5	$(63.44m)^2$	100.0	33.2	31.3	5.5	36.8
	0	$(42.19m)^2$	44.2	24.4	26.9	0.7	27.6
	1 - 5	$(47\ 37m)^2$	55.8	8.8	4.4	4.8	9.2
Summer 1982	0 - 5	$(64.15m)^2$	100.0	48.5	42.0	7.4	49.4
	0	$(45.05m)^2$	49.3	38.0	38.4	0.6	39.0
	1 - 5	$(45.66m)^2$	50.7	10.5	3.6	6.8	10.4
Fall 1982	0 - 5	$(64.43m)^2$	100.0	32.3	31.8	0.9	32.7
	0	$(43.29m)^2$	45.1	28.1	31.2	-2.4	28.8
	1 - 5	$(47.72m)^2$	54.9	4.2	0.5	3.3	3.8
Winter 1983/84	0 - 5	$(72.26m)^2$	100.0	45.8	35.7	11.2	46.8
	0	$(52.69m)^2$	53.2	39.7	33.4	7.9	41.3
	1 - 5	$(49/56m)^2$	46.8	6.2	2.3	3.2	5.5

* Harmonics fitted to 1983 data used for seasonal adjstment of 1982 foreecasts and analyses; 1982 harmonics used to adjust Winter 1983/84 data.

Table 5
Root Mean Square Wave Amplitudes

Level	Season	Wave No.	Analysis	Adjusted Forecast	RMS Ratio	MS Ratio
		1	11.09	9.39	.847	.717
		2	10.05	9.86	.939	.881
500 mb	Spring	3	10.00	10.95	1.095	1.199
	1982	4	9.76	6.50	.788	.621
		5	8.25	4.70	.570	.325
		1	11.97	9.35	.781	.611
		2	8.96	6.85	.764	.584
500 mb	Summer	3	9.57	8.10	.874	.717
	1982	4	8.48	6.24	.736	.541
		5	5.58	3.28	.588	.346
		1	13.24	10.27	.776	.602
		2	12.13	9.67	.797	.635
500 mb	Fall	3	10.54	7.02	.666	.443
	1982	4	9.16	5.88	.642	.412
		5	9.40	5.05	.537	.289
		1	18.01	16.06	.892	.795
		2	11.37	9.47	.833	.694
500 mb	Winter	3	12.81	9.94	.777	.604
	1982/83	4	11.08	7.87	.711	.505
		5	8.78	6.24	.711	.506
		1	8.40	6.97	.830	.688
		2	7.28	5.47	.752	.566
1000 mb	Spring	3	6.81	5.01	.736	.541
	1982	4	5.60	3.88	.692	.479
		5	4.76	2.29	.482	.232
		1	10.82	7.83	.724	.524
		2	8.20	6.50	.793	.629
1000 mb	Summer	3	6.96	5.96	.856	.733
	1982	4	5.16	3.38	.655	.429
		5	3.78	2.11	.558	.311
		1	12.24	11.21	.916	.839
		2	11.05	9.61	.870	.757
1000 mb	Winter	3	8.66	6.13	.707	.500
	1982/83	4	6.02	4.10	.682	.465
		5	4.62	2.90	.627	.393