

## FORECASTING OPEN OCEAN FOG AND VISIBILITY <sup>1</sup>

Lawrence D. Burroughs  
Ocean Products Center, NMC

### 1. INTRODUCTION

A large percentage of the accidents at sea occur with visibilities under one kilometer (Tremant, 1987). Although other obstructions may lower visibility below 1 km, the most prevalent obstruction is fog. Fog is only reported if the visibility is 1 km or less. Until now there has been no objective guidance available to National Weather Service forecasters for fog or lowered visibilities. The Open Ocean Fog and Visibility Forecast Guidance System was developed to fill this need. This system is designed to provide fog and visibility guidance over the North Pacific and North Atlantic during the prime season for fog and lowered visibilities (April - September). The guidance is not applicable to coastal areas.

### 2. DEVELOPMENT

The perfect prog technique (Klein *et al*, 1959) was used to develop the fog and visibility forecast equations. With this approach all data used in the development of relationships are analyzed or observed data. Usually, the predictor and predictand are concurrent in time. When the equations are used to predict, forecast values of the predictors must be obtained and substituted into the equations to give a forecast of the predictand. The name perfect prog comes from the fact that forecast data are entered into the equation as if it were equivalent to the analyzed data that were used to develop the equations (Wilson and Macdonald, 1985).

#### 2.1 Predictand Data

The predictand data were taken from ship data for the years 1980 - 1983 at 12 hour intervals. Fog data included observed fog of any kind and observed drizzle when the past weather indicated fog and the visibility was less than 1 km. The fog was categorized as no fog or fog. Visibility data were corrected to be consistent with the observed weather and the WMO code and include poor visibility due to fog. Visibility is categorized to delineate areas with a visibility of less than or greater than 3 n mi.

#### 2.2 Predictor Data

The predictor data came primarily from the analyzed fields of the Global Data Assimilation System (GDAS) (Kistler and Parrish, 1982 and Dey and Mo-

rone, 1985) or computations made by using the GDAS data. In addition a boundary layer diagnostic model was used to create air temperatures and equivalent potential temperatures at 19 m (the nominal height of most ship instrumentation). In all, 20 basic GDAS fields and 45 computed fields were used plus four climatological parameters and two location parameters. The data were interpolated to the location of the ship data to derive the forecast equations.

#### 2.3 Regions

Burroughs (1987) describes the development of open ocean fog forecasting regions based on a National Climatic data Center fog climatology (Guttman, 1971), the frequency of fronts over the open ocean and the frequency of high and low centers over the open ocean. Thirteen warm and 13 cool season regions were delineated. These regions were then modified by the inclusion of drizzle and observed fog data for 1980 - 1984. The number of regions was reduced to those shown in Figs. 1 and 2 after test equations were derived and evaluated, and the results showed the equations for the regions excluded to be too unreliable to use.

#### 2.4 Discriminant Analysis

Equations were developed for fog and visibility with discriminant techniques. A short discussion on these techniques, following Tatsuoka (1971), is given below. For additional information see Anderson (1958), Miller (1962), or Cooley and Lohnes (1971).

In meteorology, categorical events are often related to continuous predictors by the use of discriminant analysis. The events are grouped by category. Then equations are developed by using the continuous predictors which best separate the groups.

Having developed the equations, the next question is how to classify a particular observation (vector of predictor values) into a particular group. The objective is to minimize the number of misclassifications. The chi-squared statistic is used for this purpose. The smaller the value of chi-squared for a particular group the nearer the observation is to the average value for that group. Since the equations for each group differ only by the coefficients and constants used, the observation is categorized into the group whose equation has the smallest chi-squared value. This assumes that the groups have

