A UNITED
DEPARTM
COMM
PUBLIC

NOAA Technical Memorandum NWSTM PR-15

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

AN EXPERIMENT IN THE PRODUCTION OF "POP" FORECASTS USING A STATISTICAL MODEL

G. HIRATA

PACIFIC REGION
HONOLULU,
HAWAII
September 1976

NOAA TECHNICAL MEMORANDUM

National Weather Service, Pacific Region Subseries

The Technical Memorandum series provides an informal medium for the documentation and quick dissemination of results not appropriate, or not yet ready, for formal publication in the standard journals. The series is used to report on work in progress, to describe technical procedures and practices, or to report to a limited audience. These Technical Memoranda will report on investigations devoted primarily to regional and local problems of interest mainly to Pacific Region personnel, and hence will not be widely distributed.

Papers 1 and 2 are in the former series, ESSA Technical Memoranda, Pacific Region Technical Memoranda (PRTM); papers 3-8 are in the former series, ESSA Technical Memoranda, Weather Bureau Technical Memoranda (WBTM); and papers 9-14 are part of the series, NOAA Technical Memoranda NWS.

Papers 1-3 are available from the Pacific Region Headquarters, Attention: OPS, P. O. Box 3650, Honolulu, Hawaii 96811. Beginning with 4, the papers are available from the National Technical Information Service, U. S. Dept. of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, VA 22151. Papercopy Price: $3.50 per copy; Microfiche Price: $2.25 per copy. Order by accession number shown in parentheses at the end of each entry.

ESSA Technical Memoranda

No. 1 The Trade Wind Regime of Central and Western Maui. Carl M. Peterson. January 1966.

No. 2 A Meteorological Glossary of Terms Used by Forecasters in Hawaii (Revised). R. F. Shaw. November 1967.

No. 3 Utilization of Aircraft Meteorological Reports at WBFC Honolulu. E. M. Chadsey, P. R. Moore, R. E. Rush, J. E. Smith, J. Vederman. June 1967.

No. 4 Tropical Numerical Weather Prediction in Hawaii - A Status Report. E. M. Carlstead. November 1967. (PB-183-621)

No. 5 A Computer Method to Generate and Plot Streamlines. Roger A. Davis. February 1969. (PB-183-622)

No. 6 Verification of an Objective Method to Forecast Frontal Passages in the Hawaiian Islands. E. M. Carlstead. September 1969.

No. 7 Meteorological Characteristics of the Cold January 1969 in Hawaii. Richard I. Sasaki. November 1969. (PB-188-040)

No. 8 Giant Waves Hit Hawaii. Jack D. Bottoms. September 1970. (COM-71-00021)

- continued on inside back cover -

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

NOAA Technical Memorandum NWSTM PR-15

AN EXPERIMENT IN THE PRODUCTION OF PROBABILITY OF PRECIPITATION FORECASTS FOR HAWAII USING A STATISTICAL MODEL

G. Hirata

September 1976

TABLE OF CONTENTS

AN EXPERIMENT IN THE PRODUCTION OF PROBABILITY OF PRECIPITATION FORECASTS FOR HAWAII USING A STATISTICAL MODEL

I-INTRODUCTION

In November 1971, the National Weather Service in Hawaii commenced production of subjective "POP (probability of precipitation) forecasts for four Hawaiian weather stations: Hilo on the island of Hawaii; Kahului, Maui; Honolulu, Oahu; and Lihue, Kauai. At approximately 1800Z (0800 HST) and 0600Z (2000 HST), forecasts of the probability of significant precipitation (amounts greater or equal to 1/100 of an inch) for three consecutive twelve-hour time periods extending out to 36 hours were made. During November, computer produced forecasts for the same stations and for the same time periods were initiated using the Navy's computer at Fleet Weather Central, Pearl Harbor. The forecasts are now routinely produced.

The original program which was developed by Richard Jones of the University of Hawaii was modified to accept "real time" data and generate the probability forecasts.

