PM 86/4

FAA WJH Technical Center
00093125

DOT/FAA/PM-86/4

Program Engineering
and Maintenance Service
Washington, D.C. 20591

A Study of the Methodology of Low-Altitude Wind Shear Detection with Special Emphasis on the Low Level Wind Shear Alert System Concept

FEDERAL AVIATION ADMINISTRATION
FEB 20 1986
TECHNICAL CENTER LIBRARY
ATLANTIC CITY, N.J. 08405

F. Wesley Wilson, Jr.
John A. Flueck

February 1986

Final Report

This Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

Technical Report Documentation Page

1. Report No. DOT/FAA/PM-86/4	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A Study of the Methodology of Low-Altitude Wind Shear Detection with Special Emphasis on the Low Level Wind Shear Alert System Concept		5. Report Date February 1986
		6. Performing Organization Code
7. Author(s) F. Wesley Wilson, Jr., Ph.D., & John A. Flueck, Ph.D.		8. Performing Organization Report No. JAWS NCAR Report No. 02-85
9. Performing Organization Name and Address Joint Airport Weather Studies (JAWS) Project Research Application Program, NCAR P.O. Box 3000 Boulder, CO 80307		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. DTFA01-82-Y-10513
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Program Engineering and Maintenance Service Washington, D.C. 20591		13. Type of Report and Period Covered Final Report Oct. 1983 - July 1985
		14. Sponsoring Agency Code FAA/APM-650

15. Supplementary Notes
Research performed under Interagency Agreement No. DTFA01-82-Y-10513 between the National Science Foundation and the Department of Transportation, Federal Aviation Administration. Managed by the Weather Sensors Program, APM-650.

16. Abstract
This basic principles study of the Low Level Wind Shear Alert System (LLWAS) addresses the question of the degree to which a change of algorithm and network geometry can improve the performance of an anemometer-based wind shear detection system to an acceptable level. The performance of the current system seems to be deficient in that it is ineffective for the detection of microbursts and that it issues many alarms in situations where pilots do not experience hazardous wind shear.

The theoretical basis for this investigation is simulation testing. This process involves testing each combination of algorithm and geometry against a realistic mathematically generated weather model that includes wind shear events. The binary time series of (event/no event) and (detection/no detection) are compared by statistical methods. The total detection skill of each system is computed and the systems that exhibit the best skill are determined. The systems that have the best skills have high probabilities of detection and low probabilities of false alarms.

Eleven different algorithms and seven different regular geometries are tested, with several different station spacings, selected for each geometry. Satisfactory geometries and algorithms are determined. In addition, these methods are applied to the problem of designing the enhanced LLWAS for Stapleton International Airport. This problem involves an irregular geometry, and the methods that have been developed also work well in this context. The primary conclusion of this study is that it is possible to make significant improvements in the LLWAS by modification of the network geometry and the data analysis algorithm.

17. Key Words Wind Shear, Wind Shear Detection, Wind Field Divergence, Microburst, Gust Front, Simulation Testing, Probability of Detection, False Alarm Ratio, Total Skill Score	18. Distribution Statement This Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 109	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

Section	Topic	Page
	Table of Contents	iii
	List of Figures	iv
	List of Tables	vi
	Glossary	vii
	Executive Summary	viii
I.	Introduction	1
II.	Goals and Scope of the Study	2
III.	Detection Algorithms	3
IV.	Mathematical Wind Field Model	13
V.	Wind Field Filtering or Smoothing	22
VI.	Network Geometry	24
VII.	Evaluation and Test Procedures	30
VIII.	Results of the Comparative Testing	34
IX.	Recommendations for LLWAS Improvements	50
X.	Recommendations for Further Study	53
XI.	References	54
XII.	Appendices	55
	A. Estimate of the Wind Field Divergence	55
	B. The Mathematical Description of the Expected Values for the Anomaly Detection Algorithms	56
	C. The Mathematical Microburst Model	58
	D. A Review of Some Forecasting Verification Measures	64
	E. Comparative Verification Results for All Algorithms	74
	F. Study of an Irregular Grid Problem: An Expanded LLWAS at Stapleton	88

LIST OF FIGURES

Figure No.		Page
III-1	Underestimation of wind shear intensity due to station spacing.	8
III-2	Line divergence. The divergence is estimated by the vector difference of the projections onto the line of sight	9
III-3	Triangle divergence as a measure of net flux at the boundary	11
III-4	Impact of event location on triangle divergence.	12
IV-1	Gust front model	15
IV-2	Microburst model	16
IV-3	A sample microburst in an ambient wind field, generated by the mathematical wind shear model on a 19-station network.	18
IV-4	A sample gust front and microburst in an ambient wind field generated by the mathematical wind shear model on a 19-station network.	19
IV-5.A	List of events generated by the wind field model in a six-hour simulation.	20
IV-5.B	Distribution of wind shear events near Stapleton (Denver) as generated by the wind field model in a six-hour simulation	21
VI-1	Regular station network geometries	25
VI-2	The station network geometries for 6 and 7 stations, two scales, within a 5-km radius disc.	27
VI-3	The station network geometries for 11 and 13 stations, three scales, within a 5-km disc	28
VI-4	The station network geometries for 19 stations, three accompanying scales, within a 5-km radius disc	29
VIII-1	Plots of performance scores versus threshold values for TEDC algorithm under 13-station network and scale of 2.25 km.	43
VIII-2	Plots of performance scores versus network size for OLDCF algorithm. Scale = 2.25	45
VIII-3	Plots of performance scores versus network size for NEWCF algorithm. Scale = 2.25	46
VIII-4	Plots of performance scores versus network size for TEDC algorithm. Scale = 2.25	47
VIII-5	Plots of performance scores versus network size for NLR algorithm. Scale = 2.25	48
VIII-6	Plots of performance scores versus network size for LMD algorithm. Scale = 2.25	49
XIIC-1	Partition of the microburst region	59

LIST OF FIGURES (continued)

Figure No.		Page
XIIF-1	The original LLWAS at Stapleton (Geometry 0) and the rectangular protection region.	89
XIIF-2	All station sites that were considered	90
XIIF-3A	Geometry A	92
XIIF-3B	Geometry B	93
XIIF-3C	Geometry C	94
XIIF-3D	Geometry D	95
XIIF-3E	Geometry E	96
XIIF-3F	Geometry F	97
XIIF-3G	Geometry G	98

LIST OF TABLES

Table No.		Page
V-1	Sensitivity study of the recursive filter for 13 stations, scale 2.25, threshold 6.00 and "a" varying from .70 to 1.00.	23
VII-1	The general form of the 2X2 table for the wind shear verification problem.	32
VIII-1	Algorithm performance results for the 19-, 13-, 11-, and 6-station networks, scale=1.50 km simulations	36
VIII-2	Algorithm performance results for the 19-, 13-, 11-, and 6-station networks, scale=2.25 km simulations	39
VIII-3	Cross-validation of detection methods for the 13-station, scale 2.25 km network using five randomly selected wind sequences	41
VIII-4	Old centerfield algorithm cross-validation performance results for the stations=6, scale=3.00, threshold=7.5 situation compared to an optimum threshold=10.0 situation .	42
VIII-5	Threshold values versus performance values for the TEDC algorithm on the 13-station network with scale of 2.25 km..	42
XIIE-1	Simulation testing for 6-station network, scale=3.00. OLDCF represents the existing LLWAS with optimal threshold (7.5 m/s or 15 knots)	76
XIIE-2	Simulation testing for 6-station network, scale=3.00. OLDCF represents the existing LLWAS with optimal threshold (9.5 m/s or 19 knots)	77
XIIE-3	Simulation testing for 6-station network, scale=1.50. . . .	78
XIIE-4	Simulation testing for 7-station network, scale=1.50. . . .	79
XIIE-5	Simulation testing for 11-station network, scale=1.50 . . .	80
XIIE-6	Simulation testing for 13-station network, scale=1.50 . . .	81
XIIE-7	Simulation testing for 19-station network, scale=1.50 . . .	82
XIIE-8	Simulation testing for 6-station network, scale=2.25. . . .	83
XIIE-9	Simulation testing for 7-station network, scale=2.25. . . .	84
XIIE-10	Simulation testing for 11-station network, scale=2.25 . . .	85
XIIE-11	Simulation testing for 13-station network, scale=2.25 . . .	86
XIIE-12	Simulation testing for 19-station network, scale=2.25 . . .	87
XIIF-1	Comparison of the anomaly algorithms for various network geometries.	99
XIIF-2	Comparison of the identification algorithms for various network geometries.	100

GLOSSARY

EVALUATION MEASURES

CSI	Critical skill index
EL.FAR	Eligibile false alarm ratio
EL.POD	Eligible probability of detection
FAR	False alarm ratio
POD	Probability of detection
TSS	Total skill score

DETECTION ALGORITHMS

LMD	Local median
LMN	Local mean
LVD	Local vector difference
NEWCF	New centerfield
NLR	Network linear regression
NMN	Network mean
NMD	Network median
OLDCF	Old centerfield
RFCF	Recursive filter centerfield
TDC	Triangle divergence/convergence
TEDC	Triangle edge divergence/convergence

EXECUTIVE SUMMARY

This basic principles study of the Low-Level Wind Shear Alert System (LLWAS) investigates the effectiveness of the current system and the possibility of improving the system by some combination of modifying the network geometry and incorporating a new detection algorithm. Quantitative comparisons of the various proposed systems are obtained by simulation testing. Based on these tests, it is clear that substantial improvements in the LLWAS are possible. Since this is a scientific study, it does not confront such issues as the relative cost effectiveness of the various proposed improvements or their specific benefits in an operational aviation system. Thus, our recommendations are confined to feasibility and to scientific comparisons of merit. However, we believe that these principles could be used as a guide for management decisions involving operational usefulness and cost effectiveness.

The algorithms that are tested fall into three broad classes:

1. Algorithms that measure the difference between the wind field that is currently observed at each station and the wind field that is expected at that station, based on some estimate of the average behavior on the entire network, e.g., the running average of the observed values at the centerfield station or the mean wind field over the entire network.

2. Algorithms that measure the difference between the wind field that is currently observed at each station and the wind field that is expected at that station, based on a modelling of the wind field on the network, e.g., the least squares linear regression estimated wind field.

3. Algorithms that use numerical differentiation to estimate the wind field convergence and hence can identify the nature of the wind shear event.

The goal of our testing is to measure the effectiveness of each system for detecting wind shear events that occur in a disc of radius 5 km (approximately 3 miles). We test geometries that have 6, 7, 11, 13, and 19 stations and are uniformly positioned in this disc. Except for the uniform placement of the stations, the 6-station geometry with the centerfield algorithm is an idealized replication of the currently used LLWAS. The other geometries are studied to see how much improvement can be achieved by increasing the number of stations in the disc. Since we are always trying to protect the same 5-km disk, the station density increases as the number of stations increases. Therefore, our conclusions about the number of stations are really conclusions about station density, and the actual numbers of stations that would be needed to protect a different region would depend on the area of the region. Also, it is possible that a greater density might be beneficial or even necessary if some other goal were being addressed, e.g., estimation of the intensity of the wind shear event in the runway corridor.

The simulation-testing procedure used in this study is based on a statistical evaluation of the ability of each combination of geometry and algorithm to detect reliably wind shear events. In particular, a system that issues false alarms is penalized in our scoring procedure. Our principal measure of effectiveness is the Total Skill Score (TSS), an advanced statistical verification measure that has been used in the evaluation of numerous forecasting methods. In addition, we compute the Probability of Detection (POD, the probability that a wind shear event will be detected by a station that is in the proximity of the wind shear event); the Eligible Probability of Detection (EL.POD, the probability that a wind shear event will be detected by a station that is in the proximity of the wind shear event); the False Alarm Ratio (FAR, the probability that a wind shear alarm is false); and the Eligible False Alarm Ratio (EL.FAR, the probability that an alarm will be issued by a station that is not in the proximity of the wind shear event). Other statistical measures were considered; we have chosen to use the ones that provide the most useful information about our problem.

The actual testing involves using a realistic mathematical wind field model to simulate a six-hour development of wind shear events. The nature and frequency of the events are similar to patterns that have been observed at Stapleton during an active afternoon (JAWS-1982 and CLAWS-1984). As the simulated wind field evolves with time, the measurements that would have been observed by LLWAS stations are recorded and analyzed according to the principles of each of the proposed algorithms. Two time-series are generated: (1) the series of wind shear events, and (2) the series of alarms that are issued by the application of each of the algorithms. Comparison of these time-series yields the statistical measures. Each system is tuned so that it operates with optimal skill, i.e., maximum TSS, and the other parameters (POD, FAR, EL.POD, EL.FAR) are recorded as they result from this mode of operation.

The testing that we have conducted shows that LLWAS can definitely be improved, especially in its ability to detect microbursts. A new detection algorithm should be adopted and denser station deployment is very beneficial. The following table illustrates the nature of the relative comparisons for the regular geometries that we have tested. The methods considered here are:

OLDCF	Essentially the existing LLWAS (Centerfield algorithm)
NEWCF	An upgrade of the existing LLWAS (Centerfield algorithm)
NLR	Modeling method (Network Linear Regression)
TEDC	Wind field convergence and wind shear identification algorithm.

Spacing is a scale factor that relates to station density (3.00 corresponds to the density of the current LLWAS; 2.25 corresponds to the spacing that we found to perform well).

METHOD	STATIONS	SPACING	TSS	POD	FAR	EL.POD	EL.FAR
OLDCF	6	3.00	.21	.79	.45	.19	.68
NEWCF	6	2.25	.46	.63	.23	.32	.32
NLR	6	2.25	.58	.72	.17	.60	.35
TEDC	6	2.25	.72	.72	.19	.50	.25
NEWCF	13	2.25	.65	.80	.17	.67	.27
NLR	13	2.25	.77	.94	.16	.80	.22
TEDC	13	2.25	.72	.82	.11	.72	.19
NEWCF	19	2.25	.70	.78	.10	.77	.28
NLR	19	2.25	.80	.94	.14	.89	.22
TEDC	19	2.25	.81	.91	.16	.84	.22

Inspection of these results indicates that the greater the number of stations, the better the detection skills. Essentially all other methods offer a substantial improvement over the old centerfield (OLDCF) method. One should note that when there is a sizable difference between the POD and the EL.POD, or the FAR and the EL.FAR, it implies that the system is correctly detecting a wind shear event in the region, but that it is frequently unable to correctly identify the location of the event. If ATC were to use the alarm at a station to divert traffic to another runway, a large EL.FAR is an indication that they might divert the traffic from a place where there is no wind shear to a place where there is a wind shear event in progress. The study closes with nine recommendations for improving the present LLWAS and thirteen recommendations for further study.

I. Introduction

Low-altitude wind shear, or the rate of change of wind velocity between two spatial points, has been of interest to meteorologists for many years. In the classical Thunderstorm Project, Byers and Braham (1949) identified thunderstorm outflows and downdrafts; we now would classify the smaller scale downdrafts as microbursts.

In the early 1970's the Federal Aviation Administration (FAA) became interested in low-altitude wind shears and their impact on aviation safety. At that time, the primary wind shear danger to aircraft was believed to be the gust front. Based on this belief, a Low-Level Wind Shear Alert System (LLWAS) was designed (Goff, 1980). This system consists of a minicomputer and a small anemometer network located near the airport (i.e., a centerfield and five outlying sensors typically about two miles from centerfield). The computer polls these stations frequently and performs an analysis that is intended to determine automatically when there is a wind shear event (i.e., gust front) in the airport vicinity. When a wind shear detection is made, the computer issues an alarm in the control tower, and this result is relayed to pilots by Air Traffic Control (ATC).

In 1977, this system was installed at six airports: Atlanta, Denver, Houston, New York, Oklahoma City, and Tampa. By 1983, the system was in operation at 59 airports, and an additional 51 airports are scheduled for installation by early 1986. Thus, LLWAS is the present operational standard for the detection of low-altitude wind shear in the vicinity of an airport.

As early as 1976 some evidence had been presented that small-scale downbursts were related to aircraft accidents (Fujita, 1976; Fujita and Byers, 1977; Fujita and Caracena, 1977). The first detailed investigation of these ideas was undertaken in the Joint Airport Weather Studies (JAWS) Project, in the summer of 1982 in Denver, Colorado. The observation, recording, and analysis of the LLWAS during the JAWS Project allowed researchers to evaluate the performance of LLWAS and to compare it with other meteorological instruments for the measurement of low-altitude wind fields. These comparisons (Bedard et al., 1984) confirmed the suspicion that LLWAS is prone to issue numerous false alarms and to miss a number of wind shear events. A second major aviation wind shear hazard, the microburst, was shown by JAWS to be much more prevalent at Denver, Colorado, than had been anticipated, and this wind shear event was not a primary consideration during the design of LLWAS.

Lastly, most recently the National Research Council (1983) recognized the problems and limitations of the present LLWAS and has recommended that ". . . every effort should be made to assess and improve its performance." It is in response to the above problems that the FAA commissioned this study.

II. Goals and Scope of the Study

This study was undertaken to determine if it is possible to change the LLWAS design so that it will have an improved performance for the detection of both gust fronts and microbursts, while reducing the incidence of false alarms. The study will focus on two major objectives:

1. Develop a theoretical basis and a methodology for the detection and identification of microbursts and gust fronts by a meso-network of anemometers, and

2. Provide the FAA with recommendations for better ground-based anemometer wind shear detection systems based on quantitative evaluations of the effectiveness of alternative systems.

Given the above two objectives we propose that the following five tasks be undertaken:

1. Develop a realistic two-dimensional (u,v) mathematical dynamic wind shear event model which contains both microbursts and gust fronts interacting within an ambient wind field,

2. Develop a number of new algorithms for detecting both gust front and microburst wind shear events,

3. Design scoring and verification procedures for giving a quantitative evaluation of the effectiveness of the various algorithms and geometries that are to be studied,

4. Incorporate the mathematical wind shear model, the detection algorithms, and the scoring and verification procedures into a computer simulation and testing software package to assess which wind shear detection systems appear to have the best wind shear detection capabilities,

5. Evaluate, through computer simulation, the effects of algorithms and mesonet designs (including the number of sampling stations) on the relative performances of these wind shear detection systems.

The scope of the study was limited to the following principal items:

1. The protection of a disc of radius 5 km which contains the airport runways,

2. A hypothetical number of anemometer stations ranging from 6 to 19,

3. The utilization of a two-dimensional wind shear event model which has some allowance for temporal changes,

4. A preselected group of detection algorithms (including the current LLWAS algorithm),

5. Scoring and verification procedures that allow some flexibility in determining a detection or "limit" and which uses both traditional and statistical methods for assessing performance, and

6. Six hours of wind shear simulation time for each combination of algorithm and mesonet design.

III. Detection and Identification Algorithms

There are two primary parts to the operation of a low-altitude wind shear detection system: (1) the sampling of the wind field, and (2) the interpretation of the collected data. The data interpretation is performed by a computer program that is the realization of a number of detection algorithms. The algorithms in this study can be placed into two subclasses: (1) those that detect anomalous wind field behavior and (2) those that estimate wind field divergence. The former group assumes that a wind shear event is an anomaly above some background value, whereas the latter group relies on some meteorological understanding of wind shears in that a strong positive divergence is evidence of the presence of a microburst, and a strong convergence (negative divergence) is evidence of the presence of a gust front. Therefore, this second group of algorithms gives an identification of the wind shear event in addition to its detection.

It also should be noted that algorithms that rely on single-station detection, as opposed to multiple-station detection, are more prone to issuing false alarms due to local wind phenomena or measurement noise. Thus we are faced with the paradox that warnings based on single-station algorithms could include many false alarms, but for relatively large grid scales these algorithms may be the only ones that are likely to issue timely alarms. Furthermore, a single-station alarm is an indication that a dangerous event may have initiated but provides no identification of the type of event.

A. Wind Field Anomalies as a Detection Device

A wind field anomaly at a station is defined as the difference between the measured wind velocity vector and the velocity vector that is expected at that station:

measured velocity - expected velocity = wind field anomaly

The designation that a station anomaly is substantial (and thus indicative of a wind shear event) is based on the estimated anomaly being larger than some predetermined threshold value. (Note that there could be a different threshold value for each different algorithm and network geometry.)

Although we have found no theoretical basis for the choice of threshold values for the anomaly algorithms, the threshold values have been determined heuristically so that each algorithm has the best performance (i.e., the highest chance of detection without an exorbitant number of false alarms). The different anomaly detection algorithms presented in this study are based on different schemes for determining the expected or background value that is to be used at a given station. There are nine anomaly detection algorithms that have been tested in this study, and they can be briefly identified by their expected value characteristics (a mathematical description of each of these methods is presented in Appendix B).

1. (OLDCF) Old centerfield average value (current LLWAS). The expected value used for each system is the running average value of the centerfield station and is compared to each of the five "outlying" stations by computing a vector difference. Typically, the outlying stations are about 3-5 km from centerfield and the threshold is 15 knots (7.5 m/s).

2. (NEWCF) New centerfield average value. Each outlying station is compared with the centerfield running average as in OLDCF. In addition, the current centerfield reading is compared with the centerfield running average, so that it also is possible to make a wind shear detection at centerfield.

3. (RFCF) Recursive filter centerfield average value. This approach is the same as the NEWCF with the exception that the "running average value at centerfield" is replaced by another averaged value based on a recursive filter. This latter approach is based on digital filtering, and both approaches are described in detail in Section V.

4. (NMN) Network mean wind field. The expected value used for each station is computed as follows:

   i) The time averaged wind field is computed for each station,
   ii) These station data are trimmed by (possibly) deleting stations with extremely large and small wind speeds (see page 6),
   iii) The mean of the remaining average values is used as the expected value.

5. (NMD) Network median wind field. The expected value used for each station is computed as follows:

   i) The time averaged wind field is computed for each station,
   ii) The median of the averaged values is used as the expected value.

   Recall that the median of an ordered list of values is the middle value. In this case, the median (u,v) is taken to be the vector with median u-value and median v-value.

6. (LMN) Local mean wind field. The expected value used for each station is the (spatial) mean of the (time) averaged wind field for the proximate stations (see page 6).

7. (LMD) Local median wind field. The expected value used for each station is the median of the (time) averaged wind field for the proximate stations.

8. (NLR) Network linear regression model wind field. The expected value used for each station is the bivariate linear regression wind field value using the (time) averaged values for each station. This method is based on a first order least squares fit to the averaged and possibly trimmed station data.

9. (LVD) Local vector difference from proximate stations. This method is based on the computation of the extreme vector difference for each station, rather than on the computation of an expected value. The procedure is the following:

   i) Find the time averaged value for each station that is proximate to the given station (see page 6),
   ii) Compute the vector difference between the current wind field value at the given station and each of these time averaged values,
   iii) Define the anomaly to be the maximum observed vector difference.

The first five algorithms are similar in that they attempt to obtain a single wind field value that is used as the expected value for the wind field on the entire network. The last four algorithms are similar in that they seek to model locally the gross variation in the wind field over the network.

The algorithms that use a single background or expected value for all stations have a possibilty of issuing false alarms due to natural variation in the wind field over the network, especially if the network covers a sizable region. On the other hand, the methods that attempt to model the gross wind field variation may succeed in modelling the wind field so well that wind shear events are incorporated into the expected value and therefore fail to be detected. In particular, the local vector difference (LVD) method seeks the most extreme wind field variation; hence it is most likely to be fooled by subscale events and measurement noise. The quantitative scoring of the effectiveness of these algorithms will provide a way to determine which of these potential problems is most severe, and whether any of these algorithms can give satisfactory performance.

1. Smoothing of Raw Data

The original LLWAS operates with a six- to ten-second station polling period. Observations of the wind field by anemometers for such short time periods are likely to contain fairly significant measurement noise. Also

there is the problem that a wind shear signal can be a spike on the network and adversely affect the quality of the wind field modelling. Two data smoothing techniques have been introduced to protect the algorithm performance from being adversely affected by these problems.

The first is to use time-series averaging of the station data to reduce the high frequency component of the wind field variation at each station before the network or expected wind field models are constructed. Therefore, we test for anomalies between the current wind field observation at each station and the expected wind field based on the time-averaged wind at each station in the network over a time period that is significantly longer than the polling period (i.e., typically two minutes).

Two averaging methods have been studied: the running average method that is used in the current LLWAS and a comparable recursive filter technique. The details of these time-series analyses are provided in Section V.

The second technique is to trim the network data set. Stations that show extreme wind field averaged values are dropped from the data set used in the construction of the expected network wind field. (Note that this is the same as giving these extreme values a weight of zero.) However, these values are not dropped from the set of observations that are compared with the expected wind field for the purpose of detecting anomalies. To avoid introducing a bias to the data, the standard practice (e.g., Mosteller and Tukey, 1977) is to drop an equal number of high and low values. In the present case, the measurements consist of vector data (i.e., wind field speed and direction) and so there is some question as to what "high" and "low" values imply. Our trimming strategy is based on high and low wind speed, and we will trim either 0 or 1 from each end of the ranked average values, i.e., drop 0 or 2 station's values.

2. Proximate Stations

The local methods LMN, LMD, and LVD all depend on the comparison of values from stations that are near a given station. For this purpose, we define the stations that are proximate to a given station to be the set of all stations that lie within a specified radius of a given station including itself. This radius of proximity is usually selected so that a given station, which is interior to the network, will have about six to seven proximate stations.

3. Pair Detection

Each of the above eight algorithms can be extended to a two- or pair-station algorithm, which issues an alarm only if two proximate stations simultaneously detect the wind shear event. The simultaneous detection of a wind shear event by two nearby stations will significantly reduce the possibility of false alarms. However, if the station spacing is relatively large, then there is a possibilty of an increase in missed detections. The best performance of these two proximate-station algorithms occurs when a

somewhat lower detection threshold is used than for the single-station detection algorithm.

B. Wind Field Divergence as a Detection and Identification Device

As indicated earlier, gust fronts and microbursts are the wind shear hazards that are the focus of this study. A gust front can be characterized by a moving strip along which there is an intense inflow, and a microburst can be characterized by a small region from which there is an intense outflow. According to the Divergence Theorem (e.g., Courant and Hilbert, 1962), the net mass flux along the boundary of a region is equal to the product of the area of the region and the average wind field divergence in the region. With the usual sign conventions, positive divergence is associated with outflow and negative divergence (convergence) is associated with inflow. Hence, a strong divergence exists at the center of a microburst and strong convergence (negative divergence) exists along the leading edge of a gust front and at the outer boundary of a microburst.

It is well known that wind field divergence can be estimated from wind field derivatives. Furthermore, a simple cross-product formula can be used to estimate wind field derivatives (see Appendix A). By the Divergence Theorem, this estimate is equivalent to estimating the strength of the net outflow along the boundary of the region. This method has been chosen as the basis for the detection and identification algorithms.

For gust fronts, wind speed shifts can be greater than 10 m/s over a distance of approximately .4 km. The numerical derivative estimate across this distance yields an estimated divergence on the order of $-.025\ S^{-1}$ (the units of divergence). To observe this magnitude of convergence would require station spacing of .4 km or less. A station spacing of 1 km would yield an estimate of $-.010\ S^{-1}$ and a station spacing of 2 km would yield an estimate of $-.005\ S^{-1}$, (Figure III-1). Similar calculations for microbursts, using a minimum value of 10 m/s wind velocity reversal over 1 km, leads to a threshold value of $+.010\ S^{-1}$. Because of the rapid decrease of wind field intensity outside of the divergence region of the microburst, sparse station spacing can be expected to decrease detection capabilities.

The simplest estimation technique that could be employed is to compute the one-dimensional numerical derivative along the line-of-sight between two stations. For this, the wind field velocity vector at each station is projected onto the line-of-sight of the stations, and the difference quotient of these projected vectors is computed (Figure III-2). This estimate is called line divergence.

If there is divergence along the line-of-sight, then this calculation will yield an accurate estimate of its intensity. If the maximum divergence is along a line that is skewed to the line-of-sight, then this calculation will underestimate the intensity of the event. Since it is only effective in cases where the stations are properly positioned with regard to the event, this method of detection does not lead to a viable

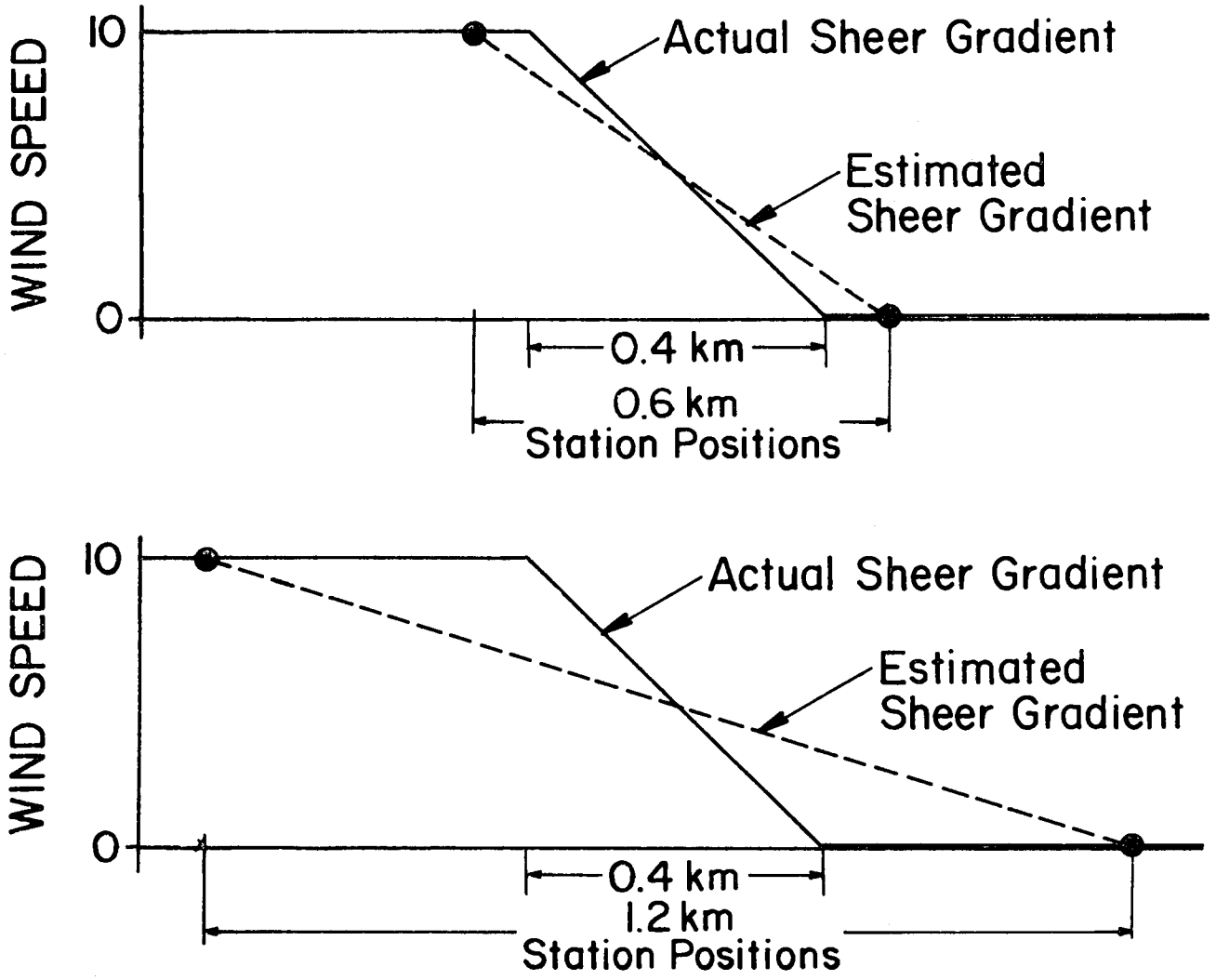


Figure III-1. Underestimation of wind shear intensity due to station spacing.

