Spatial Interpolation of Meteorological Data in Complex Terrain Using Temporal Statistics

WILLIAM PORCH* AND DANIEL RODRIGUEZ

Lawrence Livermore National Laboratory, Livermore, CA 94550

(Manuscript received 20 February 1987, in final form 6 June 1987)

ABSTRACT

Diagnostic wind field numerical models have significant difficulty developing representative wind velocities in complex terrain. A large part of this difficulty begins with the initial wind field interpolation. If this interpolated wind field does not closely represent the true winds, mass-consistent adjustments cannot retrieve the correct atmospheric flow patterns. Presently, the initial interpolation in diagnostic models is almost exclusively done using a simple one-over-separation-squared ($1/r^2$) interpolation algorithm. This algorithm uses the closest or larger number of measurements of the wind velocity closest to the interpolation location. In this paper, we explore different interpolation algorithms using not only the measurement field at the interpolation time, but also the statistical relationships between stations. These algorithms were tested with data from the 1980 Atmospheric Studies in Complex Terrain (ASCOT) sponsored by the Department of Energy. The results show that, while consistent (though in most cases marginal) improvement in interpolated wind speeds was obtained, little improvement was derived for interpolated wind direction.

1. Introduction

Between 1979 and 1981, the Department of Energy sponsored a series of atmospheric studies in complex terrain (ASCOT) whose main objective was to increase fundamental knowledge of transport and dispersion processes in complex terrain. The life cycle of nocturnal cold air drainage flows was of special interest because of its importance as a mechanism for transporting pollutants from energy-related facilities to populations and agricultural centers that are often situated in valleys (Dickerson and Gudiksen, 1983). This phenomenon was studied in the Geysers geothermal area of northern California by means of a series of intensive field experiments involving the release of perfluorocarbon and heavy methane tracers. We have used this dataset to test several spatial interpolation algorithms for wind speed and direction. Previous interpolation algorithms have concentrated on the use of historical data to improve a general weighting function rather than the square of the separation distance (Gandin, 1965). Goodin et al. (1979) extended the weighting function to a two-dimensional polynomial without using the historical record. Johnson (1982) has used the historical data to interpolate climatological functions of a region. Effects of topography have also been included in objective interpolation schemes for regions of special interest for wind power applications (Erasmus, 1986).

Examples of interpolation problems in complex terrain include the interpolation of data from two adjacent valleys or data from a valley and along a ridgeline that may have been obtained from different flow regimes. Presently, modelers have no way to account for these differences. This often leads to an arbitrary mixing of data during the interpolation phase of a calculation. Interpolation techniques that couple these kinematically different regimes in a physically realistic manner are needed to initialize wind field models better.

The efforts reported in this paper represent a preliminary attempt to find and test methods of wind interpolation that are superior to the current distance weighting approach. Specifically, a technique is explored for developing predictor equations for winds at selected locations based on statistical regression between station observations. The statistical relationships were developed and tested using the ASCOT database described in section 2. Comparisons of the impact of the interpolation differences were tested using the tracer data and MATHEW/ADPIC (Sherman, 1978; Lange, 1978), a diagnostic particle-in-cell numerical air quality assessment model.

2. Description of database

A series of field studies were conducted in the complex terrain setting of the Geysers geothermal area of northern California in 1979, 1980 and 1981. The 1979 studies were exploratory in nature; they were used to plan a more intensive series of experiments during September 1980, whose primary purpose was the acquisition of information about the spatial and temporal characteristics of drainage flows. Because of interest in drainage winds, the greatest attention was given to

* Present affiliation: Los Alamos National Laboratory, Los Alamos, NM 87545.

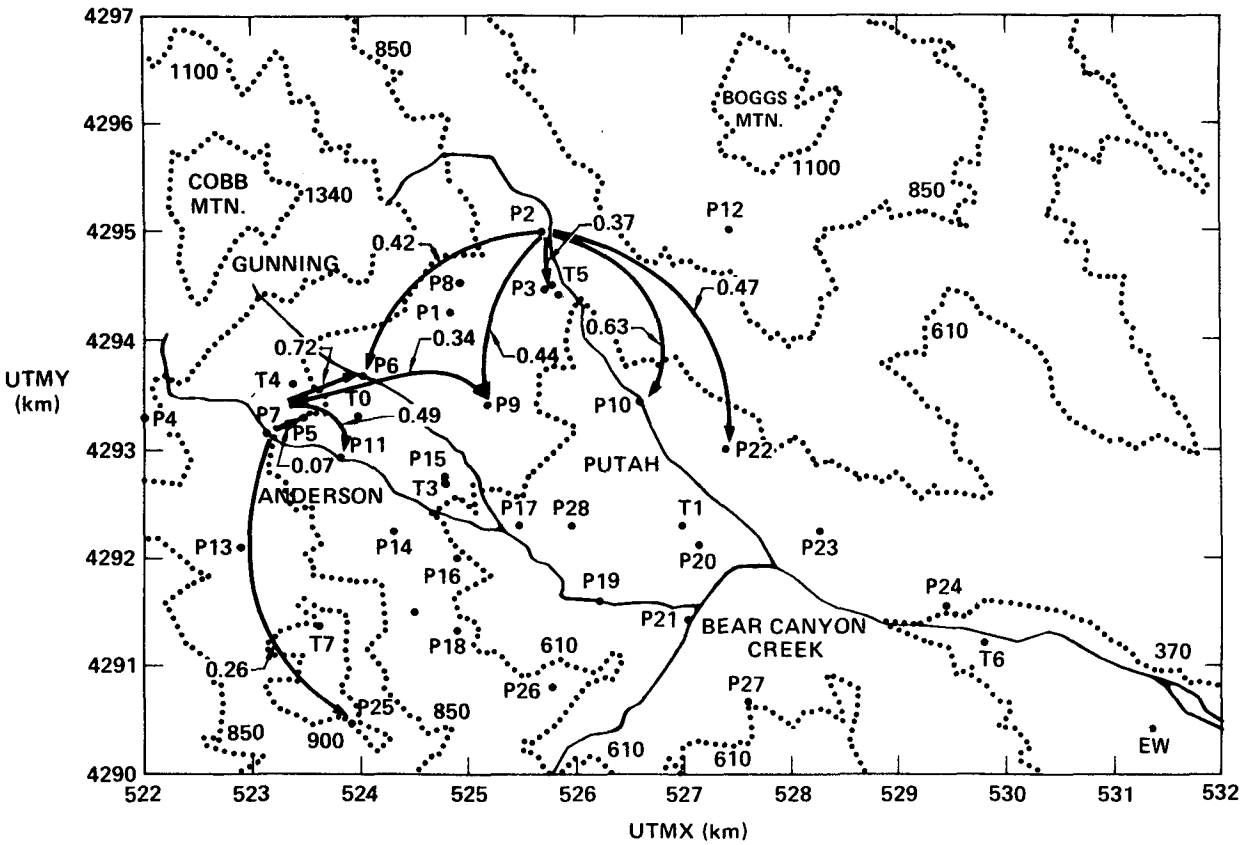

FIG. 1. A map of the Geysers region of northern California with height contours in meters and the locations of 27 PAM meteorological towers and 7 long-term towers (marked with P and T, respectively). The long-term towers were used in this paper as test locations for interpolation procedures. Stations P2 and P7 are at locations of tracer releases and arrows point to towers whose wind parameters most often entered regression equations for reproducing data at the tracer release points. Correlation coefficients of wind speeds are associated with the arrows.

nighttime conditions (Gudiksen and Walton, 1981). However, a continuous record of meteorological observations was made over an 11-day period starting 9 September 1980. This paper concentrates on a subset of the wind speed and wind direction data obtained at 5 m above ground level by a network of 27 "portable automated mesonet" (PAM) systems. These systems were deployed in a 5 km × 10 km area in the Geysers region. Stations P2 and P7 (see Fig. 1) were the focus of regression studies because they were closest to the two tracer release points; it would be reasonable to assume that an accurate prediction of winds at P2 and P7 would greatly influence the performance of diagnostic models.