The purpose of this manuscript is to describe the method used in producing these forecasts and present some of the results of the experiment.

II-METHOD

Background

Many objective methods of probability forecasting attempt to analyze past data and, by fitting some model to the data, make current forecasts using estimated parameters. The first methods that were attempted were linear regression analyses where the variable being predicted is given the numerical value "0" or "1". For example; "0" for no significant rainfall, and "1" for significant rainfall. This "0" or "1" variable is then used as the dependent variable in a linear regression analysis. Given a new set of observations on the independent variables, a forecast of the "0" or "1" variable can then be produced. This method has the weakness of taking on values greater than one or less than zero. Another problem with linear regression is that the dependent variable is not normally distributed about the regression line so estimation procedures are inefficient and tests of significance do not hold. Many of the shortcomings of using linear regression analysis, however, can be overcome by using the logit function which was first introduced by Berkson (1) and later refined and used by Brelsford and Jones (2). A detailed description of a biometric application is also described by Jones (5).

Computer programs

The computer program used for the generation of the "POP" forecasts is made up of two parts: (1) a program for estimating the parameters in the logistic function (by maximum likelihood involving non-linear iterative techniques); and, (2) prediction of the probability of precipitation using estimated coefficients on current observed data.

(1) Estimation of parameters

Using approximately 5 years of data for each of the four stations for each month of the year, estimations of the independent parameters were derived. The independent parameters decided upon after testing were: (a) sea level pressure; (b) temperature of the air; (c) dew point; and, (d) the presence or absence of precipitation at the time of observation.

(2) Prediction of the probability of precipitation

The logit model is used for estimating the probabilities for the dichotomous random variables, that is, y=1 (significant precipitation) or y=0 (no precipitation). The model may be shown as:

$$P_r\left\{y=1 \mid x_1, x_2, x_3, x_4\right\} = \left[1+\exp(a+B_1x_1+B_2x_2+B_3x_3+B_4x_4)\right]^{-1}$$

where X_1, X_2, X_3 and X_4 are 4 independent variables and B_1, B_2, B_3, B_4 are the corresponding regression coefficients. This model is discussed in detail in reference to meteorology by Brelsford and Jones (2).

III-PROCEDURE

The "input" parameters for the prediction program are derived from the hourly synoptic reports transmitted by the four stations. These parameters are the surface or sea level pressure, temperature, dew-point and whether or not it is raining at the time of observation. For the morning forecast (18Z) the 17Z report is processed while for the 06Z forecast, the 05Z report. These parameters are processed automatically from the time of receipt to the time they are used in the computer program and the probabilities produced. The computer program takes approximately 1/2 minute to produce 12 forecasts for the various time periods and stations.

A description of the data input and output is shown in Fig. (5). These are produced each time a "run" is made and the printer output is available for quality control purposes.

IV- VERIFICATION

If the forecast chance of rain is "p" and in 100 similar situations there were 100 x p instances of rain, we then say that a perfect score was attained. For example, if a forecast for 20% probability of precipitation was forecast for 100 cases, a perfect score would be obtained only if 100 x .20 or 20 cases of precipitation occurred. If precipitation was recorded in all 100 cases, this would not have been recorded as a perfect score.

The definition of verification as given by Brier and Allen (1) is "the entire process of comparing the predicted weather with the actual weather, utilizing data so obtained to produce one or more indices or scores and then interpreting these scores by comparing them with some standard depending upon the purpose to be served by the verification." The standard adopted by the National Weather Service and used extensively throughout the different regions is the abbreviated Brier Score. In this method, a perfect score is given a value of "0" and the worst one, a "1". The scores usually distribute themselves randomly between these two extremes. One of the chief reasons for the attractiveness of the Brier Score is that it cannot be "beat", "played", or "biased". For a detailed justification of this statement, the reader is referred to Brier and Allen (3) or to Hughes (4).