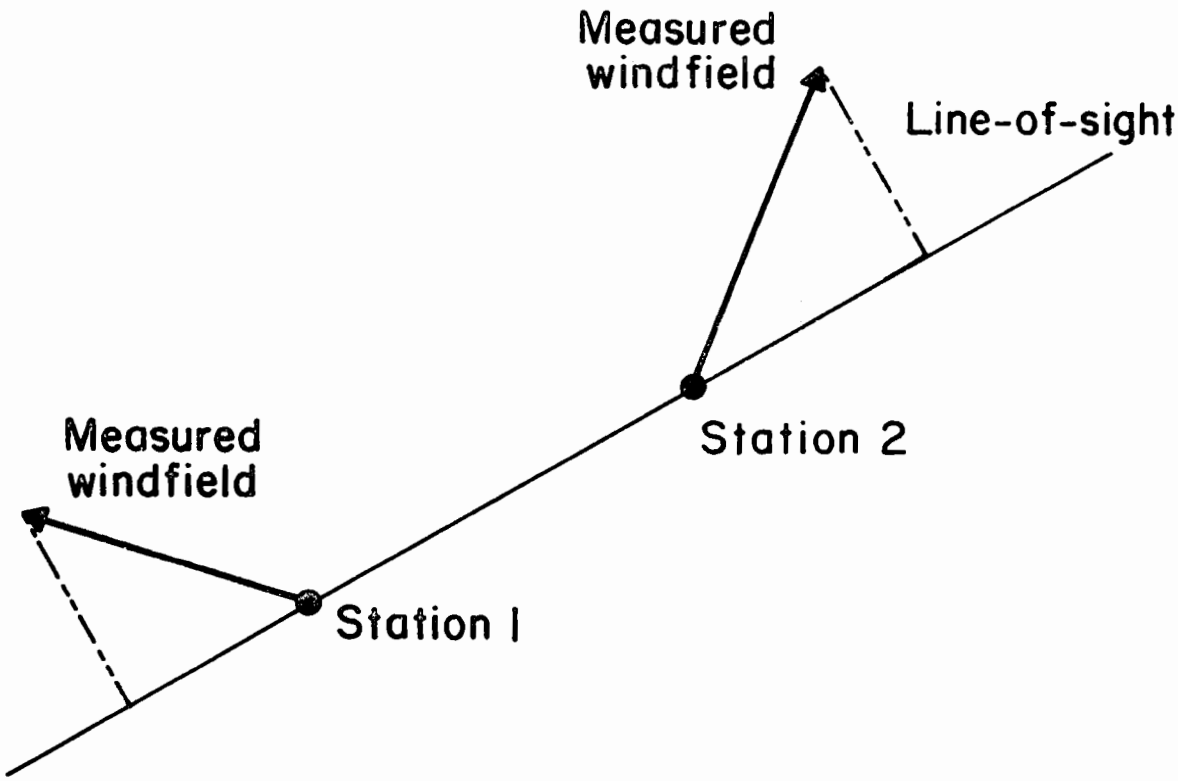


Figure III-2. Line divergence. The divergence is estimated by the vector difference of the projections onto the line-of-sight.

algorithm by itself. A further difficulty with this method is that it may be sensitive to measurement error at one of the stations, and could be prone to issuing false alarms. This also is a possibility when this method is included as part of another algorithm.

A second estimation technique is based on a triangulation of the network geometry and an estimation of the net wind field flux across the boundary of each triangle in the network. As previously discussed, this value can be efficiently computed by a vector cross-product formula, and is called triangle divergence (Figure III-3). It is the basis for the first divergence algorithm.

The triangle divergence method is not as sensitive to the orientation of the event as line divergence, but there still are difficulties. If the event is contained in the triangle and its scale is small compared with the triangle scale, then the method will underestimate the intensity of the event. Also, the location of the event is important, because an event that is partially outside a triangle does not have as severe a net flux at the boundary as an event that is centered within the triangle (Figure III-4).

A second divergence algorithm is based on the combination of the two methods described above with the additional observation that an event partially outside of a triangle (where triangle divergence will produce an underestimation of divergence) must overlap at least one edge of the triangle. This algorithm consists of computing the triangle divergence and the line divergences along each of the three edges and comparing the maximum of these four estimates to a threshold value. The increased sensitivity of this algorithm could lead both to better detection and to more false alarms.

In summary, two algorithms are used to estimate wind field divergence:

1. Triangle divergence-convergence (TDC)

   The station network is triangulated and the triangle divergence is computed for each triangle. The triangles, rather than stations, are designated as divergent or convergent.

2. Triangle and edge divergence (TEDC)

   The station network is triangulated and the triangle divergence is computed in each triangle. Then the line divergence is computed for each of the three edges of the triangle. The divergence value of the triangle is the maximum of these four estimates and the convergence value is the maximum negative value. Again, the triangles are designated as divergent or convergent based on the sign of the triangle divergence.

The divergence pattern of the triangular regions should allow one to distinguish the microburst from the gust front wind shear event.

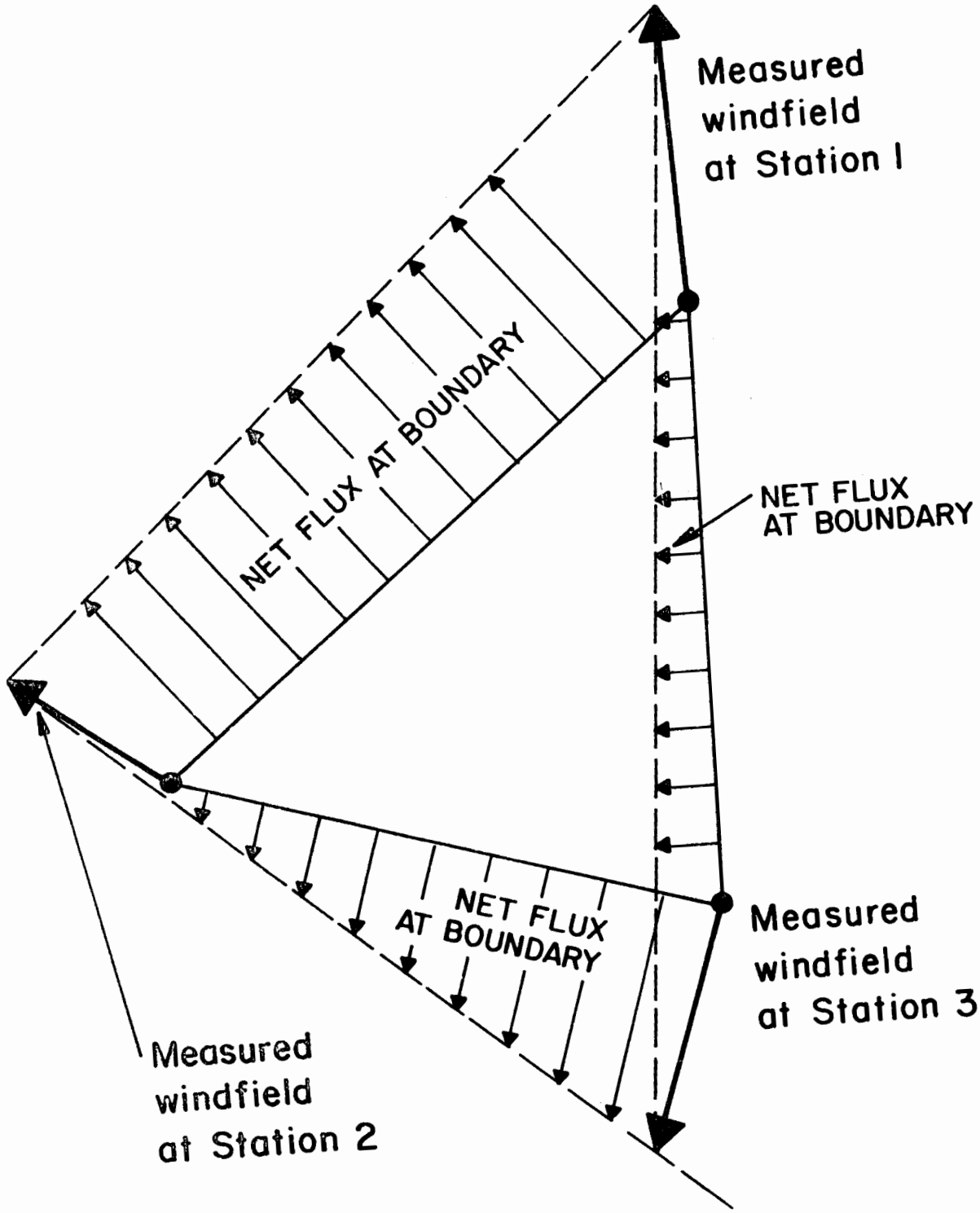


Figure III-3. Triangle divergence as a measure of net flux at the boundary.

FAA WJH Technical Center
00093125

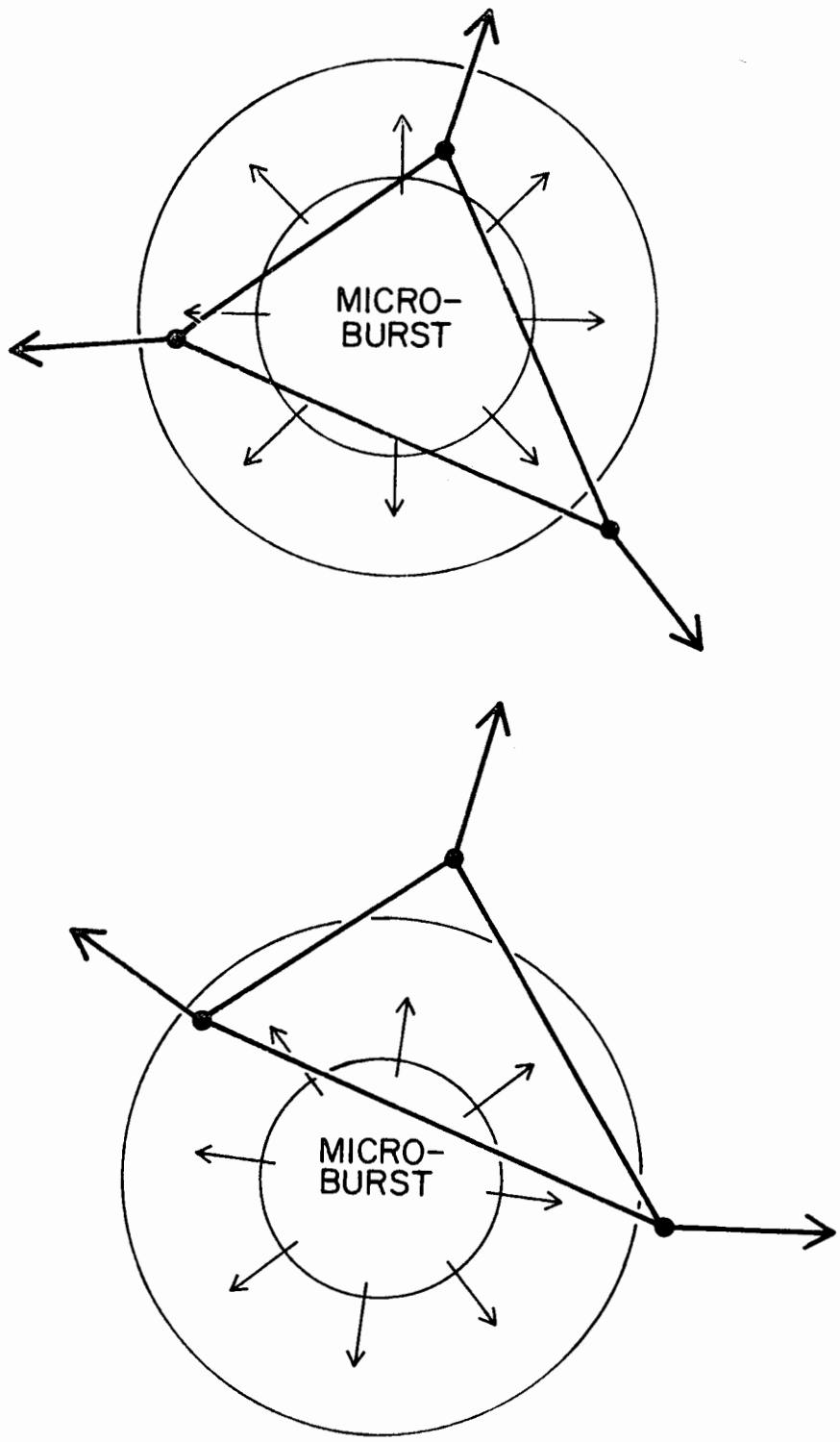


Figure III-4. Impact of event location on triangle divergence.

IV. Mathematical Wind Field Model

For the mathematical wind field model, we wanted to produce a two-dimensional model (i.e., u and v components) which would satisfy the mass continuity condition. Furthermore, we desired not only to simulate rather closely the real microburst and gust front events but to allow for an ambient wind field, a "spurious" or single station anomaly event, and measurement noise. Hence, the mathematical wind field model consists of the following components: (1) an ambient wind field, (2) a wind shear event, (3) the spurious event, and (4) measurement noise. The model wind field is the vector sum of these four component parts.

The ambient wind field satisfies a linear vector field model (i.e., a bivariate regression model) that is similar to events observed in the CLAWS[1] LLWAS data and the JAWS PAM[2] data. The wind shear event is initiated randomly in time and geographic location, and is either a simulated microburst, gust front, spurious event, or null (no) event. Lastly, random noise is imposed on each station value and is parameterized to be similar to noise observed in recorded LLWAS data.

After studying several segments of the CLAWS LLWAS data, it has been determined that the 10-second data of the current LLWAS have a noise structure that appears to be normally distributed about their mean value. The direction error does not seem to vary with speed for speeds greater than 5 kts and has a standard deviation of approximately 6.5°. When mean speed is below 5 kts (2.5 m/s), the direction error is enormous. This is similar to the range in which the National Weather Service reports winds as "light and variable." The speed error increases linearly with wind speed above 5 kts, and the data indicate a standard deviation of approximately 0.14 times the mean wind speed. These values have been used for the noise model.

Specifically, the Cartesian ambient and wind shear event wind fields are combined and converted to polar form, and then the noise is selected from the appropriate normal family (for direction, $\sigma = 6.5°$; for speed, $\sigma = 0.14\ \mu$).

The choice of a linear vector field model for the ambient wind field allows us the flexibility of starting with a translating wind, a regional convergence or divergence, or perhaps a drainage flow for the background prevailing wind field. The mathematical form of such a wind field at position (x,y) is

$$u(x,y) = u_0 + u_x(x - x_0) + u_y(y - y_0) ,$$

$$v(x,y) = v_0 + v_x(x - x_0) + v_y(y - y_0) ,$$

[1] Classify, Locate and Avoid Wind Shear experiment conducted at Stapleton Airport, Denver, CO, July 2 - August 15, 1984.

[2] Portable Automated Mesonet operated by the Field Observing Facility, NCAR.

where u_x, u_y, v_x, v_y denote partial derivatives of u and v.

Realistic linear wind fields have been developed by the analysis of the JAWS PAM data and the CLAWS LLWAS data. The superposition of a wind shear event and typical wind field "noise" on this ambient behavior provides a realistic test for the various detection and identification algorithms.

The spurious event is an attempt to simulate small-scale wind events such as dust devils, wake vortices, etc., which occasionally cause a station to exhibit anomalous behavior. For the purpose of this study, these local events are not considered to be hazardous to aircraft, although wake vortices can affect aircraft. Typically, such events affect only one station for a brief time interval. These events are modelled to occur infrequently, to have speed magnitude of 7.5 m/s, to have a random direction, and to persist for 30 seconds. Any detection of the spurious event is viewed as a false alarm.

A simple gust front model has been selected in which the front is a straight line that moves across the region with a front speed of about 10 m/s. This wind field has constant direction, wind speed of zero ahead of the front, wind speed of 10 m/s at distances of .4 km or more behind the front, and wind speed that increases linearly from 0 to 10 m/s in the .4 km band along the front (Figure IV-1). The constant wind direction may be slightly skewed from the direction of frontal propagation, and the angle of the skew is randomly selected from a small interval centered about zero. This event will persist until it crosses the entire region, at which time it will have added a constant vector value to the ambient wind field. This additive constant is then allowed to decay uniformly to zero over a fifteen-minute period. Note that a microburst may be initiated during this transition or decay period.

The most complicated event is the microburst. A radially symmetric outflow from a central position has been selected (Figure IV-2). In order to scale the radial outflow properly, it is necessary to impose mass continuity on the modelled wind field. The effect of this requirement is detailed in Appendix C. Relying upon a statistical analysis of the JAWS microburst data (Wilson et al., 1983) we have selected the following parameters to describe the modelled microburst (Figure IV-2):

1. Time for increase to maximum intensity = 6.3 min,
2. Initial radius of the divergence region = 0.9 km,
3. Final radius of the divergence region = 1.6 km,
4. Height of the outflow = .3 km,
5. Initial maximum horizontal velocity (one side), and = 5.7 m/s
6. Final maximum velocity (one side) = 11.8 m/s.

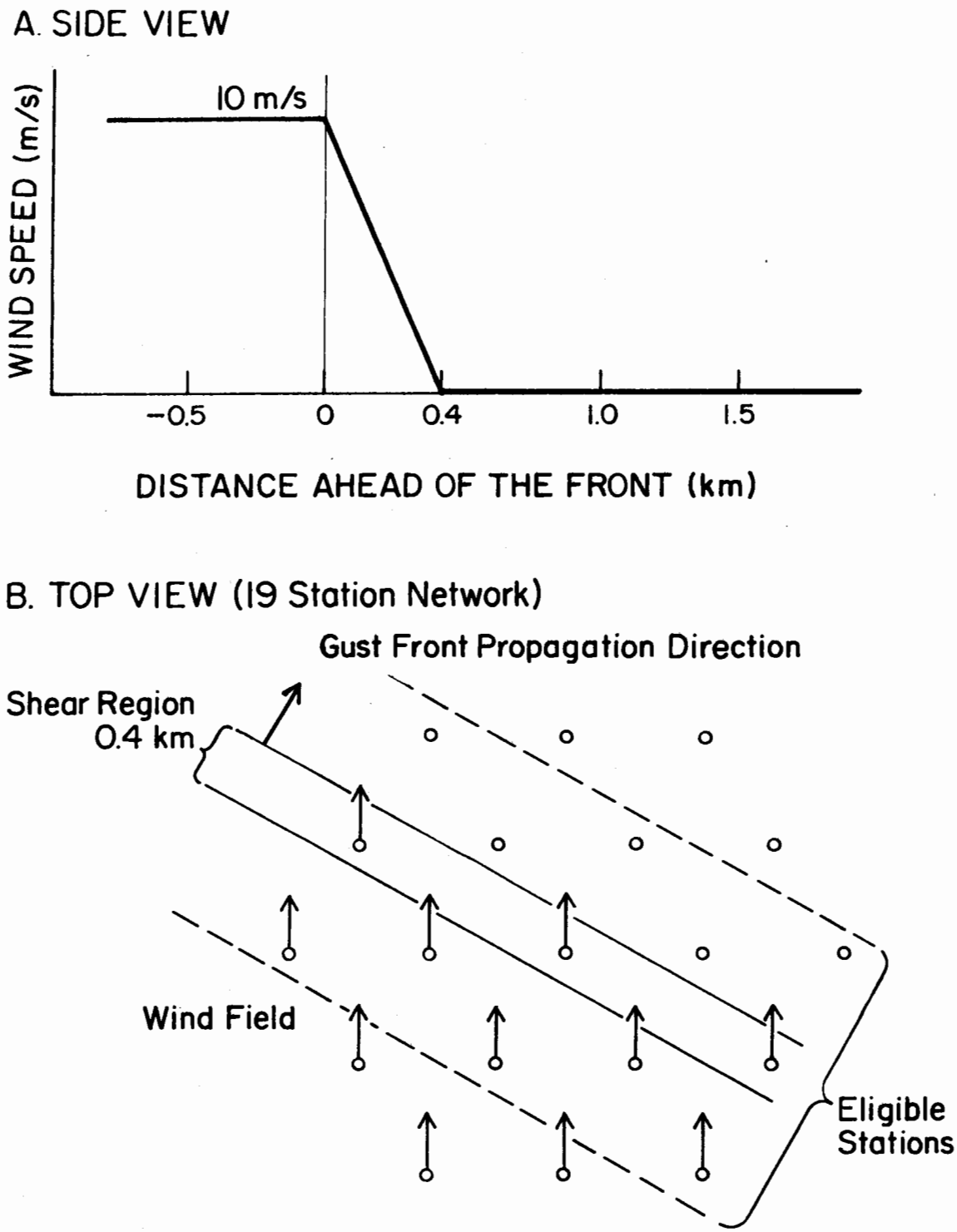
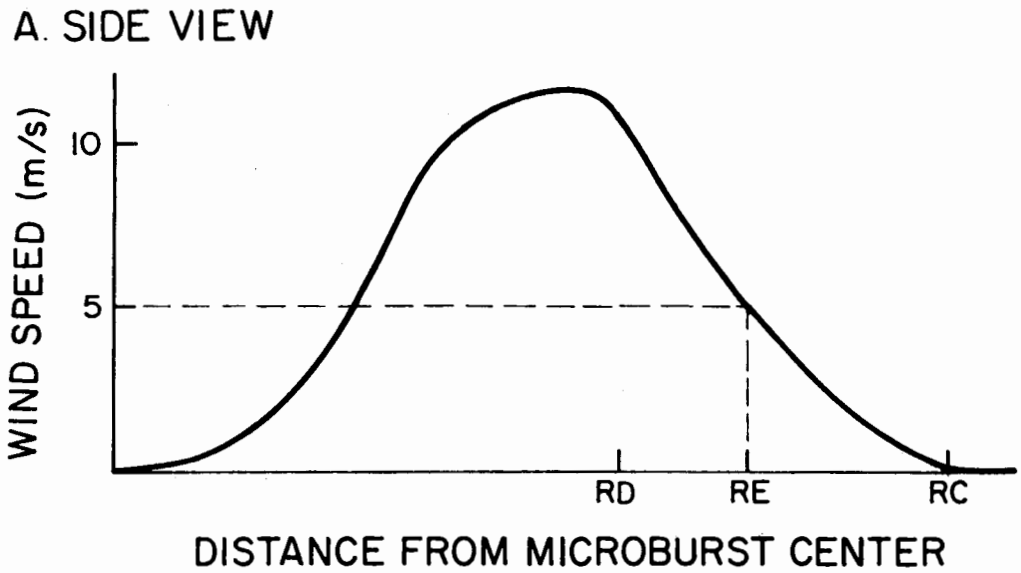


Figure IV-1. Gust front model.



RD = Radius of Divergence Region
RC = Outer Radius of Convergence Region
RE = Radius of Eligible Station Region

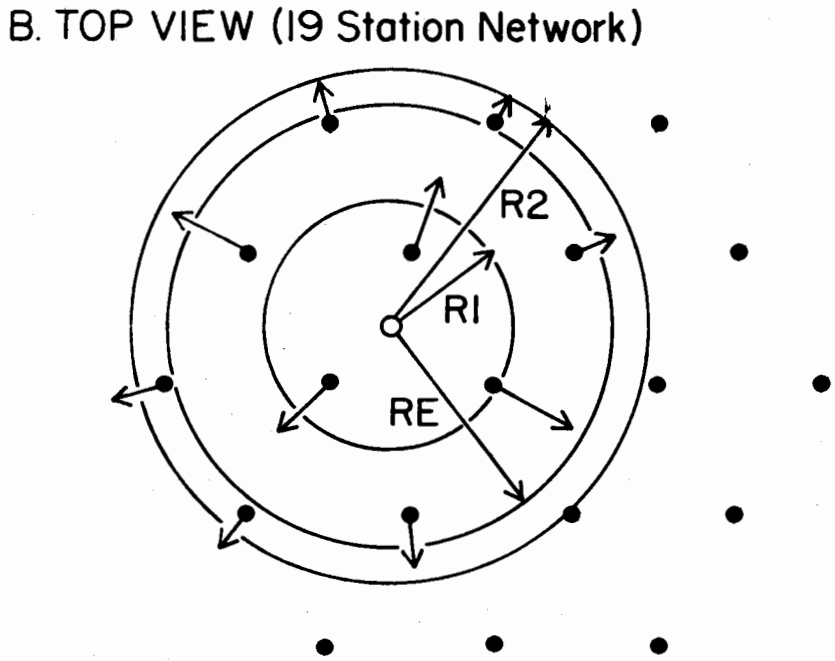


Figure IV-2. Microburst model.

It is assumed that the radius of the divergence region expands linearly and that the maximum horizontal velocity also increases linearly over this time period. The model algorithm computes the strength of the horizontal outflow and the outer radius of the convergence region as a consequence of the choice of parameters. At the end of the 6.3-minute period, the event decays into the ambient wind field by allowing the maximum horizontal velocity to decay to zero over an additional 10-minute period. During this time the radius of the divergence continues to expand linearly. Thus, the total life of the modeled microbursts is 16.3 minutes, although it typically is not hazardous for this full time period. Sample microburst and gust front events as seen by a regular grid network are shown in Figures IV-3 and IV-4.

For some of our evaluations we designated stations as eligible or ineligible to observe the wind shear event. A station is defined as eligible if it is contained in the shear region at any sampling (polling) time. If it is not in the shear region, it is defined as ineligible. For the gust front, the shear region is the .4-km band along which the wind speed is increasing from 0 to 10 m/s incremented by the 1-km band on either side of this shear (Figure IV-1). For the microburst, the shear region is the disc about the center that contains the full divergence region and the portion of the convergence region in which the microburst wind speed is greater than 5 m/s (Figure IV-2). The radius of this disc is the eligibility radius.

In a testing of the algorithms, the initiation times and locations of the wind shear events are selected randomly in the geographic region that is to be detected. Gust fronts are constrained to start at the boundary and move across the region; the centers of the microbursts remain stationary, but the radii of the divergence regions increase with time, as is described above. Figure IV-5 shows (A) a list of the events and their initiation times and (B) the geographic dispersion of the wind shear events in a hypothetical protected rectangular region, near Stapleton. This is the result of the six-hour simulation that is used in Appendix F.

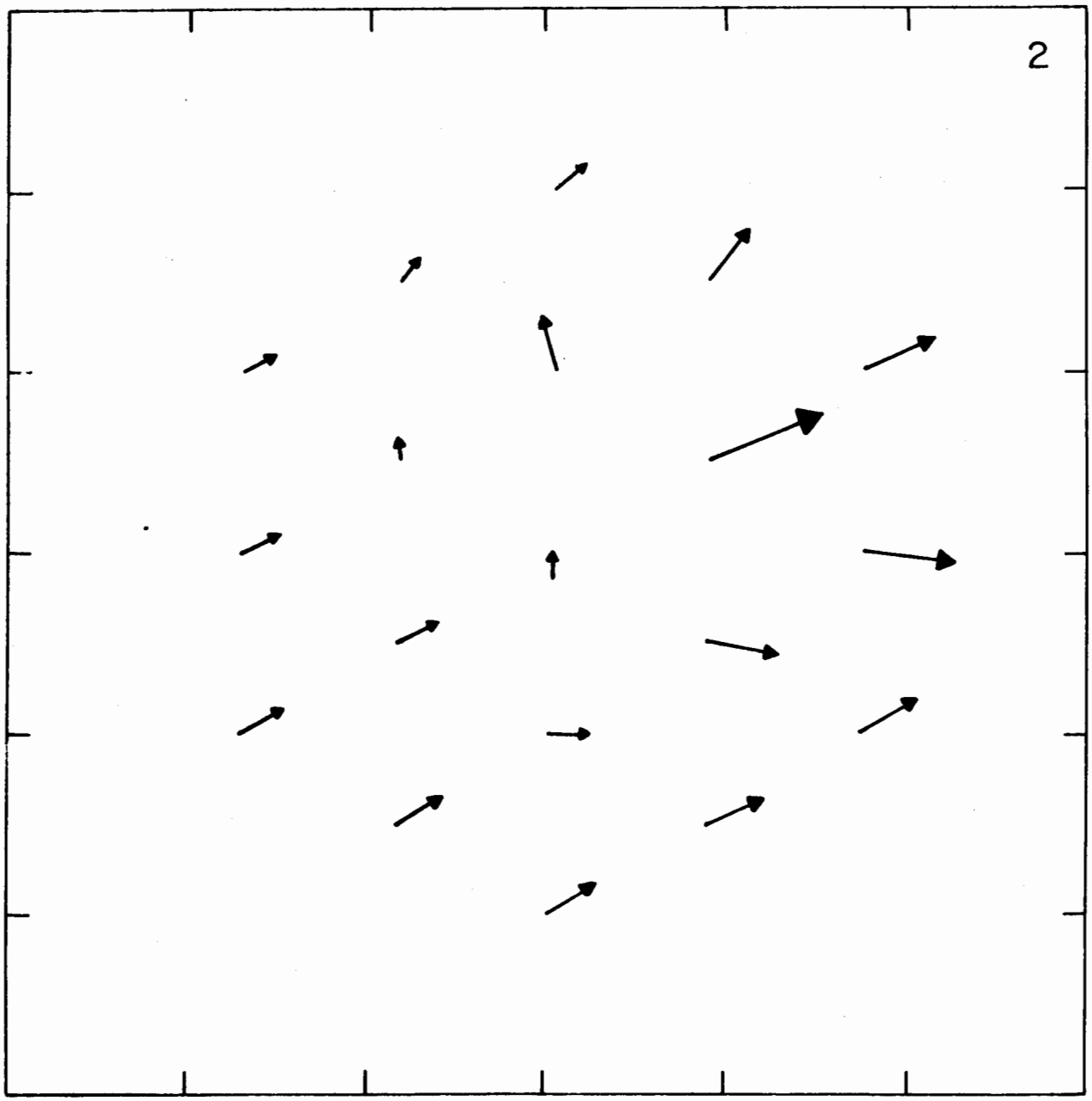


Figure IV-3. A sample microburst in an ambient wind field, generated by the mathematical wind shear model on a 19-station network.

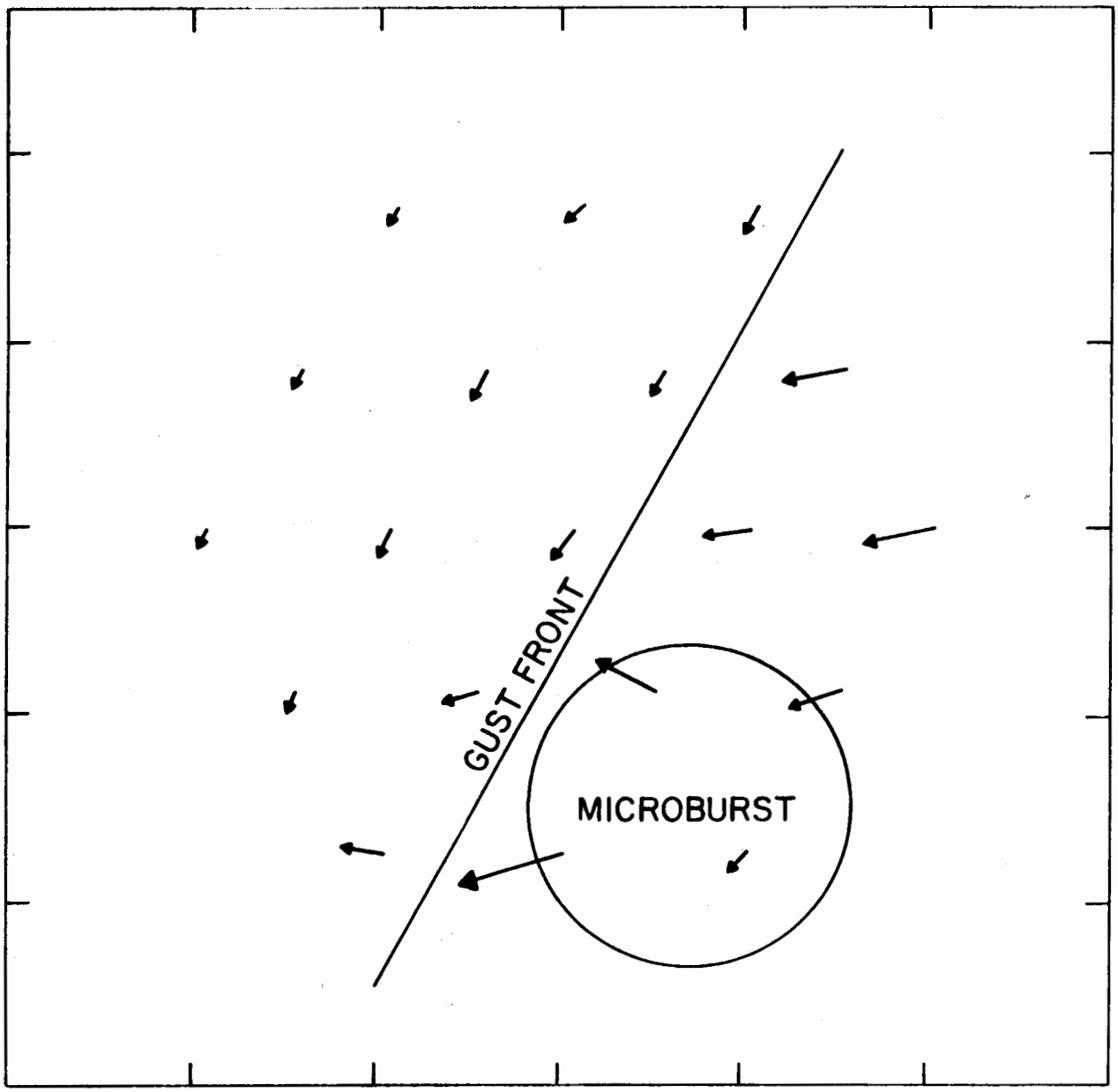


Figure IV-4. A sample gust front and microburst in an ambient wind field, generated by the mathematical wind shear model on a 19-station network.

