U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
OFFICE OF SYSTEMS DEVELOPMENT
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 85-6

TESTING MOS SURFACE WIND GUST PREDICTION EQUATIONS

Gary M. Carter and Valery J. Dagostaro

March 1985

TESTING MOS SURFACE WIND GUST PREDICTION EQUATIONS

Gary M. Carter and Valery J. Dagostaro

1. INTRODUCTION

National Weather Service (NWS) forecasters have been provided with surface wind guidance based on the Model Output Statistics (MOS) technique (Glahn and Lowry, 1972) since May 1973 (Carter, 1975). The current objective wind guidance system produces forecasts at 6-h increments for projections of 6 to 48 hours from 0000 and 1200 GMT for about 300 NWS locations throughout the conterminous United States and Alaska. These predictions are for 1-min average winds at specific times throughout each day (i.e., 0000, 0600, 1200 or 1800 GMT). Usually, strong winds are forecast by this procedure only in conjunction with well organized, synoptic scale weather systems.

Automated guidance for the magnitude of surface wind gusts, in excess of the mean wind, also would be of interest to forecasters. In the past, this additional information was not provided because the predictand data necessary for derivation of MOS forecast equations were not available. However, since 1977, the Techniques Development Laboratory (TDL) has been archiving wind gusts contained in hourly "Airways" surface observations. This report describes our efforts to develop and test a MOS system to predict surface wind gusts by making use of the data in TDL's hourly archive.

2. DEVELOPMENTAL APPROACH

We developed the experimental wind gust prediction equations by using the Regression Estimation of Event Probability (REEP) statistical model (Miller, 1964) to establish relationships among various categories of observed surface wind gusts (predictands) and forecast fields (predictors) from the Limited-area Fine Mesh (LFM) model (Newell and Deaven, 1981). Climate-related variables, such as the sine and cosine of the day of the year and of twice the day of the year, also were included as potential predictors. The predictand data were divided into two categories of ≥ 25 and ≥ 35 kt. A forward screening regression procedure was applied to select the best predictors for each equation as long as each new term contributed at least 0.75% to the reduction of variance for any one of the predictands, or until a total of 12 variables had been selected. In contrast to the traditional, single-station technique used to develop the MOS surface wind guidance system for 1-min average direction and speed, separate sets of regional wind gust prediction equations were derived by grouping the data for 236 stations into the 10 geographic areas shown in Fig. 1.

For any particular reporting time, the occurrence of a surface wind gust may be considered a rare event at most observation sites. Hence, for this study, we formed the predictand data set by obtaining the highest gust available from observations taken at three different times (2100, 0000, and 0300 GMT). Also, the mean wind speeds reported at these same hours were examined in an attempt to distinguish "missing values" from genuine "no gust reports" in the hourly data archive. A gust report of zero was set to missing whenever the corresponding mean wind speed was $\geq$ 20 kt. This was necessary because the automated surface observation data base at the National Meteorological Center--the source for TDL's archive--indicates a zero value whenever a gust is not reported, is missing, or does not occur.

Grouping stations into regions also was used to increase the developmental sample size. In order to determine regional boundaries, we calculated the observed relative frequency of surface wind gusts for each station and predictand category for only those cases in which the corresponding 24-h LFM model forecast of 850-mb wind speed was $\geq$ 15 m s^{-1}. Stations with similar relative frequencies were grouped together.

The developmental sample included four cool seasons (October through March, 1977-78 through 1980-81) of 0000 GMT cycle forecast output from the LFM model. The variables which we screened as potential predictors were those that would be expected to have a physical relationship with the occurrence of surface wind gusts. These included wind components, constant pressure level height, and temperature forecasts at various levels throughout the troposphere. Fields directly forecast by the LFM model, as well as derived variables such as geostrophic wind, vorticity, and various stability indices, were screened. The model output predictors were interpolated to the location of each station in the sample.

3. REGRESSION RESULTS

Sets of experimental equations were derived for each forecast region for the 24-h projection from 0000 GMT. Table 1 shows the MOS gust prediction equations for Region 6 (see Fig. 1). Since the equations for both categories were derived simultaneously, they are comprised of the same five predictors but, of course, the individual regression coefficients differ. In general, these equations are quite representative of those for the other regions in the sense that the low-level wind speed forecasts were the most important predictors.

The reductions of variance for each of the 10 regions are presented in Table 2. Of course, the magnitude of the reduction of variance is related to the relative frequency of the event. Hence, for all 10 regions, the values for $\geq$ 25 kt are much greater than those for $\geq$ 35 kt. The values for both categories are relatively low for locations throughout the mountainous western United States where the numerical model is unable to accurately forecast orographically induced weather patterns. In addition, the statistical approach does not fully account for orographic effects or any non-linear relationships among the model output variables and the corresponding gusts which are observed at the surface. The reduction of variance also is low for Region 10, the southern coastal area. This may be a reflection of the model's inability to accurately predict low-latitude circulation patterns.

