United States
Environmental Protection
Agency

Research and Development

Environmental Monitoring Systems
Laboratory
Las Vegas NV 89114

EPA-600/S4-83-035 Oct. 1983

EPA

Project Summary

Statistical Correlations of Surface Wind Data: A Comparison Between a National Weather Service Station and a Nearby Aerometric Monitoring Network

John E. Langstaff, Anthony D. Thrall, and Mei-Kao Liu

This report presents a statistical analysis of wind data collected at a network of stations in the Southeast Ohio River Valley. The study determines the extent to which wind measurements made by the National Weather Service (NWS) station at the Tri-State Airport can be used to estimate the wind measurements at network stations. A combined stratification/regression analysis was conducted. The analysis shows that NWS station measurements can be used to gain insight into the wind measurements at network stations, and a methodology is identified for carrying this out. With this methodology, we demonstrate that the wind data collected at the airport can be used to provide input to a complex-terrain wind model for estimating the surface wind in the study area for periods prior to the establishment of the monitoring network.

This Project Summary was developed by EPA's Environmental Monitoring Systems Laboratory, Las Vegas, NV, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

Monitoring of human exposure to toxic and hazardous chemicals released to the environment can be expensive and time consuming, particularly if extensive meteorological data must be acquired to support modelling of pollutant distributions. Cost effectiveness could be improved if data from existing networks such as National Air Surveillance Network (NASN) and/or National Weather Service (NWS) stations could be used to estimate surface wind measurements in a given study area.

Recently, the Environmental Monitoring Systems Laboratory, Las Vegas, of the U.S. Environmental Protection Agency (EPA), conducted a field measurement program in the Southeast Ohio River Valley in support of the design and development of an exposure assessment monitoring network. As part of this study, surface wind data were collected from a network of stations temporarily established in the Southeast Ohio River Valley and compared with wind measurements made by the National Weather Service (NWS) station at the Tri-State Airport.

Weather observations recorded from the NWS stations at the airport provide an extended and continuous history at a single location. Therefore, a primary objective of this study was to examine the corrections, if any, between the NWS wind measurements and those made at other locations.

If the existence of a statistically significant correlation between data measurements made at Tri-State Airport and those made at other locations could be established, the NWS data could be utilized to derive wind patterns in prior years. These synthesized wind patterns could then be used in a complex terrain wind field model

to reproduce the detailed spatial distribution of the wind field.

Study Area

The study area encompasses the tri-state junction of West Virginia, Kentucky, and Ohio along the Ohio River. The area covers over 250 square miles and contains approximately 160,000 people. The major population centers are Huntington, Ashland, and Ironton. The study area was divided into a 44 x 33 grid of 1 km x 1 km cells. The origin of the grid (i.e., the lower left hand corner) was placed at 347000 E and 4243000 N in Zone 17 of the UTM coordinate. Wind monitoring stations were installed at 12 sites in the study area. The names and UTM coordinates of the network stations (as well as the NWS stations) are given in Table 1.

Data Analysis

For each of these 12 sites, hourly average wind speeds and directions were calculated along with their standard deviations, the latter being computed from instantaneous observations every two minutes. The data used in this analysis cover the period February 1, 1980 through February 28, 1981, except for periods where data were missing. The distribution of wind speeds and directions for these 12 sites was plotted as wind roses (monthly).

Two types of analyses were conducted to examine the relationship among wind measurements collected at different monitoring sites. First, an analysis of the correlation between stations was conducted using the full data set. Second, after the data were stratified into bins based on wind direction, a regression analysis was performed to 1) determine the correlation of the NWS station measurements with those taken at the other 12 stations, and 2) provide a procedure for calculating the wind and speed directions at these 12 stations.

The results of a linear regression analysis of the NWS station on each of the 12 stations of the network are given in Table 2.

The degree of correlation evident in Table 2 indicates that prediction errors can be significantly reduced by regressing the station wind field on NWS measurements rather than by estimating the wind field solely on the basis of station averages. Further improvements are possible by stratification of the data.