Wind field models like MATHEW assume that the wind conditions persist over a specified period of time, typically 1 h. Consequently, the one minute PAM data were averaged over an hour which also made them more consistent with the tracer averaging times. Several other factors which might have affected the success of regression between stations were also considered. These factors included the time of day, the strength of large-scale (synoptically forced) winds and the down- and cross-slope components of flow. The separation of data by time of day was based on whether the measurements were taken during the night (after 1800 PST and before 0600 PST) or during the day (the remaining hours). The relative strength of the large-scale wind was estimated by calculating the median wind speed at a ridge-top station (P4). The PAM data were then separated into two groups depending on whether the hourly averaged wind speeds fell above or below the median value at P4. Finally, the wind speeds and wind directions within the four groups were expressed in terms of u and v wind components and then rotated through a locally unique angle and scaled to provide cross- and downslope components at each station. The ASCOT field experiments emphasized the documentation of cold air drainage and morning upslope wind conditions. The downslope directions were determined from the mean wind directions over a 5-h period on a night for which the drainage winds were particularly well

developed (11–12 September 1980). The winds were reasonably steady at all but six locations where, in some cases, the hour-to-hour variations in wind direction approached 100 degrees. For these latter stations, which gave every indication of being located outside the drainage layer, downslope directions were assigned on the basis of values estimated from topographic contour maps.

3. Statistical relationships of wind in complex terrain

The statistical analyses of the PAM data were performed using a statistical software package which included stepwise linear multiple regression. The process of analysis began with the identification and rejection of encoded data that signified missing observations. This step was followed by the averaging of data over hourly intervals which produced records containing as many as 264 wind speeds and wind direction values for each station. The PAM stations P2 and P7 were the focus of the regression study. Their measurements were paired in time with the remaining station observations (e.g., the wind speeds at P2 were compared against the corresponding ones at P3) to develop regression equations based on a straight-line fit. The arrows emanating from P2 or P7 in Fig. 1 (with attached correlation coefficients) point to stations selected from a five-station linear regression calculation. Obviously a simple inverse distance weighting scheme for the interpolation of data may be inadequate in complex terrain. Alternatively, these correlations suggest that statistically based weights for interpolation may offer some improvement.

The method we used for building a predictor model employs a linear regression equation for a particular response (wind speed or wind direction at P2 or P7) in terms of the independent variables (data from the remaining 25 stations). However, evaluating the final equation is not an easy process since it involves two opposing criteria. The model should include as many independent variables as possible to ensure an accurate fit to the dataset used to determine the regression. However, inclusion of too many variables can degrade the regression of an independent dataset.

The regression procedure used in this analysis was stepwise regression (Draper and Smith, 1966). To determine the optimal number of variables in the stepwise procedure for complex terrain velocity data, a random number generator was used to divide the wind speed records at all stations into two groups of equal size. These partial records were then combined to form "training" and "test" sets; the training set was used to develop regression relationships for the prediction of wind speeds at P2. The results of the stepwise analysis, expressed in terms of the correlation between measurements and predictions as a function of the number of stations incorporated into the regression equation,

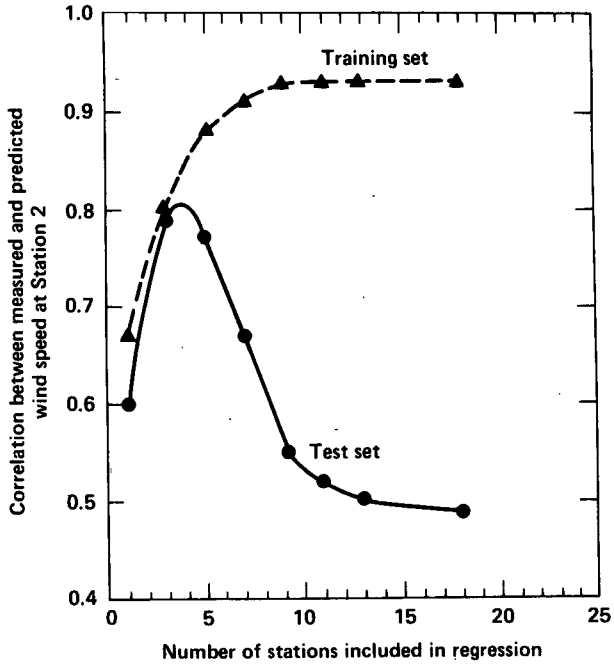

FIG. 2. Plot of the effect of increasing number of parameters used in the regression of wind speed at station P2 on the correlation of measured and predicted speed in a randomly selected training and test dataset.

are presented in Fig. 2. As might be expected, the addition of more and more stations offered a diminishing return in terms of their positive contributions to the correlation. Having developed a predictor equation for wind speeds at P2 containing 18 terms, it was then applied to the test set on a term-by-term basis starting with the most highly correlated terms. The result, shown in Fig. 2, suggested that four to five parameters should be used in the prediction of wind speed at P2.

The behavior exhibited by the test curve in Fig. 2 was found to be repeatable when other factors—time of day, strength of large-scale winds and direction of flow relative to the local (terrain) slope at each station—were considered. Table 1 shows the correlation between measured and regression-predicted wind speeds and directions at stations P2 and P7. Predictions using only the five most highly correlated terms for each of the eight categories in Table 1 were made in terms of u and v wind components. These components were transformed to wind speeds and wind directions and then plotted against the corresponding measurements in a scatter diagram. The entries in Table 1 are the computed correlations based on a linear fit through the scatter of points. The following conclusions can be drawn from this table:

1) Wind speeds are easier to predict at night than during the day at station P7 but not at station P2.

2) Regressions based on data which have been separated into high and low ridge-top wind speed groups

TABLE 1. Correlation of measured versus regression predicted values (number of hours included in regression).

	Using all data available				Using separate regressions on high and low winds at ridge and then combined			
	Using u and v components		Using cross and downslope components		Using u and v components		Using cross and downslope components	
	Day	Night	Day	Night	Day	Night	Day	Night
				Station 7				
Speed	0.70	0.82			0.74	0.91		
	(183)	(155)			(165)	(147)		
Direction	0.97	0.56	0.98	0.59	0.97	0.69	0.99	0.75
	(110)	(60)	(110)	(60)	(74)	(53)	(74)	(73)
				Station 2				
Speed	0.70	0.69			0.73	0.73		
	(166)	(149)			(165)	(138)		
Direction	0.94	0.81	0.94	0.74	0.95	0.77	0.90	0.79
	(109)	(74)	(109)	(74)	(124)	(96)	(110)	(97)

and then recombined tend to improve the predictions at P2 and P7.