The gust equations can be used to generate probabilistic predictions of each predictand category. From these, one category is selected as the "best" category. The threshold probability technique is used to transform the probabilistic estimates into best category forecasts; that is, when the forecast probability of a given category exceeds a predetermined critical value, that category is chosen. The procedure is to compare the forecast cumulative probability to the critical threshold probability which has been established for each categorical limit. The set of threshold probabilities was determined by generating probabilistic forecasts for the entire developmental sample. Several objective techniques exist for obtaining these threshold probabilities. For this study, we tested four different threshold techniques: (1) Beta model, (2) Miller and Best model, (3) R model, and (4) RC model. The latter three schemes are presented in Bermowitz and Best (1978), while the Beta model is described in Miller and Best (1981).

4. TEST RESULTS

We evaluated the test equations on independent data from the cool season of 1981-82. Probabilistic forecasts valid approximately 24 hours from 0000 GMT were generated for 236 stations and converted into categorical predictions by each of the four techniques mentioned in the previous section. For comparison, another set of categorical forecasts was produced by multiplying the traditional, 24-h (0000 GMT cycle) MOS forecast of wind speed by a factor of 1.5. This is quite similar to the method which is employed by TDL's computer worded forecast program to estimate the magnitude of surface wind gusts (Glahn, 1978; Bermowitz et al., 1980; Bermowitz and Miller, 1984).

Table 3 shows three-category contingency tables associated with the categorical forecasts produced by each threshold technique, as well as the corresponding table for the simple scheme (MOS speed x 1.5). These results indicate that the Beta, Miller and Best, and R models greatly underforecast the occurrence of surface wind gusts of $\geq$ 35 kt. Also, the Beta and Miller and Best models predicted only about half of the gusts which were observed in the 25-34 kt category. Both the RC model and the simple scheme were much better in regard to forecasting for both the 25-34 and $\geq$ 35 kt categories.

Several verification measures were computed from information contained in the contingency tables. These included the Heidke skill score, the threat score for $\geq$ 25 and $\geq$ 35 kt, and the bias-by-category. In addition, we calculated the overall Brier score for the probabilistic forecasts. The results are presented in Table 4. The verification scores for the 25-34 and $\geq$ 35 kt categories indicate the RC model and the simple scheme produced categorical forecasts which were superior to those from the other three models. Overall, the results for the simple scheme were slightly better than those for the RC model categorical predictions.

5. SUMMARY

This report describes a study of the feasibility of using the observations in TDL's hourly data archive to develop a MOS surface wind gust forecast system. Our test involved 236 stations throughout the conterminous United States. Sets of equations to predict the probability of wind gusts of $\geq$ 25 and $\geq$ 35 kt for a projection of approximately 24 hours from 0000 GMT were derived for 10 climatically homogeneous regions. The developmental sample consisted of forecast fields from the 0000 GMT cycle runs of the LFM model and gust observations from TDL's hourly data archive during the cool seasons of 1977-78 through 1980-81. These REEP equations produce probabilistic forecasts for the occurrence of surface wind gusts within $\pm$ 3 hours of 0000 GMT (i.e., 2100, 0000, or 0300 GMT). Threshold values, which are needed in order to convert the probabilistic forecasts into best category predictions, also were developed based on four different techniques.

We evaluated the test equations and threshold models on independent data from the cool season of 1981-82. For comparison, another set of categorical forecasts was produced by multiplying the traditional, 24-h MOS forecast of wind speed by a factor of 1.5. The verification results (skill score, threat score, and bias-by-category) indicate the simple scheme (MOS speed x 1.5) produced categorical forecasts which were slightly better than the forecasts associated with the RC model which was the best of the four threshold probability techniques.

To some extent, the disappointing test results are an indication the difficulty of predicting relatively rare events such as strong wind gusts, and the fact that the wind speed forecasts from the simple scheme were based on single-station MOS prediction equations. We also think these findings are related to the problems described in Section 2 of this report which are associated with distinguishing "missing values" from "no gust reports." Until other techniques can be devised to better handle these problems, we recommend that a simple scheme which consists of multiplying the traditional MOS wind speed by a factor of 1.5, be employed whenever objective surface wind gust estimates are required.

6. ACKNOWLEDGMENTS

We are grateful to Belinda Howard for typing the manuscript, and to the many other members of the Techniques Development Laboratory who contribute to the development and maintenance of the MOS system.

REFERENCES

Bermowitz, R. J., and D. L. Best, 1978: An objective method for maximizing threat score. TDL Office Note 78-15, National Weather Service, NOAA, U.S. Department of Commerce, 11 pp.