V-RESULTS

Table 1 shows the statistical results for the three forecast time periods as well as for each of the four stations for which forecasts were made. It is interesting to note that the best skill scores are shown for Hilo and Honolulu in the first and second forecast periods. These skill scores show a downward trend as forecasts are made for periods greater than 24 hours. Figs 1 through 3 show how reliable the forecasts were. In almost all of the forecast periods, the probabilities were mainly in the lower and upper ranges, from 20-40% and 70-90% respectively, and these were quite reliable. In the mid-levels, however, the computer had a tendency to overforecast as indicated by the plotted graphs. (Figs 1-4)

VI-REMARKS

This preliminary study was mainly instituted to see if the determination of probabilities of precipitation by purely objective techniques is feasible. The technique presented by this manuscript, using the logit function, shows definite promise. Further study is suggested and perhaps the next step would be to compare forecasts produced by human forecasters with those generated by the computer.

VII-ACKNOWLEDGMENTS

Grateful appreciation is expressed for help rendered by Dr. Richard Jones of the University of Hawaii in the form of suggestions and computer programs. The author would also like to thank the Commanding Officer, Fleet Weather Central, Pearl Harbor for the use of the Central's computer facilities.

BIBLIOGRAPHY

(1) Berkson, J., "Maximum Likelihood and Minimum X^2 Estimates of The Logistic Function." J. Amer. Statist. Assoc. 50,130-162. 1944.

(2) Brelsford, W.M. and Jones, R.H., "Estimating Probabilities", Mon. Wea. Rev., 95, 570-576. 1967.

(3) Brier, G.W. & Allen, R.A., "Verification of Weather Forecasts Comp. of Met., 841-848

(4) Hughes, L.A., "On the Use and Misuse of the Brier Verification Score", ESSA Tech. Memo. WBTM CR-18, Aug 67.

(5) Jones, R.H., "Probability Estimation Using A Multinomial Logistic Function", Journal of Statistical Computation and Simulation, Vol 4, 1975.

APPENDIX I

FORECAST VERIFICATION STATISTICS USED IN THIS STUDY

F = Number of probability forecasts

P = Number of precipitation cases

P/F = Precipitation frequency

C_i = Monthly climatological probability at a station

P_{mi} = Monthly number of precipitation cases at a station

F_{mi} = Monthly number of probability forecasts at a station

B_{Ci} = Climatological Brier Score for 1 station for 1 month

$$B_{Ci} = (C_i - P_{mi} / F_{mi})^2 + P_{mi} / F_{mi}(1 - P_{mi} / F_{mi})$$

B_C = Climatological Brier Score derived from the weighted average of the individual climatological Brier Scores for each station and each month.

$$B_C = \left(\sum_{i=1}^{n} \sum_{j=1}^{k} F_{mij} B_{Cij}\right) / F$$

where n = the number of months
k = the number of stations

B = Brier Score

K_i = Probability value of the 13 probability classes

F_{Ki} = Number of forecasts in each of the 13 probability classes

P_{Ki} = Number of precipitation cases in each of the 13 probability classes.

B_i = Brier Score for 1 probability class for 1 station for 1 month.

$$B_i = \left[P_{Ki}(K_i - 1)^2 + (F_{Ki} - P_{Ki})K_i^2\right] / F_{Ki}$$

B = Total Brier Score weighed and summed over the 13 classes is:

