D10-32
486 3P.

N87-10521

8.5.2 THE TUKEY ALGORITHM FOR ENHANCING MST RADAR DATA

19936

S. A. Bowhill

Aeronomy Laboratory
Department of Electrical and Computer Engineering
University of Illinois
Urbana, Illinois 61801

One of the most troublesome features in MST velocity measurements is the determination of unwanted scatterers whose velocity is different from that of the surrounding atmosphere. Aircraft seen in the sidelobes of the antenna are the principal problem.

Because coherent integration essentially eliminates echoes with line-of-sight velocities greater than 10 or 20 m/s, aircraft are seen only when their flight path is almost perpendicular to the line-of-sight. Then, they give large returns whose velocities may be positive or negative, and certainly different from that of the surrounding air. These "glitches" in the minute-by-minute velocity records are quite troublesome in that they may distort the statistics of the velocity. Table 1 illustrates a simulation where a fairly smooth velocity profile was generated by applying a 20-point moving average to a sequence of random values whose standard deviation is unity. Standard deviation of those points should be .224. The lines A on Table 1 show sample standard deviations of 20 sequences of 60 points generated in this way and represent hourly data. The mean standard deviation is .216, reasonably close to the theoretical value. The considerable variation in these hourly standard deviations illustrates the problems in trying to determine accurate statistics on a time series such as this.

Lines B represent standard deviations of the simulated hourly data with randomly added 1-minute glitches, each having a standard deviation of unity (or about 4.5 x the time series standard deviation). The probability that a glitch will appear in a given minute of data varies from .01 to 0.3. Mean standard deviations for the glitched data are as high as .520 for a glitch probability of 0.3.

An objective way is therefore needed to remove sporadic points of this kind. For this purpose, the Tukey algorithm is appropriate and has some advantages over averaging.

The Tukey algorithm, applied to a data array, uses for each data point the median of it and the two points surrounding it. If the three points form a monotonically increasing or decreasing sequence, the original point is copied without change. However, if the central data point is remote from the other two, it is replaced by whichever of the two surrounding points is closest in value.

In Table 1, lines C and D show the results of applying the Tukey algorithm to the original and to the glitched data. As can be seen, almost all of the glitches are successfully removed by the use of the Tukey algorithm, and even the original data are not much affected as far as standard deviation is concerned.

The greatest effect of the Tukey algorithm is on data where the successive points are uncorrelated. Table 2 shows the results of applying the Tukey algorithm to five samples of a 1000-element random array with population standard deviation unity. Its standard deviation is reduced by a factor of approximately .687, so serious errors can be produced by applying the algorithm to such data.

TABLE 1

Glitch probability	Velocity	Standard Deviations x 1000					Mean
0.01	A	174	158	217	179	254	197
	B	192	158	217	179	333	216
	C	171	159	205	178	249	192
	D	172	159	205	178	251	193
0.03	A	367	237	212	201	183	240
	B	367	345	222	272	184	278
	C	365	230	207	188	180	234
	D	365	231	209	189	181	235
0.1	A	238	194	274	202	196	221
	B	268	332	377	328	303	322
	C	235	189	267	199	199	218
	D	236	188	262	233	197	223
0.3	A	206	175	143	271	150	189
	B	527	536	474	563	501	520
	C	193	172	142	271	143	184
	D	213	207	188	282	144	207

A is simulated wind, theoretical standard deviation = .224
B is A with added glitches, each with s.d. = 1
C is A smoothed with Tukey algorithm
D is B smoothed with Tukey algorithm

TABLE 2

Effect of Tukey algorithm on 1000-element random array of theoretical standard deviation unity.

Sample No.	Standard Deviations Random Array	Standard Deviations Tukeyed array	Ratio
1	1.0110	0.6952	0.6876
2	0.9731	0.6751	0.6938
3	1.0049	0.6845	0.6812
4	1.0019	0.6954	0.6941
5	0.9946	0.6755	0.6792

Average ratio = 0.6872 ± .0005

Table 3 shows another simulation in which 1000-element arrays were generated with Gaussian correlation functions, with E-folding times of 0, 0.7, 1.4, and 3.1 min. Lines A represent the correlation function or the original winds and lines B represent the correlation functions after Tukey smoothing. Acceptable results are found of the correlation times of 1.4 min or more.

The Tukey algorithm has been successfully used in smoothing stratosphere MST velocity data (BOWHILL and GNANALINGAM, this volume) and the results were found to be of good quality.

TABLE 3

Correlation time (min)	Velocity	Correlation x 1000 vs Lag (min) 0	1	2	3	4	5	6	7
0	A	1000	-26	44	10				
	B	1000	542	325	43				
0.7	A	1000	679	200	37	37	31		
	B	1000	766	394	145	58	37		
1.4	A	1000	812	433	162	62	24	-14	
	B	1000	838	523	248	106	44	1	
3.1	A	1000	907	679	420	254	83	17	-5
	B	1000	915	707	463	256	116	38	0

A is simulated wind, Gaussian autocorrelation
B is A smoothed with Tukey algorithm