3) Separate consideration of winds at the ridge and revision of the coordinate system for down- and cross-slope components usually, but not always, improves the prediction of wind direction. This improvement is usually less than the effect of inclusion of ridge speed grouping.

Examples of the scatter plots of predicted and measured wind speeds and directions at P2 for the nighttime high and low wind speed category (whose correlations are documented in Table 1) are provided in Fig. 3. If we assume that the distribution of points in Fig. 3 is typical of the unpredictable nature of wind direction in complex terrain, and that the spread about the regression line represents four standard deviations, then variability in wind direction is approximately 10° (one standard deviation). Errors of this magnitude can have a significant effect on the performance of numerical models that simulate the transport of pollutants. Since the negative effect is most pronounced when the errors in wind direction are introduced near the point of effluent release (King and Bunker, 1984), it becomes critically important to interpolate these data correctly at P2 and P7 because of their proximity to the tracer sources in the Geysers area.

4. Statistically based interpolation

From the temporal statistics whose description was given in section 3, there now exists both measured wind speed and direction and statistically predicted values at each measurement location. The basic hypothesis of our interpolation is that the statistically predicted values become more representative at greater distances from these locations. This is similar to (but not identical with) the "optimal interpolation" procedure described by Gandin (1965). Figure 4 illustrates a simple method for accomplishing this weighting as a function of distance between an arbitrary point and the closest two measurement locations. The predicted and measured values of wind speed or wind speed components at the closest measurement station are represented by P_1 and M_1, respectively; P_2 and M_2 usually correspond to the next closest measurement station; R_c is introduced as a critical distance at which the measured and predicted values at the two stations are weighted equally. Several combinations of these interpolation parameters were tested with linear and distance squared weighting. The most successful of the algorithms used the properties of the hyperbolic tangent function for the weighting. (However, we are certain many other weighting schemes could do as well.) The interpolation formula then becomes for the parameter of interest at an arbitrary point o:

$$V_p = \Big\langle \sum_i \{P_i * \tanh(0.549 r_i/R_c) + M_i[1 - \tanh(0.549 r_i/R_c)]\} 1/r_i^n \Big\rangle \Big/ \sum_i 1/r_i^n, \quad (1)$$

where for the cases tested $i = 2$ and $n = 1$. The value of R_c was chosen as 2 km based on the smaller scale peak in the two-dimensional Fourier transform of the terrain height database for the Anderson Creek region in the Geysers (Porch, 1982).

This interpolation algorithm was then tested for its ability to predict wind speed and direction at six tower locations whose data were withheld from the original statistics based on the 27 PAM stations. The interpolated values were tested against measured values at the six towers and compared for overall correlation against simple one-over-distance-squared interpolation. Tables 2 and 3 show a comparison of the overall correlations of the two interpolation techniques with the measured wind speed under conditions of higher- and lower-than-

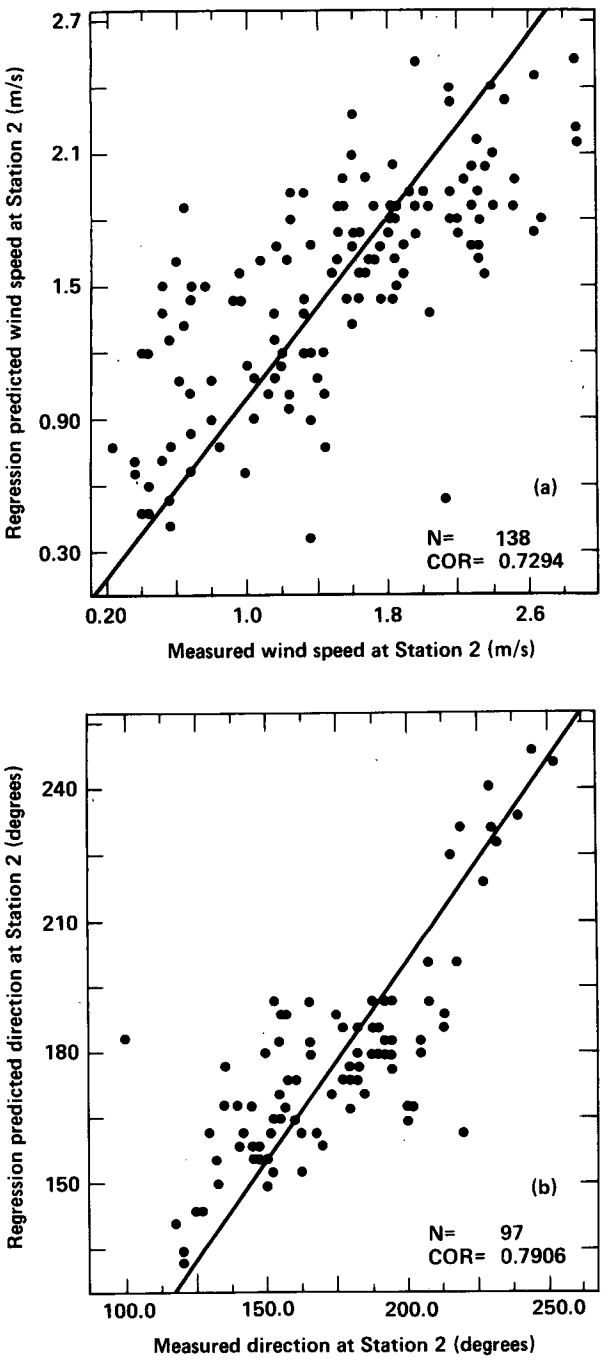

FIG. 3. Scatter plots of (a) predicted versus measured wind speed and (b) direction for nighttime values at station P2 including combination of separate regressions on groups with high and low ridge-top wind speeds.

mean speed observed at the ridge location P4. The data were stratified to separate high and low synoptic wind conditions. As shown in Table 1, this significantly improves the statistical predictions.