$$B = \left(\sum_{i=1}^{13} FB_i\right) / F$$

S = Improvement over long term climatology

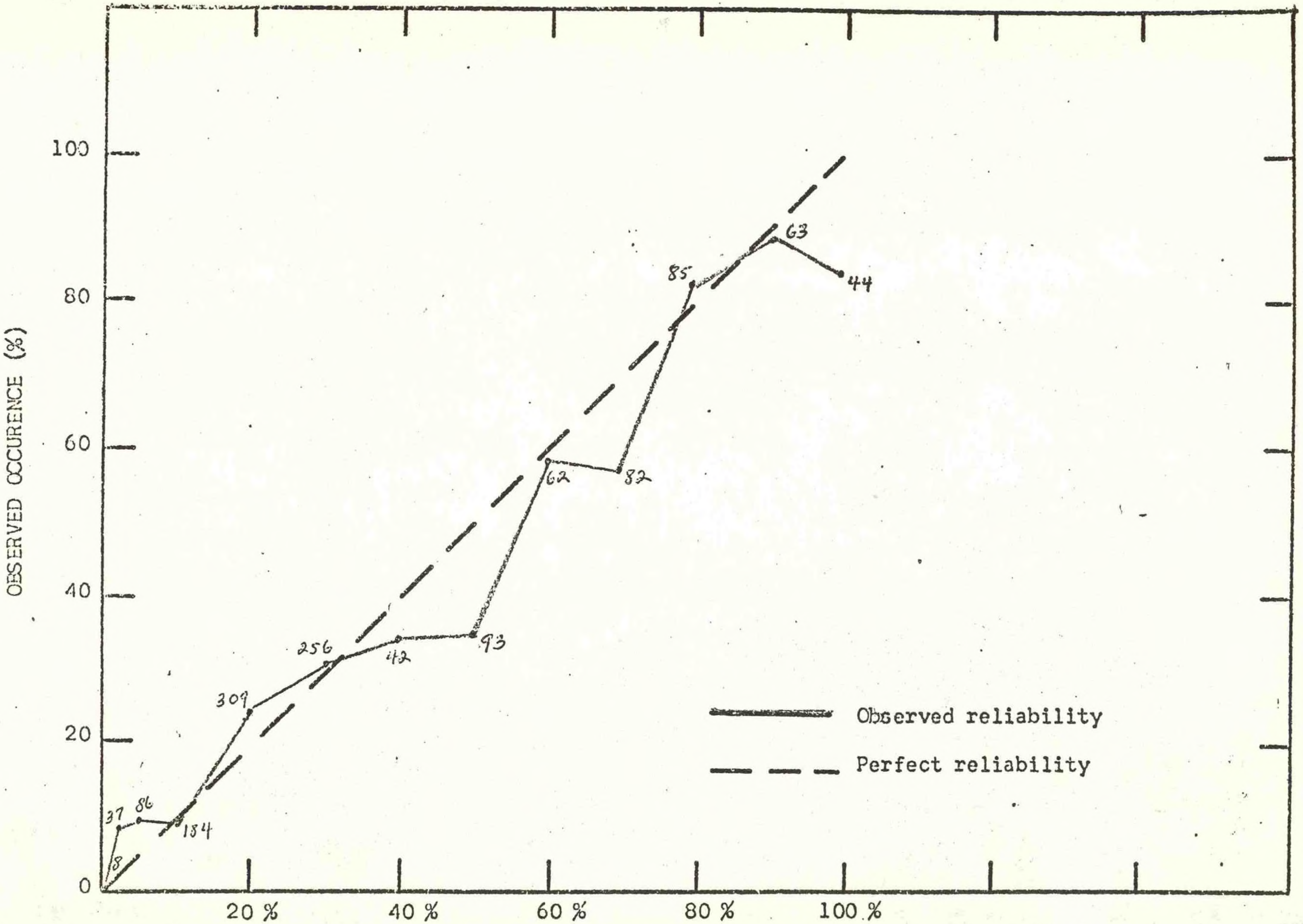
$$S = \left[(B_C - B)/B_C\right] \times 100$$