LABEL	EVENT	INITIATION TIME
MB 1	Microburst	1.800
	Spurious Event	12.000
	Spurious Event	13.000
GF 1	Gust Front	17.000
MB 2	Microburst	18.800
MB 3	Microburst	41.000
GF 2	Gust Front	53.600
MB 4	Microburst	61.400
	Spurious Event	72.000
	Spurious Event	78.000
	Spurious Event	80.600
	Spurious Event	87.400
MB 5	Microburst	88.800
	Spurious Event	104.600
	Spurious Event	114.600
MB 6	Microburst	118.200
GF 3	Gust Front	127.600
	Spurious Event	140.800
MB 7	Microburst	141.800
	Spurious Event	149.800
	Spurious Event	156.600
	Spurious Event	157.400
	Spurious Event	159.600
	Spurious Event	161.800
	Spurious Event	163.400
	Spurious Event	169.600
MB 8	Microburst	170.600
	Spurious Event	179.600
MB 9	Microburst	195.200
	Spurious Event	203.600
	Spurious Event	205.600
	Spurious Event	206.400
	Spurious Event	208.400
	Spurious Event	211.200
	Spurious Event	216.600
GF 4	Gust Front	224.000
MB 10	Microburst	228.400
	Spurious Event	240.600
	Spurious Event	257.800
	Spurious Event	272.400
MB 11	Microburst	282.000
	Spurious Event	303.800
GF 5	Gust Front	306.000
MB 12	Microburst	313.200
	Spurious Event	332.200
MB 13	Microburst	333.000
MB 14	Microburst	357.400

Figure IV-5.A. List of events generated by the wind field model in a six-hour simulation.

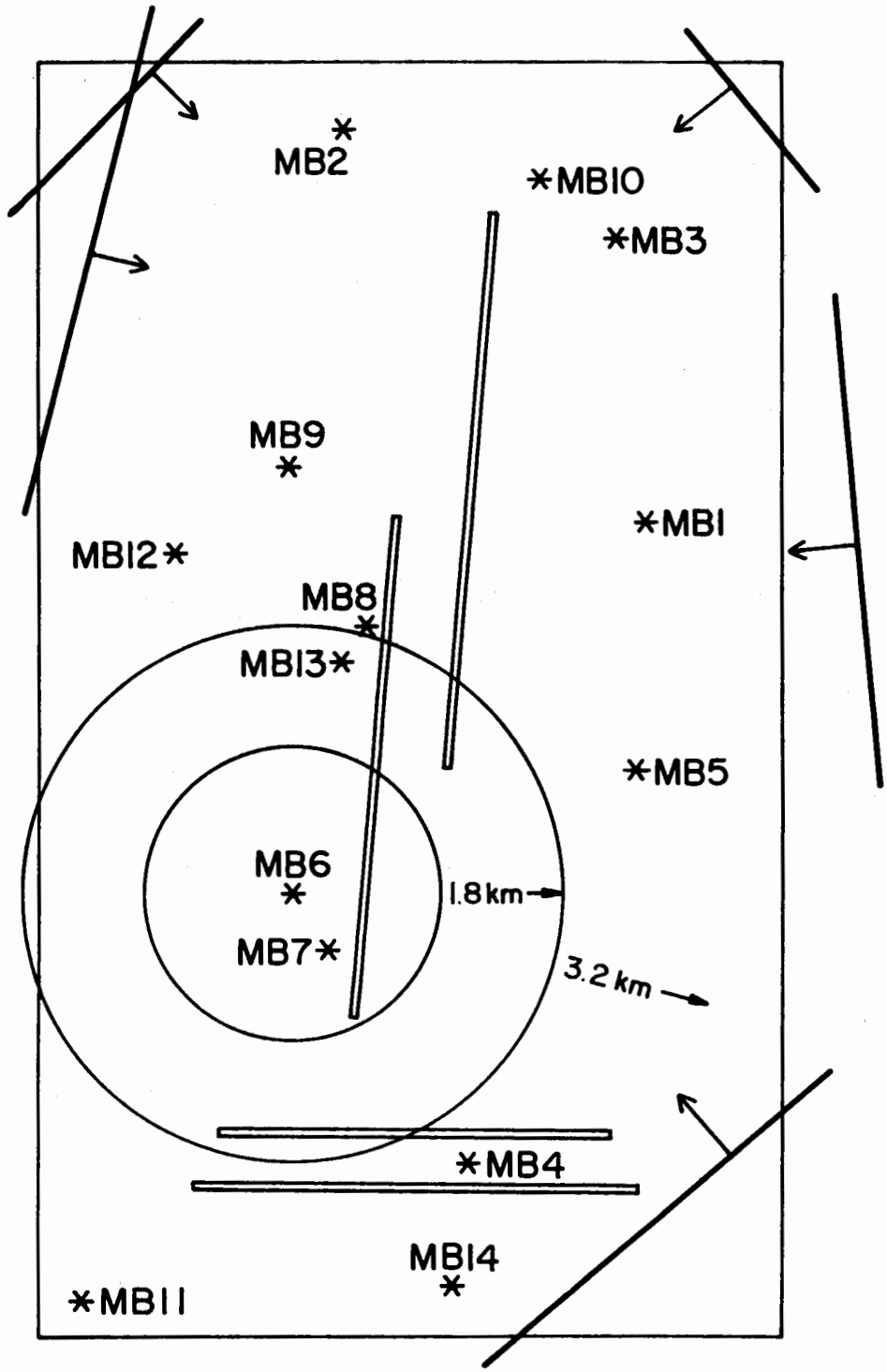


Figure IV-5.B. Distribution of wind shear events near Stapleton (Denver) as generated by the wind field model in a six-hour simulation.

V. Wind Field Filtering or Smoothing

As mentioned earlier, LLWAS typically operates with a six- to ten-second station polling period, which introduces a significant measurement noise in the wind field observations. In the current LLWAS, a two-minute running average is computed at centerfield and used as the background or expected value that is compared with the most current value from each of the five outlying stations.

Time-series averaging of a station's data, of which the running average is one example, is a way to reduce the high frequency component of the wind field variation at each station before the network wind field models are constructed. This smoothing operation is assumed to give a better estimate of the true background wind field. Therefore, we test for anomalies between the current wind field observation at each station and an expected wind that is based on spatial averaging of the time-averaged wind measurements at each station in the network over an arbitrary time period (e.g., LLWAS is typically two minutes).

Two averaging methods have been studied: the running average method used in the current LLWAS and a comparable recursive filter method. Viewed as a digital filter, the running average has several weaknesses. Although it is intended to be a low-pass filter, the running average actually allows many of the higher frequency components of the signal to leak through with reasonable power (Hanning, 1977, page 34). In addition, when there is a linear shift in the intensity of the wind field, the two-minute running average has a time lag of one minute in that the averaged wind speed matches the actual wind speed one minute previous to the current time. The comparable time lag for the recursive filter (a = .6) is about 15 seconds.

The digital filter that we have applied in this study is a single pole recursive filter of the form

$$Y_k = (1-a)\, Y_{k-1} + a\, X_k \qquad (1)$$

where the sequences X_k and Y_k are the sequences of measured data and filtered data, respectively, and a is a weighting factor that is selected between 0 and 1. This compares with the running average of size n, which is defined by

$$Y_k = \frac{1}{n}\left[X_k + \ldots + X_{k-n+1} \right] \qquad (2)$$

$$= Y_{k-1} + \frac{1}{n} X_k - \frac{1}{n} X_{k-n}$$

To successively compute the first formula, the last Y-value and the current X-value are stored, while to successively compute the second formula, the last n X-values must be stored. The two methods require comparable computational work, and for a 10-second polling rate, the running average has n = 12, and the comparable recursive filter has a = .6.

For the anomaly methods in this report, three smoothing approaches are used to generate the background or expected value:

1. Recursive filter of wind field speed and direction with a = .6,
2. Recursive filter of wind field components E-W (i.e., u) and N-S (i.e., v) with a = .6, and
3. A running average with a period of two minutes.

Differences due to the smoothing method were very small compared with the differences based on choice of algorithm and station spacing.

The divergence methods require a different approach. First, there is no comparison of individual station values with an expected model, and hence it is not necessary to use the severe smoothing that is used for the anomaly methods. Second, the divergence methods involve the numerical estimation of wind field derivatives, and these computations are sensitive to measurement noise. Thus, in the divergence studies smoothing was obtained by using a = .85 in the recursive filter.

To test the sensitivity of the selection of a = .85, several simulations were performed with the only variation being the selection of the filter constant "a." Results varied little with changes in this parameter. For example, the following values were obtained for 13 stations with 2.25 km spacing, TDC and TEDC methods, threshold = 4.00 (10^{-3} s^{-1}) and a = .70, .75, ..., 1.00 (Table V-1). We observe a mild improvement for modest smoothing over not smoothing at all (a = 1.00).

Table V-1. Sensitivity study of the recursive filter for 13 stations, scale 2.25, threshold 6.00 and "a" varying from .70 to 1.00.

Scoring Method[1]	a = .70	.75	.80	.85	.90	.95	1.00
				(TDC)			
TSS	.67	.67	.67	.67	.66	.66	.66
POD	.78	.78	.78	.78	.78	.78	.79
FAR	.13	.13	.14	.14	.15	.16	.16
EL.POD	.71	.71	.71	.71	.71	.71	.71
EL.FAR	.17	.17	.17	.18	.19	.19	.19
				(TEDC)			
TSS	.68	.68	.68	.68	.67	.66	.66
POD	.79	.79	.79	.79	.79	.79	.79
FAR	.13	.13	.14	.14	.15	.16	.16
EL.POD	.71	.71	.71	.72	.72	.72	.72
EL.FAR	.17	.18	.18	.19	.19	.20	.20

[1]See Section VII for definitions.

VI. Network Geometry

Given an area that is to be monitored for severe wind shear events by a network of anemometers, the question becomes what sampling points or station network geometry is desired. There are three principle factors that greatly influence the detection effectiveness of a network: (1) the number of stations, (2) the spacing between nearby stations, and (3) the degree of irregularity of the station spacing. With regard to the last factor, we have followed the mathematical coverage theory (Ripley, 1981) and attempted to utilize a regular grid network (i.e., proximate stations have equal distance from each other). We typically achieve this result by placing the stations at the vertices of a tessellation of a planar region by equilateral triangles. Most of this study deals with the effectiveness of various detection and identification algorithms in the presence of a regular station geometry or at least a nearly regular station geometry.

1. Network Geometry

Networks with 6, 7, 11, 13, and 19 stations have been considered (Figure VI-1). Each of these networks has a station at its center that is used for the centerfield station in the centerfield algorithms. The regular 7-station network has its outer stations at the vertices of a hexagon and, with the center station, produces a triangulation by equilateral triangles. However, the regular 6-station network has its outer stations at the vertices of a pentagon and triangulation results in isosceles triangles (not equilateral) since the edges that lie on the boundary of the region are somewhat longer than the interior edges due to the larger angle (72 degrees) at the center station.

The regular 11-station and the 13-station geometries are constructed by using two concentric pentagonal and hexagonal station patterns, respectively. The regular 19-station network is based on two concentric hexagonal rings and a complete triangulation by equilateral triangles.

Because there is a large gap between the number of stations in the true regular geometries (i.e., 7 and 19), it was decided to include the intermediate, almost regular, 13-station geometry. This geometry has equilateral triangles except for the thin triangles (slivers) along the boundary. The 6-station network is studied because it is most representative of the network that is used in the current LLWAS. Lastly the 11-station geometry is an idealized version of the network that has been deployed at New Orleans and will be deployed at Denver in the summer of 1985.

2. Scaling

In order to test the effect of station spacing, a scaling parameter has been introduced. If the center station is assumed to have coordinates (0,0), then each station is moved radially outwards when its position coordinates are multiplied by a factor larger than 1 (e.g., Figure VI-2).

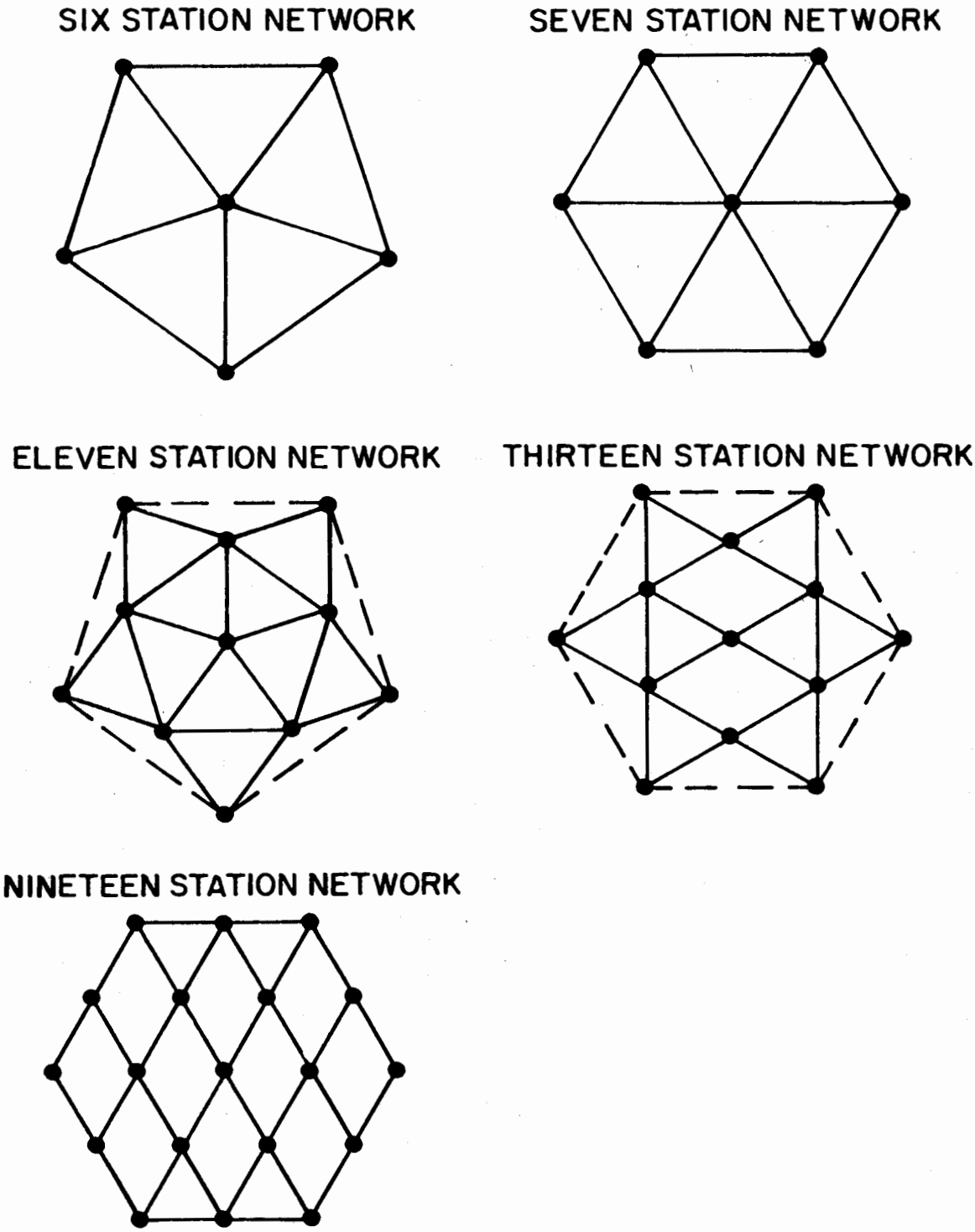


Figure VI-1. Regular station network geometries.

The original 11-, 13-, and 19-station networks were constructed with a 1-km distance between the proximate stations. When it was discovered that comparable detections could be achieved with a more widely spaced network (and hence monitor a larger region), it was decided to utilize scale factors of 1.5, 2.25, and 3.0 for these cases.

For the smaller numbers of stations (6 and 7), the station spacing is taken so that the outer boundaries of the networks have the same positions as the outer boundary of the 11-, 13-, and 19-station networks, respectively. Scale factors of 1.5 and 2.25 are used in this case, since system performance is uniformly poor for scale factors of 3 or larger. Figures VI-2, -3, -4 show the scales and the resultant regular station geometry for the 6, 7, 11, 13, and 19 station networks monitoring a 5-km radius disc.

3. Irregular Spacing

When an irregularly spaced geometry is used, numerical instability may be introduced into the performance of the detection algorithms by strongly skewed triangles or "slivers." This instability is most noticeable for the divergence algorithms. This problem is illustrated by the effects of the slivers in the 11-station and the 13-station networks. For each of these networks, the divergence algorithms were run with and without the boundary sliver triangles. Again, a 5-km radius disc was "protected", as in the other testing. The results of these tests are shown in Table VI-1.

Table VI-1. Example of the numerical instability produced by "sliver" triangles.

(SCALE = 2.25)

NO.	METHOD	THRES[1]	SLIV	TSS[2]	POD[2]	FAR[2]	EL.POD	EL.FAR
13	TDC	4.00	YES	.72	.81	.11	.71	.17
13	TDC	3.00	NO	.69	.81	.14	.69	.20
13	TEDC	4.00	YES	.73	.82	.11	.71	.18
13	TEDC	3.00	NO	.73	.87	.15	.74	.22
11	TDC	5.00	YES	.55	.69	.18	.58	.20
11	TDC	3.00	NO	.55	.72	.20	.61	.22
11	TEDC	5.00	YES	.56	.70	.18	.59	.21
11	TEDC	3.00	NO	.60	.77	.20	.67	.23

[1] units of $10^{-3}S^{-1}$
[2] see Section VII for definitions

Note that the 13-station network typically has a degraded performance when the slivers are removed, but that the system can be somewhat improved using a slightly lower threshold. This suggests that the outer triangles

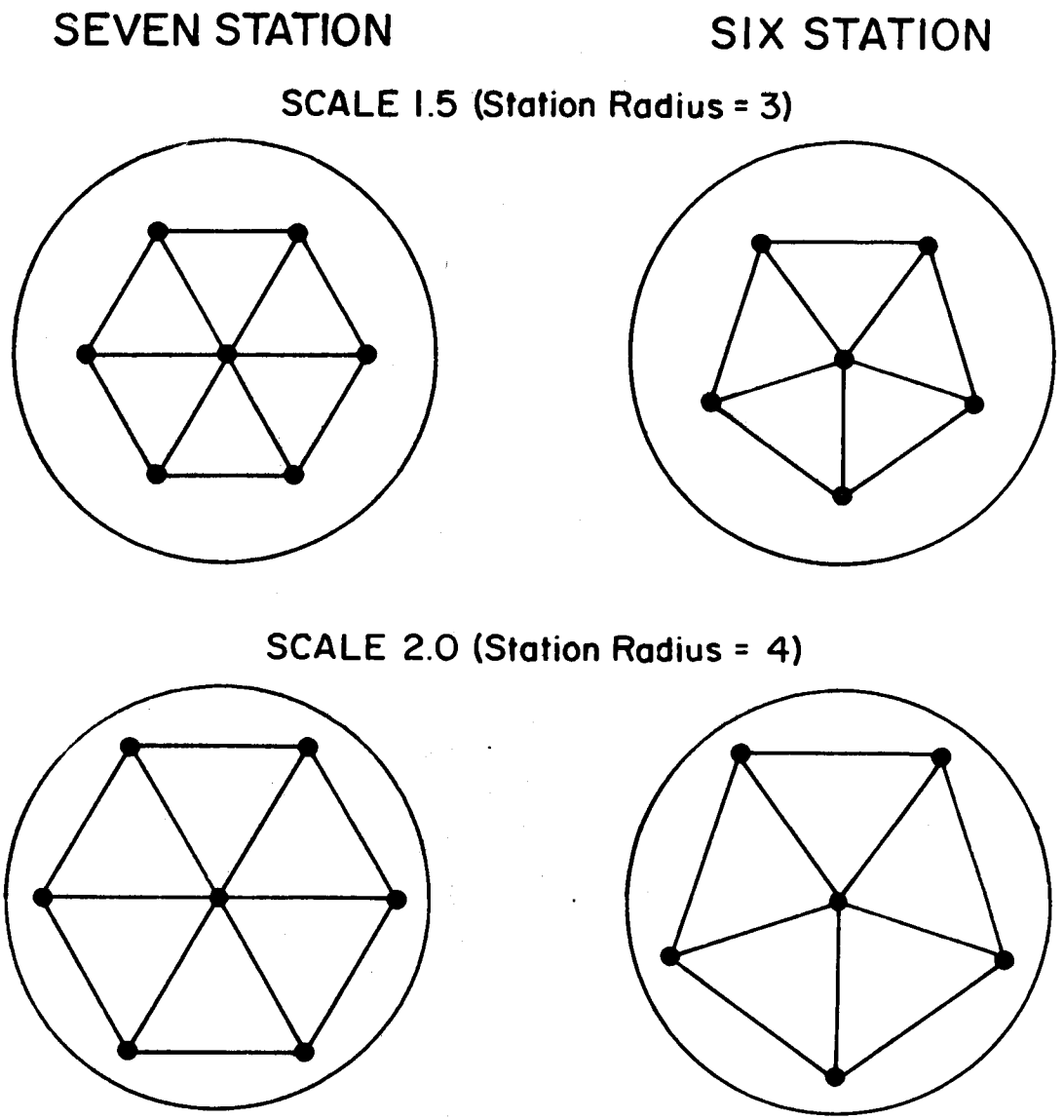


Figure VI-2. The station network geometries for 6 and 7 stations, two scales, within a 5-km radius disc.

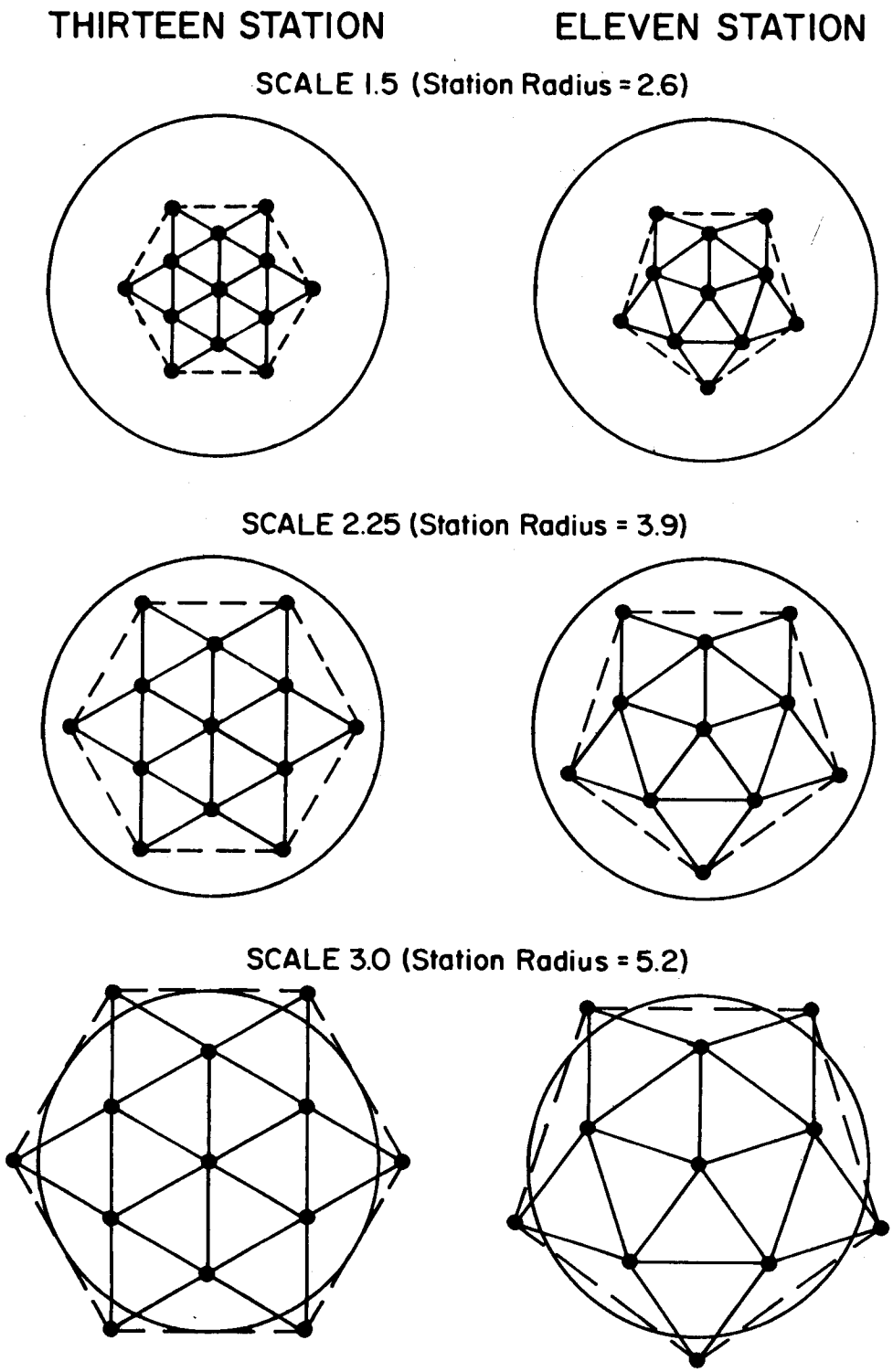


Figure VI-3. The station network geometries for 11 and 13 stations, three scales, within a 5-km disc.

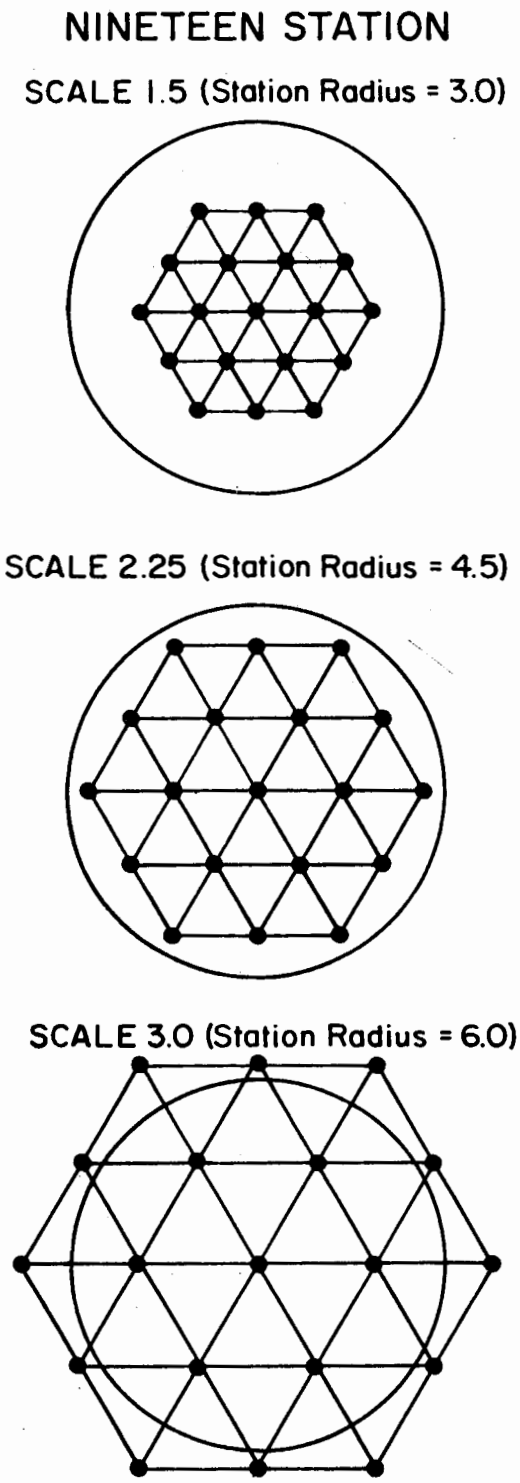


Figure VI-4. The station network geometries for 19 stations, three accompanying scales, within a 5-km radius disc.

(slivers) are providing a useful detection service, but that the numerical algorithm has become somewhat noisy and that to compensate it is necessary to raise the threshold. Of course, taking the higher threshold does degrade the sensitivity of the algorithm on the interior triangles.

The effects that are noted above seem more pronounced for the 11-station network. Examination of the geometries shows that the slivers are considerably thinner in this case and the system performance appears to reflect the effect of these slivers. Now there is a bigger difference in the optimal thresholds. The TDC algorithm without slivers gives a slightly inferior performance, presumably because of missed or delayed detections at the boundary of the region. When the edge divergence is added, the system appears to perform equally well with or without slivers. It is important to note a substantially lower threshold is used without slivers (i.e., 7.00 versus 4.00).

For an irregular geometry, the divergence algorithm can be sensitive to very thin triangles. One solution is to set the threshold higher when there are very thin triangles, realizing this will degrade the algorithm's performance in the region where the geometry is more regular. Another solution is to set the threshold separately for each triangle, depending on its degree of thinness. These results indicate that additional understanding of the sliver problem is important in the implementation of these methods for irregular grids.

VII. Evaluation and Test Procedures

Given the known mathematical wind field model (Section IV), the network geometry (Section VI), and the numerous detection algorithms (Section III), we now need an evaluation and testing procedure to properly compare the performance of each wind shear detection algorithm. In the terminology of forecasting research, we need a testing and verification system. The verification technique used in this study is based on the 2x2 contingency table, and thus it requires the dichotomization of both the time series of wind shear events and the time series of algorithm detections. A more detailed discussion of the verification techniques can be found in Appendix D.

A. Dichotomization of Wind Shear Events

We assume that the simulation time period (i.e., total run time) is known and is partitioned into polling intervals of 12 seconds duration. This is the basic unit of the wind shear event time series as generated by the mathematical wind field model. Thus, for each polling interval and a given station there is (1) or is not (0) a wind shear event at that station. Now the dichotomized or binary simulated wind shear data series can thus be summarized by a finite sequence such as

$$(0,0,0,0,1,1,1,0,0,0,0,0,0,1,1,1,1,1,1,0,0,\ldots,0).$$

This approach can be extended to a network of stations by simply coding "1" or "0," depending on whether there is a wind shear event anywhere on the network during the polling interval.

