

**ESTIMATING VISIBILITY OVER THE NORTH PACIFIC OCEAN
USING MODEL OUTPUT STATISTICS**

Robert J. Renard (1) and
William J. Thompson (2)
Department of Meteorology
Naval Postgraduate School
Monterey, California 93943

ABSTRACT

The method of model output statistics (MOS) is used to develop multiple linear regression equations for forecasting the probability of marine visibility in five categories (0-.49, 0.5-1.9, 2-9.9, 10-19 and 20-50 km) at 24-h intervals to 48-h, for the summer season, North Pacific Ocean area. Further manipulation of the scheme yields categorical visibility forecasts for three (0-1.9, 2-9.9, 10-50 km) and two (0-9.9, 10-50 km) visibility categories. Dependent and independent tests are verified using percentage correct, bias, Heidke skill score and threat score. The experiment establishes the credibility of MOS applications over open ocean areas, with levels of skill commensurate to those for MOS visibility forecasts over land.

1. INTRODUCTION

Although fog and visibility forecast schemes abound for coastal locations, the open ocean has been largely ignored. These kinds of forecasts are of particular importance in order to safely execute maritime shipping and naval sea/air operations. Maritime casualties due to fog-related low visibility are highest in the summer months (Figure 1) when the combination of extent and density of fog is at a maximum (3,4). Since the ongoing computerized atmospheric prediction models do not output visibility directly, a reasonable approach to forecasting visibility is through the use of Model Output Statistics (MOS) (5). For the experiment reported on here, the North Pacific Ocean (30-60N, 145E-130W) was selected as the test basin, with various Fleet Numerical Oceanography Center (FNOC), Monterey, CA analysis and prediction models supplying the basic Model Output Parameters (MOP) from a 23x12 section of FNOC's Northern Hemisphere 63x63 polar stereographic grid. Verification of the developed MOS forecast scheme is compared to that using visibility climatology (3), visibility persistence, and a limited sample of National Weather Service MOS visibility forecasts for the continental United States (6).

2. DATA/PARAMETERS

The surface ship observational data from the North Pacific Ocean were obtained from the Naval Oceanography Center Detachment, Asheville NC, which is co-located with the National Climatic Data Center (NCDC). These data, Tape Data Family-11 (TDF-11), which are filtered to exclude duplications

and erroneous reports, are a compilation of information from ships' logs, ships' weather reporting forms, published ship observations, automatic observing buoys, teletype reports and data purchased from foreign meteorological services. The quality varies from those observations taken by a deckhand to those of a trained observer. Data at 0000 GMT (local daylight) for the summer months July/August 1979 served as the dependent/independent data set. Over 4000 synoptic ship reports were available for each month.

The basic set of MOP's consists of 24 diagnostic-prognostic parameters generated from FNOC's Mass Structure Analysis Model and the Primitive Equation, Marine Wind and OceanWave Prediction Models. An additional 79 interactive and derived dynamic and thermal parameters, continuous and binary, were obtained from this set. Appendix (A) is a selected list of those model output and climatology parameters used in developing the MOS equations.

3. PROCESSING THE DATA AND DEVELOPMENT OF REGRESSION EQUATIONS

The first step consisted of interpolating the MOP's and derived parameters (via a curvilinear bi-cubic spline routine) from the FNOC grid to each ship position, where they were matched to the respective visibility code. These interpolated parameters (predictors) were then used in the stepwise multiple linear regression program BMDP2R (7) to derive five equations, the predictands of which are parameters indicating the five visibility ranges shown in Table I.

REGRESSION EQUATION (Visibility category)	VISIBILITY RANGE	SYNOPTIC OBSERVATION CODE
1	0.0- 0.49 km	90-92
2	0.5- 1.9 km	93-94
3	2.0- 9.9 km	95-96
4	10.0-19.0 km	97
5	20.0-50.0 km	98-99

Table I. Visibility categories

A comparison of open ocean visibility forecasting using MOS, in one case with a categorical predictand (8) and in the other case with a probabilistic predictand (9,10), indicated the desirability of the latter approach. The remainder of this paper will focus on the probabilistic visibility approach. Table II indicates the predictand values assigned to each ship observation as a function of reported visibility, for each of the five regression equations developed.