APPENDIX II

Tables & Figures

Period	Station	Number of Forecasts	Number of Precip. Cases	Frequency	Brier Score	Climat. Brier Score	% Imp. Over Climat
One	Hilo	363	235	.650	.194	.232	16.4
	Kahului	363	55	.152	.139	.144	3.9
	Honolulu	363	90	.248	.134	.163	17.6
	Lihue	362	124	.343	.218	.225	2.8
Two	Hilo	363	236	.650	.201	.231	12.7
	Kahului	363	55	.152	.131	.135	3.3
	Honolulu	363	71	.196	.159	.161	1.5
	Lihue	362	124	.343	.225	.224	-0.3
Three	Hilo	363	235	.647	.222	.231	3.9
	Kahului	363	55	.152	.165	.147	-12.3
	Honolulu	363	71	.196	.170	.161	-5.3
	Lihue	362	123	.340	.225	.224	-0.3

TABLE (1) STATISTICAL SUMMARY PERIOD NOVEMBER 1971 THROUGH APRIL 1972. 18Z and 06Z FORECASTS COMBINED.



FORECAST PROBABILITY

FIRST PERIOD

FIG 1. Forecast probability vs. observed relative frequency of occurence by computer for November 1971 through April 1972. Number of forecasts is given next to each data point. All stations combined. 06Z and 1800Z forecasts combined.