The data used for the comparisons shown in Tables 2 and 3 were also stratified by comparing the correlations using the closest two stations and the closest station and the station whose wind speeds provided the highest correlation. Finally, the correlations were separated based on all the appropriate data and for situations where the predictability of the wind speed at the two stations was above and below the mean. The measure of predictability was the sum of the absolute values of the difference between the predicted and measured values at the two locations divided by the sum of these values. Tables 2 and 3 show that the increase in correlation is very small. The largest increases at station T5 are only statistically significant at the 25% to 50% level. This is due, in part, to the difficulty of interpolation in complex terrain and the fact that most of the six test stations were in close proximity to at least one of the 27 stations used to develop the statistical interpolation. There is usually some improvement associated with interpolations based on the closest station and the one most correlated to it than the closest two stations. This is true for both the statistically based as well as the simple $1/r^2$ interpolation. The differences are greater when the data are stratified by the predictability of wind speed at the measurement locations (though the database is only half as large). However,

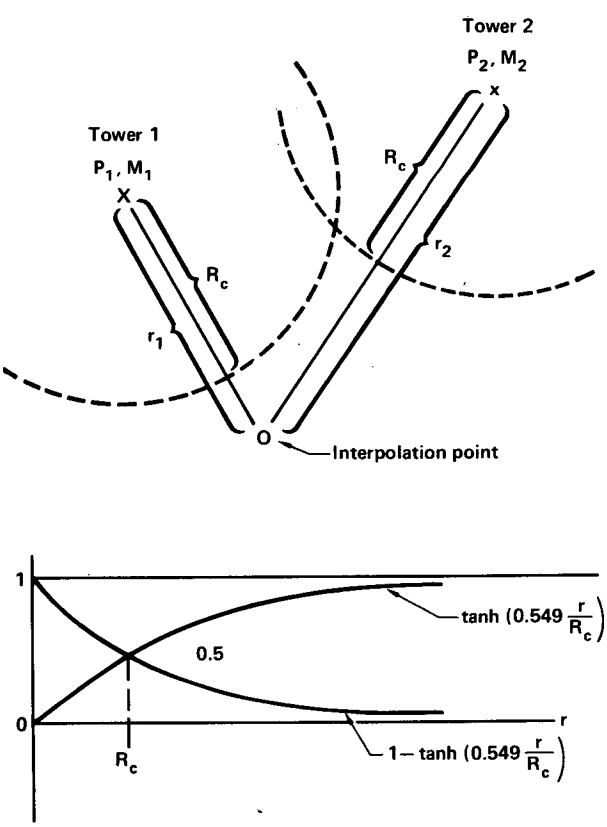

FIG. 4. Illustration of geometry of statistical interpolation between two stations each with predicted and measured wind parameters P_1, M_1 and P_2, M_2 along with the hyperbolic tangent algorithm incorporating the critical separation where each are weighted equally.

TABLE 2. Comparison of correlations with measured values of wind speed at six stations not included in statistical dataset with statistical and one-over-distance-squared interpolation for lower-than-mean wind speeds at the ridge.

Independent data station	Interpolation criteria	Interpolation stations		Method	Total data correlation (#)	Predictability above mean correlation (#)	Predictability below mean correlation (#)
T1	Closest 2	P20	P22	$1/R^2$	0.57 (113)	0.63 (57)	0.52 (56)
		R = 0.31 km	0.58 km	Statistical	0.59 (113)	0.63 (57)	0.54 (56)
	Closest + Most Correl.	P20	P15	$1/R^2$	0.58 (167)	0.61 (84)	0.50 (83)
		R = 0.31	2.25	Statistical	0.63 (167)	0.67 (84)	0.55 (83)
T2	Closest 2	P18	P16	$1/R^2$	0.47 (139)	0.41 (69)	0.51 (70)
		R = 0.44	0.64	Statistical	0.47 (139)	0.42 (69)	0.55 (70)
	Closest + Most Correl.	P18	P14	$1/R^2$	0.54 (148)	0.58 (74)	0.43 (74)
		R = 0.44	0.77	Statistical	0.54 (148)	0.56 (74)	0.46 (74)
T3	Closest 2	P15	P14	$1/R^2$	0.64 (155)	0.59 (77)	0.70 (78)
		R = 0.05	0.66	Statistical	0.65 (155)	0.60 (77)	0.71 (78)
	Closest + Most Correl.	P15	P18	$1/R^2$	0.63 (157)	0.58 (79)	0.63 (78)
		R = 0.05	1.38	Statistical	0.64 (157)	0.59 (79)	0.64 (78)
T4	Closest 2	P5	P6	$1/R^2$	0.62 (157)	0.63 (79)	0.56 (78)
		R = 0.18	0.66	Statistical	0.62 (157)	0.63 (79)	0.57 (78)
	Closest + Most Correl.	P5	P18	$1/R^2$	0.62 (157)	0.52 (79)	0.68 (78)
		R = 0.18	2.73	Statistical	0.62 (157)	0.54 (79)	0.70 (78)
T5	Closest 2	P3	P2	$1/R^2$	0.61 (117)	0.73 (58)	0.57 (59)
		R = 0.07	0.51	Statistical	0.61 (117)	0.74 (58)	0.57 (59)
	Closest + Most Correl.	P3	P10	$1/R^2$	0.60 (88)	0.74 (44)	0.52 (44)
		R = 0.07	1.08	Statistical	0.61 (88)	0.76 (44)	0.54 (44)
T6	Closest 2	P24	P23	$1/R^2$	0.24 (116)	0.46 (58)	0.07 (58)
		R = 0.50	1.85	Statistical	0.31 (116)	0.51 (58)	0.19 (58)
	Closest + Most Correl.	Same		$1/R^2$ Statistical	Same		

the improvement associated with conditions of higher interstation predictability occurs in only two-thirds of the comparisons. This ratio is larger in the higher ridge wind situations shown in Table 3. Figures 5 and 6 illustrate the results shown in Tables 2 and 3 as a function of station separation.

Table 4 shows the comparisons for wind direction interpolations for lower-than-mean ridge wind speeds. Figure 7 illustrates the results shown in Table 4 as a function of station separation. These correlations are based on wind direction values adjusted so that wind directions differ by less than 360°. This causes the correlation values to be larger because of data clusters around 0° and 360°. The magnitudes of the correlations do not show a definite improvement using the statistically based interpolation algorithms. In fact, there were several correlations which were slightly degraded. The differences were at most statistically significant at the 50% level.

In spite of the fact that the correlations were similar, the absolute values can be quite different. Tables 5 and 6 show that wind speed and direction differences from the two techniques were quite large (mean speed over 1 m s^{-1} and mean direction over 10°) during a tracer release. Tables 5 and 6 also show that a relatively high value for the power law exponent is required to explain the difference in wind speed between the data stations at 4 m and the interpolation test stations at 10 m. Adjustments for height were made by using the following power-law profile:

$$\bar{U}_{10} = \bar{U}_4\left(\frac{10}{4}\right)^E \quad 0 \leqslant E \leqslant 1. \tag{2}$$

Though values from 0.1 to 0.3 are commonly used for neutral to stable conditions, it is obvious from Tables 5 and 6 that a larger value is required. Haltiner and Martin (1957) suggest a value of 0.8 for cases of extreme stability which seems to produce a better over-

TABLE 3. As in Table 2 except for higher-than-mean wind speeds at the ridge.