VISIBILITY FORECAST EQUATION					SYNOPTIC OBSERVATION CODE
(1)	(2)	(3)	(4)	(5)	
100	25	0	0	0	90
100	50	0	0	0	91
100	75	25	0	0	92
75	100	50	0	0	93
50	100	75	25	0	94
25	75	100	50	25	95
0	50	100	75	50	96
0	25	75	100	75	97
0	0	50	75	100	98
0	0	25	50	100	99

Table II. Visibility probability (%) (= predictand) assigned to each synoptic ship observation as a function of reported visibility, for each of the five regression equations developed.

For example, in deriving the equation for specifying visibility category 3 (see Table I), observations coded as 95 or 96 were assigned a predictand value of 100%, those with codes 94 and 97 a value of 75%, codes 93 and 98 a value of 50%, and so forth. Ideally, the predictand used in developing that equation should be 100% for all observations in codes 95 and 96 and 0% for all other visibility codes (i.e. 90 to 94 and 97 to 99). But, it is commonly accepted that visibility observations at sea are inexact at best (i.e. code 95 may be reported when in fact code 94 was observed, etc.). The ideal approach was tried first but it was not as successful as assigning to the predictand percentages other than 0% to visibility codes outside of the category to which the equation applies, in this case category (3). Several variations for predictand assignment were tried, such as 80% for code 94, 60% for code 93, 30% for code 92; and similarly for codes 97, 98 and 99. Considering all equations, it was most methodical and the success of the technique was best when using the quartile reduction approach, that is reducing the predictand value by 25% increments in either direction from the codes defining the category. Table II entries should not be viewed horizontally--only vertically, and, of course, the percentages should not add up to 100% or any other prescribed value. This is an experimental quantitative approach to an

existing problem in working with visibility observations at sea.

Three sets of five equations each: a diagnostic set ($\tau = 0$ h) and two prognostic sets ($\tau = 24$ and 48 h), were derived (10) from the July 1979 data set (Tables III, IV and V). Only those predictors that contributed at least 0.5% to the explained variance of the predictand were retained. The evaporative heat flux (EHF) is prominent in all equations. The majority of explained variance was determined by this one parameter whenever it was the leading parameter. Negative (positive) EHF implies that the moisture flux is directed downward toward (upward from) the sea and is associated with low (high) visibility. It is evident that the visibility class 2-9.9 km is the most difficult to specify from the available FNOC predictor parameters.

VIS CODE GROUP	VISIBILITY PROBABILITY	R ² (percent)
90-92 (0-.49 km)	-35.1586 -0.9191 EHF 43.9857 FTER 0.0039 RASTDX 0.0048 RHRSQ 0.5606 BVISX 0.0255 RASTDR	18.6 2.6 1.3 1.0 0.9 0.6 25.0
93-94 (0.5-1.9 km)	356.8071 -1.6095 EHF -1.1414 BVISR 28.4439 FTER 0.4441 BVISX 0.0047 U925 -0.3126 PS	19.0 6.2 1.2 0.7 0.5 0.5 28.1
95-96 (2-9.9 km)	129.1194 -0.9573 BVISX -0.6316 RHX -0.4581 ASTDRX	5.0 1.2 1.4 7.6
97 (10-19 km)	75.6061 0.5649 EHF -38.1213 FTER -0.9247 BVISX 0.7383 BVISR -0.0237 RASTDR 0.0041 U925	14.8 2.2 1.7 1.4 0.8 0.5 21.4
98-99 (20-50 km)	57.5600 1.9054 EHF 1.4265 BVISR -40.5343 FTER -0.6891 BVISX 0.0056 U925	22.8 5.4 1.6 1.2 0.6 31.6

Table III. Regression equations for estimating visibility probability, by visibility code-group for the North Pacific Ocean 30-60N 145E-130W, $\tau = 0$ h (4079 observations, July 1979). Variables for initial time listed in order of selection. R² specifies variance explained by each predictor. See Appendix A for parameter description.