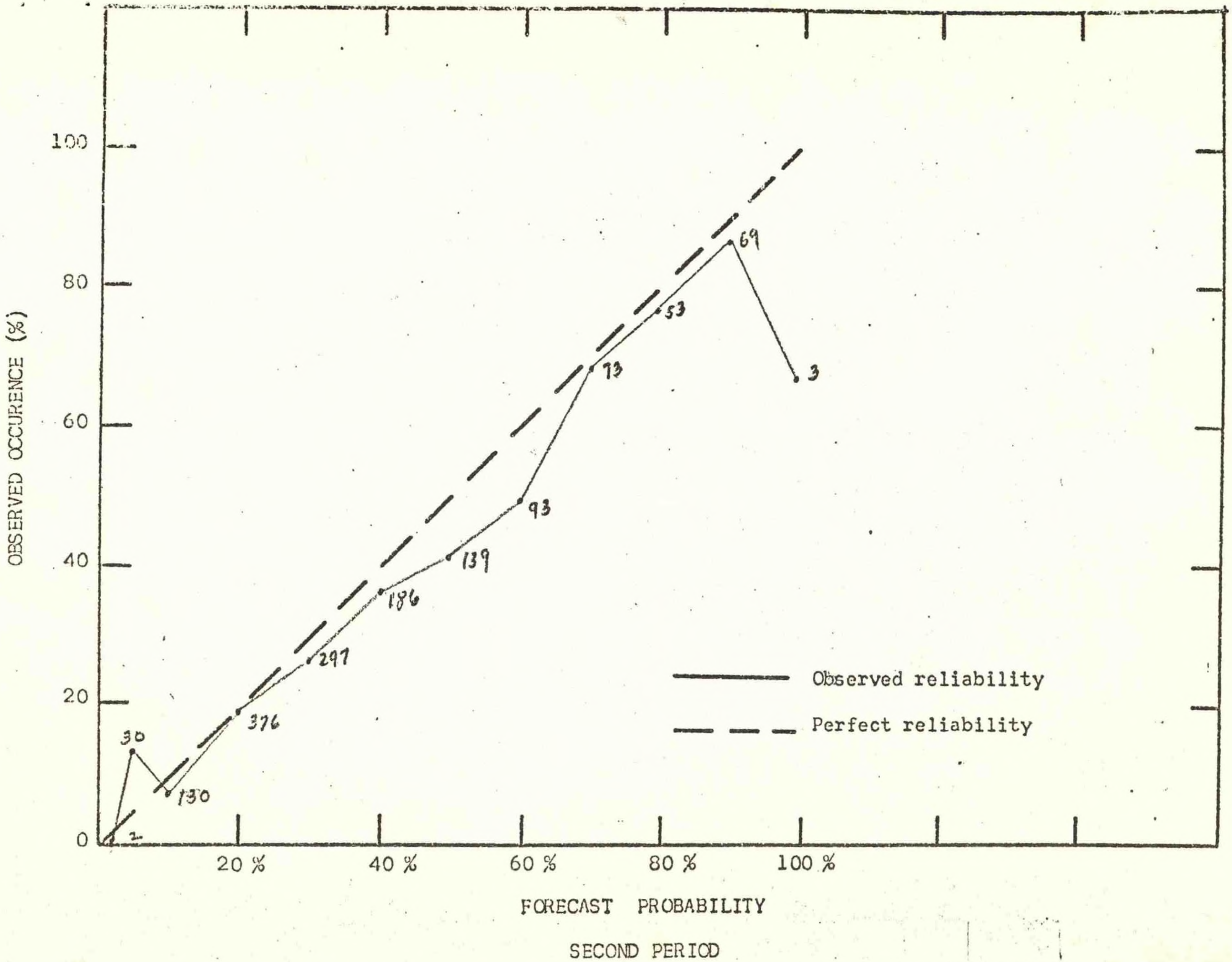


FIG 2. Forecast probability vs. observed relative frequency of occurency by computer, November 1971 through April 1972. Number of forecasts is given next to each data point. All stations combined. 06Z and 1800Z forecasts combined.