Independent data station	Interpolation criteria	Interpolation stations		Method	Total data correlation (#)	Predictability above mean correlation (#)	Predictability below mean correlation (#)
T1	Closest 2	P20	P22	$1/R^2$	0.74 (154)	0.86 (77)	0.57 (77)
		R = 0.31 km	0.58 km	Statistical	0.74 (154)	0.86 (77)	0.54 (77)
	Closest + Most Correl.	P20	P15	$1/R^2$	0.71 (156)	0.77 (78)	0.63 (78)
		R = 0.31	2.25	Statistical	0.74 (156)	0.79 (78)	0.67 (78)
T2	Closest 2	P18	P16	$1/R^2$	0.78 (168)	0.78 (84)	0.56 (84)
		R = 0.44	0.64	Statistical	0.80 (168)	0.80 (84)	0.64 (84)
	Closest + Most Correl.	P18	P14	$1/R^2$	0.81 (169)	0.74 (84)	0.62 (85)
		R = 0.44	0.77	Statistical	0.82 (169)	0.73 (84)	0.69 (85)
T3	Closest 2	P15	P14	$1/R^2$	0.71 (174)	0.73 (87)	0.61 (87)
		R = 0.05	0.66	Statistical	0.73 (174)	0.75 (87)	0.63 (87)
	Closest + Most Correl.	P15	P18	$1/R^2$	0.71 (170)	0.67 (86)	0.57 (84)
		R = 0.05	1.38	Statistical	0.72 (170)	0.68 (86)	0.59 (84)
T4	Closest 2	P5	P6	$1/R^2$	0.76 (169)	0.78 (84)	0.72 (85)
		R = 0.18	0.66	Statistical	0.75 (169)	0.77 (84)	0.71 (85)
	Closest + Most Correl.	P5	P18	$1/R^2$	0.76 (169)	0.64 (84)	0.73 (85)
		R = 0.18	2.73	Statistical	0.76 (169)	0.64 (84)	0.73 (85)
T5	Closest 2	P3	P2	$1/R^2$	0.73 (155)	0.70 (78)	0.77 (77)
		R = 0.07	0.51	Statistical	0.73 (155)	0.70 (78)	0.76 (77)
	Closest + Most Correl.	P3	P10	$1/R^2$	0.73 (154)	0.73 (77)	0.76 (77)
		R = 0.07	1.08	Statistical	0.74 (154)	0.74 (77)	0.77 (77)
T6	Closest 2	P24	P23	$1/R^2$	0.32 (132)	0.72 (65)	0.12 (67)
		R = 0.50	1.85	Statistical	0.44 (132)	0.77 (65)	0.30 (67)
	Closest + Most Correl.	Same		$1/R^2$ Statistical	Same		

all comparison for both 1-over-distance-squared and statistically based interpolation.

5. Incorporation of interpolations in models

The ASCOT release of PDCH on the night of 19 September 1980 reveals that the tracer material was largely entrapped within a fairly shallow cold air drainage layer. The construction of a three-dimensional field of gridded wind values begins with the raising or lowering of a spatially irregular distribution of surface observations to a common level. This normalization to a reference height, typically the median height of the measuring instruments, is accomplished through the application of a stability dependent power law formula which operates on the measured wind speeds, but not the directions. The next step is the assignation of speed and direction values at the intersecting points of a regular mesh at the reference height. Normally, their derivation is based on an inverse distance squared weighting of the three nearest normalized observations. For this analysis, we used the nearest two observations to simplify the comparison of interpolation algorithms.

Prior to the incorporation of one or more vertical soundings into the 3-D lattice, the wind observations in the vertical are regularized by assuming that speed and direction vary linearly between data levels. This allows values to be picked off at selected heights including the equal intervals of the computational grid. The procedures just described for the horizontal interpolation of data at the near-surface reference level are then repeated at elevations in the layer stretching from the top of the surface layer to the maximum vertical extent of the grid.

An extrapolation method that allows the controlled melding of surface layer winds and winds through the mixed layer is used to complete the field prescription.

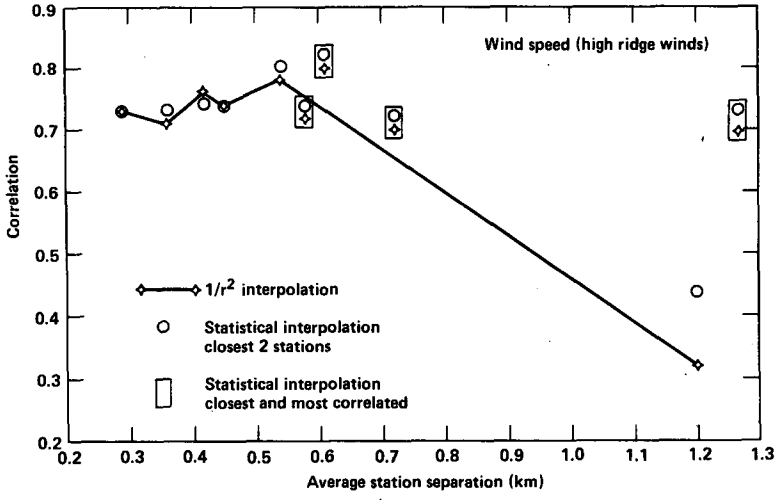

FIG. 5. Correlation of measured and interpolated wind speed values at six towers not used in the statistical database as a function of average database station distance to interpolation point under the condition of lower-than-mean ridge-top wind speeds.

This method assumes that surface observations, by virtue of their relative abundance, provide the most reliable picture of the winds at lower levels of the atmosphere. Similarly, upper air soundings generally provide the best wind information near the top of the mixed layer and above.

Based on these assumptions, assignments at intermediate points are made by creating a synthesized profile that both preserves the intrinsic shape of an actual (designated) profile and matches the interpolated winds at the tops of the surface and mixed layers along each column of grid points.

Three-dimensional arrays of u- and v-component winds, which result from the reexpression of wind speeds and directions, are the final products of these operations. Terrain in blocked form is then introduced which causes the component winds to be shifted upward by an amount equal to the local terrain elevation. Because of these shifts, there can be certain inconsistencies in the final composition, vis-a-vis, the orientation of the velocity vectors in relation to the model terrain. The purpose of MATHEW, our diagnotic wind adjustment model, is to remove these inconsistencies.

6. Model descriptions

The theoretical framework for the MATHEW model equations was developed by Sasaki (1958), who pioneered the use of variational techniques in a meteorological context. One of his proposed formalisms de-

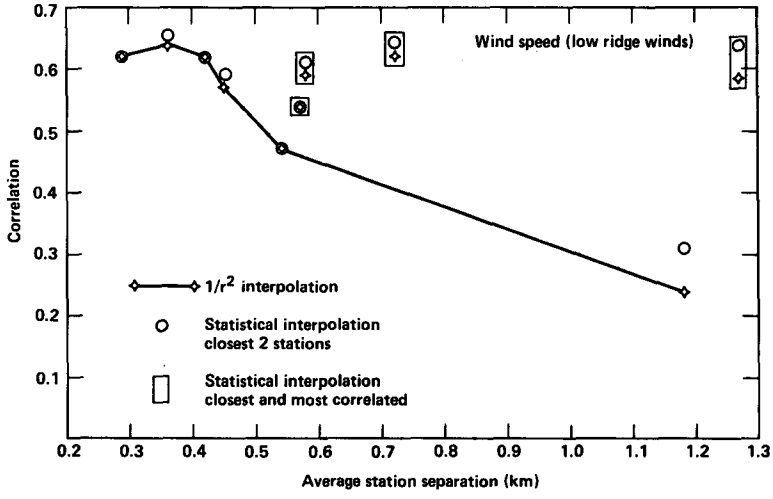

FIG. 6. As in Fig. 5 except for higher-than-mean ridge-top wind speeds.

TABLE 4. As in Table 2 except for values of wind direction.