_____, M. M. Heffernan, and H. R. Glahn, 1980: Computer worded forecasts: An update. Preprints Eighth Conference on Weather Forecasting and Analysis, Denver, Amer. Meteor. Soc., 453-456.

_____, and R. L. Miller, 1984: A field evaluation and some recent changes in the computer worded forecast. Preprints Tenth Conference on Weather Forecasting and Analysis, Clearwater Beach, Amer. Meteor. Soc., 92-97.

Carter, G. M., 1975: Automated prediction of surface wind from numerical model output. Mon. Wea. Rev., 103, 866-873.

Glahn, H. R., 1978: Computer worded public weather forecasts. NOAA Technical Memorandum NWS TDL-67, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 24 pp.

_____, and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. J. Appl. Meteor., 11, 1203-1211.

Miller, R. G., 1964: Regression estimation of event probabilities. Tech. Report No. 1, Contract CWB-10704, The Travelers Research Center, Inc., Hartford, Conn., 153 pp.

______, and D. L. Best, 1981: A Beta classification model. TDL Office Note 81-8, National Weather Service, NOAA, U.S. Department of Commerce, 14 pp.

Newell, J. E., and D. G. Deaven, 1981: The LFM-II model--1980. NOAA Technical Memorandum NWS NMC-66, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 20 pp.

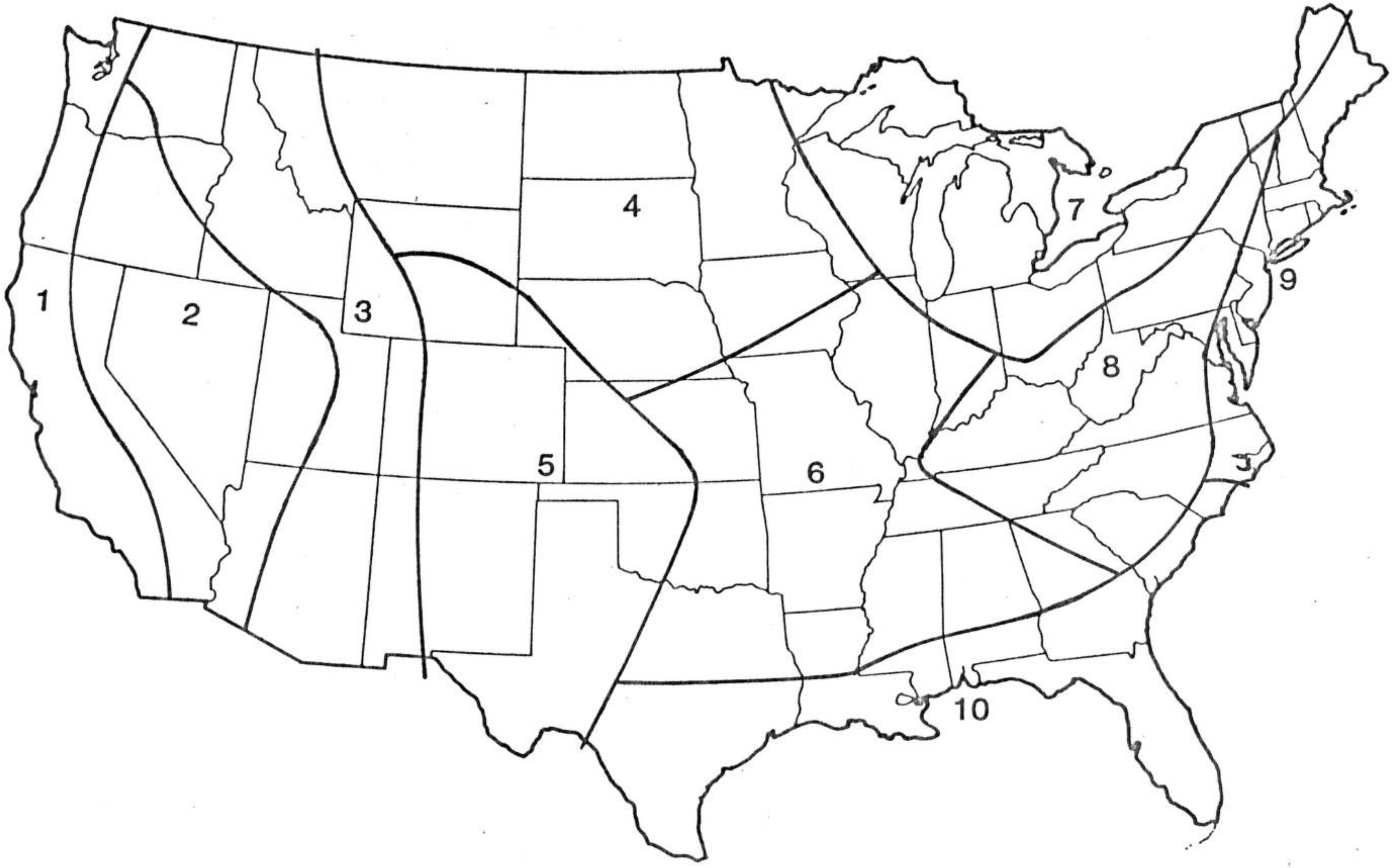

Figure 1. The 10 regions which were used to develop experimental MOS surface wind gust prediction equations.