<sup>1</sup> OPC contribution no. 34

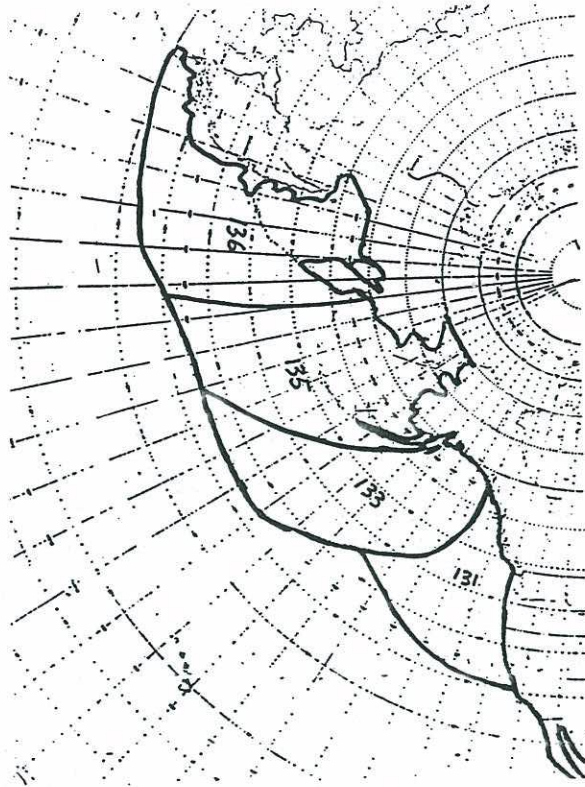


Figure 1: Warm season fog and visibility regions for the North Pacific.

multivariate normal distributions with equal dispersion matrices, and, therefore, that a pooled covariance matrix can be used to compute chi-squared for each group.

When the covariance matrix cannot be pooled, the group equations are corrected for the differences in dispersion by using the individual group covariance matrices.

Thus far the assumption has been that the relative frequencies of all the groups are equal. For fog and visibility, this is not the case. Chi-squared can be further adjusted to take into account the prior probabilities (relative frequencies of group membership) which have been determined from the dependent predictand samples.

We have already assumed that the group dispersions are multivariate normal; if we further assume that an observation fits into one of the groups and does not fall outside the ensemble of groups, then chi-squared can be related to the posterior probability that an observation belongs to a particular group. The event forecast is assigned to the group whose posterior probability is greatest.

## 2.5 Inflation and Thresholding

If the discrimination was perfect, no further adjustments would be necessary. This, generally, is not the case, and two other procedures can be used to help minimize the number of misclassifications: inflation (Klein *et al*, 1959 and Miller, 1988) and thresholding.

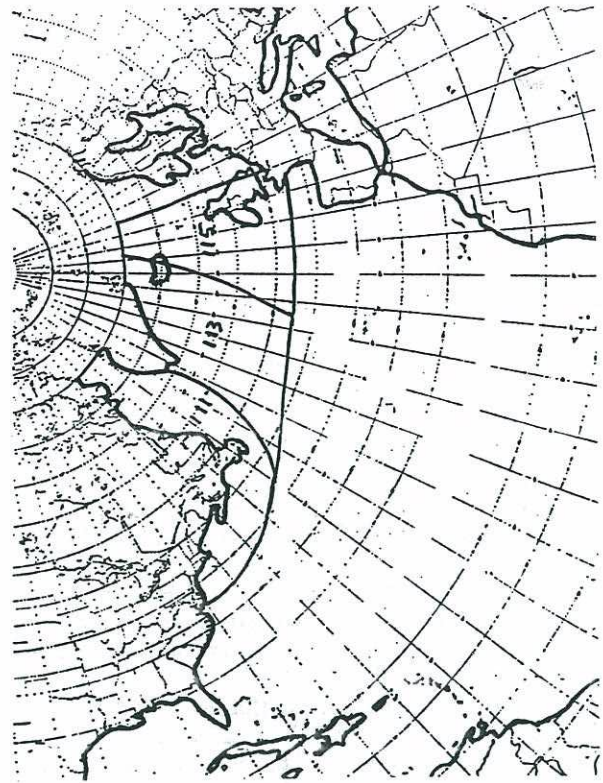


Figure 2: Warm season (April - September) fog and visibility regions for the North Atlantic.

Inflation adjusts the forecasts so that the variance of the forecasts and the variance in the climatology are approximately equal. This procedure is applied to the fog and visibility posterior probabilities prior to classification.

Thresholding means that a given category may not be chosen unless a certain predictand value is reached (in this case a given inflated posterior probability forecast). These values are normally determined empirically. The thresholds for both the fog and visibility were determined by an iterative procedure which optimized the equations for each projection, so that for each category the bias was as close to 1.0 as possible. (A value of 1.0 means the forecasts in a given category are neither over nor under forecast.) By optimizing the equations for each projection, the thresholds become model dependent. Thresholding partially accounts for differences between the developmental data set and the model output being used as input to the equations.

When a model changes significantly, or the equations are used with output from a different model, the thresholds must be recomputed. In a sense the equations, through the use of this kind of thresholding, become MOS (model output statistics)-like. Because of the periodic necessity to recompute the thresholds, there is a cost in time. But this cost is less than that required to redevelop the entire equation set periodically as would be necessary with MOS, and a stationary numerical model (one that is not greatly changing over long periods of time) is not required. The benefits of this type of thresholding in terms of better forecasts