Independent data station	Interpolation criteria	Interpolation stations		Method	Total data correlation (#)	Predictability above mean correlation (#)	Predictability below mean correlation (#)
T1	Closest 2	P20	P22	$1/R^2$	0.85 (107)	0.89 (52)	0.83 (55)
		R = 0.31 km	0.58 km	Statistical	0.86 (107)	0.89 (52)	0.85 (55)
	Closest + Most Correl.	P20	P15	$1/R^2$	0.84 (139)	0.90 (68)	0.80 (71)
		R = 0.31	2.25	Statistical	0.84 (139)	0.90 (68)	0.81 (71)
T2	Closest 2	P18	P16	$1/R^2$	0.67 (96)	0.71 (48)	0.64 (48)
		R = 0.44	0.64	Statistical	0.67 (96)	0.73 (48)	0.63 (48)
	Closest + Most Correl.	P18	P14	$1/R^2$	0.76 (96)	0.78 (48)	0.74 (48)
		R = 0.44	0.77	Statistical	0.74 (96)	0.76 (48)	0.72 (48)
T3	Closest 2	P15	P14	$1/R^2$	0.84 (120)	0.77 (58)	0.86 (62)
		R = 0.05	0.66	Statistical	0.83 (120)	0.76 (58)	0.86 (62)
	Closest + Most Correl.	P15	P18	$1/R^2$	0.86 (104)	0.90 (52)	0.84 (52)
		R = 0.05	1.38	Statistical	0.86 (104)	0.90 (52)	0.84 (52)
T4	Closest 2	P5	P6	$1/R^2$	0.88 (148)	0.94 (72)	0.80 (76)
		R = 0.18	0.66	Statistical	0.90 (148)	0.94 (72)	0.84 (76)
	Closest + Most Correl.	P5	P18	$1/R^2$	0.87 (102)	0.87 (51)	0.87 (51)
		R = 0.18	2.73	Statistical	0.87 (102)	0.87 (51)	0.87 (51)
T5	Closest 2	P3	P2	$1/R^2$	0.93 (96)	0.97 (45)	0.90 (51)
		R = 0.07	0.51	Statistical	0.93 (96)	0.97 (45)	0.91 (51)
	Closest + Most Correl.	P3	P10	$1/R^2$	0.95 (136)	0.98 (68)	0.92 (68)
		R = 0.07	1.08	Statistical	0.95 (136)	0.98 (68)	0.92 (68)
T6	Closest 2	P24	P23	$1/R^2$	0.63 (120)	0.78 (59)	0.56 (61)
		R = 0.50	1.85	Statistical	0.65 (120)	0.79 (59)	0.60 (61)
	Closest + Most Correl.	Same		$1/R^2$ Statistical	Same		

fines an integral function whose extremum minimizes the variance of the difference between observed and analyzed values subject to certain physical constraints. In MATHEW a solution is sought that represents a minimal perturbation to a set of "observed" wind components in three dimensions and produces a mass-consistent field. The key phrase here is "minimal perturbation"; it underscores the need to exercise care when developing the initial field since the formulation in MATHEW is premised on the assumption that a reasonable, first-guess field of wind velocities can be supplied by a subsidiary (preprocessing) code.

Once a field, or a set of fields, has been adjusted for mass-consistency, the wind component arrays are passed to ADPIC, our advection-diffusion particle-in-cell code. ADPIC releases thousands of Lagrangian marker particles to simulate the transport and dispersion of a pollutant cloud. As appropriate, each particle is tagged with a combination of mass or activity. The model solves the flux-conservative form of the advection-diffusion equation which includes a pseudo-velocity (the sum of the advection and diffusion velocities). The advection velocities are supplied by MATHEW and the diffusion velocities are a function of particle distribution and horizontal and vertical diffusivities whose forms are based on accepted diffusion theory. Concentrations are calculated at selected times by summing the particle masses or activities in each grid cell and dividing through by the individual cell volumes.

7. Model results

Several experiments in the Geysers area were conducted in September 1980 to obtain information on the behavior of effluents as they dispersed along three major paths of cold air drainage. We chose a release along one of these paths, Anderson Creek, to evaluate

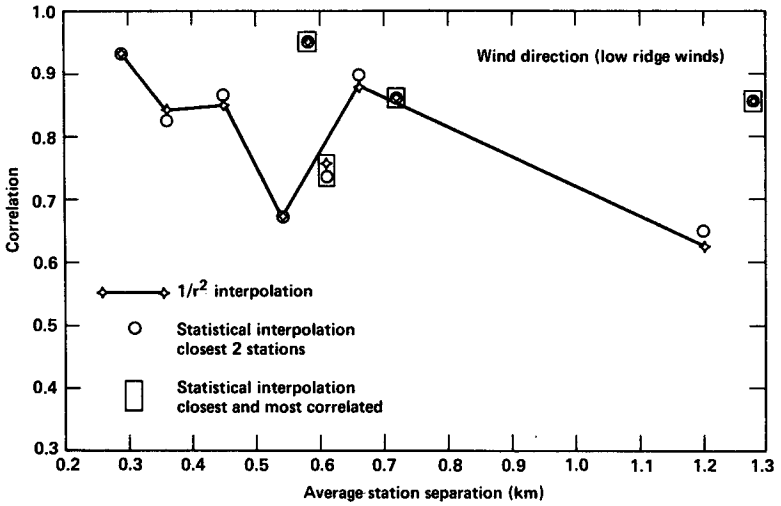

FIG. 7. As in Fig. 5 except for wind direction.

our statistically based interpolation algorithm using MATHEW and ADPIC. The selected episode involved the vaporization and near-surface (5 m) injection of a perfluorocarbon (PMCH) into the atmosphere at a uniform rate starting at 2300 PST on 19 September and lasting for 1 h. Since the dispersion characteristics of the effluent were unknown under drainage conditions, the surface samplers were quasi-uniformly spread throughout the pooling region of the basin.

The surface concentration pattern of PMCH in units of ppt, which were averaged over the first 2 h after the initiation of the release, is shown in Fig. 8a. A simulation of the PMCH plume by ADPIC entailed the creation and constant release over 1 h of 15 000 particles (each tagged with the appropriate mass fraction) into the hydrodynamic flow field. The model-generated patterns of surface concentration, corresponding in time to the observed pattern in Fig. 8a, are shown in

TABLE 5. Mean values from 0000 to 0400 (PDT) hourly estimates during drainage wind conditions on tracer release experiment night using different tower height exponents for the wind speed.