B. Dichotomization of the Algorithm Output

Each detection algorithm, whether the anomaly or divergence type, produces output for each station in the form of a computed real number that is converted to a second binary sequence when a predetermined threshold value is chosen. If the algorithm's station value is less than the threshold, then the detection algorithm is credited with no detection of a wind shear event (i.e., "0") at that station; and if it is equal to or greater, it is credited with a detection of a wind shear event (i.e., "1") at the given station.

Of course this dichotomization can be extended to polling intervals and the entire station network. In short, if a polling interval has any network station at or above its threshold value, then the algorithm will claim a detection for the entire network.

The conversion of continuous information (i.e., algorithm output) to binary information (i.e., "0" or "1") is necessarily somewhat arbitrary since very small changes in the wind field may push the output of the analysis past the threshold value if the computed value is close to the threshold. Lowering the threshold will increase the number of both valid and invalid alarms and raising the threshold will decrease them. The choice of the detection threshold for each algorithm is a critical factor in the performance of the algorithm.

In our simulation study, we have attempted to "fine tune" the threshold value for each of the algorithms by adjusting the threshold to maximize the TSS. This is the same as maximizing the slope between the observed wind shear events and the detected wind shear events in a 2x2 table. Thus each of the algorithms is tested at its own predetermined threshold that provides the "best skill" for detection and nondetection in the given situation. In an operational setting, the criteria for the choice of the threshold value may differ from the one used in this simulation study in order to accommodate operational conditions.

The above algorithm dichotomization has allowed any network station to enter into the detection of a wind shear event. Thus, the actual wind shear event could be on one side of the network and a station on the other side of the network could accidentally produce a value above its threshold value and the algorithm would claim a wind shear detection. To handle this problem we have employed the eligible-ineligible classification of stations in the anomaly and the divergence detection methods. Recall that stations are defined as eligible or ineligible to detect the wind shear event (Section IV), depending on whether or not they are in the shear region of the wind field.

For the divergent methods it is necessary to define when a triangle is eligible or ineligible. We have adopted the rule that if at least one vertex station of a triangle is eligible, then the entire triangle is considered to be an eligible triangle. If none of the stations is eligible, then the triangle is considered ineligible. Of course one could require more than one vertex station to be eligible to classify as an eligible triangle, but we have found this restriction to be of little use in this study.

C. Contingency Table and Scoring

Having dichotomized the two variables (i.e., wind shear event and algorithm detection event) on the same wind shear simulation and polling interval, we now can cross-classify each polling interval by employing the traditional 2x2 contingency table commonly used in statistics (e.g., Snedecor and Cochran, 1967). There are four possible joint events in the 2x2 table: (1) there is a wind shear event and it is detected (i.e., a correct detection, CD); (2) there is a wind shear event and it is not detected (i.e., a missed detection, MD); (3) there is no wind shear event and there is a claimed detection (i.e., incorrect detection ID); and (4) there is no wind shear event and no claimed detection (i.e., correct nondetection CND). Table VII-1 presents the general form of the 2x2 contingency table for this situation. The four interior cells record the number of joint events and the margins of the table record the row and column totals. As an example, n_{CD} is the number of CDs in the total number of trials (n_T) and $n_{CD} + n_{ID}$ is the total number of claimed detections.

Table VII-1. The general form of the 2x2 table for the wind shear verification problem.

		Wind Shear Event: Shear	Wind Shear Event: No Shear	
Detection Event	Detect	n_{CD}	n_{ID}	$n_{CD} + n_{ID}$
	No Detect	n_{MD}	n_{CND}	$n_{MD} + n_{CND}$
		$n_{CD} + n_{MD}$	$n_{ID} + n_{CND}$	n_T

In the simulations, each trial (i.e., polling interval) is scored as obtaining one of the four joint events. However, this is an unrestricted method of scoring in that a wind shear event anywhere in the network and any station's claimed detection in the same polling interval will produce a successful result (CD).

D. Evaluation Measures

Using a separate contingency table for each algorithm, one can perform quantitative evaluations and comparisons of the various algorithms. The principal quantitative measures that were selected for the comparative evaluations of the wind shear detection algorithms are:

1. The Probability of Detection (POD)
   $POD=n_{CD}/(n_{CD}+n_{MD})$,
   which simply is the probability of correctly detecting the wind shear events (i.e., $0 \leq POD \leq 1$).

2. The False Alarm Ratio (FAR)
   $FAR=n_{ID}/n_{CD}+n_{ID})$,
   which is the probability of a detection being false (i.e., $0 \leq FAR \leq 1$).

3. The Critical Success Index (CSI)
   $CSI=n_{CD}/(n_T-n_{CND})$
   which is a conditional measure of prediction skill (i.e., $0 \leq CSI \leq 1$).

4. The True Skill Score (TSS)
   $TSS=[(n_{CD} \cdot n_{CND})-(n_{MD} \cdot n_{ID})]/[n_{CD}+n_{MD}).(n_{ID}+n_{CND})]$
   which is a measure of prediction skill beyond that due to simple guessing or chance (i.e., $-1 \leq TSS \leq +1$).

Two other measures which are derived from the above measures will be used in evaluating the comparative testing of the algorithms. They are eligible POD (EL.POD) and eligible FAR (EL.FAR). The former simply is the POD restricted only to the eligible stations rather than the entire network. The latter is the analogue for the ineligible stations in that an ineligible station has claimed a detection when it should not do so.

E. Inferences from Forecast Verification Scores

The standard forecast verification measures (e.g., POD, FAR, CSI, etc.) simply are indices and as such only are population summary measures. Thus, if detection rule D_1 gives a POD of .75 for the population P of wind shear events and another rule D_2 gives a POD of .65 for the same population, then we report that D_1 is a better detection rule than D_2 conditional on P. The question of how would D_1 and D_2 behave on a

different population of wind shear events appears not to be formally addressed in the forecast verification literature. Perhaps P is naively presumed to be large and variable enough to serve for all other possible P's.

An elementary statistical method for answering the above "strength of results" question is to randomly generate k independent subsets of wind shear events of size n and use the binomial probability statement for assessing the probability that D_1 is better than D_2 in all possible future subsets of size n. Thus, for example, if D_1 has a higher POD than D_2 in all five randomly generated subsets, then on the basis of POD the D_1 rule is better than D_2 with only a 0.03 probability that this result is due to chance (i.e., luck). This process is a special case of what has become known as "cross-validation" (e.g., Mosteller and Tukey, 1977). We will use this approach to inference in this study.

TSS differs from the previous measures in that it has a statistical basis and hence is theoretically amenable to statistical inference based on a single random sample. However, in practice some of the assumptions that underlie the inference procedure (e.g., normality, constant variance, etc.) may not be met, and thus caution generally is needed in interpreting these results. Again the cross-validation method is applicable and will be used.

VIII. Results of the Comparative Testing

A number of simulation trials were conducted for the eleven detection algorithms under various parameter combinations to explore the study's multidimensional parameter space. The principle dimensions of this space were (1) number of stations in the detection network; (2) the relative frequencies of the microbursts, gust fronts, and spurious events; (3) the number of stations trimmed (Section III.A.1) from the ranked station wind speed values; (4) the scale or spacing of the regular grid; (5) the radius of the protected disc; (6) the minimum number of simultaneously detecting stations required to produce an alarm; and (7) the minimum detection wind shear value or threshold needed to produce an alarm.

After some initial experiments, we decided that a fixed area should be used for all simulation trials. We selected a 5-km radius disc because it most closely represented the area of protection needed at a typical airport. We also established that a trim of 1 (one station from each end of the ranked station wind data) or 0 was most appropriate for our problem. The trim of 1 was used with the 19-, 13-, and 11-station networks and 0 with the 7- and 6-station networks.

The relative frequencies of microbursts, gust fronts, and spurious events were set to give ample numbers of events and also realistic numbers of events during the active period of the day, as seen at Stapleton during a very active day (CLAWS 1984). In particular, this requires a six-hour simulation and produces an event on the network about 50 percent of the time.

As mentioned earlier, the simulation results suggest that only the 1.5 and the 2.25 scales are effective. Larger scales (e.g., 3.0) degraded the performance measures too much and were omitted after some preliminary simulation experiments.

We determined from initial simulation results that single-station detection produces a better skill performance than requiring simultaneous detection by two proximate stations. Thus, all results presented in this chapter are based on single-station detection. For comparison, Appendix E presents some results for the at least "two proximate station" detection rule.

We have selected four algorithms for comparison and discussion in this section. They are (1) new centerfield (NEWCF), (2) local median (LMD), (3) network wide linear regression (NLR), and (4) the triangle and edge divergence (TEDC). The first algorithm (NEWCF) is presented as an improved LLWAS standard, and the last three typically have better performance over all parameter situations. The comparative results for all eleven algorithms are given in Appendix E.

In addition, we have created a stylized version of the original LLWAS system (OLDCF). This system has five outlying stations and a centerfield station at a spacing of 3 km, threshold of 7.5 m/s, a 2-min moving average of centerfield values, and does not issue an alarm at the centerfield station.

A single-simulation trial for a given algorithm proceeds as follows: (1) generating a synthetic ambient wind field with embedded microburst, gust front, and spurious events using the mathematical model; (2) sampling the wind field during one polling interval (e.g., 12 secs) at the selected station positions; (3) applying the algorithm to each station to see if a wind shear detection is claimed; and (4) recording the joint detection-wind shear event that occurs in a 2x2 contingency table. This procedure then is repeated for a large number of polling intervals or trials (e.g., 1800) to form one simulation experiment. The resulting contingency table for the experiment is evaluated by the methods previously described in Section VII.D.

The following verification scores initially are viewed as population results. Thus, an algorithm with a higher score is presumed to be better than one with a lower score. In Section VIII.C we revisit this question from a sample viewpoint.

A. Results for Scale 1.5

Table VIII-1 presents the comparative performance simulation results for the four above-mentioned algorithms, for five network geometries and a scale of 1.5 km. Also included for comparison are the results for the current 6-station (OLDCF) LLWAS network. Appendix E presents the results for all eleven algorithms. It should be noted that with this scale, the

Table VIII-1. Algorithm performance results for the 19-, 13-, 11-, 7- and 6-station networks, scale=1.50-km simulations.

Network	Algorithm	TSS	CSI	POD	FAR	EL.POD	EL.FAR
A. 19 Stations							
	1. NEWCF	.67	.70	.81	.15	.81	.30
	2. LMD	.71	.73	.83	.14	.81	.18
	3. NLR	.72	.75	.86	.15	.85	.20
	4. TEDC	.68	.71	.82	.16	.74	.22
B. 13 Stations							
	1. NEWCF	.63	.67	.79	.18	.76	.29
	2. LMD	.67	.70	.80	.15	.75	.20
	3. NLR	.71	.73	.84	.15	.78	.29
	4. TEDC	.65	.68	.77	.14	.70	.18
C. 11 Stations							
	1. NEWCF	.60	.65	.78	.20	.68	.28
	2. LMD	.64	.68	.81	.18	.68	.22
	3. NLR	.63	.67	.77	.17	.65	.24
	4. TEDC	.49	.55	.64	.21	.56	.18
		(.54)[1]	(.60)	(.73)	(.22)	(.63)	(.21)
D. 7 Stations							
	1. NEWCF	.60	.63	.71	.14	.64	.24
	2. LMD	.66	.70	.83	.18	.72	.18
	3. NLR	.67	.71	.84	.18	.70	.30
	4. TEDC	.62	.66	.77	.17	.70	.17
E. 6 Stations							
	1. NEWCF	.53	.61	.77	.25	.57	.30
	2. LMD	.56	.63	.76	.22	.61	.20
	3. NLR	.55	.63	.78	.24	.59	.32
	4. TEDC	.57	.62	.74	.20	.62	.21
F. Old Centerfield, six stations, 3.0 km spacing, threshold 7.5							
	1. OLDCF	.21	.48	.79	.45	.19	.68

[1]The TEDC performance results when the five "sliver" triangles are removed.

network does not cover the full 5-km protection disc (cf. Figures VI-2, -3, -4), and we shall see that overall, this performance is somewhat inferior to the performance with scale = 2.25.

1. 19-Station Network

Panel A of Table VIII-1 presents the comparative results for the 19-station network. One quickly notes that the NEWCF (i.e., the OLDCF with a number of improvements) has the poorest overall performance. NLR has the best performance and LMD is close behind, followed by TEDC. Note that EL.POD, with one exception, is very similar to POD, and that EL.FAR compared to FAR accentuates the comparative difference.

2. 13-Station Network

Panel B of Table VIII-1 presents the 13-station network comparative results. Again, NEWCF typically exhibits the worst performance and NLR the best. However, the rather large EL.FAR for NLR is bothersome. Again EL.POD presents a similar performance picture to that of POD, and with the exception of NLR, the EL.FAR presents a similar comparative picture.

3. 11-Station Network

Panel C of Table VIII-1 presents the comparative results for the 11-station network. Now, TEDC typically is the poorest performer (e.g., TSS is .49) and NEWCF is the second poorest. This is a major change in the relative performance for TEDC, and upon looking closely at the 11-station geometry (Figure VI-3) one notes that it has five exterior triangles which are rather small (i.e., slivers); we suspect this creates problems with the divergence calculations.

If these five triangles are removed, the TEDC performance values improve and are given in the parentheses (e.g., TSS = .54). However, TEDC remains the worst performer in all but the EL.FAR category.

Comparing the performance results between the three networks (i.e., 19, 13, and 11), we note that for the four unconditional performance measures, the decrease in performance is virtually linear (in the median) with decreasing size of network. However, this is not true for the two conditional performance measures (i.e., EL.POD and EL.FAR).

4. 7-Station Network

Panel D of Table VIII-1 presents the comparative results for the 7station network. TEDC continues its relatively poor performance in that it is only best in the EL.FAR category. However, NEWCF typically continues to be the poorest performer, and the NLR and LMD now appear to be the better algorithms. However, the NLR has, as we have seen elsewhere, the highest eligible EL.FAR (i.e., .30).

5. 6-Station Network

Panels E and F of Table VIII-1 present the comparative results for the 6-station network. Now TEDC's performance is improved since it leads in performance in three categories (i.e., TSS, FAR, and EL.POD). LMD and NLR appear to be the better performers with the exception of the relatively large EL.FAR (i.e., .32) for NLR. However, compared to the larger sized networks, the performance of the 6-station network is rather poor.

For comparison's sake, the OLDCF values are present in Panel F, and they are greatly inferior to the values from the other four algorithms. There appears to be considerable opportunity for improvement.

B. The 2.25 Scale

The 2.25 km scale is our larger scale in the simulations. In general, the comparative results are similar to those in the 1.50 scale but the differences in performance are somewhat more pronounced. The complete result tables can be found in Appendix E.

1. 19-Station Network

Panel A of Table VIII-2 presents the simulation performance results for the 19-station network under the 2.25-km spacing. Again, as in the 1.5-km scale case, the NEWCF algorithm has the poorest performance. Now, NLR and TEDC are rather similar in their performances (e.g., .80 and .81 for TSS, respectively). Furthermore, their performance is distinctly separated from the NEWCF performance.

2. 13-Station Network

Panel B of Table VIII-2 presents the relevant comparative performance results. Again NEWCF is the worst performer. Now NLR appears to be the best, and LMD and TEDC appear to be rather similar in performance. Note that NLR has a decidedly better performance than LMD (e.g., TSS = .77 vs. .69).

3. 11-Station Network

Panel C of Table VIII-2 presents the comparative simulation results for this network. Now the TEDC algorithm has dramatically changed in its performance and is the worst. However, if the "sliver" triangles again are removed (see 11-station network, scale=1.50, for more details), the results improve slightly (i.e., values in parentheses). NLR generally is the best performer and LMD is the second best.

4. 7-Station Network

Panel D of Table VIII-2 presents the comparative simulation results for this network. Again, the NEWCF algorithm typically is the worst performer and the NLR again is the best. However, compared to the larger size networks (i.e., 19, 13, and 11), this network's typical performance is uniformly inferior.

Table VIII-2. Algorithm performance results for the 19-, 13-, 11-, 7- and 6-station network, scale=2.25-km simulations.

Network	Algorithm	TSS	CSI	POD	FAR	EL.POD	EL.FAR
A. 19 Stations							
	1. NEWCF	.70	.71	.78	.10	.77	.28
	2. LMD	.76	.78	.88	.13	.87	.19
	3. NLR	.80	.81	.94	.14	.89	.22
	4. TEDC	.81	.82	.91	.16	.84	.21
B. 13 Stations							
	1. NEWCF	.65	.69	.80	.17	.67	.27
	2. LMD	.69	.72	.84	.16	.75	.21
	3. NLR	.77	.79	.94	.16	.80	.22
	4. TEDC	.72	.74	.82	.11	.72	.19
C. 11 Stations							
	1. NEWCF	.59	.66	.81	.22	.69	.33
	2. LMD	.65	.69	.82	.18	.74	.26
	3. NLR	.67	.72	.88	.20	.64	.20
	4. TEDC	.55	.60	.71	.20	.60	.22
		(.59)[1]	(.65)	(.78)	(.21)	(.67)	(.21)
D. 7 Stations							
	1. NEWCF	.52	.56	.62	.15	.38	.32
	2. LMD	.54	.61	.75	.23	.56	.26
	3. NLR	.63	.69	.84	.21	.69	.33
	4. TEDC	.58	.65	.81	.24	.63	.28
E. 6 Stations							
	1. NEWCF	.46	.53	.63	.23	.32	.32
	2. LMD	.44	.55	.70	.28	.40	.19
	3. NLR	.58	.62	.72	.17	.60	.35
	4. TEDC	.58	.62	.72	.19	.50	.25
F. Old Centerfield, six stations, 3.0 km spacing, threshold 7.5							
	1. OLDCF	.21	.48	.79	.45	.19	.68

[1]The TEDC performance results when the "sliver" triangles are removed.

5. 6-Station Network

Panels E and F of Table VIII-2 present these comparative simulation results. Again, the NEWCF performance is worst and the NLR and TEDC are the best. OLDCF performance values are presented in Panel F, and we see that its results are still greatly inferior to the other four algorithms.

C. Cross-Validation

To indicate the "strength" of the above simulation results from a sample theory viewpoint, Table VIII-3 presents the cross-validation output for the 13-station network with scaling of 2.25 km. Whether one prefers TSS, POD, EL.POD, etc., it is quickly seen that the relative ranking of algorithms from sample to sample is virtually unchanged. It also is interesting to note how little the values change from sample to sample for most detection methods (e.g., the POD values for NLR are .94, .89, .92, .92, and .88).

Thus, for the 13-station network, the network linear regression (NLR) typically is the best, and the triangle convergence-divergence with edge effects (TEDC) appears to be second best. The local median and local mean are in the middle (i.e., LMN and LMD), and the new centerfield method (NEWCF/RA) is one of the worst. (Remember that the performance of OLDCF is never better than the NEWCF and typically is far worse.) Excluding the EL.FAR, the vector difference method is uniformly worst. Thus, these cross-validation results are strongly in support of the results presented in Table VIII-2.

D. The Old Centerfield Case

Table VIII-4 presents some simulation results for the old centerfield (OLDCF) idealized situation. Recall that this is our attempt at simulating the original LLWAS system currently utilized at most of the nation's airports.

For the currently employed threshold of 7.5 m/s, the relative performance of OLDCF is very poor (e.g., TSS = .21, FAR = .45, EL.POD = .09, EL.FAR = .67, etc.). This performance is unchanged as seen in the five cross-validation runs. Even if one is allowed to "fine-tune" the threshold value in order to maximize the TSS (i.e., 10.0 m/s), the results still are very poor. This "tuning" generally appears to trade POD for FAR.

Comparing the original OLDCF cross-validation results with the "fine-tuned" results for a scale of 2.25, one readily can see the potential improvements that are available in a revised LLWAS.

Table VIII-3. Cross-validation of detection methods for the 13-station, scale 2.25 km network using five randomly selected wind sequences.

SAMPLE NUMBER	1	2	3	4	5
A. TSS/CSI					
NEWCFD/RA (NEWCF)	.65/.69	.66/.69	.67/.69	.66/.70	.67/.69
CF/OTHER (RFCF)	.68/.71	.68/.72	.70/.72	.69/.71	.69/.72
MEAN (NMN)	.65/.69	.65/.68	.66/.70	.66/.70	.66/.69
MEDIAN (MMD)	.66/.70	.65/.69	.67/.70	.66./70	.66/.70
L.MEAN (LMN)	.70/.74	.70/.74	.70/.74	.71/.74	.71/.74
L.MEDIAN (LMD)	.69/.72	.70/.73	.70/.73	.70/.73	.70/.73
LN.REG. (NLR)	.77/.79	.76/.78	.75/.78	.76/.78	.77/.78
VECT.DIFF. (LVD)	.42/.48	.43/.50	.51/.56	.51/.56	.50/.56
TRIANGLES (TDC)	.72/.74	.74/.76	.70/.72	.72/.74	.74/.75
TRIANGLE/EDGE (TEDC)	.73/.74	.75/.76	.70/.72	.72/.74	.74/.75
B. POD/FAR					
NEWCF/RA	.80/.17	.80/.16	.74/.08	.80/.16	.77/.13
CF/OTHER	.83/.16	.82/.16	.79/.11	.79/.12	.79/.12
MEAN	.79/.16	.79/.16	.81/.16	.81/.17	.80/.16
MEDIAN	.81/.17	.81/.17	.81/.17	.81/.17	.83/.18
L.MEAN	.88/.18	.87/.17	.88/.18	.89/.18	.88/.18
L.MEDIAN	.84/.16	.83/.14	.88/.19	.89/.19	.88/.19
LN.REG.	.94/.16	.89/.17	.92/.17	.92/.16	.88/.12
VECT.DIFF.	.54/.20	.57/.21	.65/.19	.65/.19	.64/.19
TRIANGLES	.81/.11	.83/.10	.80/.13	.82/.12	.83/.11
TRIANGLE/EDGE	.82/.11	.84/.10	.81/.13	.82/.12	.83/.11
C. EL.POD/EL.FAR					
NEWCF/RA	.60/.31	.59/.31	.54/.27	.60/.34	.56/.30
CF/OTHER	.61/.33	.60/.33	.56/.30	.55/.31	.57/.31
MEAN	.59/.23	.58/.24	.60/.26	.59/.27	.59/.27
MEDIAN	.61/.25	.61/.25	.60/.27	.61/.27	.60/.29
L.MEAN	.68/.30	.68/.30	.69/.31	.69/.32	.69/.32
L.MEDIAN	.63/.24	.65/.24	.69/.31	.70/.31	.70/.32
LN.REG.	.74/.31	.72/.25	.74/.32	.72/.29	.70/.23
VECT.DIFF.	.31/.19	.30/.22	.36/.22	.36/.22	.36/.23
TRIANGLES	.71/.17	.70/.19	.66/.22	.69/.22	.69/.21
TEDC TRIANGLE/EDGE	.71/.18	.71/.19	.67/.23	.69/.22	.70/.21
D. NUMBER OF WIND SHEAR EVENTS					
MB/GF/SP	12/5/26	13/5/27	12/5/27	12/5/26	12/5/27

Table VIII-4. Old centerfield algorithm cross-validation performance results for the stations=6, scale=3.00, threshold=7.5 situation compared to an optimum threshold=10.0 situation.

	Threshold	TSS	CSI	POD	FAR	EL.POD	EL.FAR
(1)	7.5	.21	.48	.79	.45	.19	.68
(2)	7.5	.21	.48	.79	.45	.19	.66
(3)	7.5	.23	.49	.80	.44	.21	.67
(4)	7.5	.23	.49	.80	.44	.20	.67
(5)	7.5	.23	.49	.81	.44	.20	.67
	10.0	.42	.46	.50	.15	.15	.26

E. Fine-Tuning the Threshold

As we have indicated elsewhere, the threshold value of each algorithm has been fine-tuned to give the best performance (i.e., maximize the TSS) for each of the given networks and station spacings. Thus, with the exception of OLDCF (7.5 m/s), we have tried to allow each algorithm to be compared under its best performance conditions. (Remember if TSS = 1.0, then all wind shear events are correctly predicted and hence no false alarms or missed detections occur.)

To see how the performance measures (e.g., POD, FAR, CSI, etc.) respond to changes in the threshold values, Table VIII-5 presents the performance values versus threshold values for the TEDC algorithm under a 13-station network and scale of 2.25. Figure VIII-1 presents a graph of these values.

Table VIII-5. Threshold values versus performance values for the TEDC algorithm on the 13-station network with scale of 2.25 km.

Performance Index	Threshold Values						
	1.0	2.0	3.0	4.0	5.0	6.0	7.0
TSS	.00	.42	.70	.72	.66	.53	.38
POD	1.00	.97	.91	.82	.71	.55	.39
FAR	.52	.38	.20	.11	.07	.03	.02
CSI	.48	.61	.74	.74	.68	.54	.39
EL.POD	.86	.84	.79	.72	.61	.49	.37
EL.FAR	.98	.60	.31	.19	.12	.05	.01

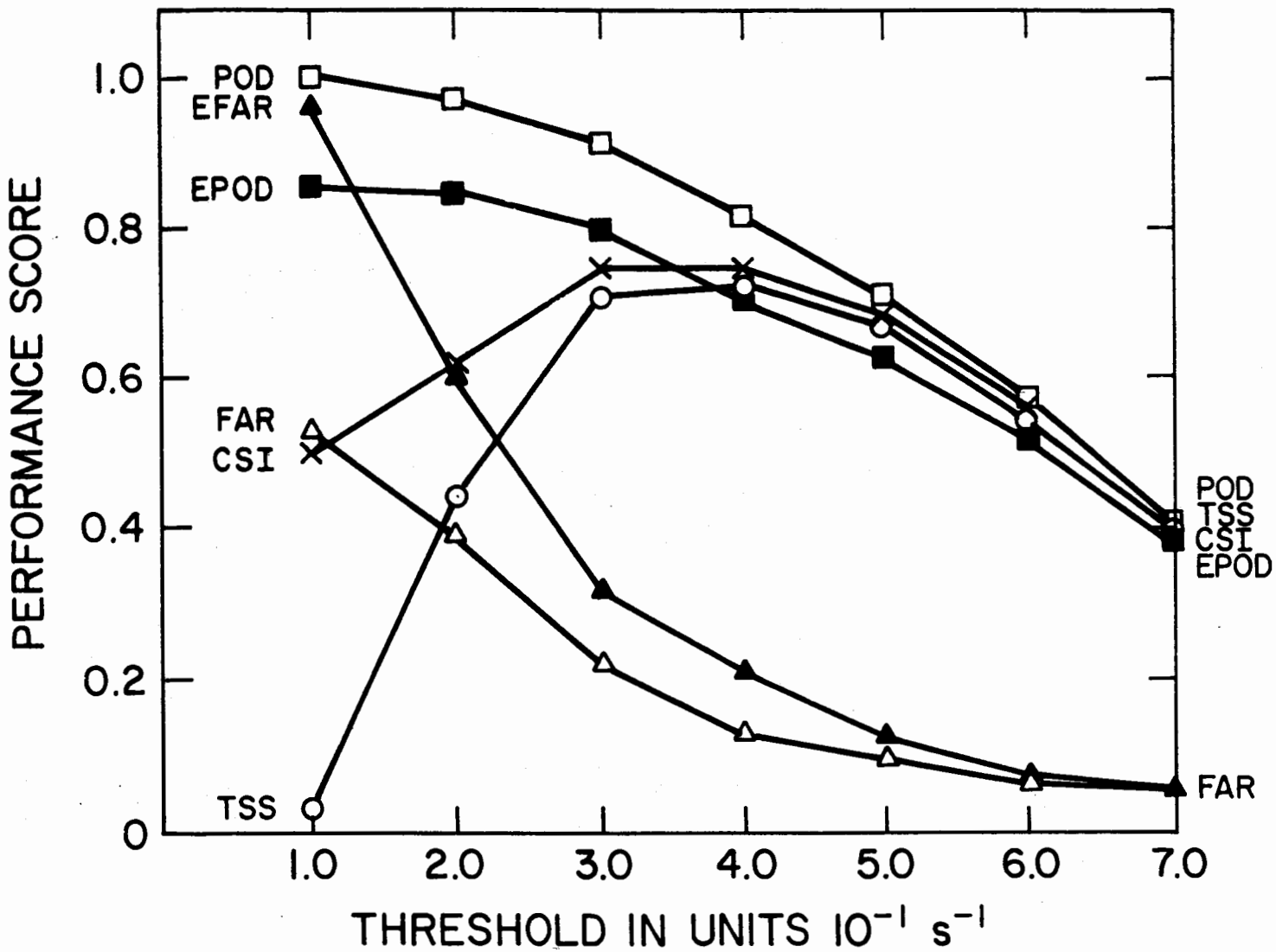


Figure VIII-1. Plots of performance scores versus threshold values for the TEDC algorithm under 13-station network and scale of 2.25 km.

We quickly note that with a very low threshold (e.g., 1.0) POD and EL.POD are high (i.e., virtually every difference initiates an alarm), as are FAR and EL.FAR. In short, we are overpredicting wind shear events. As we increase the threshold, the TSS and CSI scores increase and the POD, EL.POD, FAR, and EL.FAR decrease. At a threshold value of 4.0 we have the maximum TSS and have substantially reduced the FAR and EL.FAR. At the same time we have not reduced the POD and EL.POD that much (e.g., .86 to .71 for EL.POD).

However, if we continue to increase the threshold value, we will lose much more in POD (e.g., .82 to .38) and gain little in FAR (e.g., .11 to .02). Thus, the selection of the maximization threshold not only attempts to maximize TSS and CSI, but it essentially maximizes the difference between POD and FAR. It is this latter difference that is easily tracked in Figure VIII-1 across changing threshold values.

F. Selection of Network Size

Many factors enter into the selection of the size of the ground-based sampling network including costs, access, performance, etc. To aid this process we have constructed some charts based on Table VIII-2. Figures VIII-2 to -6 present the performance scores as a function of sampling network size for the above four algorithms and the OLDCF algorithm. The latter is only given for the 6-station situations, but should be used as a base-line for the other four algorithms.

We can quickly see that the 6-station network is inadequate in all algorithms. The improvements in scores typically are substantial as the network size increases from the initial 6, and thus are best at 19 stations. For NEWCF and NLR the typical improvement from 13 to 19 stations is less than for LMD and TEDC.

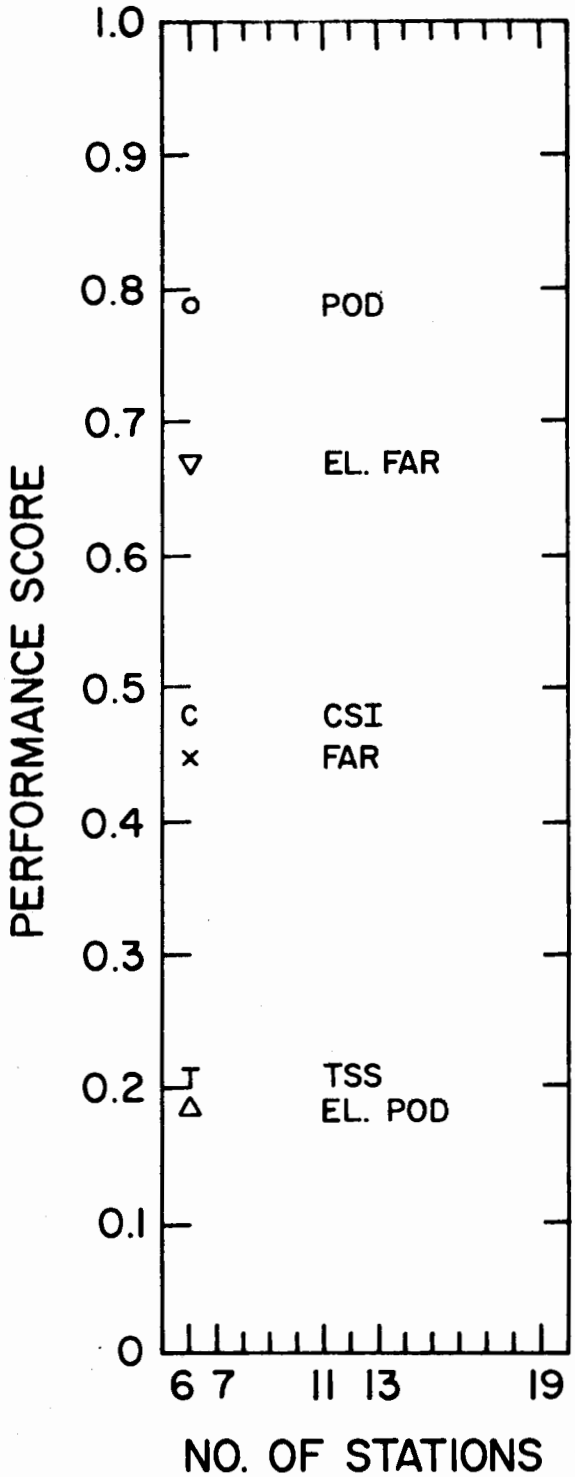


Figure VIII-2. Plots of performance scores versus network size for OLDCF algorithm. Scale = 2.25.