Table 1. LFM predictor variables, cumulative reductions of variance, and equation coefficients for forecasting surface wind gusts 24 hours after 0000 GMT for Region 6 (see Fig. 1) during the cool season months of October through March.

Predictor (Units)	Projection (h)	Cum. Reduction of Variance $\geq$ 25 kt	Cum. Reduction of Variance $\geq$ 35 kt	Coefficient $\geq$ 25 kt	Coefficient $\geq$ 35 kt
1000-mb Geostrophic Speed ($m\ s^{-1}$)	24	0.115	0.019	0.0121	0.0014
1000-mb Geostrophic V Wind ($m\ s^{-1}$)	24	0.162	0.025	-0.0041	-0.0004
850-mb Speed ($m\ s^{-1}$)	18	0.186	0.026	0.0095	0.0005
700-mb Relative Vorticity x 10^5 (s^{-1})	24	0.197	0.028	0.0151	0.0019
Cosine Day of Year	--	0.205	0.029	-0.0763	-0.0050
		Regression Constant		-0.1010	-0.0124

Table 2. Reduction of variance (%) for the 24-h surface wind gust forecast equations (0000 GMT cycle) for each of the 10 regions shown in Fig. 1. The developmental sample was comprised of the cool seasons of 1977-78 through 1980-81.

Region	$\geq$ 25 kt	$\geq$ 35 kt
1	9.3	1.8
2	8.4	2.3
3	8.3	2.3
4	24.4	8.7
5	22.2	9.8
6	20.5	2.9
7	22.5	6.8
8	17.7	3.8
9	31.1	11.8
10	12.4	1.4
Overall Average	17.7	5.2

Table 3. Contingency tables for 0000 GMT cycle 24-h categorical forecasts of surface wind gusts made from various types of threshold probabilities for 236 stations in the conterminous United States during the cool season of 1981-82.

			Beta Thresholds						Miller and Best Thresholds			
			FCST						FCST			
			< 25	25-34	≥ 35	TOTAL			< 25	25-34	≥ 35	TOTAL
LFM-based		< 25	34464	561	42	35067		< 25	34367	652	48	35067
MOS	OBS	25-34	1981	568	61	2610	OBS	25-34	1920	626	64	2610
Probabilities		≥ 35	229	201	66	496		≥ 35	216	208	72	496
		TOTAL	36674	1330	169	38173		TOTAL	36403	1486	184	38173
			R Thresholds						RC Thresholds			
			FCST						FCST			
			< 25	25-34	≥ 35	TOTAL			< 25	25-34	≥ 35	TOTAL
LFM-based		< 25	33594	1387	86	35067		< 25	33290	1664	113	35067
MOS	OBS	25-34	1641	878	91	2610	OBS	25-34	1544	946	120	2610
Probabilities		≥ 35	153	255	88	496		≥ 35	141	248	107	496
		TOTAL	35388	2520	265	38173		TOTAL	34975	2858	340	38173
									FCST			
									< 25	25-34	≥ 35	TOTAL
LFM-based MOS			---	---	---	---		< 25	33692	1280	95	35067
Speed x 1.5							OBS	25-34	1602	849	159	2610
								≥ 35	151	217	128	496
								TOTAL	35445	2346	382	38173

Table 4. Verification results for 0000 GMT cycle 24-h forecasts of surface wind gusts for 236 stations in the conterminous United States during the cool season of 1981-82.

Verification Measures	LFM-based MOS Probabilities	LFM-based MOS Speed x 1.5
Brier Score	.124	--
	Beta Thresholds	
Skill Score	.300	--
Threat Score/Bias		
--/< 25	--/1.046	--
≥ 25/25-34	.242/.510	--
≥ 35/≥ 35	.110/.341	--
	Miller & Best Thresholds	
Skill Score	.315	--
Threat Score/Bias		
--/< 25	--/1.041	--
≥ 25/25-34	.255/.569	--
≥ 35/≥ 35	.118/.371	--
	R Thresholds	
Skill Score	.342	--
Threat Score/Bias		
--/< 25	--/1.009	--
≥ 25/25-34	.287/.966	--
≥ 35/≥ 35	.131/.534	--
	RC Thresholds	
Skill Score	.345	.357
Threat Score/Bias		
--/< 25	--/.997	--/1.011
≥ 25/25-34	.291/1.095	.302/.899
≥ 35/≥ 35	.147/.686	.171/.770