Independent data station	Interpolation criteria	Average wind speed measured ($m\ s^{-1}$)	Average difference wind speed measured and predicted $P = 0.8$ (0.3)		Average wind direction measured (deg)	Average difference wind direction measured and predicted	
			$1/R^2$ ($m\ s^{-1}$)	Statistical ($m\ s^{-1}$)		$1/R^2$ (deg)	Statistical (deg)
T1	Closest 2	1.46	0.24 (0.70)	0.19 (0.63)	293.43	25.32	36.94
	Closest and most correlated		0.32 (0.75)	0.37 (0.77)		14.34	12.80
T2	Closest 2	1.02	0.29 (0.31)	0.17 (0.31)	267.86	41.31	34.63
	Closest and most correlated		0.27 (0.33)	0.35 (0.22)		42.58	34.93
T3	Closest 2	0.90	0.20 (0.37)	0.19 (0.34)	267.44	47.24	43.18
	Closest and most correlated		0.20 (0.37)	0.20 (0.36)		47.24	46.15
T4	Closest 2	1.58	0.67 (0.84)	0.53 (0.84)	233.68	62.29	57.78
	Closest and most correlated		0.69 (0.83)	0.63 (0.83)		63.42	62.31
T5	Closest 2	4.96	1.37 (1.03)	0.82 (1.37)	2.38	7.41	6.43
	Closest and most correlated		1.45 (0.98)	1.28 (1.08)		7.54	7.36
T1	Closest 2	0.90	0.68 (0.61)	0.47 (0.46)	171.15	90.69	90.66
	Closest and most correlated		Same	Same		Same	Same

TABLE 6. Measured and predicted values at 0000 PDT (beginning of tracer release) using different tower height exponents for the wind speed.

Independent data station	Interpolation criteria	Wind speed measured ($m\ s^{-1}$)	Wind speed predicted $P = 0.8$ (0.3)		Wind direction measured (deg)	Wind direction predicted	
			$1/R^2$ ($m\ s^{-1}$)	Statistical (m^{-1})		$1/R^2$ (m^{-1})	Statistical ($m\ s^{-1}$)
T1	Closest 2	2.13	1.95 (1.21)	2.13 (1.32)	304.06	319.17	331.43
	Closest and most correlated		1.71 (1.06)	1.61 (1.00)		301.43	302.56
T2	Closest 2	1.14	1.63 (1.01)	1.21 (0.75)	282.09	294.03	287.65
	Closest and most correlated		1.51 (0.94)	1.55 (0.96)		299.40	272.00
T3	Closest 2	0.91	0.90 (0.56)	0.95 (0.59)	325.27	301.14	293.86
	Closest and most correlated		0.90 (0.56)	0.90 (0.56)		301.14	299.36
T4	Closest 2	2.43	1.13 (0.70)	1.24 (0.77)	277.10	258.37	262.00
	Closest and most correlated		1.16 (0.72)	1.22 (0.76)		257.91	270.00
T5	Closest 2	5.55	6.12 (3.80)	5.81 (3.61)	4.03	0.15	1.99
	Closest and most correlated		6.17 (3.83)	6.12 (3.80)		359.85	359.85
T1	Closest 2	1.67	0.82 (0.51)	1.09 (0.68)	265.89	393.02	347.20
	Closest and most correlated		Same	Same		Same	Same

Fig. 8b for near-surface wind assignments based on the statistical properties of the flow over several days and in Fig. 8c for wind assignments determined exclusively by their distance from observations.

The similarities between the three sets of contours are evident with Figs. 8b and c showing only minor differences. In all three cases, the reader can mentally visualize the downslope progression of the effluent, followed by a general broadening of the plume as it pools in the basin. The more northerly disposition of the calculated concentrations suggest that the transporting winds near the surface have a fairly large (and erroneous) cross-slope component. This, in turn, suggests that the upper level (synoptically induced) winds are disproportionately influential in describing the movement of ADPIC particles, perhaps due to a poor adjustment of the wind fields by MATHEW. The separation of the calculated 1 ppt contours near the confluence of Anderson and Putah creeks indicates a lofting of the plume which is not in evidence in the observed pattern. Detachment of the plume from the surface primarily illustrates the tendency by MATHEW to underestimate the downslope component of the drainage flour when the gradients in terrain are especially steep.

As a means of quantitatively rating the performance of the models using the two interpolation techniques, we have chosen a method based on the formation of ratios between the observed and calculated concentrations. The procedure begins by considering the possible pairings between the two sets of concentrations matched in space and time. When the actual and simulated samplers both show a zero detection, they are disregarded. When the measured (M) concentration is zero but the calculated (C) concentration is nonzero, background is added to both M and C; this addition takes the lower limit of instrument detection into account. Whenever M is greater than zero, background is added only to C.

The cumulative ratios, $R = M/C$ (or C/M if C is greater than M), for the concentrations averaged over the first 2 h following the start of the PMCH release are presented in Fig. 9; the filled and unfilled circles indicate the method, statistically based or distance weighted, by which the near-surface winds were interpolated to a regular mesh. As might be expected from the similarities noted in the concentration fields (see Fig. 8), the quantitative differences that emerged from using the two different methods of wind interpolation are minor and, we believe, inconsequential.

Concerning the overall performance of the models, it is helpful to remember the complexities of the terrain setting and the meteorological conditions under which the field experiment was conducted. The injection and transport of a tracer material within a drainage layer in the rugged Geysers area constitute the most challenging set of conditions yet attempted in a MATHEW/ADPIC evaluation. Only about 40% of the predicted values are within a factor of 10 of observed concentrations. While analyzing the 1980 ASCOT data, Dickerson (1984) discovered variations in 2-h measured concentrations greater than a factor of 20 for samplers