First, the data were stratified into bins on the basis of the NWS wind speed and direction. Then, for each bin and each of the 12 network stations, a linear regression was performed using NWS wind speed

Table 1. *Names and Coordinates of Meteorological Stations*

Site	UTM Coordinates East	UTM Coordinates North
Ashland Business College (ABC)	353,841	4,260,707
Ashland Synthetic Fuels (ASF)	360,049	4,249,390
Ashland City Building (ASH)	357,195	4,259,951
Bamer Residence (BAM)	355,458	4,273,024
Condit Elementary School (CON)	356,915	4,257,829
Fire Station No. 2 (FIR)	359,756	4,255,098
Flatwoods (FLA)	352,012	4,266,646
Huntington Water Corporation (HUN)	376,500	4,254,412
KEN Department of Human Resources (KEN)	354,524	4,257,122
Ohio Department of Transportation (ODT)	355,049	4,263,500
Sunrise Hill (SUN)	365,500	4,260,085
Worthington (WOR)	348,670	4,268,183
Tri-State Airport (NWS)	363,750	4,247,500

Table 2. *Regression Results for Each Network Station, Based on NWS Station Data*

(a) Wind Speed

STATION	CORRELATION	SLOPE	INTERCEPT	MEAN	STANDARD DEVIATION
ABC	0.70	0.71	-0.12	2.1	1.60
ASF	0.66	0.57	0.34	2.4	1.39
ASH	0.70	0.85	-0.53	2.5	1.91
BAM	0.79	0.49	-0.11	1.5	1.00
CON	0.76	0.89	-0.38	2.4	1.84
FIR	0.55	0.37	-0.06	1.0	1.02
FLA	0.77	0.78	-0.56	1.8	1.60
HUN	0.78	0.65	-0.27	1.6	1.20
KEN	0.77	0.70	0.54	2.7	1.43
ODT	0.76	0.62	0.15	2.1	1.28
SUN	0.63	0.65	1.09	3.2	1.63
WOR	0.75	0.81	-0.37	2.3	1.76

(b) Wind Direction

STATION	SLOPE	INTERCEPT
ABC	1.00	-11.00
ASF	1.00	1.74
ASH	1.00	-13.05
BAM	1.00	0.22
CON	1.00	-16.55
FIR	1.00	-13.50
FLA	1.00	-7.00
HUN	1.00	-6.37
KEN	1.00	-5.09
ODT	1.00	-8.66
SUN	1.00	14.94
WOR	1.00	-22.59

(c) u^1

STATION	CORRELATION	SLOPE	INTERCEPT	MEAN	STANDARD DEVIATION
ABC	0.75	0.59	-0.34	-0.5	1.80
ASF	0.74	0.61	0.09	-0.1	1.90
ASH	0.72	0.52	0.01	-0.1	1.73
BAM	0.80	0.37	0.06	0.1	1.09
CON	0.79	0.69	-0.12	-0.2	1.98
FIR	0.81	0.51	0.07	-0.0	1.47
FLA	0.74	0.49	0.21	0.1	1.52
HUN	0.70	0.40	0.11	-0.0	1.25
KEN	0.84	0.74	-0.22	-0.4	2.01
ODT	0.78	0.55	0.20	0.1	1.60
SUN	0.75	0.71	-0.26	-0.4	2.25
WOR	0.61	0.45	-0.01	-0.1	1.71

Table 2. (continued)

	(d) v^1				
STATION	CORRELATION	SLOPE	INTERCEPT	MEAN	STANDARD DEVIATION
ABC	0.80	0.55	-0.36	-0.8	1.73
ASF	0.61	0.35	-0.64	-1.1	1.65
ASH	0.67	0.58	-0.22	-0.8	2.54
BAM	0.87	0.44	-0.11	-0.5	1.40
CON	0.75	0.66	-0.07	-0.7	2.25
FIR	0.52	0.13	0.21	0.1	0.63
FLA	0.76	0.57	-0.10	-0.5	1.87
HUN	0.85	0.56	-0.07	-0.4	1.57
KEN	0.87	0.74	-0.24	-0.8	2.15
ODT	0.81	0.59	0.09	-0.4	1.81
SUN	0.81	0.80	-0.24	-1.1	2.65
WOR	0.78	0.65	-0.02	-0.6	2.28

[1] u and v are orthogonal wind velocity components.

and direction as independent variables. This approach was adopted in order to improve the performance of the linear model.