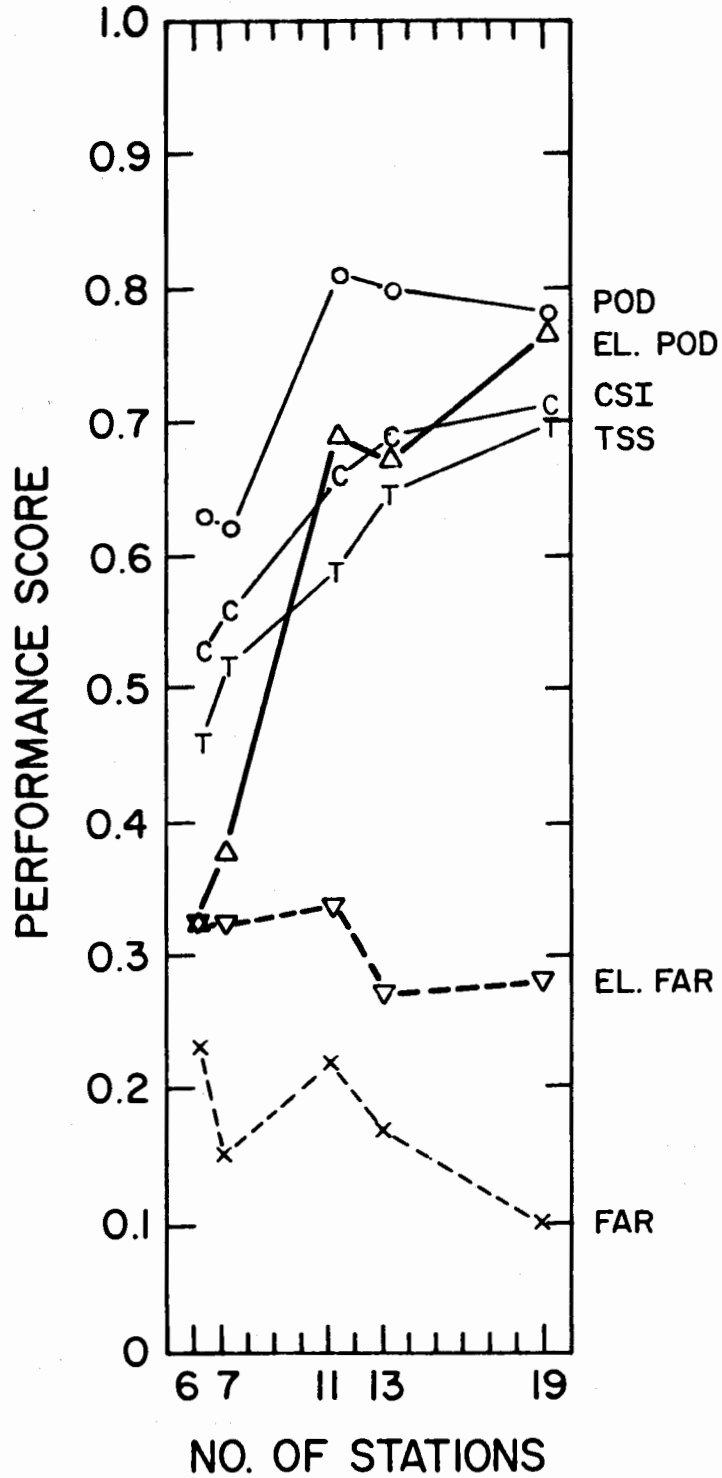


Figure VIII-3. Plots of performance scores versus network size for NEWCF algorithm. Scale = 2.25.

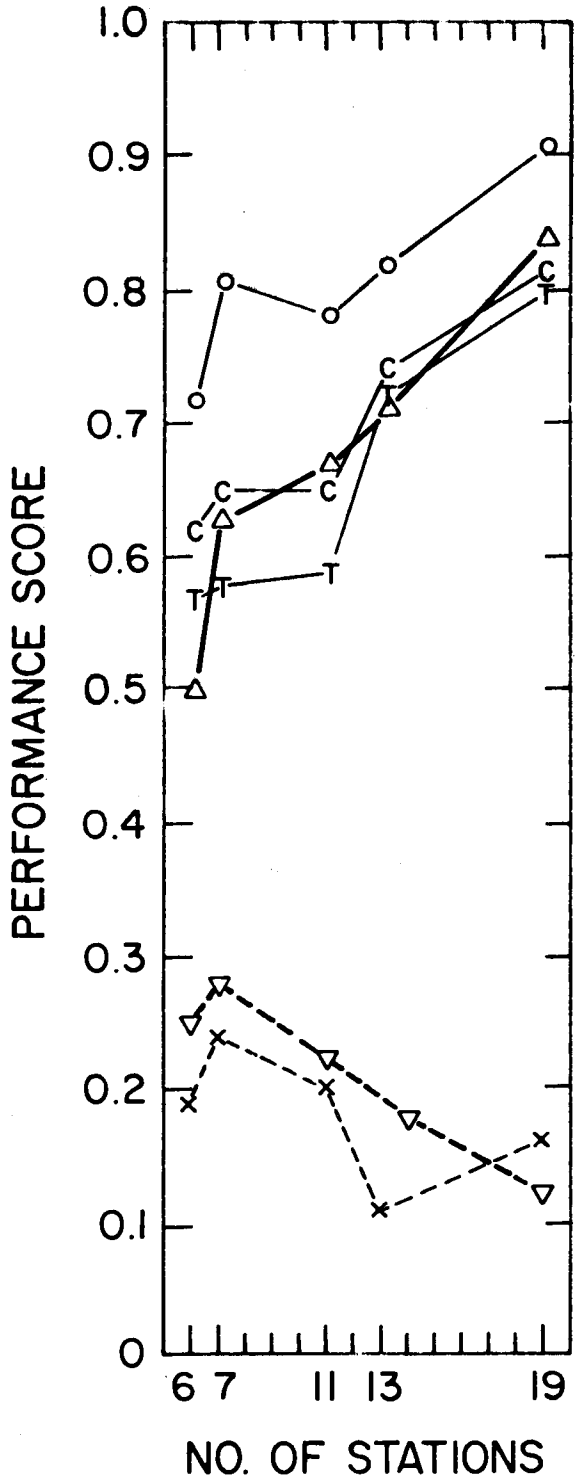


Figure VIII-4. Plots of performance scores versus network size for TEDC algorithm. Scale = 2.25.

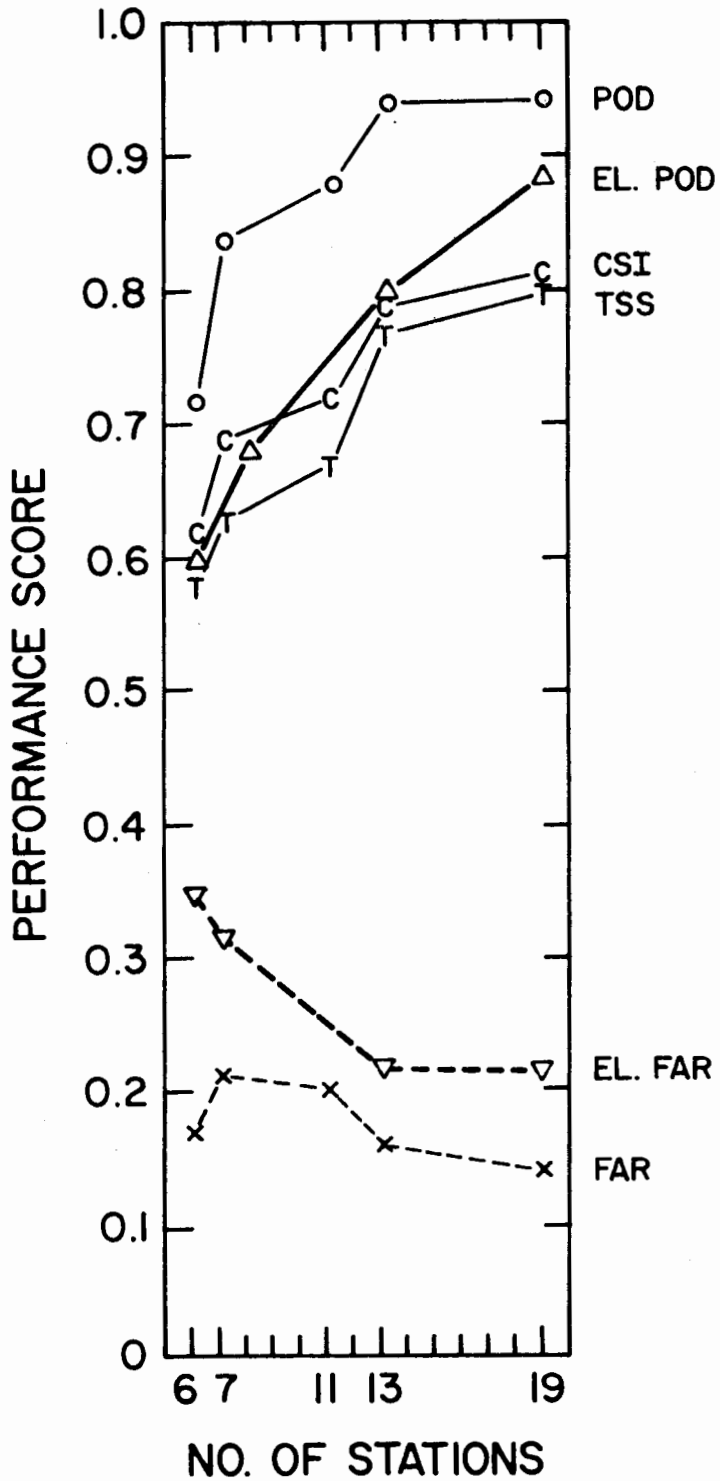


Figure VIII-5. Plots of performance scores versus network size for NLR algorithm. Scale = 2.25.

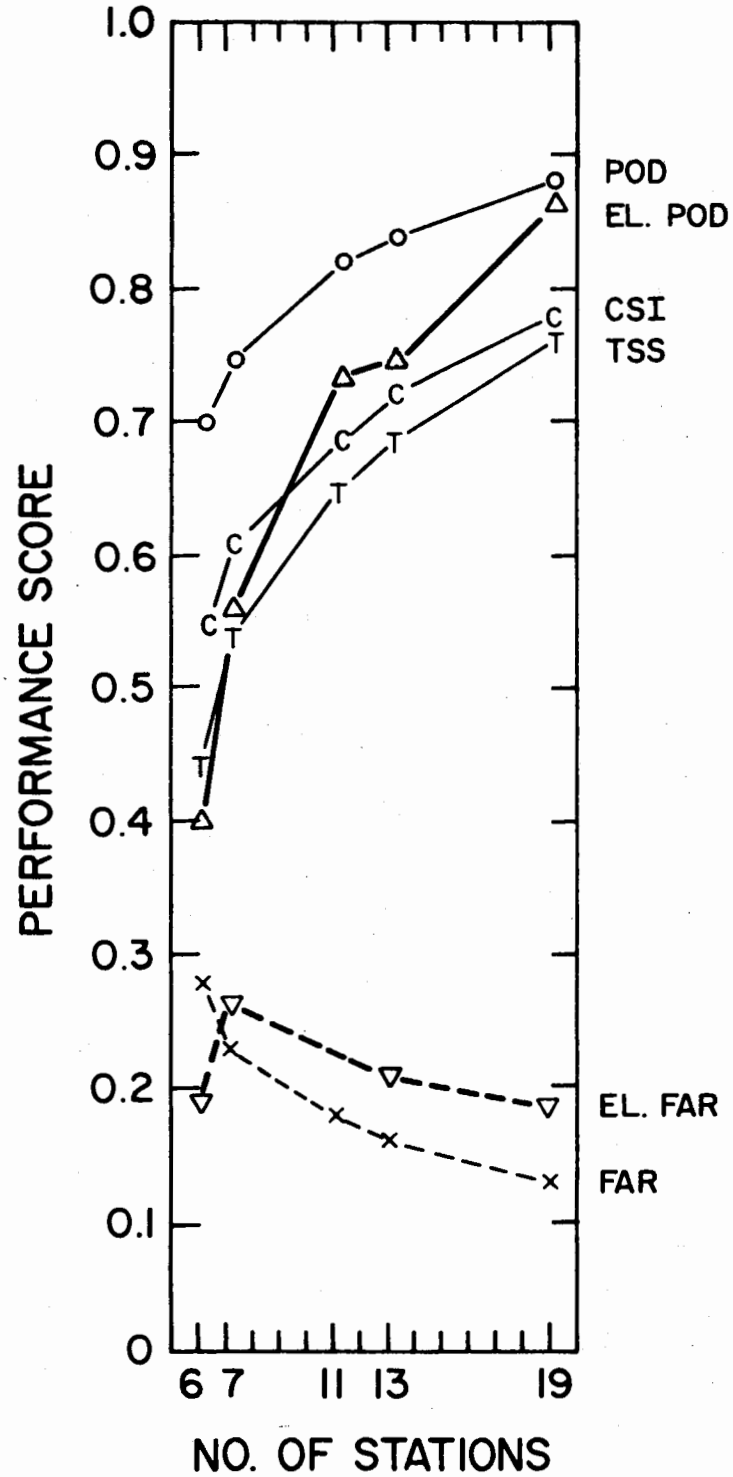


Figure VIII-6. Plots of performance scores versus network size for LMD algorithm. Scale = 2.25.

IX. Recommendations for LLWAS Improvements

Having presented and discussed the results of our simulated wind shear comparative experiments in the previous sections, we believe a number of changes are needed in the present LLWAS. Hence, we offer the following recommendations for improving the present LLWAS.

1. We recommend that the centerfield anomaly algorithms currently used in the LLWAS be replaced.

   Justification:
   a. It is well known in statistical theory and practice (e.g., Snedecor and Cochran, 1967) that pooling of relevant time and spatial data generally gives a better measure of location or background than using a single sampling point. Furthermore, the omission of the current (single-polling) value of the centerfield station as a detector of a wind shear event is ignoring pertinent information.
   b. In comparing the results of Section VIII, we found that the centerfield type algorithms (i.e., OLDCF and NEWCF) typically produced the poorest performance of all of the investigated algorithms. For all numbers of stations that were considered (6, 7, 11, 13, 19), NEWCF has a TSS that is on the order of .10 less than NLR, a POD that is 10 to 15 percent less than NLR with comparable or worse false alarm values (Table VIII-2).
   c. The change of the algorithm is an improvement that can be made with small cost compared with the introduction of additional stations, and our testing shows that, even with an existing 6-station geometry, there is a substantial benefit over the original centerfield algorithms (Table VIII-2 and Table XIIF-1).

2. We recommend that a uniform spacing of 2 to 2.5 km be utilized in the detection network.

   Justification:
   a. The scale of the microburst event is such that a spacing greater than 2.5 km would allow a microburst to impact and "walk through" the network without detection.
   b. The simulation results typically show a substantial loss of performance when the spacing becomes greater than 2.5 km. The change in algorithm performance between 1.5- and 2.5-km spacing is much less severe.

3. We recommend that the replacement wind shear algorithm be selected from one of the following candidates:
   a. First preference is the trimmed network-wide linear regression anomaly (NLR);
   b. Second preference is the triangle and edge divergence-convergence method (TEDC);
   c. Third preference is the local median anomaly method (LMD).

Justification:
For scale = 2.25 and every network, NLR consistently has the best performance (cf. Table VIII-2). TEDC is usually second best and LMD is always second or third. For example, with 13 stations, the TSS values are .77 (NLR), .72 (TEDC), and .69 (LMD).

4. We recommend that the wind shear alarm continue to be based on the "at least one station" rule, rather than on requiring simultaneous detection at multiple stations.

   Justification:
   a. Although the FAR measure typically was improved when one used "at least two stations," (Appendix E), the POD generally was substantially decreased, particularly for the more plausible networks (e.g., 13 stations) and the likely spacing (e.g., 2.25).
   b. If one is restricted to a relatively small number of stations in the network (e.g., 7), then the detection performance of the "at least two stations" method is greatly inferior.

5. We recommend that the wind shear threshold for each selected airport network be "fine tuned" by simulation testing to maximize the TSS.

   Justification:
   The simulation results indicate a strong dependence on the threshold value. Choosing a threshold either too high or low greatly degrades the performance (e.g., the relation between prediction and observation).

6. We recommend that a station density and uniformity similar to that of the 19-station geometry with scale = 2.25 be used to sample wind shear events, and the stations be uniformly spread throughout the the area of protection.

   Justification:
   a. The simulation results for our parameter values indicate that the 19-station network has the best performance. As an example, TSS becomes maximum at about .80 for TEDC and NLR. The corresponding POD's are .91 and .94 and the EL.FAR's are at a minimum (e.g., .21 for TEDC and .23 for NLR).
   b. The 13-station network has a somewhat weaker performance, but still gives very good results. For example, looking at TSS, NLR now has .77 versus .80 and TEDC has .72 versus .81.
   c. In every case, better results are obtained if the number of stations is increased (i.e., density is increased). The evaluations that we have produced (Figures VIII-2 to -6) could be used as a basis for cost effectiveness and aviation community needs studies, which are beyond the scope of this report. Without such detailed analysis, it is clear from these figures that increasing the number of stations from 6 to 7, from 7 to

11, and from 11 to 13 gives substantial improvement for each increase, and this is an indication that using 13 or more stations will be justified by subsequent analysis.
d. When the protection region is modified in shape or size, the actual number of stations that are needed will vary. Irregular geometries will also affect this choice. An example of these effects is given in Appendix F.

7. We recommend that the wind shear detection system attempt to distinguish between gust front and microburst events; in this situation our preference is in favor of the TEDC method.

Justification:
a. We believe that gust fronts and microbursts provide different types of aviation hazards. With training and education regarding the meteorological events and their effects on the performance of their aircraft, pilots will be able to use the knowledge of the nature of the wind shear event as they make flight decisions.
b. At this stage in our research, only the TDC and TEDC algorithms are able to provide a useful distinction between the two types of wind events, and TEDC provides the best performance.

8. We recommend that the FAA use the results of this study and the accompanying computerized simulation-testing program, to guide them both in the modifications of the existing LLWAS systems and in the design of new systems.

Justification:
a. The computerized simulation-testing program model provides a useful formal structure in which one can give an a priori evaluation of the theoretical performance of alternative station networks and algorithms.
b. The computerized wind shear model can be expanded to capture vertical motions and thus provide further insight into wind shear events.

9. We recommend that a new computing system be installed at all LLWAS operations in order to provide sufficient computing capacity for retrieving and processing data from the recommended larger network of stations (e.g., 13) and the proposed new algorithms (e.g., NLR or TEDC).

Justification:
In order to implement the above recommendations on network size and algorithms, and based on our knowledge of the LLWAS computer at Stapleton, it is our opinion that new computing capability is required.

X. Recommendations for Further Studies

Having completed this initial study on the theory and methodology of low-altitude wind shear detection, we find that a number of further studies need to be considered. We believe that this study only has "skimmed-the-cream" in this interesting detection and classification problem. Hence, we briefly list some of the problems that we believe merit future attention.

1. Are some algorithms better on gust fronts and others on microbursts? How sensitive are the comparative results to the type of wind shear?

2. We wonder how sensitive the comparative results are to the frequencies of gust front and microburst events.

3. How sensitive are the comparative results to different speeds of the wind shear events?

4. How sensitive are the comparative results to irregular geometries?

5. Determine what general modifications in the algorithms are needed to accommodate "missing data."

6. Investigate the benefits to be derived from using conditional station thresholds.

7. Investigate the possibility of obtaining information about the location of a wind shear event from the algorithm output.

8. Modify the present wind shear detection model so that one can evaluate the performance of protecting more general geographic shapes and grids.

9. Design a procedure for evaluating a detection system (i.e., algorithm and geometry) for "Timeliness of First Detection" of a microburst.

10. Investigate the possibility of using different thresholds in the triangle-divergence algorithms to compensate for the effects of "sliver" triangles when using an irregular network geometry.

11. Explore alternative visual displays for presenting the type and location of wind shear events in the ATC tower.

12. What effect does the sheltering of some stations have on the performance of the various algorithms?

13. Ascertain the value to ATC and pilots of the ability to distinguish between the types of wind shear events.

It is our belief that the pursuit of these recommendations will have an important impact on future aviation safety.

XI. References

Bedard, A.J., Jr., J. McCarthy, and T. Lefebvre, 1984: Statistics from the operation of the Low-Level Wind Shear Alert System (LLWAS) during the Joint Airport Weather Studies (JAWS) Project: An interim report from the JAWS Project at NCAR. JAWS NCAR Report No. 01-83. Final Report submitted to FAA, in press. Available from Research Applications Program (RAP), NCAR, P.O. Box 3000, Boulder, Colo. 80307.

Byers, H.R., and R.R. Braham, Jr., 1949: The Thunderstorm. U. S. Government Printing Office, Washington, D.C.

Courant, R., and D. Hilbert, 1962: Methods of Mathematical Physics. Interscience, Vol. 2, New York.

Fujita, T.T., 1976: Spearhead echo and downburst near the approach end of a John F. Kennedy Airport runway, New York City. SMRP Res. Paper 137, University of Chicago, 51 pp.

Fujita, T.T., and H.R. Byers, 1977: Spearhead echo and downbursts in the crash of airliner. Monthly Weather Review, 105, 129-146.

Fujita, T.T., and F. Caracena, 1977: Analysis of three weather-related aircraft accidents. Bulletin of American Meteorological Society, 58, 1164-1181.

Goff, R.C., 1980: The Low-Level Wind Shear Alert System (LLWSAS). FAA System Research and Development Service, FAA-RD-80-45, Washington, D.C.

Hanning, R., 1977: Digital Filters, Prentice Hall, Englewood Cliffs, N.J.

Mosteller, F., and J.W. Tukey, 1977: Data Analysis and Regression. J. Wiley and Sons, New York, N.Y.

National Research Council, 1983: Low-altitude wind shear and its hazard to aviation. National Academy Press, Washington, D.C., 109 pp.

Ripley, B.D., 1981: Spatial Statistics. Wiley and Sons, New York, N.Y., 252 pp.

Snedecor, W., and W.C. Cochran, 1967: Statistical Methods, Sixth Edition, Iowa State Press, Ames, Iowa.

Wilson, J., R. Roberts, C. Kessinger, and J. McCarthy, 1983: Microburst wind structure and evaluation of Doppler radar for airport wind shear detection. Journal of Climate and Applied Meteorology, 23, 898-915.

XII. Appendices

Appendix A. The Estimate of the Wind Field Divergence

If (u,v) describes a wind field in Cartesian coordinates, then the divergence of this wind field is defined as

$$\text{div } (x,y) = u_x\ (x,y) + v_y\ (x,y) \qquad (1)$$

where the subscript designates partial differentiation with respect to the subscripted variable. The Divergence Theorem states that the total change of mass in a region (integral of div (x,y)) is equal to the net mass flux at the boundary of the region (cf. Courant and Hilbert, 1962). In particular, if there is a microburst whose center is interior to a triangle, then there will be a net outflow at the boundary and hence the 2-dimensional wind field (u,v) will appear to be losing mass in that triangle. We can detect the presence of a microburst by using the wind field that is measured at the vertices to estimate whether there is a net loss of mass. This can be done either directly by linear interpolation of the wind field along each side of the triangle and direct estimation of the net outflow along that side, or indirectly by estimating div (x,y) at the barycenter of the triangle and using this as an estimate for average divergence on the triangle. These two different approaches to estimating the net mass flux at the boundary give identical results and estimating div (x,y) is significantly simpler numerically.

To compute div (x,y), it is only necessary to estimate the two numerical derivatives. For this we use the following method for estimating the derivatives of a function f that is defined on a triangle T and whose values are known at the vertices (x(i),y(i)), i = 0,1,2 of the triangle.

Lemma. Let z = f(x,y) be a continuously differential function that is defined on the triangle T and suppose that the values z(i) = f(x(i),y(i)) are known. Define the vectors

$$V(i) = (\ x(i),Y(i),z(i)\ ) - (\ x(0),y(0),z(0)\ ) \qquad (2)$$

for i = 1,2. Then there is a constant A such that the following vector cross product formula holds

$$V(1) \times V(2) = A * (f_x,\ f_y,\ -1\ ) \qquad (3)$$

Proof. The function F(x,y,z) = f(x,y) - z has the property that the graph z = f(x,y) is described by the relation F(x,y,z) = 0. The points [x(i),y(i)], i = 0,1,2 lie on this graph and the vector cross product V(1) x V(2) is normal to the triangle that they determine and therefore approximates a normal vector to the graph. On the other hand, the normal to the graph is given precisely by the gradient of F, which is

$$\text{grad } F = (\ f_x,\ f_y,\ -1) \qquad (4)$$

Therefore, these vectors are equal up to a scalar multiple A, cf. equations (3) and (4).

Appendix B. The Mathematical Descriptions of the Expected Values for the Anomaly Detection Algorithms.

Each of the anomaly detection methods is based on the difference of the observed wind field value at a station from the wind field value that is expected at that station. The different methods are based on different schemes for using the mesonet data to predict the wind field value that is expected. To eliminate the variation that occurs in data that are measured at very short time intervals, time-averaged data are always used to compute the expected wind field value. The time averaging is described in Section V.

Five methods use the mesonet data to predict a single vector value, which is used as the expected value at every station in the mesonet. Three methods use the centerfield value (there are three centerfield methods), the network mean value, and the network median value. Three other methods build a model for the wind field on the network and use this model to predict the expected value for each station. These expected values will usually vary from station to station. These methods use the local mean value, the local median value, and a bivariate linear model (regression) for the wind field. This appendix contains mathematical descriptions of the methods that are used to compute these expected values.

Two centerfield methods (OLDCF and NEWCF) use the two-minute running average value from the measurements at a centrally located station as the value that is expected at every station in the network. OLDCF is the method that was used in the original LLWAS design and is in operation at most LLWAS installations. The third centerfield method (RFCF) uses the value that is computed from the measurements at this central station by use of the recursive filter averaging that is described in Section V.

The network mean (NMN) wind field is computed by taking the network average of the time-averaged wind fields from all of the stations. The network averaging of a wind field could be interpreted to mean that the wind field Cartesian coordinates (u,v) are averaged separately or that the wind field polar coordinates (speed and direction) are averaged separately. Because of the ambiguity caused by the folding of the direction angle, it is not possible to unambiguously define the average direction over the network, unless the variation over the network is small, and so we compute the network mean by averaging the Cartesian coordinates.

When dealing with data that have occasional spikes, it has been found that the sample mean can be a somewhat misleading estimate of the expected value. Statisticians have developed two techniques for dealing with this problem. One is to use the median instead of the sample mean as described in the next paragraph. The other technique is based on trimming the data. For scalar data, the idea is simply to drop an equal number of highest values and of lowest values from the data set before computing the average value. This will prevent a spike in the data from creating a rapid movement of the estimated mean value. For vector data, it is not immediately clear what is meant by a high or a low value. Since we are

concerned with wind shears, we decided that the spikes will usually be indicated by spikes in the wind speed, and so our trimming strategy is to trim the station values that correspond to the highest and lowest wind speeds. When there are few stations (i.e., 6 or 7), we discovered that trimming had little effect, and so no trimming is employed. For 11-19 stations the method has the best performance when two stations are trimmed (one high and one low). This is the trimming strategy that is used for NMN throughout this report.

The network median (NMD) wind field is computed by taking the median of the time-averaged wind fields at each of the stations. For scalar data, the median value is the middle value of the data values when they are arranged in increasing order (or the average of the two middle values when there is an even number of data points). For vector data, there is again the question of whether to use the Cartesian coordinates or the polar coordinates of the wind field, and again, because of the directional ambiguity, only the median of the Cartesian coordinates is unambiguous. There is also a problem of ranking the vector data so that the middle value can be determined. We have done this by computing the medians of the scalars u and v separately. An alternative would be to rank the stations according to wind speed (as in trimming) and to use the (u,v) value of the middle station. We have found that this latter approach has slightly poorer skill than the method we have chosen.

The local mean (LMN) and the local median (LMD) are computed for a given station by considering only the stations that are proximate to that station. For this, a radius of influence is prescribed, and the proximate stations are defined to be those whose distance from the given station is less than the radius of influence. In particular, each station is proximate to itself. In practice, the radius of influence is chosen so that most stations have 6 or 7 proximate stations. Once the proximate stations have been determined, the mean or median of the time-averaged wind field values for this set of stations is computed and used as the expected value at the given station. Notice that this procedure provides a modeling of the wind field that is in the spirit of objective analysis.

The network linear regression (NLR) is another method that provides a modeling of the wind field over the network. In this case, the model is chosen to have the form

$$u(x,y) = a + bx + cy$$

$$v(x,y) = d + ex + fy$$

and the coeficients a, b, c, d, e, and f are determined by the solution of the linear least squares problem. Once the model has been determined, the expected value at each station is computed by evaluating the model equations at the coordinates (x,y) of the station. This method is a theoretical extension of the idea of the network mean, and has a similar sensitivity to spikes in the data. Therefore, a trimming strategy seems to be advisable. Experience shows that the same trimming strategy that is used for NMN also gives the best performance for the NLR.

Appendix C. The Mathematical Microburst Model

This is a simplified mathematical model of a microburst that satisfies the mass continuity condition. This model is chosen for the test model for the LLWAS study because it is believed that this kind of model will exhibit realistic radial outflow at ground level. Comparisons with the JAWS data are favorable.

Our simplification is to assume that the microburst is radially symmetric about its center. Mathematically, this means that the mass continuity condition can be described by ordinary differential equations instead of by partial differential equations. Because of this, it is possible to determine a closed form for the equations that describe the wind velocities in this microburst model.

Figure XIIC-1. illustrates the geometry chosen in a radial slice from the center of the microburst. The four different regions are determined by the nature of the forcing (boundary condition) imposed at the top of the model. This forcing is indicated by arrows in the figure and is decribed explicitly later in the text. Regions I and II are the divergence region and regions III and IV are the convergence region. The wind velocity is 0 in the static region.

The model is assumed to be radially symmetric about the orgin. The wind field velocity has a radially horizontal component, V(R,Z), and a vertical component, W(R,Z). Mass continuity is required in each thin cylindrical shell of height H and radius R about the orgin.