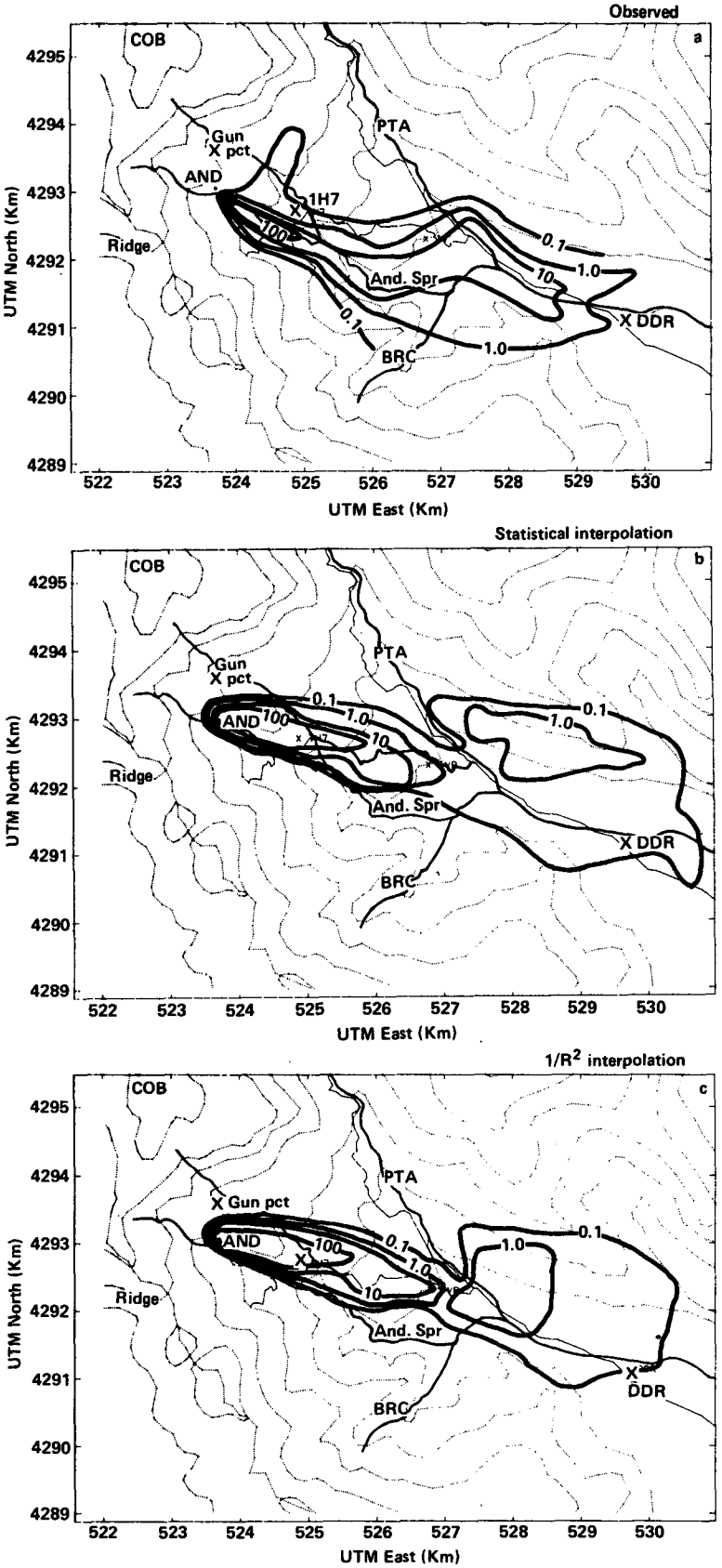

FIG. 8. Contours of (a) measured tracer concentrations in ppt compared with MATHEW/ADPIC simulations using (b) statistical and (c) $1/r^2$ interpolation of wind speed components.

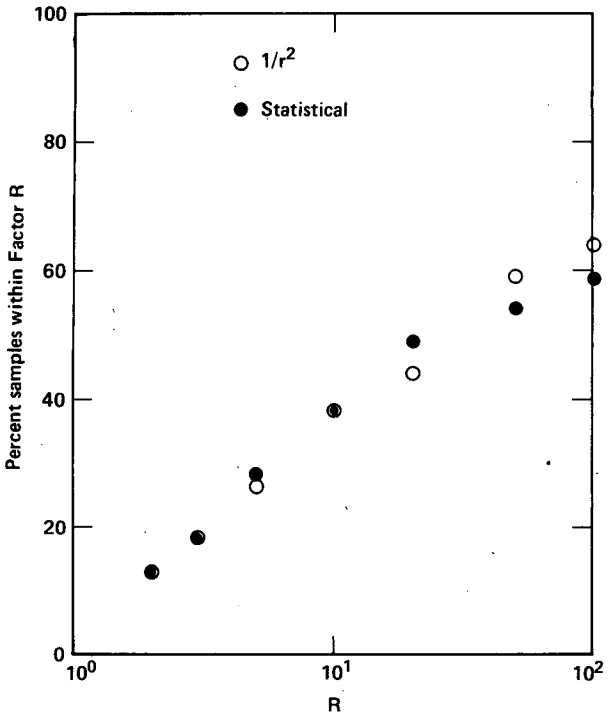

FIG. 9. Comparison of numerical transport and diffusion from MATHEW/ADPIC and measured tracer concentrations on a point-by-point ratio using $1/r^2$ and statistical interpolated winds.

placed 50 to 60 m apart and located 1 km from the source points. Given these extraordinarily large concentration gradients, it is easy to understand why the models found it difficult to resolve the effluent distributions.

8. Conclusions

In this paper we have compared standard 1-over-distance-squared interpolation with a statistically based interpolation algorithm. The subjective interpretation of these comparisons can vary depending on whether or not the modeler of complex terrain wind fields has the data resources to develop a statistical database. In this respect we were particularly fortunate to have such a dense data network on a 10 km scale. With the ASCOT database, the statistical interpolation comparisons did show a marginal improvement in the interpolated wind field. The difference in the wind fields produced from mass-consistent wind field processing of the $1/r^2$ and statistically interpolated wind fields gives a measure of the confidence ascribable to the predicted wind field. Also, the statistics developed for application of the statistical interpolation is useful for data quality control and missing data filling in emergency situations (Porch and Walton, 1980). However, if one does not have the database with which to develop the temporal statistics for wind predictions, it is obvious from our analysis that standard techniques work just about as well for the case of high spatial density of wind measurements in complex terrain. This is more true for wind direction than wind speed. Because pollution transport model evaluation is so strongly dependent on wind direction diagnostic wind field model performance in complex terrain will continue to be dominated by the difficulty of wind direction interpolation. A hope for the future might be to improve the directional interpolation by incorporating the results of prognostic numerical models into the wind direction estimates.

Acknowledgments. The authors would like to thank M. Dickerson and P. Gudiksen for their support in this project and D. Bennett for her help with the statistical analysis. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

REFERENCES

Dickerson, M. H., and P. H. Gudiksen (Eds.), 1984: Atmospheric studies in complex terrain. Technical Progress Report FY-1979 through FY-1983, Lawrence Livermore National Laboratory Report UCID-19851, ASCOT 84-1.

——, K. T. Foster and P. H. Gudiksen, 1984: Experimental and model transport and diffusion studies in complex terrain with emphasis on tracer studies. Lawrence Livermore National Laboratory Report UCRL-90053.

Draper, N. R., and H. Smith, 1966: *Applied Regression Analysis.* Wiley & Sons.

Erasmus, D. A., 1986: A model for objective simulation of boundary-layer winds in an area of complex terrain. *J. Climate Appl. Meteor.*, **25,** 1832–1841.

Gandin, L. S., 1965: Objective analysis of meteorological fields. U.S. Dept. of Commerce, Springfield, VA, TT 65-50007, Israel Program for Scientific Translations, Jerusalem, 242 pp.

Goodin, W. R., G. J. McRae and J. H. Seinfeld, 1979: A comparison of interpolation methods for sparse data: Application to wind and concentration fields. *J. Appl. Meteor.*, **18,** 761–771.

Gudiksen, P. H., and J. J. Walton, 1981: Categorization of nocturnal drainage flows in Anderson Creek Valley. Lawrence Livermore National Laboratory Report UCRL-85569.

Haltiner, G. J., and F. L. Martin, 1957: *Dynamic and Physical Meteorology.* McGraw-Hill, p. 232.

Johnson, G. T., 1982: Climatological interpolation functions for mesoscale wind fields. *J. Appl. Meteor.*, **21,** 1130–1136.

King, D. S., and S. S. Bunker, 1984: Application of atmospheric transport models for complex terrain. *J. Climate Appl. Meteor.*, **23,** 239–246.

Lange, R., 1978: ADPIC—A three-dimensional particle-in-cell model for the dispersal of atmospheric pollutants and its comparison to regional tracer studies. *J. Appl. Meteor.*, **17,** 320–329.

Porch, W. M., 1982: Implication of spatial averaging in complex-terrain wind studies. *J. Appl. Meteor.*, **21,** 1258–1265.

Porch, W. P., and J. J. Walton, 1981: Development and application of computer codes to fill missing meteorological data. Lawrence Livermore National Laboratory Report UCID-19151.

Sherman, C. A., 1978: MATHEW: A mass-consistent model for wind fields over complex terrain. *J. Appl. Meteor.*, **17,** 312–319.