The predicted values were obtained using the regression equation for each bin. The RMSE and the interquartile range provide measures of the degree of spread of observed values about the predicted values. For the prediction to be reliable, there must be a sufficient number of measurements in each bin. This constrains the number of bins that can be used. Figure 1 illustrates the flow of data in processing and analysis.

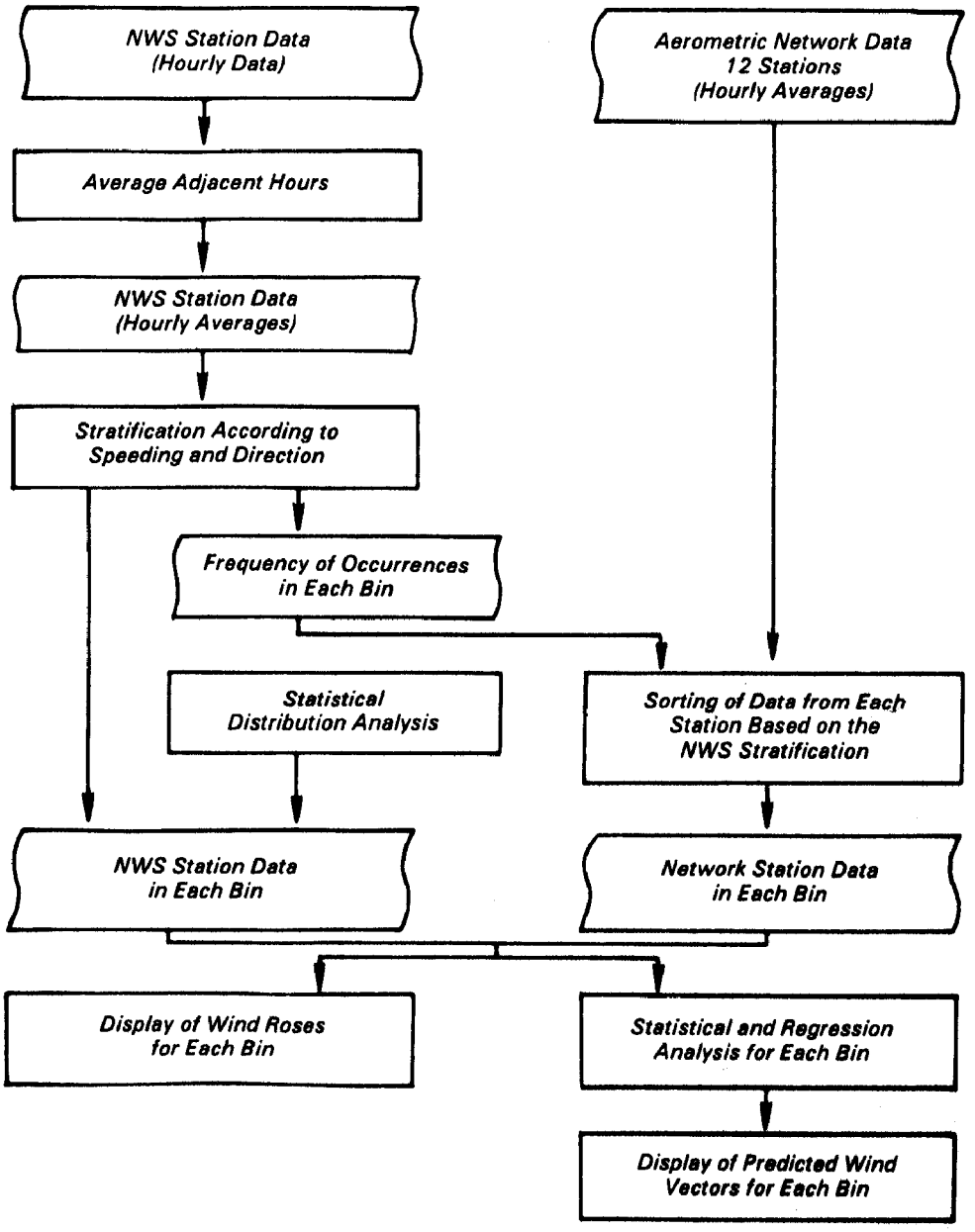

Figure 1. *Flow chart of data processing and data analysis.*