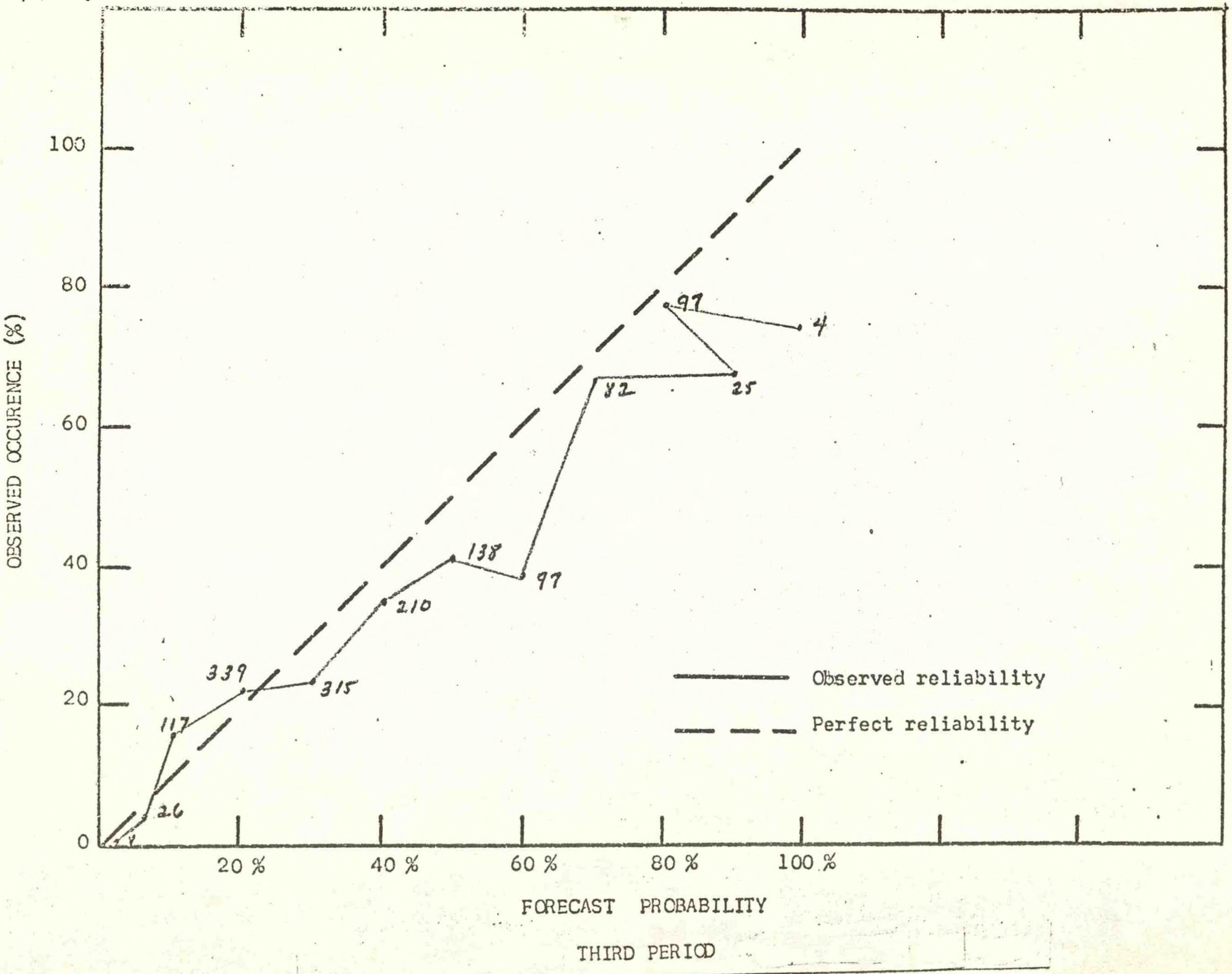


FIG 3. Forecast probability vs. observed relative frequency of occurence by computer, November 1971 through April 1972. Number of forecasts is given next to each data point. All stations combined. 06Z and 1800Z forecasts combined.