At the top of the cylinder (the upper boundary), it is assumed that the wind field direction is negatively vertical in region I, becomes parallel by a sine-cosine transformation across region II, continues the sine-cosine tranformation to positively vertical across region III, and is vertical in region IV. The wind speed at the top of the cylinder is constant in regions I and II and tends linearly to zero over regions III and IV. We specify the vertical velocity by the following formulas:

VERTICAL WIND SPEED

$$W(R, Z) = \begin{cases} (Z/H) \quad w \quad S(R) & 0 < R < R1 \\ (Z/H) \quad w \quad S(R) \quad (R-R2) / (R1-R2) & R1 < R < R2 \end{cases}$$

where w is a constant representing the maximum vertical wind speed at the center of the microburst. We require that the horizontal velocity has the form:

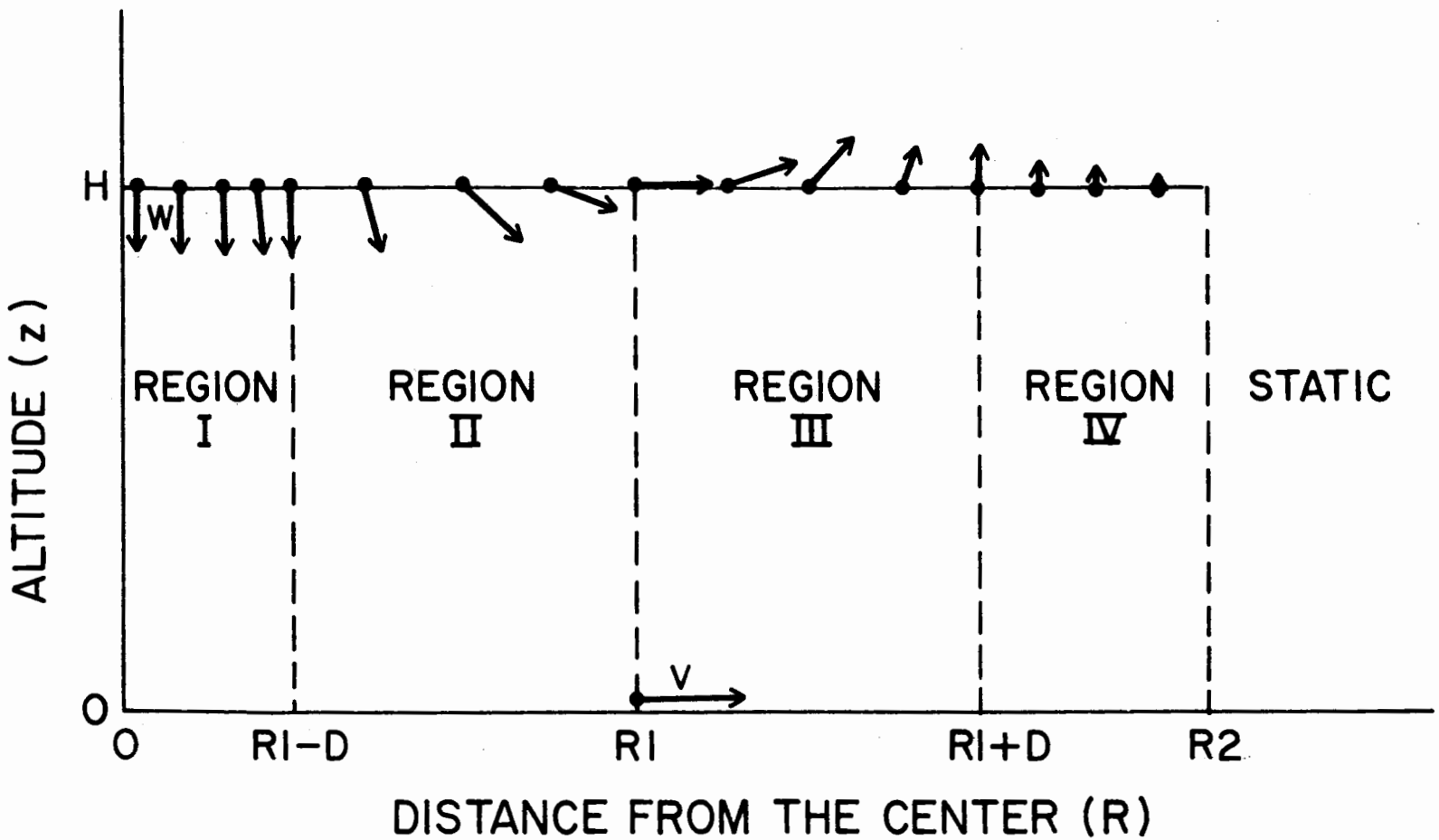


Figure XIIC-1. Partition of the microburst region.

HORIZONTAL WIND SPEED

$$V(R,Z) = \begin{cases} (Z/H) \quad w \quad C(R) + v\,A(R) \quad (-Z/H)EXP(1+E) & 0 < R < R1 \\ (Z/H) \quad w \quad C(R) \quad (R-R2)\,/\,(R1-R2) + v \quad A(R) \quad (1-Z/H)EXP(1+E) & R1 < R < R2 \end{cases}$$

where

$$S(R) = \begin{cases} -1 & 0 < R < R1 - D \\ SIN\,[\,(R-R1)\,\pi\,/\,2D\,] & R1-D < R < R1+D \\ +1 & R1+D < R < R2 \end{cases}$$

$$C(R) = \begin{cases} 0 & 0 < R < R1 - D \\ COS\,[\,(R-R1)\,\pi\,/\,2D\,] & R1-D < R < R1+D \\ 0 & R1+D < R < R2 \end{cases}$$

INPUT PARAMETERS:

R1	Radius of the divergence region
D	Radius of the transition region
H	Height of the model region
w	wind speed at Z = H, R = 0
v	wind speed at Z = 0, R = R1

OUTPUTS:

A(R)	Scalar function describing the relative strength of the outflow at the surface
R2	Radius of the outer boundary
E	Exponent of the scaling factor for the horizontal wind speed as a function of Z

The continuity requirement forces a horizontal distribution of the horizontal wind velocity that is described by the scaler function A(R) [i.e., V(R,0) = vA/R], and a vertical distribution of the horizontal wind velocity that is described by the value G, which is the measure of the net outward flux of the horizontal wind field that is described by the exponent E.

THEORETICAL DERIVATION

It is assumed that the model satisfies continuity in each cylindrical shell of height H, inner radius R, and thickness dR. Computing the integrals for the total mass flux on each face of the cylindrical shell, requiring that the total mass flux be zero on the surface of the cylindrical shell, and taking the limit as dR tends to zero, we obtain differential equations for A(R) (separately defined in each of the four regions):

<u>REGION I</u> (0 < R < R1 - D).

$$\frac{d}{dR}\,[\,R \quad A(R)\,] = \frac{1}{G} \quad \frac{w}{v} \quad R$$

there G is the integral of (1-Z/H)EXP(1.+E) from 0 to H. This differential equation is integrable and we compute

$$R \quad A(R) = \frac{1}{G} \quad \frac{w}{v} \quad \frac{R^2}{2}$$

which satisfies the required initial condition A(0) = 0. Evaluating this function at (R1-D), we obtain the initial value for solving the differential equation in region II.

<u>REGION II</u> (R1 - D < R < R1).

$$\frac{d}{dR} [R \quad A(R)] =$$

$$- \frac{1}{G} \quad \frac{w}{v} \quad [(H/2) \; C(R) + (1 - \frac{H}{4} \frac{\pi}{D}) \; R \quad S(R)]$$

This differential equation is integrable and we have the initial condition for A(R1-D) from region I; we compute

$$R \quad A(R) = \frac{1}{G} \quad \frac{w}{v} \quad [\frac{(R1-D)^2}{2} + (\frac{2\,D}{\pi} - \frac{H}{2}) \quad R \quad C(R)$$

$$- (\frac{2\,D}{\pi})^2 \quad (S(R) + 1)]$$

<u>REGION III</u> (R1 < R < R1 + D).

$$\frac{d}{dR} [R \quad A(R)] =$$

$$- \frac{1}{G} \quad \frac{w}{v} \quad \{ R \quad S(R) \quad (R-R2) / (R1-R2)$$

$$+ (H/2) \quad \frac{d}{dR} [R \quad C(R) \quad (R-R2) / (R1-R2)] \}$$

This differential equation is integrable and we have the initial condition A(R1) from region II [from the definition of A(R), we know that A(R1) = 1]; we compute

$$R \; A(R) = R1 -$$

$$\frac{1}{G} \; \frac{w}{v} \; \frac{1}{R1-R2} \quad \{ (\frac{H}{2} - \frac{2\,D}{\pi}) \; [R1 \; (R1-R2) - R \; (R-R2) \; C(R)]$$

$$+ (\frac{2\,D}{\pi})^2 \quad (2\,R-R2) \quad S(R) - 2 \quad (\frac{2\,D}{\pi})^3 \quad (1-C(R)) \}$$

which satisfies the required initial condition A(R1) = 1 at R1.

The parameter R2 is defined to be the value of R at which A(R2) = 0. It is not necessary that R2 > R1 + D as is indicated in Figure XIIC-1; however, this is true in all examples, in which we have computed R2. In our formulation, R2 is computed by applying a root finder to A(R). It is necessary to have an a priori estimate of the magnitude of R2, so that we can determine which formula for A(R) (region III or region IV) is to be used in the minimization.

Let B(R,X) be the function that uses the formula for A(R) (region III) when X is substituted for R2. Then A(R) = B(R,R2). Let T(R) = B(R,R). The following Lemma shows that T(R1+D) can be used to test the relative sizes of R2 and R1+D.

Lemma. If R < R2, then T(R) > A(R); if R > R2, then T(R) < A(R). In particular, if T(R1+D) > 0, then R2 > R1+D, and if T(R1+D) < 0, then R2 < R1+D.

Proof. For fixed R, B(R,X) is an decreasing function of X, since increasing X increases the strength of the vertical outflow at the top of the cylinder, and consequently decreases the required strength of the outflow at the side with radius R+dR. Therefore, T(R) > A(R) for R < R2 and T(R) < A(R) for R > R2. In particular,

$$T(R1+D) < A(R1+D) < 0 \qquad \text{for } R2 < R1+D$$

and

$$T(R1+D) > A(R1+D) > 0 \qquad \text{for } R2 > R1+D.$$

This implies the statement of the lemma.

REGION IV (R1 + D < R < R2) .

$$\frac{d}{dR}\left[R \quad A(R) \right] = - \frac{1}{G} \quad \frac{w}{v} \quad \left[R \quad \frac{R - R2}{R1-R2} \right]$$

Since this differential equation is integrable, we compute

$$R \quad A(R) = (R1+D) \quad A(R1+D) - \frac{1}{G} \quad \frac{w}{v} \quad \frac{R^2 \ (2\ R-3\ R2)}{6\ (R1-R2)}$$

$$+ \frac{1}{R\ G} \quad \frac{w}{v} \quad \frac{(R1+D)^2 \ [2\ (R1+D)-3\ R2]}{6\ (R1-R2)}$$

which satisfies the required initial condition at R1+D.

A Simplification.

1. There is a recurring factor $F = 2\ D\ /\ \pi$.
2. In the formula for A(R) in Region II, the value of the bracket at R = R1 is

$$B = \frac{(R1-D)^2}{2} + \left(F - \frac{H}{2} \right) \quad R1 - F \quad (H-F)$$

This value is used to determine the value of G by the relationship

$$1 = \frac{1}{G\ R1} \quad \frac{w}{v} \quad B$$

and hence,

$$G = \frac{1}{R1} \quad \frac{w}{v} \quad B$$

But the formula for A(R) uses the factor $\frac{1}{G} \quad \frac{w}{v}$ repeatedly, and we see this factor is equal to R1/B. The importance of this observation is that it both simplifies the computation of A(R) and shows that A(R) is independent of the wind speed ratio. Since A(R) does not depend on the ratio w/v , neither does the the outer radius R2, and so we can determine a reasonable model for the surface winds for use in the LLWAS analysis without needing to understand the intricacies of the vertical wind intensity and its relation to the maximum observed horizontal intensity. Also, we have simplified the formulas for A(R), which provides for more efficient computation.

<u>REGION I</u> ($0 < R < R1 - D$).

$$A(R) = \frac{R1}{B} \quad \frac{R}{2}$$

<u>REGION II</u> ($R1 - D < R < R1$).

$$A(R) = \frac{R1}{R\ B} \quad \left[\frac{(R1-D)^2}{2} + \left(F - \frac{H}{2} \right) \quad R \quad C(R) \right.$$

$$\left. + (H-F) \quad F \quad (S(R) + 1.) \right]$$

<u>REGION III</u> ($R1 < R < R1 + D$).

$$A(R) = \frac{R1}{R} - \frac{R1}{R\ B} \quad \left\{ \left(\frac{H}{2} - F \right) \left[\left(\frac{R-R2}{R1-R2} \right) R\ C(R) - R1 \right] \right.$$

$$\left. + F^2 \quad \left[\left(\frac{2\ R-R2}{R1-R2} \right) S(R) + 2 \left(\frac{1-C(R)}{R1-R2} \right) \quad F \right] \right\}$$

<u>REGION IV</u> ($R1 + D < R < R2$) .

$$A(R) = \frac{R1+D}{R} \quad A(R1+D) - \frac{R1}{R\ B} \quad \frac{R^2\ (2\ R-3\ R2)}{6\ (R1-R2)}$$

$$+ \frac{R1}{R\ B} \quad \frac{(R1+D)^2\ [2\ (R1+D)-3\ R2]}{6\ (R1-R2)}$$

Appendix D: Discussion of Some Forecast Verification Methods

1. Background

Given the computerized mathematical wind shear model and the eleven alternative algorithms, one needs an evaluation system to properly compare the performance of each detection algorithm. In the terminology of forecasting research, we need a verification plan.

We will adopt an elementary approach to scoring the detection of wind shear events by an algorithm. We will assume that the relevant time period of observation is known (e.g., 6 hours) and can be partitioned into equal intervals of length t* (e.g., 12 seconds). We will further assume that at most one wind shear event (e.g., microburst) can occur within any given interval. Thus, the observed "data sequence" will look as follows:

.....	0	0	1	0	0	0	1	1	1	0	.....
	t*										t

More succinctly, the sequence has the form

(. . ., 0, 0, 1, 0, 0, 0, 1, 1, 1, 0, ...),

where: 0 = no wind shear event occurred (WS') in the given time interval,

1 = a wind shear event occurred (WS) in the given time interval.

Lastly, we will assume that the detection algorithm's output similarly can be quantified to indicate the presumed presence (1) or absence (0) of a wind shear event. Hence, its data sequence also can be summarized as,

(. . ., 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, ...),

where: 0 = no predicted wind shear event (PWS') in the given time interval,
1 = a predicted wind shear event (PWS) in the given time interval.

Because the two data sequences both are dichotomized on the same wind shear event set, we can construct the traditional 2 x 2 contingency table commonly used in statistics. Table XII.D.1 gives the general form of this table with the cross-tabulation of all of the $n_{..}$ observations by the two factors or variables.

Table XII.D.1 The Traditional 2 x 2 Contingency Table for Cross-Classified Events (n_{ij} is the total count of the occurrence of the particular joint wind shear event).

Observed Windshear Event

		WS	WS'	
Predicted Windshear Event	PWS	n_{11}	n_{12}	$n_{1.}$
	PWS	n_{21}	n_{22}	$n_{2.}$
	Totals	$n_{.1}$	$n_{.2}$	$n_{..}$

Now, we desire a summary measure which will indicate how well our detection (i.e., forecast) algorithm performs. We would like the summary measure to have the following desirable properties:

a. Be easy to understand in the particular context,
b. Have a convenient fixed range of values,
c. Use all of the available information,
d. Be able to provide a measure of the degree of association between the two factors,
e. Be able to provide a measure of the variability of the degree of association, and
f. Generally be able to distinguish signal from noise (e.g., assess the probability that association exists).

We will examine some standard forecasting verification measures to see how well they fulfill these desired properties.

2. Some Standard Forecasting Verification Measures

The following three well-known forecast verification measures are applicable to our situation.

a. The Probability of Detection (POD)

$$POD = n_{11}/(n_{11} + n_{21}) = n_{11}/n_{.1} \qquad (1)$$

This is a conditional column measure of "success" with range [0,1]. However, it only conditions on the first column of the Contingency Table and thus omits information (e.g., information present in the second column). Correspondingly, one could use the second column and define the probability of false detection (POFD) as $n_{12}/n_{.2}$.

b. The False Alarm Ratio (FAR)

$$FAR = n12/(n11 + n12) = n12/n1. \qquad (2)$$

This is a conditional row measure of "failure" with range [0,1]. One can provide similar information in terms of 1 - FAR (i.e., the correct alarm ratio, CAR). However, FAR and CAR only condition on the first row, and hence they also omit information.

c. The Critical Success Index (CSI)

$$CSI = n_{11}/(n_{11} + n_{12} + n_{21}) = n_{11}/(n - n_{22}) \tag{3}$$

This measure also is a conditional measure of success (i.e., conditioning on three of the four possible cells or joint events) with range [0,1]. Hence, it also omits information, and furthermore it does not have the same direct interpretation as the two previous ones.

The CSI appears to have been first defined and utilized by Gilbert (1884). Furthermore, he also presented the other two standard measures (i.e., POD and FAR) as well.

All three of these standard verification measures or indices fail to satisfy at least the last four desired properties, and the CSI does not appear to satisfy any of the six properties. Hence we need to investigate other measures of verification in hopes of satisfying more of the desired properties.

3. Some Alternative Statistical Measures

In an attempt at satisfying the above six desired properties for a verification measure, we will examine some association measures that have a statistical basis.

a. The Observed-Expected Statistic (OE)

One of the oldest classical measures of evidence of association in the 2 x 2 contingency table is the observed-expected statistic (OE) and its' resulting Chi-Square test (sometimes termed the "Goodness-of-Fit" test). The test focuses on the question of whether the two population factors used to cross-classify the sample observations are statistically independent. The test is based on the general idea that each cell count (n_{ij}) can be partitioned into two terms: (1) a count due to an underlying model (M_{ij}), and (2) a count due to residual effects (e_{ij}). Hence,

$$n_{ij} = M_{ij} + e_{ij}. \tag{4}$$

As an example, Table XII.D.2 presents a hypothetical 2 x 2 contingency table for a sample of 100 wind shear observations cross-classified by observed wind shear and predicted wind shear.

Table XII.D.2 Wind Shear Observations Cross-classified by Observed and Predicted Wind Shear

		Observed Windshear		
		OM	OM'	
Predicted Windshear	PM	10	10	20
	PM'	20	60	80
		30	70	100

In this presumed randomly selected sample of size 100 from some underlying population there are 30 observed microbursts (M) and 70 non-microbursts (M'). Correspondingly, there are 20 predicted microbrust (PM) and 80 predicted non-microburst (PM') forecasts. The general problem is to partition the four n_{ij}'s into model and residual effects on the basis of some acceptable theory.

Given the assumption of statistical independence between the two cross-classification factors (i.e., observed and predicted microbursts), the partitioning model is relatively easy. We simply estimate the total expected counts (E_{ij}) in each of the four cells on the basis of the model of statistical independence between the the two factors, and these E_{ij}'s are equated to the corresponding M_{ij}'s. The calculations proceed as follows:

i. Under the statistical independence model,
P_r (PM and M jointly occurring) = Pr(PM) · Pr(M)

ii. We will estimate the unconditional or "marginal" probabilities (e.g., Pr (M)) by using frequency data from the margins of the 2 x 2 table. As an example,

$$\Pr(\text{PM and M}) = \Pr(\text{PM}) \cdot \Pr(\text{M}) = (n_{1.}/n_{..})(n_{.1}/n_{..})$$

iii. Finally, the expected count in a cell is simply the sample size ($n_{..}$) times the joint probability,

$$E_{11} = (n_{1.}/n_{..})(n_{.1}/n_{..})(n_{..}) = (n_{1.})(n_{.1})/n_{..} \qquad (5)$$

The resulting expected values for our example of Table XII.D.2 are:

$E_{11} = 6$, $E_{12} = 14$, $E_{21} = 24$, and $E_{22} = 56$.

One can then calculate the difference between the n_{ij} and E_{ij} (i.e., the residual effect, e_{ij}) for each cell and construct the OE statistic,

$$OE = \sum_{i=1}^{2} \sum_{j=1}^{2} (n_{ij} - E_{ij})^2/E_{ij} \quad . \qquad (6)$$

Under the two assumptions of random selection of the $n_{..}$ observations and variable marginal totals (e.g., $n_{i.}$'s), the OE statistic has an approximate Chi-Square(χ^2) distribution with one degree of freedom. Thus, we can calculate the statistical significance of the OE statistic and thereby assess the probability that the two classification factors of Table XII.D.2 are unrelated (i.e., statistically independent). One also can perform an "exact test" of the counts in Table XII.D.2 based on the hypergeometric distribution. For further details the reader should see Brownlee (1965) or Afifi and Azen (1979).

It should be noted that for the 2 x 2 table, Eq. (6) can be re-written as

$$OE = \frac{(n_{11} \times n_{22} - n_{21} \times n_{12})^2}{(n_{.1})(n_{.2})(n_{1.})(n_{2.})/n_{..}} \quad . \tag{7}$$

Furthermore, it has been recognized that the approximation of the OE statistic to the χ^2 distribution is improved if the so-called "Yates Continuity Correction Factor" is applied to the numerator of Eq. (7). The corrected OE statistic is:

$$OE_c = \frac{(|n_{11} \times n_{22} - n_{21} \times n_{12}| - n_{..}/2)^2}{(n_{.1})(n_{.2})(n_{1.})(n_{2.})/n_{..}} \quad . \tag{8}$$

It also is suggested that the E_{ij} should be about 5 or larger to have a satisfactory approximation to the χ^2 distribution.

The OE statistic in this current form does not satisfy many of the desired properties of Section 1. As an example, it does not provide any useful information on the degree or magnitude of association between the two factors. However this can be rectified if one transforms the uncorrected OE statistic to the directly related Pearson's Phi (ϕ) statistic,

$$\phi = (OE/n_{..})^{1/2} \quad . \tag{9}$$

The ϕ statistic is interpreted as a measure of association between the two factors and a value close to "0" suggests little association while one near "1" suggests strong positive or negative association. Furthermore, the significance test of the null hypothesis, $E(\phi) = 0$, is the same as the OE test for statistical independence.

Returning to our windshear example of Table XII.D.2 we can calculate the ϕ value for this situation. Using Eq. (6) we get:

$$OE = \frac{(10 - 6)^2}{6} + \frac{(10 - 14)^2}{14} + \frac{(20 - 24)^2}{24} + \frac{(60 - 56)^2}{56}$$

$$= \frac{4^2}{6} + \frac{(-4)^2}{14} + \frac{(-4)^2}{24} + \frac{4^2}{56}$$

$$= (4 \cdot 16)\left(\frac{1}{6} + \frac{1}{14} + \frac{1}{24} + \frac{1}{56}\right)$$

$$= 64(.2978) = 19.05$$

Then using Eq. (9), the association between the two factors is

$$\phi = (OE/n..)^{1/2} = (19.05/100)^{1/2} = (.1905)^{1/2} = \underline{.44} \quad .$$

Thus the evidence in the table suggests a .44 association, and the Chi-Square test of whether this value is "significantly" different from zero is affirmed at the .001 level of significance (i.e., $P_r(\chi^2 \geq 10.8) = .001$ and hence a smaller than .1% chance of getting a value as large or larger than .44 when in fact the association is zero in the underlying population).

However, the ϕ statistic does not satisfy a number of the previously listed desired properties (e.g., easy to understand).

b. Other Measures

There are a number of other statistical measures that are used to assess association in contingency tables, but none of them appear to fully satisfy the previous list of desired properties. They include Yule's Q, Yule's Y, a number of Goodman and Kruskal λ statistics, Somer's D, etc. For further details on these and other measures the reader is invited to see Afifi and Azen (1979).

4. The True Skill Score (TSS)

The general goal here is to allow all four joint event cells in the 2 x 2 table to contribute to our verification measure (as did the ϕ) and achieve all of the previously listed six desired properties. Furthermore, the measure should remove the chance results and only focus on the remaining skill component.

We initially proceed as in the OE statistic and estimate from the margins of the 2 x 2 table the expected count in each cell due to the model of statistical independence (i.e., chance). Then, using Eq. (4), we can form the matrix of residual effects, or skill counts (SC), for our Table XII.D.2 example as

$$SC = \begin{bmatrix} 4 & -4 \\ -4 & 4 \end{bmatrix} ,$$

where: $sc_{11} = e_{11} = 10 - 6 = 4$, etc.

Note that the matrix SC is a symmetric matrix.

Next we can sum the skill counts for the correct decisions, those on the major diagonal (i.e., calculate the trace of SC), and get +8. This is a measure of true forecast skill, but needless to say one that is not

easy to understand. Thus we need a baseline or standard which can be compared with our observed (+8) result.

A simple approach to constructing this standard is to calculate the maximum possible trace given the observed marginals. The maximum possible trace occurs only for the perfect forecast situation, and thus this new measure will be termed the True Skill Score (TSS) and defined as:

$$TSS = \frac{\text{Observed Skill Count Trace}}{\text{Perfect Skill Count Trace}} \quad . \tag{10}$$

In our example, the perfect skill count situation occurs when the 2 x 2 table of observed and predicted counts is as follows:

Table XII.D.3 The 2 x 2 contingency table for the perfect forecast of windshear events for the example of Table XII.D.2. The expected counts due to the independence model are given in parentheses.

		Observed		
		M	M'	
Predicted	M	30(9)	0(21)	30
	M'	0(21)	70(49)	70
		30	70	100

The perfect skill (PS) count matrix for this perfect forecasting situation is:

$$SC_{PS} = \begin{bmatrix} 21 & -21 \\ -21 & 21 \end{bmatrix}$$

where: $ps_{11} = 30 - 9 = 21$, etc. Again the matrix is symmetrical.

Hence the perfect skill trace is 42, and thus

$$TSS = \frac{8}{42} = .19 \quad .$$

In general, using the notation of Table XII.D.1, TSS can be rewritten as

$$TSS = \frac{(n_{11} - E_{11}) + (n_{22} - E_{22})}{(n_{11} - E_{11})_{PS} + (n_{22} - E_{22})_{PS}}$$

$$= \frac{(n_{11} + n_{22}) - [(n_{1\cdot} \times n_{\cdot 1}/n_{\cdot\cdot}) + (n_{2\cdot} \times n_{\cdot 2}/n_{\cdot\cdot})]}{(n_{1\cdot} + n_{2\cdot}) - [(n_{1\cdot}^2/n_{\cdot\cdot}) + (n_{2\cdot}^2/n_{\cdot\cdot})]} \tag{11}$$

$$= \frac{(n_{11} + n_{12}) - (n_{1\cdot} \times n_{\cdot 1} + n_{2\cdot} \times n_{\cdot 2})/n_{\cdot\cdot}}{(n_{\cdot 1} + n_{\cdot 2}) - (n_{\cdot 1}^2 + n_{\cdot 2}^2)/n_{\cdot\cdot}} \quad .$$

Upon further algebra, this becomes[1],

$$TSS = \frac{(n_{11})(n_{22}) - (n_{12})(n_{21})}{(n_{\cdot 1})(n_{\cdot 2})} \quad , \tag{12}$$

where the range is ($-1 \leq TSS \leq +1$), and TSS = 0 under the statistical independence model (i.e., zero correlation between the observed and predicted factors). This form of TSS well illustrates that the cell counts from all four cells enter the calculations (i.e., the cross-product of "successes," n_{11} n_{22}, minus the cross-product of "failures," n_{12} n_{21}). It also is interesting to note that the denominator is the cross-product of the two observed event column totals.

Furthermore, it should be noted that if shear events are coded "1" and non-shear events coded "0", then TSS also can be shown to be the slope of the least-squares fitted linear line segment between the observed and predicted shear events. Thus, perfect prediction skill gives a slope of +1, no prediction skill a slope of 0, and perfect incorrect prediction skill a slope of -1.

Lastly, Eq. (12) can be further simplified to give:

$$TSS = (n_{11}/n_{\cdot 1}) - (n_{12}/n_{\cdot 2}) \quad . \tag{13}$$

This form is simply the conditional probability of detection (POD) minus the conditional probability of false detection (POFD). In short, the POD is adjusted by the POFD to produce the true skill score (TSS).

Given that TSS is the slope coefficient (b) between the quantified observed and predicted events, the variance of TSS can be estimated as follows:

---

[1]A review of the forecasting verification literature indicates that Eq. (12), or its counter-part, has appeared before. It was first proposed by Pierce (1884) in verification of Finley's tornado data, later by Clayton (1938), and still later by Hanssen and Kuipers (1965). With this repeated exposure it is surprising that the measure has not become more widely used.

$$S_b^2 = \frac{\sigma_e^2}{\Sigma(x_i - \bar{x})^2} = \frac{SS_{yy} - bSS_{xy}}{n - 2} \Big/ (\Sigma x^2 - (\Sigma x)^2/n)$$

$$= \frac{1}{n - 2} \frac{(n_{11} + n_{21})(n_{12} + n_{22})}{(n_{11} + n_{12})(n_{21} + n_{22})} - b^2 \quad . \tag{14}$$

Thus, under the assumption of random selection of wind shear events we can compare two values of TSS for statistical significance. However, in practice some of the assumptions that underly this inference procedure (e.g., normality, constant variance, etc.) may not be met, and hence we prefer to "cross-validate" (e.g., Mosteller and Tukey, 1977) the TSS scores when possible.

In summary, TSS does address each of the six desired properties presented in Section 1. It is easy to understand, particularly in the form of Eq. (13). It has a convenient fixed range of possible values ($-1 \leq TSS \leq +1$) with zero as the no skill value. Furthermore, it uses the skill data from all four cells and thus is an unconditional measure of skill, and TSS is a measure of association in the given context (i.e., if desired, one can transform TSS into the Pearson Product Moment Correlation Coefficient). Finally, the variance of TSS is known, Eq. (14), and thus under particular assumptions one can test for "statistically" significant differences between alternative values of TSS.

REFERENCES

Afifi, A.A. and S.P. Azen, 1979: Statistical Analysis, a Computer Oriented Approach. 2nd editions, Academic Press, New York, NY, 442 pp.

Brownlee, K.A., 1965: Statistical Theory and Methodology in Science and Engineering. 2nd edition, J. Wiley & Sons, New York, NY, 590 pp.

Clayton, H.H., 1938: On verifying weather forecasts. Bull. Am. Meteor. Soc., 19, 341- .

Gilbert, G.K., 1884: Finley's tornado-predictions. Am. Meteor. Jour., 1, 166.

Hanssen, A.W. and W.J.A. Knipers, 1965: On the relationship between the frequency of rain and various meteorological parameters. Meded. Verhand., Royal Netherlands Meteoro. Institute, 81, 2-15.

Mosteller, F. and J.W. Tukey, 1977: Data Analysis and Regression. Addison-Wesley, Reading, MA, 588 pp.

Pierce, C.S., 1884: The numerical measure of the success of predictions. Science, 4, 453-454.

Appendix E. Comparative Verification Results for all Algorithms

This section contains a tabulation of the comparative testing verification results for all eleven algorithms on regularly spaced networks of 6, 7, 11, 13, and 19 stations. The measures used are the TSS, POD, FAR, CSI, EL.POD, and EL.FAR. The detection threshold that provides optimum performance in each case also is displayed.

Three different time series averaging methods were tested for each of the anomaly algorithms, and it was found that this variation had a minor impact on the results. These methods are:

1. Recursive filter of the wind field speed and direction with a weighting factor of .6 (comparable to a 2-minute running average) and with wind field values set equal to zero when the wind speed was smaller than 2 m/s.

2. Recursive filter of the wind field (u,v) components with a weighting factor of .6 and with wind field values set equal to zero when the windspeed was smaller than 2 m/s.

3. Two-minute running average (existing LLWAS strategy).

The "smooth type" (1, 2, or 3) indicates which type of time series averaging produced the best results.

The threshold was selected to produce the highest true skill score (TSS). For all methods, low thresholds produce a high POD and a high FAR and both of these values decline as the threshold is increased. For good methods, the FAR declines more quickly than the POD does, and so there is a threshold interval in which there is a relatively high POD and low FAR. It is in this region that the TSS has its maximum value.

The detection methods are described in Section III and are identified in the following tables as:

Anomaly detection methods:

OLDCF	Old Centerfield (Running Average)
NEWCF	New Centerfield (Running Average)
RFCF	New Centerfield (Recursive Filter)
NMN	Network Mean
NMD	Network Median
LMN	Local Mean
LMD	Local Median
NLR	Network Linear Regression
LVD	Local Vector Difference

Wind shear identification methods:

TDC	Triangle Divergence/Convergence
TEDC	Triangle Edge Divergence/Convergence

For the larger networks (11, 13, or 19 stations), station trimming was used for the computation of the network mean, network median, and network linear regression.