VIS CODE GROUP	VISIBILITY PROBABILITY	R ² (percent)
90-92 (0-.49 km)	18.6298 -1.9898 EHF 24 0.0213 RASTDX 00 19.9026 FTER 00 -0.5685 VVWW 36 17.9254 FTER 24	23.0 2.0 1.0 0.6 <u>0.5</u> 27.1
93-94 (0.5-1.9 km)	32.5351 -2.0482 EHF 24 -0.5285 BVISR 00 0.0204 RASTDX 00 18.4725 FTER 24	23.2 2.2 1.3 <u>0.7</u> 27.4
95-96 (2-9.9 km)	137.1898 -1,2913 BVISX 00 -19.4424 FTER 00 -0.5658 RHX 00 -5.8802 EHF 24 -0.6511 VVWW 00	3.0 1.6 1.2 0.9 <u>0.7</u> 7.4
97 (10-19 km)	61.9611 1.5293 EHF 24 -0.0210 RASTDX 00 -14.9147 FTER 00 0.5736 VVWW 36 -16.5002 FTER 24	18.2 2.2 0.7 0.7 <u>0.5</u> 22.3
98-99 (20-50 km)	63.5259 2.8336 EHF 24 -0.0245 RASTDX 00 0.5113 BVISR 00 -21.7912 FTER 24	28.3 1.8 1.2 <u>0.7</u> 32.0

Table IV. Same as Table III except Tau 24 h (4095 observations). Number following parameter indicates initial time (00) or prediction interval (12, 24, 36, 48) in h.

VIS CODE GROUP	VISIBILITY PROBABILITY	R ² (percent)
90-92 (0-.49 km)	-428.6230 -1.8534 EHF 36 27.3651 FTER 00 25.8898 FTER 48 -48.3218 GGHTA 36 0.4235 PS 36 0.4132 MBVIS 48	20.9 2.0 1.1 1.0 1.0 <u>0.6</u> 26.6
93-94 (0.5-1.9 km)	-353.1233 -1.0305 EHF 36 0.2561 CLIMO 00 22.6730 FTER 48 -0.4162 BVISR 00 0.3658 PS 24 0.0146 RASTDX 00	19.3 1.7 1.3 0.9 0.6 <u>0.6</u> 24.4
95-96 (2-9.9 km)	145.7690 -1.3323 BVISX 00 -1.1001 VVWW 00 -0.6430 RHX 00 2.4041 SSANOM 00 0.3604 VVWW 36	1.5 1.7 2.0 0.8 <u>0.6</u> 6.6

97 (10-19 km)	497.9680 1.5811 EHF 36 -19.7227 FTER 00 0.4588 UCOMP 48 40.4127 GGHTA 36 -18.4243 FTER 48 -0.4210 PS 48 -0.1205 ASDXSQ 00	16.0 1.2 1.0 0.8 0.7 0.6 <u>0.7</u> 21.0
98-99 (20-50 km)	560.3628 0.8640 EHF 36 -26.1329 FTER 00 -21.1406 FTER 48 5.0147 TSEA 00 -3.7253 EX 48 -0.4837 PS 36	23.7 1.6 1.4 1.2 1.2 <u>0.8</u> 29.9

Table V. Same as Table IV except Tau 48 h (4102 observations).

The forecast goal is to identify the one most likely category of visibility at any location for tau 0, 24 and 48 h. However, a number of comparisons of the predictand probabilities (P) computed from each of the five regression equations indicated a less-than-desirable focusing of the most likely visibility category (i.e. the one category to be forecasted). For example, the highest computed P among the five categories did not necessarily exceed the optimal threshold probability (P_t) for that category. Here P_t (Table VI), is defined, for each visibility category and time interval, as that predictand probability which best separates forecasts of occurrence and nonoccurrence of the categorical visibility event. The P_t used here maximizes the threat score (Appendix B) for each category. These considerations led to the definition of a decision ratio as a function of P, P_t (Table VI) for each regression equation (visibility category). In the experimental form shown here, P²/P_t acts to suitably identify the most likely visibility category when P ≥ P_t; Const P_t in the denominator serves to finely tune the decision ratio for best verification.

REGRESSION EQUATION (visibility category)	DECISION RATIO	THRESHOLD VALUE TAU 0, 24, 48 h
1	P ² /P _t	57, 54, 62
2	P ² /1.1 P _t	59, 55, 60
3	P ² /0.9 P _t	45, 34, 33
4	P ² /1.1 P _t	42, 47, 39
5	P ² /P _t	49, 45, 42

2) For P/P_t < 1, use P/P_t.

Table VI. The most likely visibility category at a location is that one category which is identified by the maximum decision ratio.