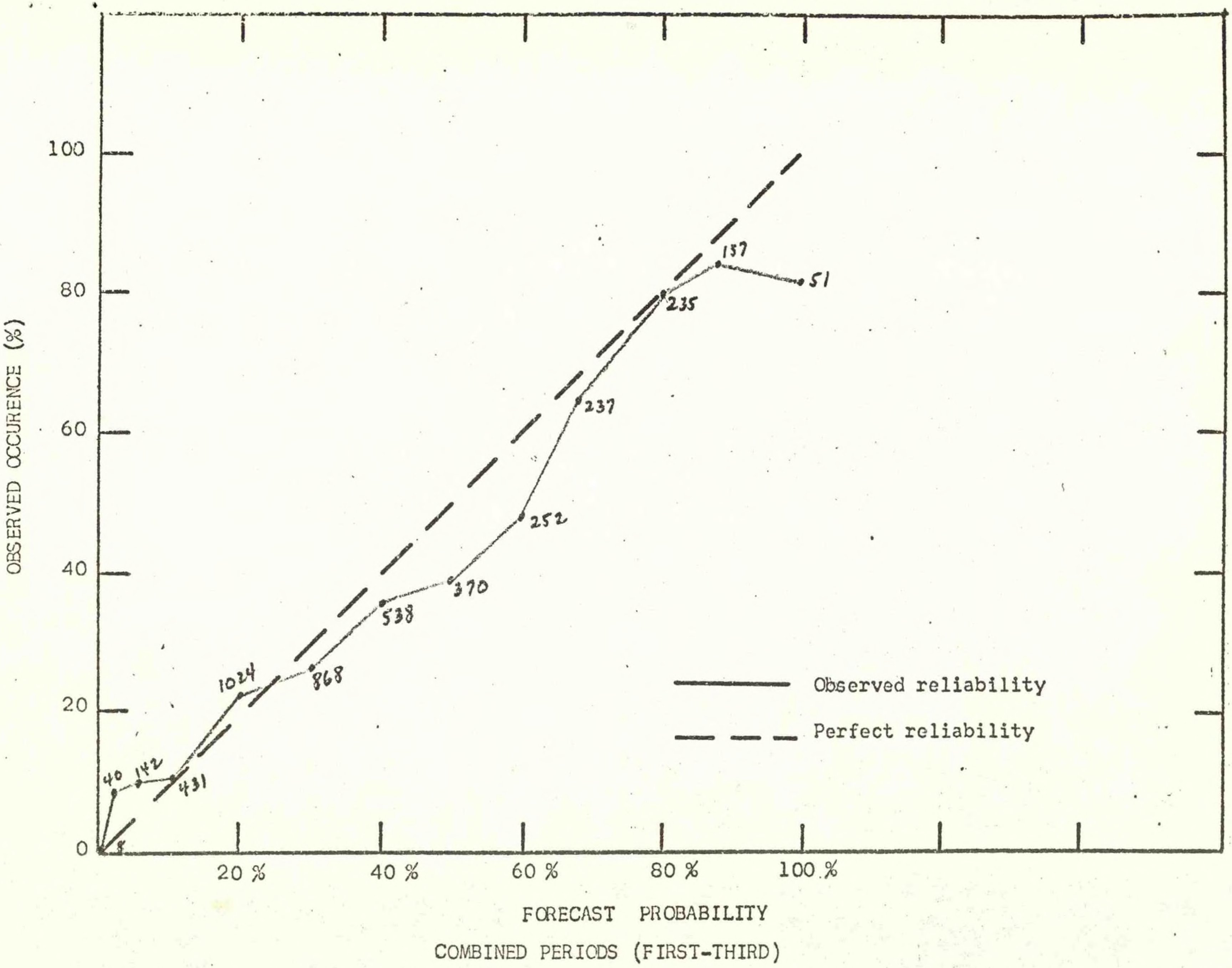


FIG 4. Forecast probability vs. observed relative frequency of occurency by computer for November 1971 through April 1972. Number of forecasts is given next to each data point. All stations combined. 06Z and 1800Z forecasts combined.

```
**************************************************HPOP DATA  DATE/TIME 05160475
SAHW1 PHNL 1605
                         PHLI 23%50%15 186/72/65/0815/009/ RWU SW ALG MTNS RB40E
                         PHNL E29,45,25 171/74/61/0512/004/RB03E15B36 OCNL R
                         PHOG 25%E48,20≠ 166/74/56/0819G25/002/ WND 050V
                         PHKO 40%E65,10 1606
                         PHTO E16,40[10 187/68/62/2406/008/RWU ALQDS RE20 BT
 5  4 16 75
          285 187 68 62 0   4
          190 166 74 56 0   4
          182 171 74 61 0   4
          165 186 72 65 0   4

          ****HPOP4A**

FXHW PHNC 160500
ATTN--WEATHER FCSTR
```

PROB OF PRECIP FOR FOLLOWING STNS FROM 06Z 16 APR 75

STATION	12HR FCST	24HR FCST	36HR FCST
HILO, HAWAII(285)	80	80	80
KAHALUI, MAUI(190)	2	10	20
HONOLULU, OAHU(182)	20	10	20
LIHUE, KAUAI(165)	50	20	50

JOB DONE

FIG 5 SAMPLE OF THE INPUT PARAMETERS (UPPER PORTION OF SHEET) AND THE OUTPUT OF PROBABILITY FORECASTS (LOWER PORTION OF SHEET).

NOAA Technical Memoranda NWS

No. 9 Tropical Numerical Weather Prediction in Hawaii - 1971. E. M. Carlstead. March 1971. (COM-71-00494)

No. 10 Climatology of Rainfall Probabilities for Oahu, Hawaii. A. N. Hull and Jon Pitko. April 1972. (COM-73-10242)

No. 11 A Cirrus Climatology for Honolulu. Clarence B. Lee and Wesley Young. April 1974. (COM-74-11244).

No. 12 Straight Line Wind Variability Over Selected Stations on Leeward Oahu. Michael J. Morrow. July 1974. (COM-74-11669)

No. 13 Forecasting Hurricanes in the Central Pacific. Paul Haraguchi. October 1975. (PB 248-371)

No. 14 Trade Wind Speed Estimation at Selected Stations on Oahu Using Honolulu Wind Observations, A Pilot Study. Michael J. Morrow. (PB 251-685)