For the divergence/convergence methods, it is necessary to select both a convergence threshold and a divergence threshold. It generally appears that the best strategy is for them to have equal magnitudes (they have opposite signs).

Table XII E-1. Simulation testing for 19-station network, scale = 1.50.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	7.00	0.66	0.82	0.17	0.70	0.78	.29
RFCF	1	7.50	0.68	0.80	0.14	0.70	0.75	.26
NMN	2	7.00	0.68	0.80	0.14	0.71	0.75	.18
NMD	3	7.50	0.68	0.77	0.11	0.70	0.73	.16
LMN	2	5.00	0.63	0.79	0.18	0.67	0.71	.18
LMD	3	5.50	0.68	0.81	0.15	0.71	0.75	.15
NLR	3	5.00	0.68	0.84	0.17	0.71	0.78	.22
LVD	3	14.00	0.46	0.66	0.25	0.54	0.52	.19

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	7.00	0.59	0.63	0.06	0.60	0.57	.14
RFCF	2	7.00	0.59	0.63	0.06	0.61	0.58	.16
NMN	3	7.00	0.54	0.56	0.03	0.55	0.50	.04
NMD	3	7.00	0.58	0.60	0.03	0.59	0.54	.05
LMN	3	5.00	0.52	0.54	0.05	0.53	0.46	.03
LMD	3	5.00	0.58	0.60	0.04	0.58	0.52	.03
NLR	3	3.50	0.71	0.83	0.14	0.73	0.69	.20
LVD	2	8.00	0.60	0.65	0.08	0.61	0.57	.17

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	5.00	0.66	0.80	0.15	0.70	0.72	.22
TEDC	5.00	0.68	0.82	0.16	0.71	0.74	.22

Table XII E-2. Simulation testing for 19-station network, scale = 2.25.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	9.00	0.70	0.81	0.13	0.72	0.77	.28
RFCF	2	9.00	0.70	0.82	0.14	0.72	0.77	.29
NMN	3	8.00	0.67	0.82	0.17	0.70	0.77	.24
NMD	1	8.00	0.68	0.86	0.19	0.72	0.80	.27
LMN	3	6.00	0.74	0.89	0.15	0.77	0.84	.19
LMD	3	6.00	0.75	0.90	0.16	0.77	0.87	.19
NLR	3	5.50	0.77	0.93	0.16	0.79	0.89	.22
LVD	3	15.00	0.42	0.67	0.29	0.53	0.52	.24

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	7.00	0.61	0.70	0.13	0.64	0.54	.24
RFCF	1	7.00	0.63	0.73	0.13	0.66	0.54	.25
NMN	3	7.00	0.56	0.61	0.08	0.58	0.46	.11
NMD	1	7.00	0.57	0.63	0.09	0.59	0.47	.12
LMN	3	5.00	0.67	0.71	0.06	0.68	0.58	.05
LMD	3	5.00	0.68	0.72	0.06	0.69	0.60	.05
NLR	2	4.00	0.75	0.84	0.11	0.77	0.72	.11
LVD	2	8.00	0.69	0.79	0.12	0.71	0.68	.28

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	4.00	0.78	0.89	0.12	0.79	0.81	.21
TEDC	4.00	0.80	0.91	0.14	0.81	0.84	.22

Table XII E-3. Simulation testing for 13-station network, scale = 1.50.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	6.50	0.64	0.81	0.19	0.68	0.76	.29
RFCF	2	7.00	0.64	0.80	0.18	0.68	0.74	.28
NMN	3	6.00	0.66	0.84	0.19	0.70	0.77	.25
NMD	3	6.00	0.66	0.85	0.20	0.70	0.77	.26
LMN	3	5.00	0.65	0.81	0.17	0.69	0.73	.20
LMD	3	5.00	0.66	0.82	0.18	0.70	0.75	.20
NLR	3	4.50	0.67	0.86	0.20	0.71	0.78	.29
LVD	3	9.50	0.05	0.96	0.51	0.48	0.69	.75

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	5.00	0.55	0.65	0.14	0.59	0.56	.18
RFCF	2	5.00	0.58	0.70	0.16	0.62	0.56	.22
NMN	3	5.00	0.58	0.63	0.07	0.60	0.54	.08
NMD	1	5.00	0.59	0.65	0.08	0.61	0.56	.08
LMN	3	3.50	0.60	0.70	0.13	0.63	0.57	.13
LMD	3	4.00	0.56	0.62	0.09	0.58	0.51	.07
NLR	3	3.50	0.61	0.70	0.12	0.64	0.58	.13
LVD	2	6.00	0.62	0.75	0.16	0.66	0.60	.32

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	5.00	0.64	0.76	0.14	0.67	0.69	.18
TEDC	5.00	0.65	0.77	0.14	0.68	0.70	.18

Table XII E-4. Simulation testing for 13-station network, scale = 2.25.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	8.50	0.64	0.77	0.15	0.67	0.67	.27
RFCF	2	8.00	0.65	0.84	0.20	0.70	0.72	.35
NMN	3	7.50	0.65	0.82	0.19	0.69	0.73	.26
NMD	3	7.50	0.66	0.84	0.19	0.70	0.74	.27
LMN	3	5.50	0.68	0.88	0.20	0.72	0.80	.28
LMD	3	6.00	0.67	0.83	0.17	0.71	0.75	.21
NLR	1	5.00	0.74	0.89	0.16	0.76	0.80	.22
LVD	3	15.00	0.37	0.61	0.30	0.49	0.40	.24

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	5.50	0.52	0.66	0.19	0.58	0.50	.25
RFCF	2	5.50	0.55	0.73	0.21	0.61	0.51	.30
NMN	3	5.00	0.55	0.69	0.18	0.60	0.50	.20
NMD	1	5.50	0.54	0.63	0.14	0.57	0.47	.14
LMN	3	4.50	0.57	0.63	0.09	0.59	0.48	.09
LMD	3	5.00	0.48	0.52	0.08	0.49	0.42	.05
NLR	3	3.50	0.65	0.76	0.14	0.68	0.58	.17
LVD	2	8.00	0.52	0.59	0.12	0.55	0.46	.22

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	4.00	0.71	0.81	0.11	0.73	0.71	.19
TEDC	4.00	0.72	0.82	0.11	0.74	0.72	.19

Table XII E-5. Simulation testing for 11-station network, scale = 1.50.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	6.00	0.58	0.75	0.20	0.63	0.68	.28
RFCF	1	6.00	0.58	0.76	0.21	0.63	0.68	.31
NMN	2	6.00	0.59	0.72	0.17	0.63	0.65	.18
NMD	1	6.00	0.60	0.75	0.18	0.64	0.67	.21
LMN	3	4.50	0.57	0.80	0.23	0.64	0.68	.29
LMD	3	5.00	0.61	0.77	0.19	0.65	0.68	.22
NLR	3	4.50	0.59	0.75	0.19	0.64	0.65	.24
LVD	3	14.00	0.44	0.58	0.21	0.50	0.41	.17

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	5.00	0.50	0.56	0.10	0.53	0.48	.13
RFCF	2	5.00	0.50	0.59	0.14	0.54	0.48	.17
NMN	3	4.50	0.52	0.57	0.08	0.54	0.49	.08
NMD	3	5.00	0.50	0.54	0.07	0.52	0.47	.05
LMN	3	4.00	0.51	0.56	0.09	0.53	0.46	.07
LMD	3	4.00	0.54	0.59	0.08	0.56	0.48	.07
NLR	1	3.00	0.52	0.62	0.14	0.56	0.46	.14
LVD	3	7.00	0.49	0.52	0.07	0.50	0.44	.14

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	7.00	0.49	0.64	0.21	0.55	0.56	.18
TEDC	7.00	0.49	0.64	0.21	0.55	0.56	.18
TDC/NO SLIVERS	4.00	0.52	0.70	0.20	0.58	0.59	.20
TEDC/NO SLIVERS	4.00	0.54	0.73	0.22	0.60	0.63	.21

Table XII E-6. Simulation testing for 11-station network, scale = 2.25.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	7.00	0.60	0.78	0.20	0.65	0.69	.33
RFCF	1	7.00	0.61	0.80	0.21	0.66	0.70	.35
NMN	3	6.50	0.59	0.80	0.22	0.65	0.69	.28
NMN	2	7.00	0.58	0.77	0.22	0.64	0.68	.26
LMN	1	5.50	0.62	0.81	0.21	0.67	0.70	.25
LMD	3	5.50	0.63	0.83	0.21	0.68	0.74	.26
NLR	1	5.00	0.62	0.75	0.17	0.66	0.64	.20
LVD	3	14.00	0.39	0.62	0.29	0.50	0.43	.25

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	5.00	0.51	0.63	0.18	0.56	0.46	.22
RFCF	2	5.00	0.53	0.67	0.18	0.58	0.46	.26
NMN	1	5.00	0.53	0.59	0.10	0.55	0.45	.11
NMD	3	6.00	0.48	0.51	0.07	0.49	0.40	.06
LMN	3	5.00	0.47	0.50	0.07	0.48	0.40	.05
LMD	3	5.00	0.49	0.53	0.07	0.51	0.42	.04
NLR	3	4.00	0.54	0.59	0.09	0.56	0.44	.09
LVD	1	7.00	0.61	0.69	0.12	0.63	0.50	.30

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	5.00	0.54	0.70	0.20	0.60	0.59	.22
TEDC	5.00	0.55	0.71	0.20	0.60	0.60	.22
TDC/NO SLIVERS	3.00	0.55	0.73	0.21	0.61	0.61	.21
TEDC/NO SLIVERS	3.00	0.59	0.78	0.21	0.65	0.67	.21

Table XII E-7. Simulation testing for 7-station network, scale = 1.50.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	7.00	0.61	0.73	0.15	0.65	0.64	.24
RFCF	2	7.00	0.63	0.78	0.17	0.67	0.66	.27
NMN	3	6.00	0.61	0.76	0.18	0.65	0.67	.21
NMD	3	6.50	0.62	0.75	0.15	0.66	0.68	.17
LMN	2	4.50	0.65	0.82	0.19	0.69	0.74	.26
LMD	1	5.00	0.64	0.80	0.18	0.68	0.72	.18
NLR	2	3.50	0.63	0.77	0.17	0.67	0.70	.30
LVD	1	8.00	0.64	0.79	0.16	0.68	0.73	.43

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	5.00	0.45	0.55	0.16	0.50	0.26	.18
RFCF	1	5.00	0.49	0.60	0.17	0.54	0.25	.20
NMN	2	5.00	0.44	0.47	0.06	0.46	0.31	.06
NMD	2	5.00	0.43	0.49	0.11	0.46	0.29	.06
LMN	2	3.00	0.62	0.77	0.18	0.66	0.46	.21
LMD	1	3.50	0.51	0.57	0.11	0.54	0.36	.07
NLR	3	3.00	0.59	0.72	0.16	0.63	0.42	.18
LVD	3	7.00	0.68	0.76	0.11	0.70	0.48	.29

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	2.50	0.59	0.72	0.17	0.63	0.65	.15
TEDC	2.50	0.62	0.77	0.17	0.66	0.70	.17

Table XII E-8. Simulation testing for 7-station network, scale = 2.25.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	8.00	0.54	0.71	0.21	0.60	0.38	.32
RFCF	1	8.00	0.54	0.73	0.22	0.61	0.40	.34
NMN	2	7.00	0.45	0.66	0.26	0.54	0.46	.30
NMD	3	7.50	0.45	0.64	0.24	0.53	0.46	.26
LMN	1	5.00	0.54	0.75	0.24	0.61	0.59	.33
LMD	1	5.50	0.54	0.73	0.22	0.61	0.56	.26
NLR	2	3.50	0.59	0.78	0.21	0.65	0.69	.33
LVD	3	9.50	0.61	0.74	0.16	0.65	0.61	.38

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	6.00	0.40	0.51	0.19	0.45	0.05	.22
RFCF	1	6.00	0.41	0.55	0.22	0.48	0.05	.26
NMN	2	5.50	0.32	0.40	0.18	0.37	0.07	.14
NMD	2	5.50	0.33	0.46	0.24	0.40	0.06	.18
LMN	1	3.50	0.50	0.65	0.20	0.56	0.13	.25
LMD	3	4.00	0.34	0.44	0.20	0.39	0.10	.14
NLR	3	3.00	0.47	0.62	0.21	0.53	0.14	.22
LVD	3	9.00	0.60	0.66	0.08	0.62	0.13	.26

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	1.50	0.54	0.72	0.21	0.60	0.54	.23
TEDC	1.50	0.58	0.81	0.24	0.65	0.63	.28

Table XII E-9. Simulation testing for 6-station network, scale = 1.50.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
OLDCF	3	7.50	0.47	0.56	0.15	0.51	0.40	.21
NEWCF	3	6.00	0.52	0.71	0.22	0.59	0.52	.30
RFCF	1	6.00	0.55	0.74	0.22	0.61	0.52	.32
NMN	3	6.00	0.48	0.61	0.18	0.54	0.47	.18
NMD	1	6.00	0.50	0.65	0.20	0.56	0.51	.19
LMN	2	4.50	0.56	0.70	0.18	0.61	0.56	.24
LMD	3	4.50	0.56	0.76	0.22	0.63	0.60	.26
NLR	2	2.50	0.55	0.78	0.24	0.63	0.57	.46
LVD	1	7.50	0.58	0.72	0.17	0.62	0.57	.40

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
OLDCF	3	7.50	0.25	0.27	0.06	0.26	0.10	.10
NEWCF	3	6.00	0.33	0.37	0.09	0.35	0.13	.15
RFCF	2	6.00	0.36	0.40	0.09	0.38	0.13	.16
NMN	3	6.00	0.25	0.26	0.04	0.26	0.12	.04
NMD	3	6.00	0.26	0.27	0.04	0.26	0.13	.04
LMN	3	4.00	0.43	0.46	0.08	0.45	0.19	.09
LMD	3	4.00	0.39	0.43	0.09	0.41	0.19	.08
NLR	3	2.50	0.54	0.67	0.18	0.59	0.21	.26
LVD	2	7.50	0.56	0.68	0.16	0.60	0.23	.32

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	1.50	0.53	0.79	0.27	0.61	0.62	.30
TEDC	2.00	0.57	0.74	0.20	0.62	0.62	.21

Table XII E-10. Simulation testing for 6-station network, scale = 2.25.

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
OLDCF	3	7.50	0.46	0.63	0.23	0.53	0.19	.33
NEWCF	3	7.50	0.46	0.63	0.23	0.53	0.19	.33
RFCF	2	7.50	0.48	0.66	0.22	0.55	0.20	.34
NMN	3	6.50	0.36	0.59	0.30	0.47	0.26	.31
NMD	3	6.50	0.36	0.64	0.32	0.49	0.28	.36
LMN	2	5.00	0.46	0.65	0.24	0.54	0.36	.30
LMD	3	5.00	0.44	0.70	0.28	0.55	0.37	.33
NLR	2	3.00	0.58	0.72	0.17	0.62	0.47	.30
LVD	3	9.00	0.53	0.68	0.19	0.59	0.37	.38

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
OLDCF	3	7.50	0.24	0.26	0.08	0.25	0.00	.12
NEWCF	3	6.00	0.32	0.40	0.19	0.37	0.00	.22
RFCF	2	6.00	0.35	0.46	0.20	0.41	0.00	.24
NMN	3	6.00	0.18	0.21	0.13	0.20	0.00	.07
NMD	3	6.00	0.18	0.24	0.21	0.23	0.00	.10
LMN	1	4.00	0.41	0.47	0.13	0.44	0.02	.14
LMD	3	4.00	0.30	0.37	0.18	0.34	0.02	.13
NLR	3	2.50	0.53	0.68	0.20	0.58	0.03	.29
LVD	3	7.50	0.54	0.68	0.18	0.59	0.02	.34

III) DIVERGENCE

METHOD	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
TDC	1.50	0.56	0.65	0.14	0.59	0.43	.19
TEDC	1.50	0.57	0.72	0.19	0.62	0.50	.25

Table XII E-11. Simulation testing for 6-station network, scale = 3.00. OLDCF represents the existing LLWAS with optimal threshold (7.5 m/s or 15 knots).

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
OLDCF	3	7.50	0.21	0.81	0.44	0.49	0.19	.68
NEWCF	3	9.50	0.43	0.57	0.21	0.50	0.14	.31
RFCF	2	9.50	0.44	0.59	0.22	0.51	0.14	.33
NMN	2	9.00	0.31	0.42	0.23	0.37	0.16	.22
NMD	3	8.50	0.32	0.56	0.32	0.44	0.19	.36
LMN	3	6.00	0.34	0.60	0.32	0.47	0.29	.37
LMD	3	6.50	0.31	0.51	0.29	0.42	0.25	.28
NLR	2	3.50	0.46	0.58	0.19	0.51	0.40	.24
LVD	3	11.50	0.43	0.58	0.21	0.50	0.34	.33

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
NEWCF	3	7.00	0.38	0.51	0.22	0.44	0.03	.25
RFCF	2	6.50	0.41	0.63	0.28	0.51	0.02	.35
NMN	2	6.00	0.24	0.41	0.31	0.35	0.03	.21
NMD	2	6.00	0.24	0.52	0.37	0.40	0.03	.31
LMN	1	4.00	0.39	0.58	0.26	0.48	0.05	.27
LMD	1	4.00	0.35	0.52	0.27	0.44	0.04	.23
NLR	1	2.50	0.39	0.55	0.24	0.47	0.06	.25
LVD	3	9.50	0.47	0.59	0.18	0.52	0.06	.30

Table XII E-12. Simulation testing for 6-station network, scale = 3.00. OLDCF represents the existing LLWAS with optimal threshold (9.5 m/s or 19 knots).

I) ANOMALIES, 1 STATION

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
OLDCF	3	9.50	0.43	0.57	0.21	0.50	0.14	.31
NEWCF	3	9.50	0.43	0.57	0.21	0.50	0.14	.31
RFCF	2	9.50	0.44	0.59	0.22	0.51	0.14	.33
NMN	2	9.00	0.31	0.42	0.23	0.37	0.16	.22
NMD	3	8.50	0.32	0.56	0.32	0.44	0.19	.36
LMN	3	6.00	0.34	0.60	0.32	0.47	0.29	.37
LMD	3	6.50	0.31	0.51	0.29	0.42	0.25	.28
NLR	2	3.50	0.46	0.58	0.19	0.51	0.40	.24
LVD	3	11.50	0.43	0.58	0.21	0.50	0.34	.33

II) ANOMALIES, 2 STATIONS

ANOM. TYPE	SMOOTH TYPE	THRESH.	TSS	POD	FAR	CSI	ELIG. POD	ELIG. FAR
OLDCF	3	7.00	0.37	0.50	0.22	0.44	0.02	.25
NEWCF	3	7.00	0.38	0.51	0.22	0.44	0.03	.25
RFCF	2	6.50	0.41	0.63	0.28	0.51	0.02	.35
NMN	2	6.00	0.24	0.41	0.31	0.35	0.03	.21
NMD	2	6.00	0.24	0.52	0.37	0.40	0.03	.31
LMN	1	4.00	0.39	0.58	0.26	0.48	0.05	.27
LMD	1	4.00	0.35	0.52	0.27	0.44	0.04	.23
NLR	1	2.50	0.39	0.55	0.24	0.47	0.06	.25
LVD	3	9.50	0.47	0.59	0.18	0.52	0.06	.30

Appendix F. Study of an Irregular Grid Problem: An Expanded LLWAS at Stapleton.

To determine the degree to which these ideas can be applied to the design of an actual airport LLWAS, and to begin to understand the impact that an irregular geometry has on the efficiency of the algorithms, we have applied the methods of this report to the problem of designing the expanded LLWAS at Stapleton International Airport in Denver. The current Stapleton LLWAS has six stations, CF, SW, SE, NW, NE, and N (Figure XIIF-1). CF appears to be somewhat sheltered by the new fire station. The plan is to add five more stations, for a total of eleven, and to move CF to a more favorable site. For economic reasons, it has been decided to make every effort to avoid moving the other five stations.

In designing the tests for determining the best geometry and algorithm, the first problem is to determine the geographic region that is to be protected. In the main report, protecting a disc of radius 5 km is taken as the goal for the regular geometry simulation testing. For Stapleton, protecting such a disc would barely protect the ends of the runways and would protect large regions in which it is not expected that there would be low altitude air traffic. Therefore, it was decided that a better plan would be to protect a rectangle that provides a 1-km buffer about the runways. This rectangle does not provide protection of the full areas of vulnerability, which extend for up to two miles beyond the ends of the runways, but does give full coverage of the airport itself and a lesser buffer beyond the airport. A significant benefit is that this region can be protected with almost all of the LLWAS sites being located on either airport or Rocky Mountain Arsenal property, so that extensive real estate negotiations are not required. It is also unlikely that it would be possible to protect a significantly larger region adequately without using more stations. Therefore the goal for the Stapleton LLWAS design is to protect a rectangle which is centered at the geographic center of the triangle that is defined by the runway extremities and whose sides have lengths 8.4 km (N-S) and 5.1 km (E-W) (cf. Figure XIIF-1). Aside from this modification, the simulation testing procedure is identical to the one that was used previously.

In addition to evaluating the relative performances of the various algorithms, we want to understand the influence of changing the geometry. In particular, it is important to see how sensitive the performance is to modest alterations of the station positions. In a real application, such as the design of the Stapleton LLWAS, various practical constraints are placed on the selection of station sites. If the performance of some algorithms is extremely sensitive to site selection, then those algorithms will be more difficult to implement. The study of locating the new stations at Stapleton helps us to gain an understanding of the sensitivity of the various algorithms in this practical setting.

Figure XIIF-2 shows the locations of all of the possible sites that were considered for this study. The numbering has been selected so that the sites that were selected for actual installation are numbered as they

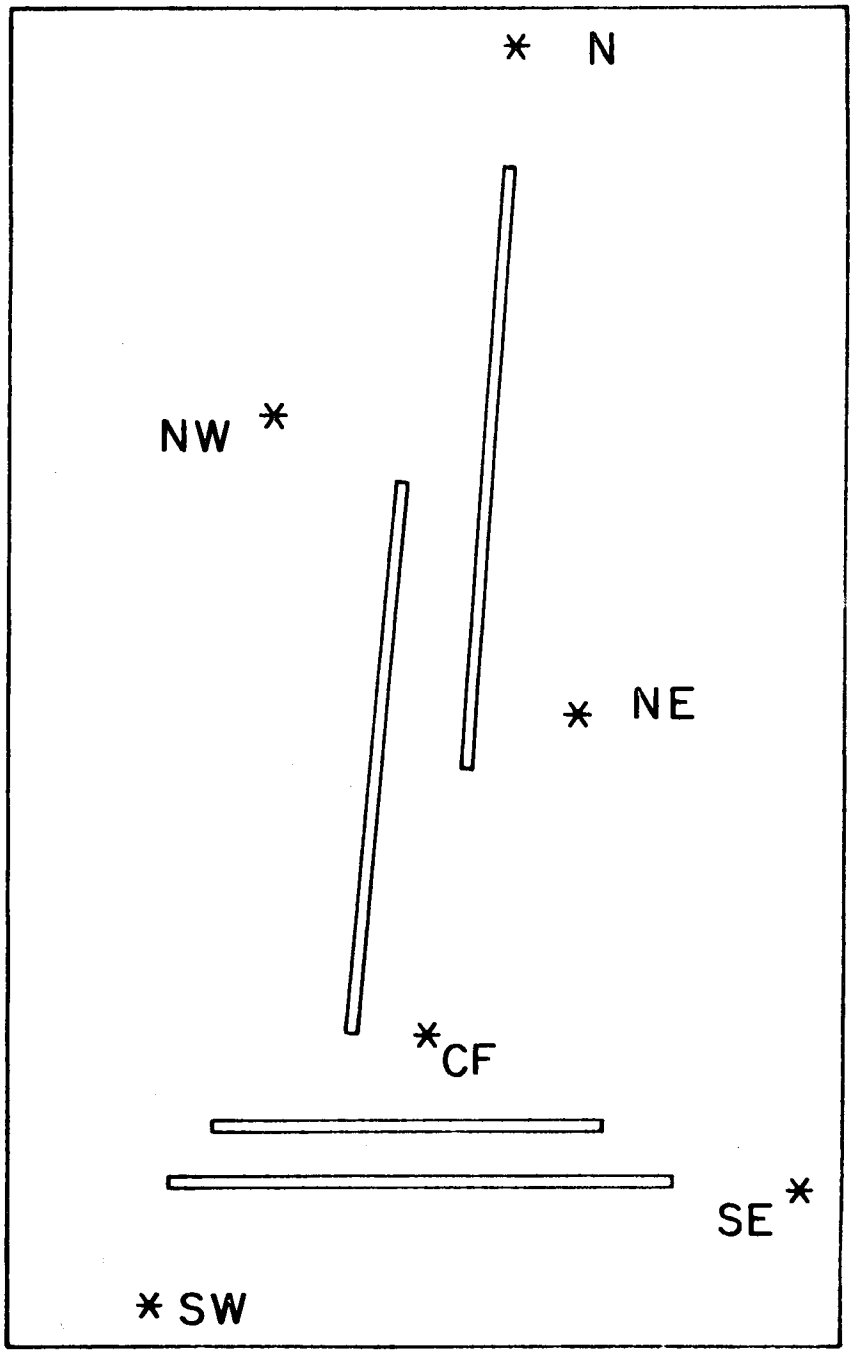


Figure XII F-1. The original LLWAS at Stapleton (Geometry 0) and the rectangular protection region.

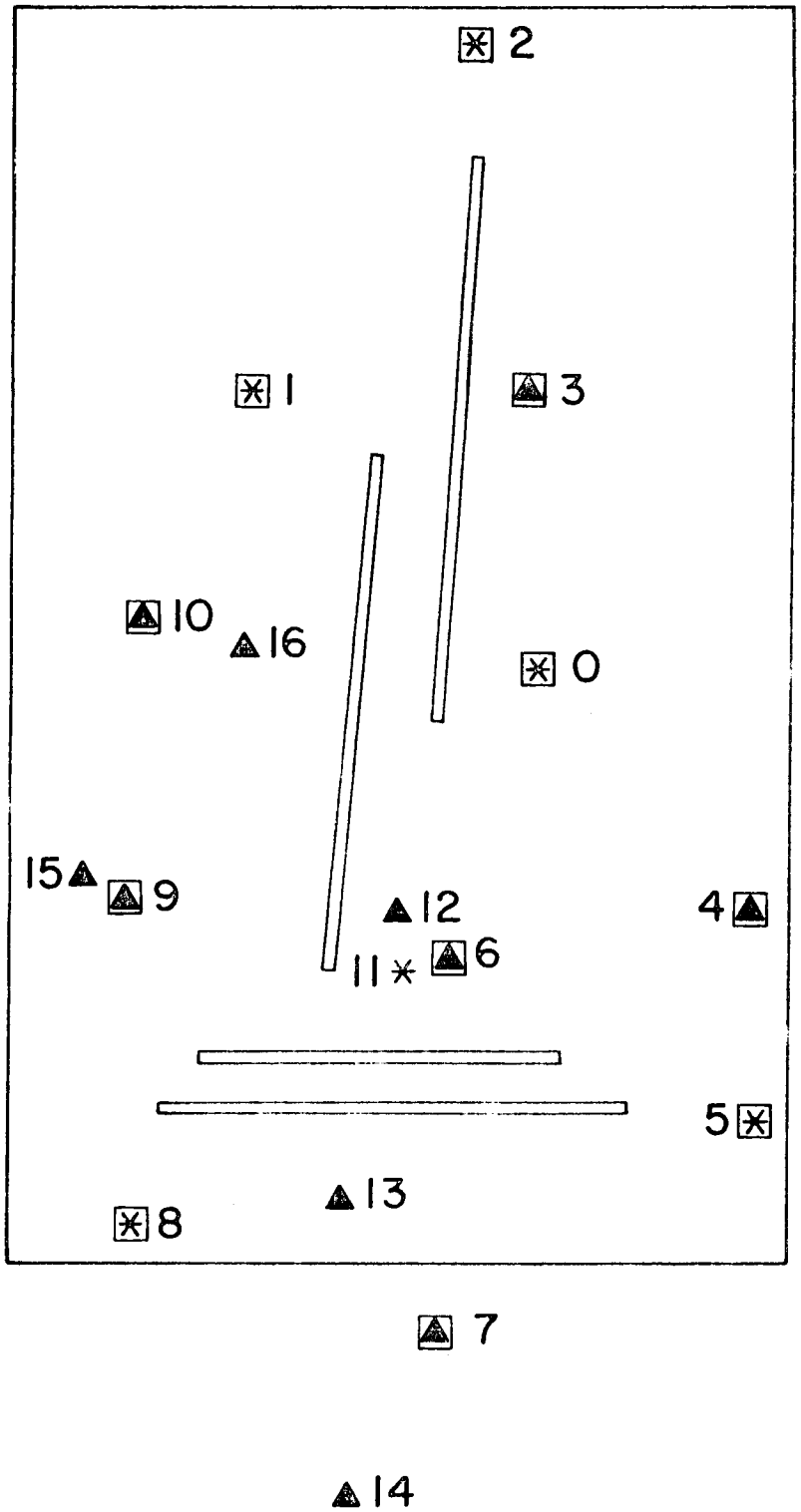


Figure XIIF-2. All station sites that were considered.

will be designated for operations ([0] to [10]). These sites are designated by a small square on the figure. The sites of the original LLWAS are designated by *.

The 11-station network that was first suggested is shown in Figure XIIF-3A. This geometry was chosen in a round-table discussion between representatives of the FAA and NCAR. Based on the requirement that the existing outlying stations should not be moved, positions for the new stations were determined so that the station spacing would be reasonably uniform. Station [12] was discussed as a geometrically satisfactory alternate site for centerfield.

A site survey revealed the partial sheltering of the old centerfield [11], a tree problem at site [7], and potential sheltering and other difficulties due to construction near [15] and [16]. The site survey also showed that [12] is sheltered from the south by a large bluff. Site [6] is the relocation for the old centerfield station that was selected at the time of the site survey. Alternate sites that were tested: [6] or [12] for site [11], [13] or [14] for site [7], [9] for [15], and [10] for [16]. The only old site that is not part of the new system (F) is the old centerfield site [11].

Figures XIIF-2B and -2C show the geometries that are obtained when the alternate locations for [7] were tested. Neither of these performed as well as Geometry A (Tables XIIF-1 and XIIF-2). It was decided to try to use site [7] and to construct a fairly high tower (70'). Figures XIIF-3D and -3E are obtained from Figures XIIF-3A and -3B by moving the old centerfield site [11] to the north ([12]). When the second site survey showed that [12], [15], and [16] are also unsuitable, the new locations [6], [9], and [10] were selected. This geometry is shown in Figure XIIF-3F. The difference between Geometries F(1) and F(2) is the designation of the centerfield station for the centerfield algorithm; for F(1), station [0] is used, and for F(2), station [6] is used. For comparison, we also tested a regular 10-station geometry (Figure XIIF-3G) and the original 6-station LLWAS (Figure XIIF-1 and Geometry 0). In all, nine geometries are tested:

0. The existing 6-station network.
A. The 11-station network that was selected by the committee without site inspections.
B. Geometry A with the south station moved up to the edge of the airport property.
C. Geometry A with the south station moved an additional kilometer to the south.
D. Geometry A with centerfield moved to the north.
E. Geometry B with centerfield moved to the north.
F(1). The 11-station network that will be installed with the most central station used as centerfield.
F(2). The 11-station network that will be installed with the centerfield station chosen to be nearest the terminal.
G. A uniform 10-station network.

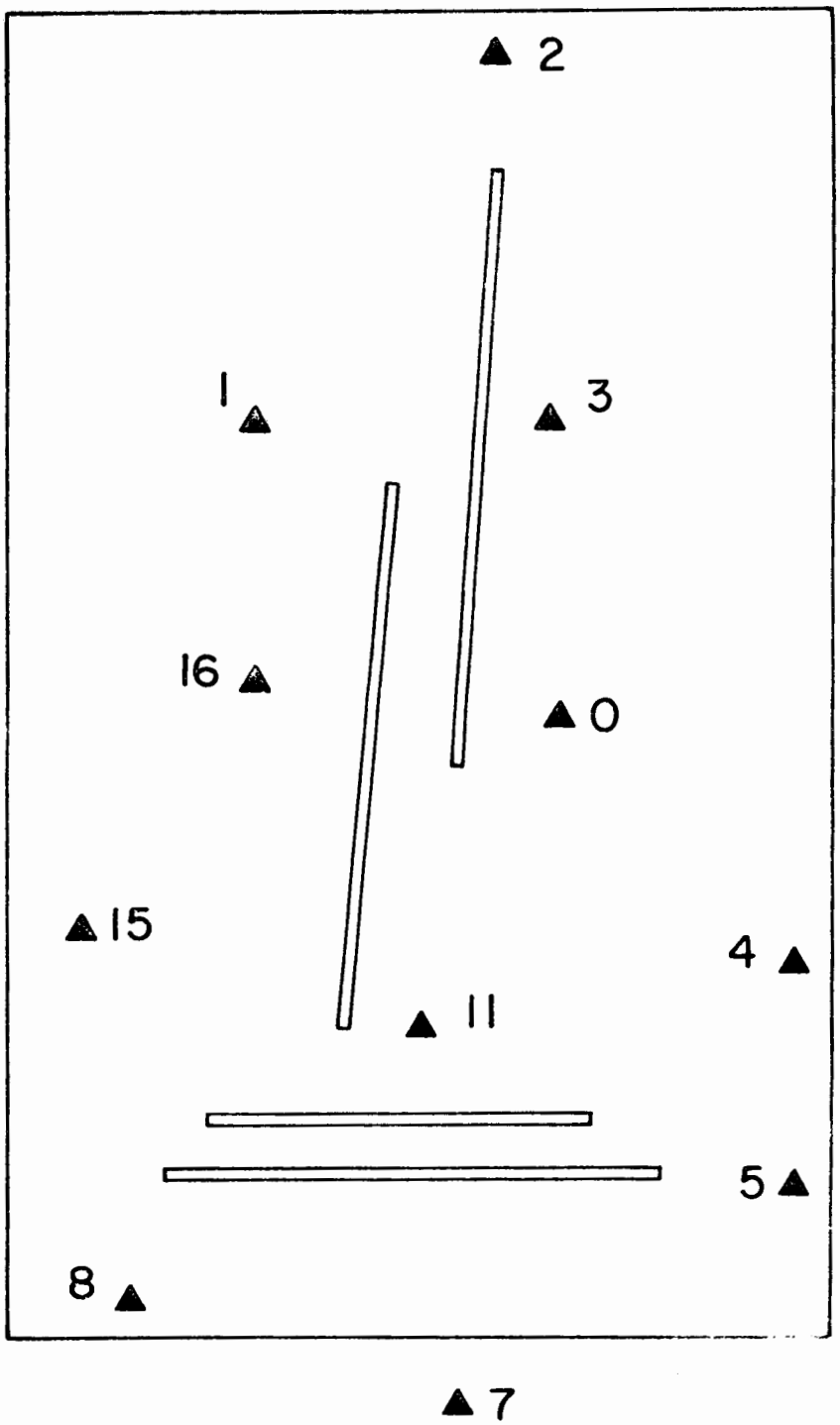


Figure XIIF-3A. Geometry A.

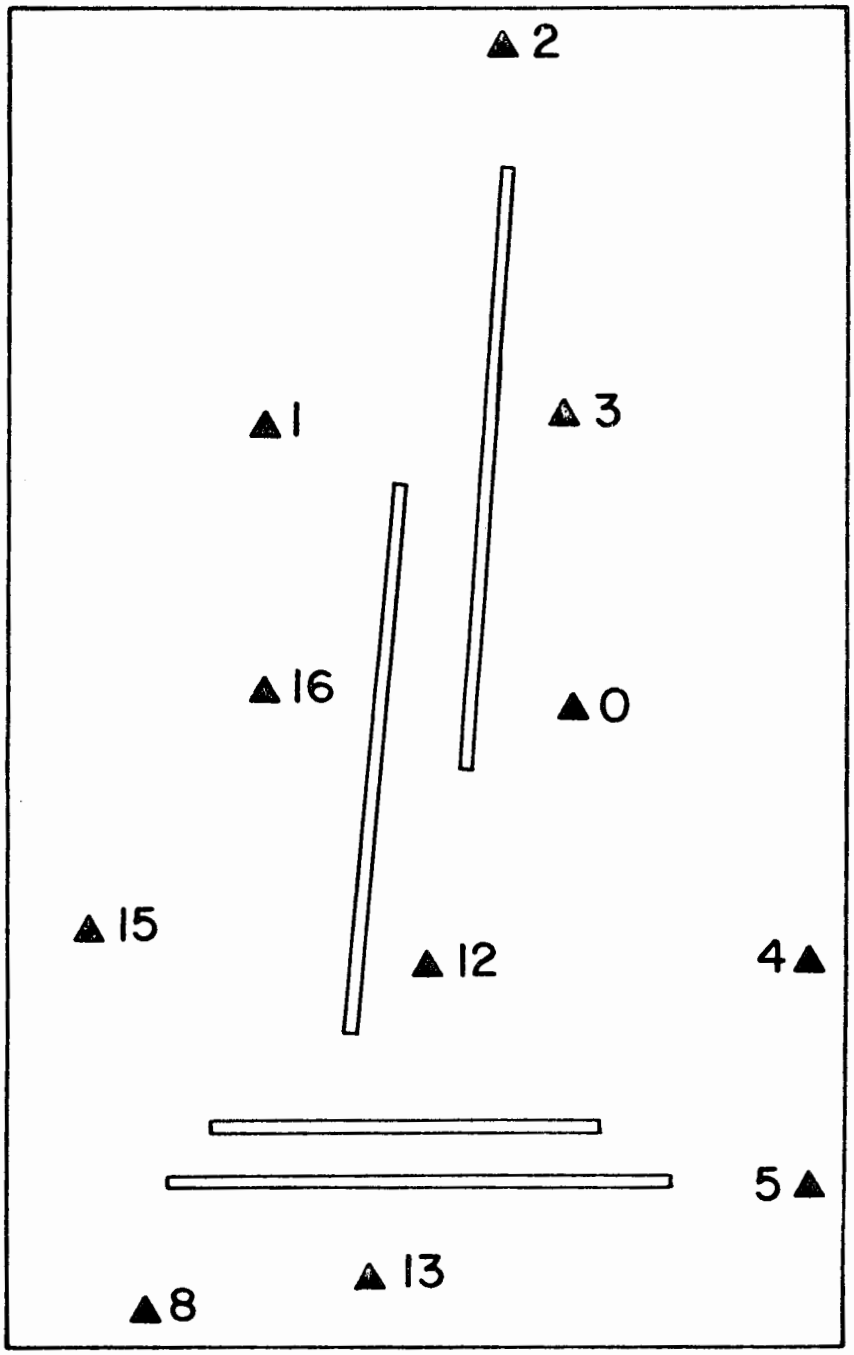


Figure XII F-3B. Geometry B.

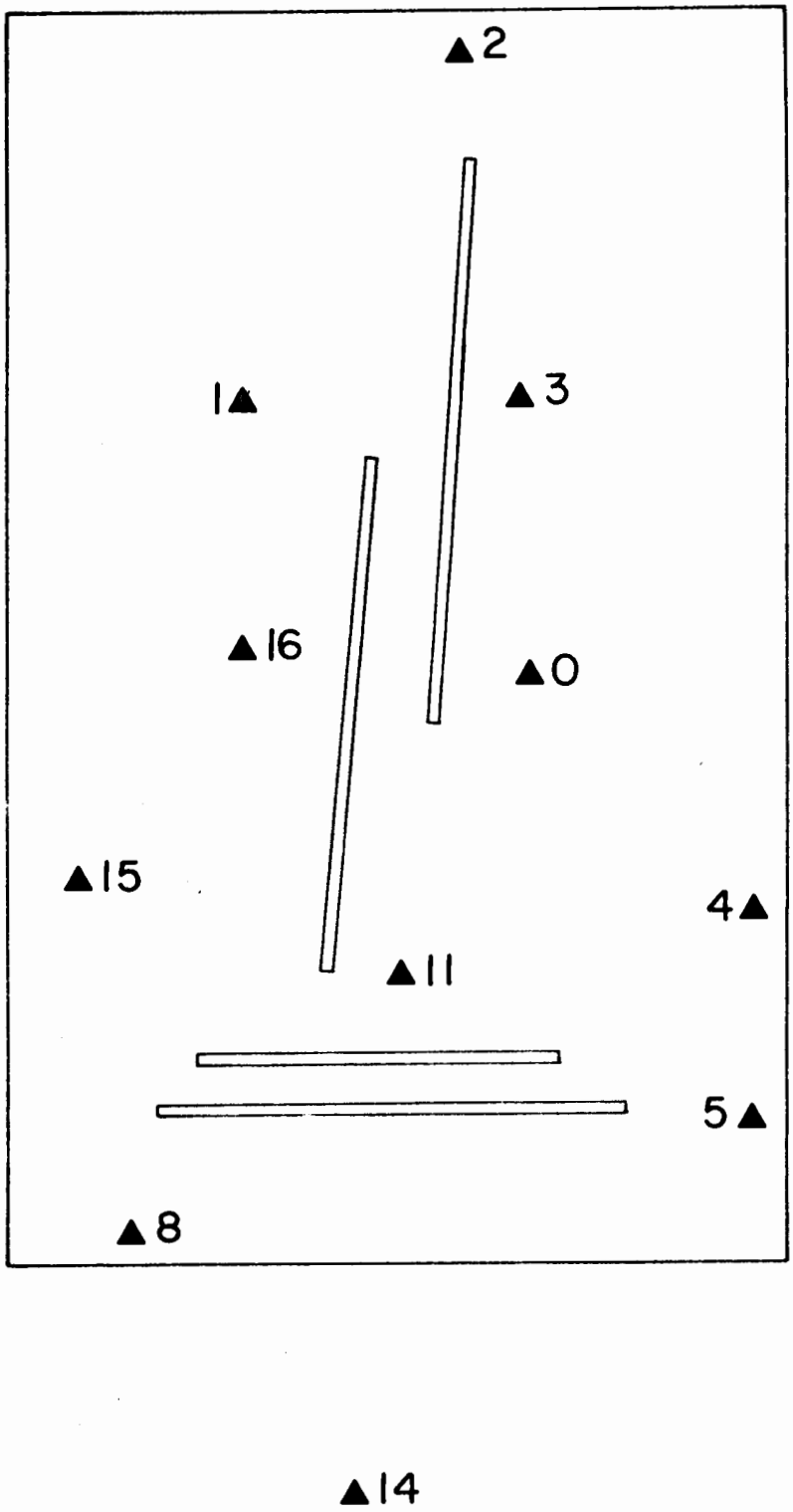


Figure XII F-3C. Geometry C.

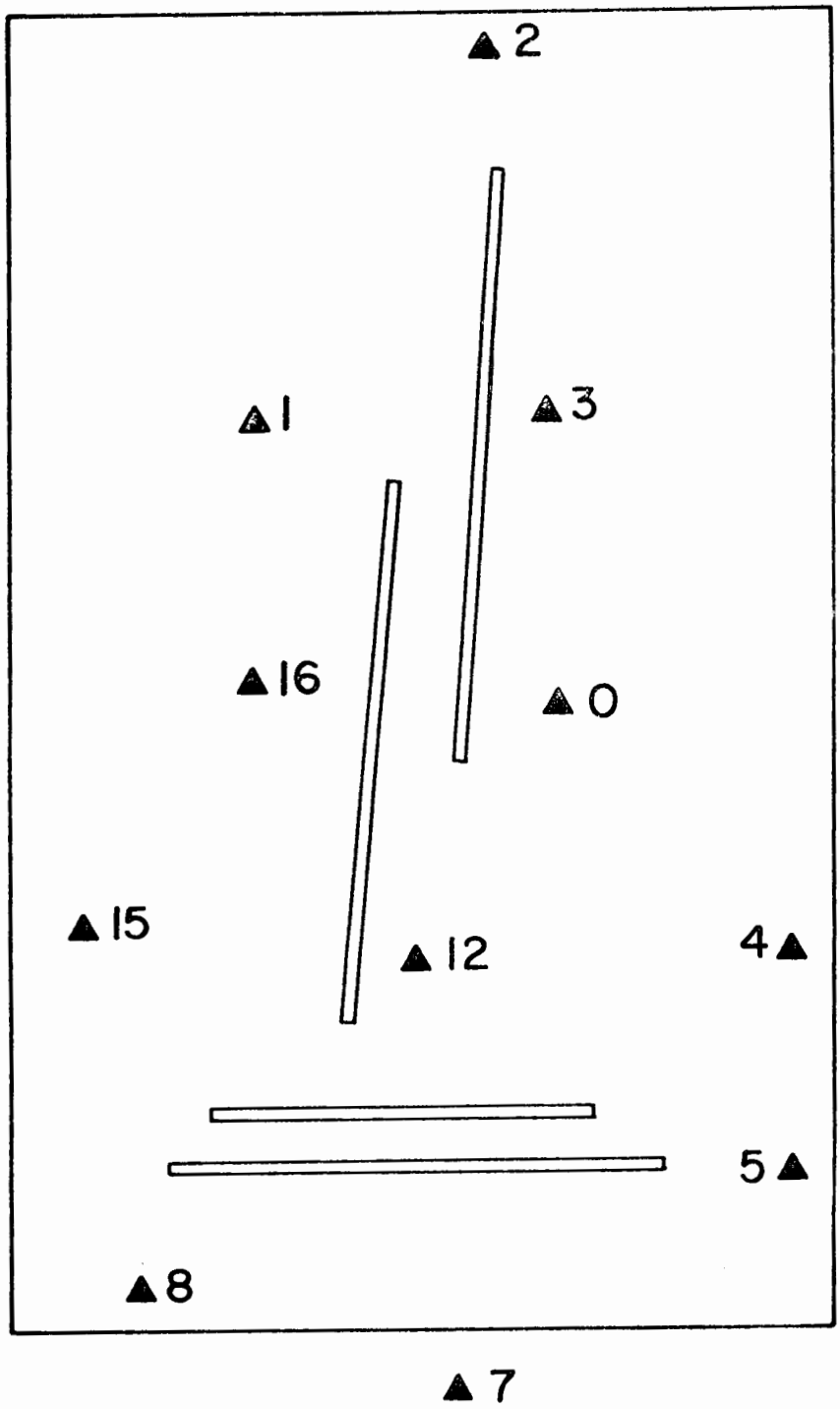


Figure XII F-3D. Geometry D.

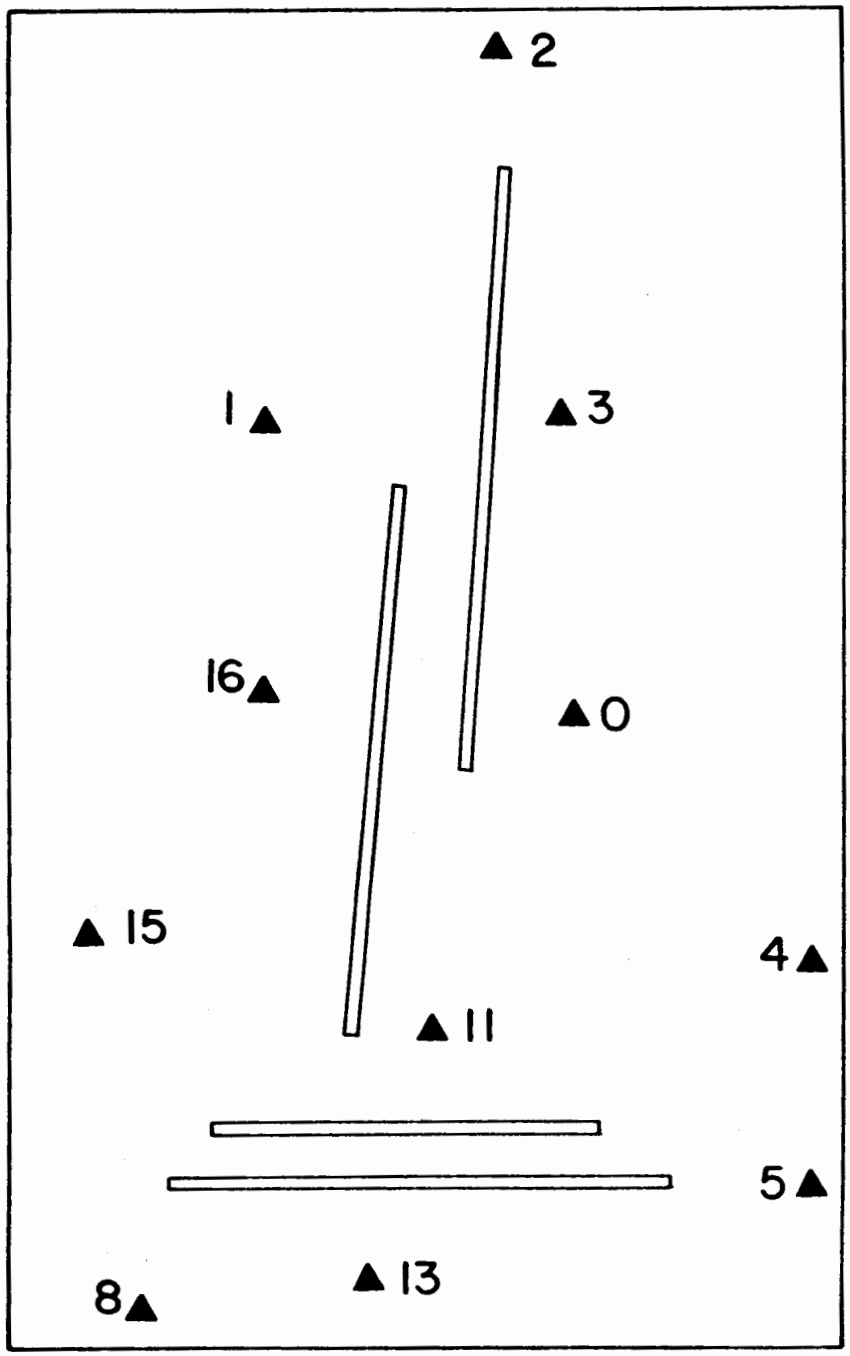


Figure XII F-3E. Geometry E.

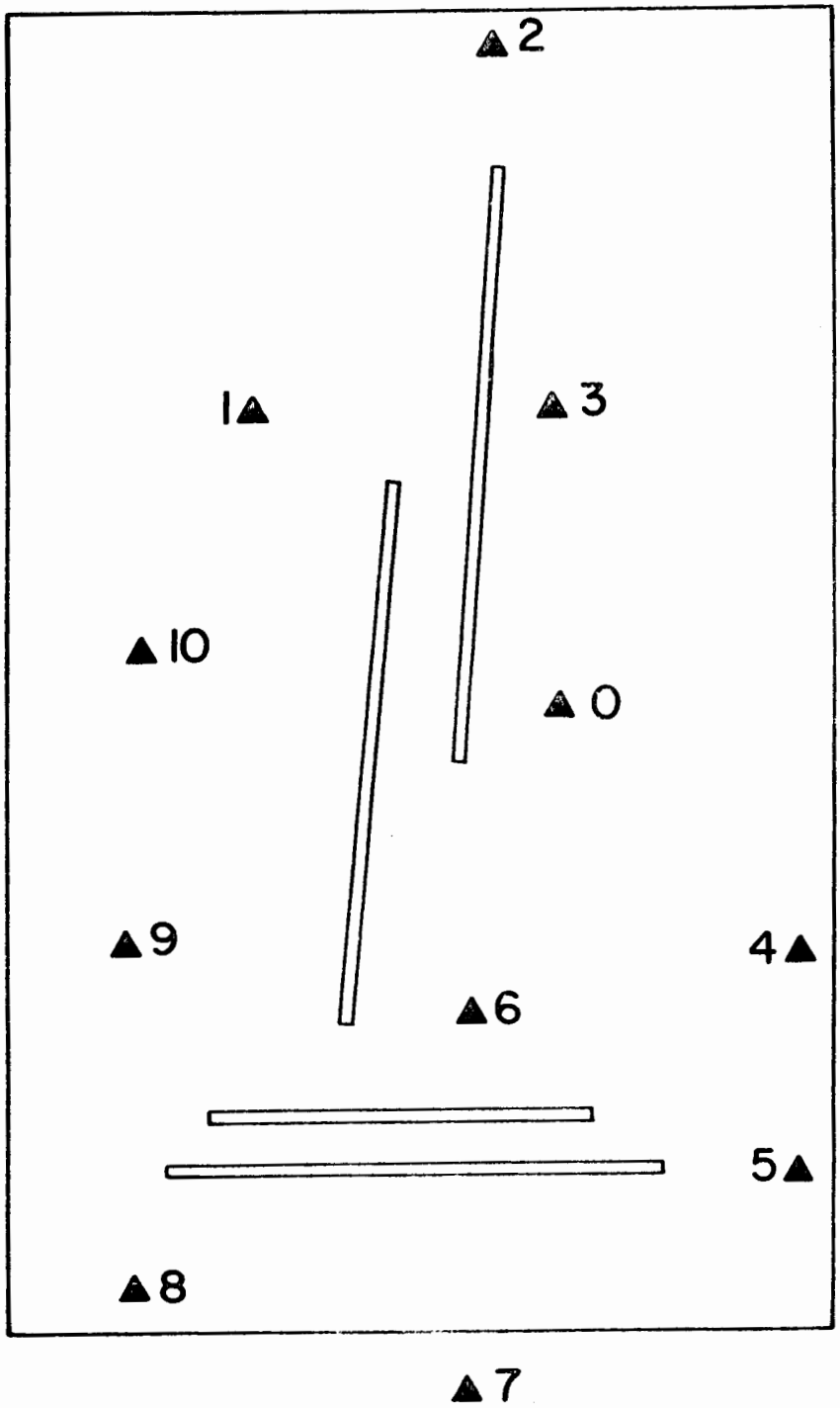


Figure XII F-3F. Geometry F.

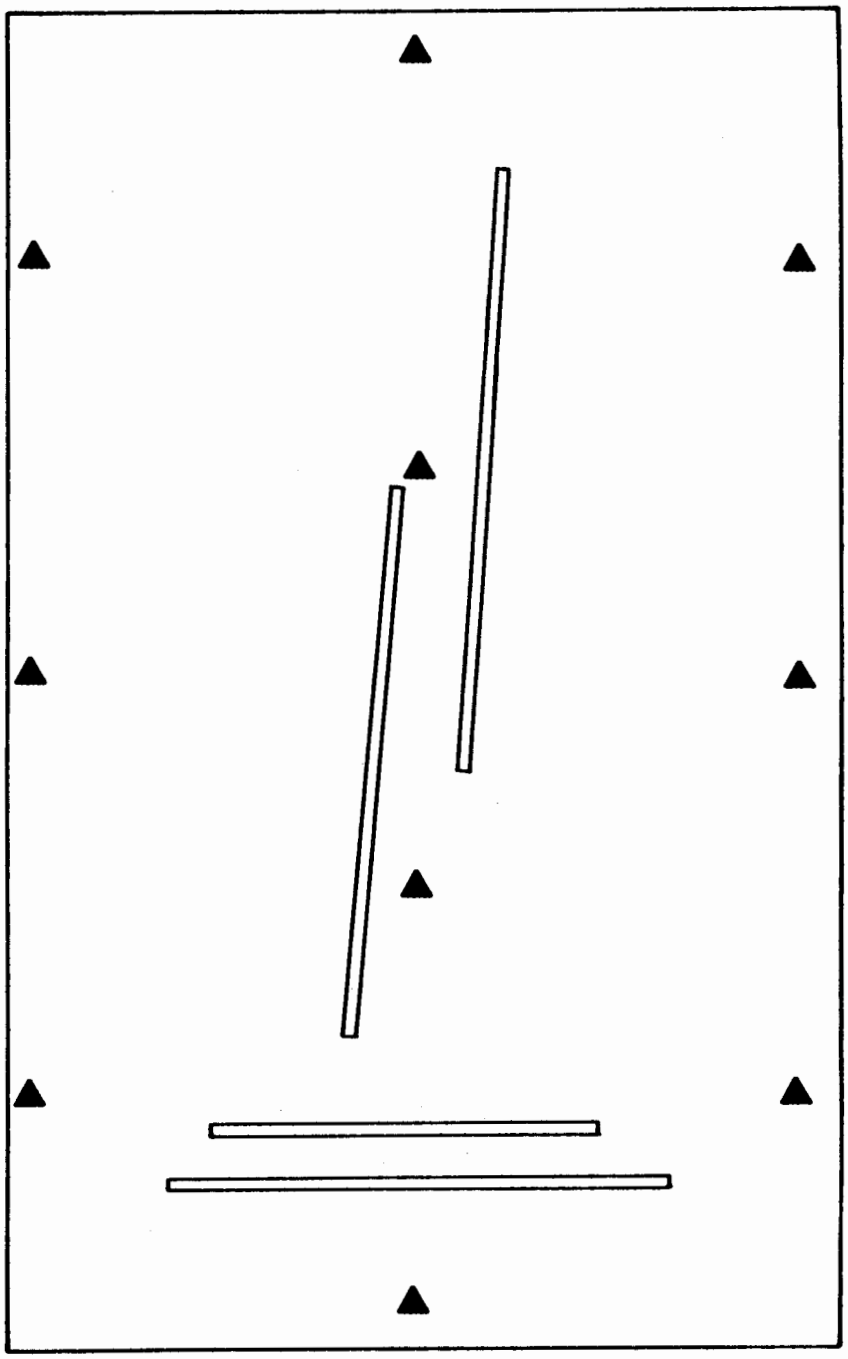


Figure XII F-3G. Geometry G.

The following tables give the results of the verification testing for the new centerfield algorithm (NEWCF) and the network linear regression algorithm (NLR). Recall that the TSS is a measure of the true skill of the network for detecting wind shear events, the CSI is the traditional meteorologist's measure of detection (prediction) skill, POD is the probability that a detection is made by some station in the network when there is an event somewhere on the network, and FAR is the probability that an issued alarm is false, i.e., a detection is made by some station, but there is no wind shear event on the network. For determining if the stations in the wind shear region are the ones that are issuing the alarms, we have also introduced the EL.POD, the probability that a station that is in the wind shear region will make a detection, and the EL.FAR, the probability that a station that is not in the wind shear region will make a (false) detection. As before, each of the algorithms is evaluated with its threshold selected to give the optimal TSS performance, i.e., detection by the network has the highest skill.

Table XIIF-1. Comparison of the anomaly algorithms for various network geometries.

NEW CENTERFIELD ALGORITHM (USING OPTIMAL THRESHOLD)

GEOMETRY	THRESHOLD	TSS	CSI	POD	FAR	EL.POD	EL.FAR
0	6.00	.60	.64	.72	.15	.47	.24
A	7.00	.68	.71	.81	.15	.71	.33
B	6.00	.66	.69	.85	.21	.76	.39
C	7.00	.67	.70	.80	.15	.69	.33
D	7.00	.69	.71	.81	.15	.73	.32
E	6.00	.68	.72	.87	.20	.77	.39
F(1)	8.00	.66	.68	.75	.12	.66	.28
F(2)	6.00	.62	.68	.85	.24	.77	.35
G	8.00	.63	.66	.75	.12	.58	.28

LINEAR REGRESSION ALGORITHM (BEST ANOMALY ALGORITHM)

GEOMETRY	THRESHOLD	TSS	CSI	POD	FAR	EL.POD	EL.FAR
0	3.00	.63	.68	.81	.20	.50	.39
A	5.00	.74	.75	.85	.14	.78	.22
B	5.00	.71	.75	.83	.14	.75	.20
C	5.00	.73	.75	.85	.14	.76	.22
D	5.00	.74	.76	.86	.13	.78	.20
E	5.00	.73	.75	.85	.14	.77	.20
F	5.00	.73	.75	.85	.14	.78	.21
G	5.00	.69	.72	.83	.16	.68	.23

Table XIIF-2. A comparison of the identification algorithms for various network geometries.

TRIANGLE DIVERGENCE METHOD

GEOMETRY	THRESHOLD	TSS	CSI	POD	FAR	EL.POD	EL.FAR
A	3.50	.71	.73	.83	.14	.81	.08
B	3.50	.71	.73	.85	.16	.82	.13
C	3.50	.71	.73	.82	.14	.81	.08
D	3.50	.71	.73	.83	.14	.82	.08
E	3.50	.70	.73	.82	.14	.80	.08
F	3.50	.70	.73	.82	.15	.80	.09
G	3.00	.70	.73	.84	.16	.83	.10
	(3.50)	.69	.71	.79	.12	.78	.06

TRIANGLE AND EDGE DIVERGENCE METHOD

GEOMETRY	THRESHOLD	TSS	CSI	POD	FAR	EL.POD	EL.FAR
A	3.50	.74	.76	.87	.14	.85	.09
B	3.50	.70	.73	.86	.17	.83	.14
C	3.50	.74	.75	.86	.14	.84	.09
D	3.50	.72	.74	.85	.15	.84	.09
E	3.50	.71	.73	.83	.15	.82	.09
F	3.50	.72	.74	.85	.16	.83	.09
G	3.00	.72	.74	.86	.15	.84	.10
	3.50[1]	.70	.73	.81	.11	.80	.06

[1] These suboptimal results are provided for comparison.

From these results, we see that Geometries A and D have the best performance for both the NEWCF and the NLR algorithms, but that the advantage over the other geometries is slight. In particular, Geometry F, the one under construction at Stapleton, tests very well. For NEWCF, F(1) is quite a bit better than F(2), and so we recommend relocating the centerfield station to the position nearest to the geometric center of the runways, site [O]. It is interesting to note that the old LLWAS 6-station geometry with the NLR algorithm has a performance that is nearly as good as the 11-station geometries with the NEWCF algorithm. This suggests that an upgrade of the computer and the algorithm at the existing LLWAS installations at other airports may provide the quickest way to get a substantial LLWAS improvement. Of course, the 11-station geometry and the NLR algorithm give an additional improvement over all the 6-station geometry.

When we consider the question of how well the position of the wind shear danger is located by the LLWAS, we find that the results are less satisfactory. NEWCF is seen to issue an alarm at a station that is not near the wind shear over 30 percent of the time. These alarms could be interpreted by pilots as false alarms, since they might not be aware that

there is a wind shear somewhere else on the network. NLR issues detections that are not near the wind shear about 20 percent of the time. While this is much better than 30 percent, it is still rather high when we speculate that it could be the basis for ATC advising an aircraft to leave an area where there is possibly no wind shear and to enter an area where there is a wind shear.

Our testing shows that the TDC and TEDC algorithms have skill levels that are nearly comparable to the NLR algorithm. When we compare POD with EL.POD and FAR with EL.FAR, we observe that, unlike the anomaly algorithms, there is very little difference when the eligibility restriction is imposed, indicating that TEDC is more reliable for prediction of the location of the wind shear. Also, the detection ability is virtually unaffected by the minor changes in the geometries (A-F).

Geometry G is a uniform 10-station network that was tested to determine the amount of degradation that occurs due to the irregular geometry. Its performance is slightly inferior to the 11-station networks when an identification algorithm (TEDC) is used. This suggests that the TEDC algorithm is not too severely penalized by grid irregularity, except that one additional station is needed to get comparable results. Our initial study of the sliver problems indicates that this difficulty may be dealt with by refinements of this method.

In conclusion, the geometry (F) that has been chosen for installation at Stapleton is expected to give a vastly improved wind shear detection system, even with the new centerfield algorithm (NLR). The TEDC algorithm provides an improvement in locating the wind shear and is able to identify the nature of the event. The new LLWAS computer that is being installed at Stapleton seems to have ample capacity to handle the additional computing requirements for these new algorithms.