United States
Department of
Agriculture

Forest Service

**Technology &
Development
Program**

5700 Aviation
September 2002
0257-2826-MTDC

Estimating Methods, Variability, and Sampling for Drop-Test Data

Ann Suter
Statistician

**USDA Forest Service
Technology and Development Program
Missoula, MT**

9E92P32—Technical Service, Aerial Delivery

September 2002

Contents

Introduction

For nearly six decades, the Forest Service has been using a procedure known as drop testing to analyze the ground patterns made by aerial drops of fire retardants or suppressants (Suter 2000). The procedure involves dropping firefighting chemicals from an airtanker flying over open cups arranged in a regularly spaced grid (figure 1). The ground patterns allow operators and managers to:

- ❑ Compare the performance of aerial delivery systems
- ❑ Compare the drop characteristics of firefighting chemicals
- ❑ Determine whether an aerial delivery system is suitable for contracting
- ❑ Investigate the effect of changes in drop height, drop speed, volume, flow rate, and other factors

Figure 1—TBM (Avenger) aircraft dropping fire retardant over a test grid.

After more than 25 aerial delivery systems were tested during the early 1990s, three concerns arose:

- ❑ Estimation methods
- ❑ Variability
- ❑ Sampling

The first concern deals with the process of making estimates using the data collected from the grid. It is not feasible to collect every gram of retardant that hits the ground. Instead, the drop is sampled at regular intervals with estimates made for points in between. Historically, linear interpolation was used to estimate between sample values. Linear interpolation assumes uniform change between points, an assumption that may be inadequate for drop data.

The second concern relates to the variability of estimates and of the test. Any time a quantity is estimated, the variability associated with the estimate should be provided.

Replicate drops can help investigators obtain a measure of the variation inherent in the test. Replication also reduces the variability of mean line length for each drop type. Because of the cost of other testing constraints, replicate drop tests are usually not conducted, making it impossible to estimate reliably the error variance due to the test.

The third concern, sampling, pertains to grid arrangement and cup placement. Although hundreds of drops have been conducted over grids, little testing has been done to determine the appropriate cup spacing and grid dimensions (Suter 2000). Usually, the length and width of the grid are estimated based on flow rate, volume, and ground speed. Some steps have been taken to achieve greater consistency. For instance, cups are placed at a uniform height and spaced in a regular pattern. For most drop tests, a denser area has been constructed in the middle of the grid where the majority of retardant is expected to fall. Constraints on time, budgets, and labor must be taken into consideration when developing a sampling scheme.

This report uses data collected from six airtanker drops to investigate estimation methods, variability, and sampling.

Procedures

Grid Layout

Figure 2 illustrates a 600- by 155-foot grid. A total of 544 stakes were driven into the ground so their tops were 4 feet high. The stakes were staggered to reduce the distance between known and unknown points.

Examination of hundreds of past drops showed that the rate of change in coverage level was often greater crossrange (perpendicular to the flight path) than downrange (in the direction of the flight path), especially for drops at high speeds. How close should the cups be placed to capture this feature of ground patterns? To answer this question, the spacing was decreased from 10 feet to 5 feet for three crossrange rows (figure 2). The effectiveness of the 5-foot spacing was compared to the 10-foot spacing.

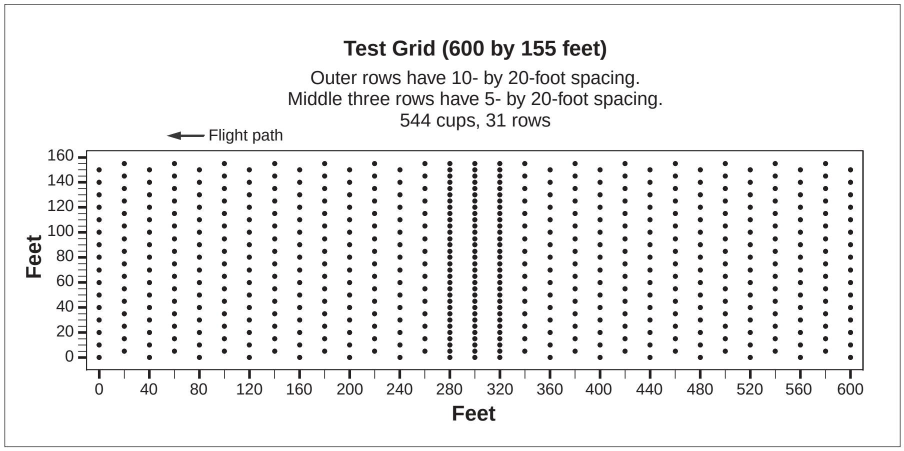

Figure 2—Diagram of the test grid.

Grid Collection Method

After a drop is made, any cup with moisture in it is capped. The row and column numbers are written on the lid, identifying the location of the cup in the grid. All the cups with lids on them are removed and taken to be weighed. Clean cups are put back out on the grid for the next drop (figure 3).

Figure 3—Grid workers gathering cups after a drop.

Weighing and Calculating

During the weighing process, the weight and coordinates of each cup are entered into a computer. The weight of the empty plastic cup and lid is subtracted from the total weight and the weight of the liquid in grams is converted to gallons per hundred square feet (gpc) using the density of the liquid. For retardant with a density of 1.095 grams/cubic centimeter, the equation is:

$$gpc = \left[\left(\frac{x}{453.6\ \textit{grams / pound}}\right) \div 9.13707\ \textit{pounds / gallon}\right] \div 0.001944\ \textit{square feet / cup}$$

where x is the amount of retardant in the cup in grams.

Figure 4 is an example of the computer output after drop samples have been weighed. Cups that were not picked up are assumed to be 0 gpc. This value is included in the array. This array is used to create a map of the drop (figure 5).

Cup position Row	Cup	gpc
0	0	0.02
0	10	0.02
0	20	0.02
0	30	0.01
0	40	0.02
0	50	0.05
0	60	0.21
0	70	0.70
0	80	0.66
0	90	0.13
0	100	0.01
0	110	0.00
0	120	0.00
0	130	0.00
0	140	0.00
0	150	0.00
20	5	0.00
20	15	0.00
20	25	0.01
20	35	0.08
20	45	0.32
20	55	0.65
20	65	0.92
20	75	0.99
20	85	0.96
20	95	0.01
20	105	0.00
20	115	0.00
20	125	0.00
20	135	0.00
20	145	0.00
20	155	0.00
40	0	0.00
40	10	0.04
40	20	0.04
40	30	0.04
40	40	0.06
40	50	0.20
40	60	0.84
40	70	1.30
40	80	1.52
40	90	0.88
40	100	0.05
40	110	0.00
40	120	0.00
40	130	0.00
40	140	0.00
40	150	0.00
60	5	0.00

Figure 4—Data array with cup position indicated in columns 1 and 2 and gpc values in column 3.

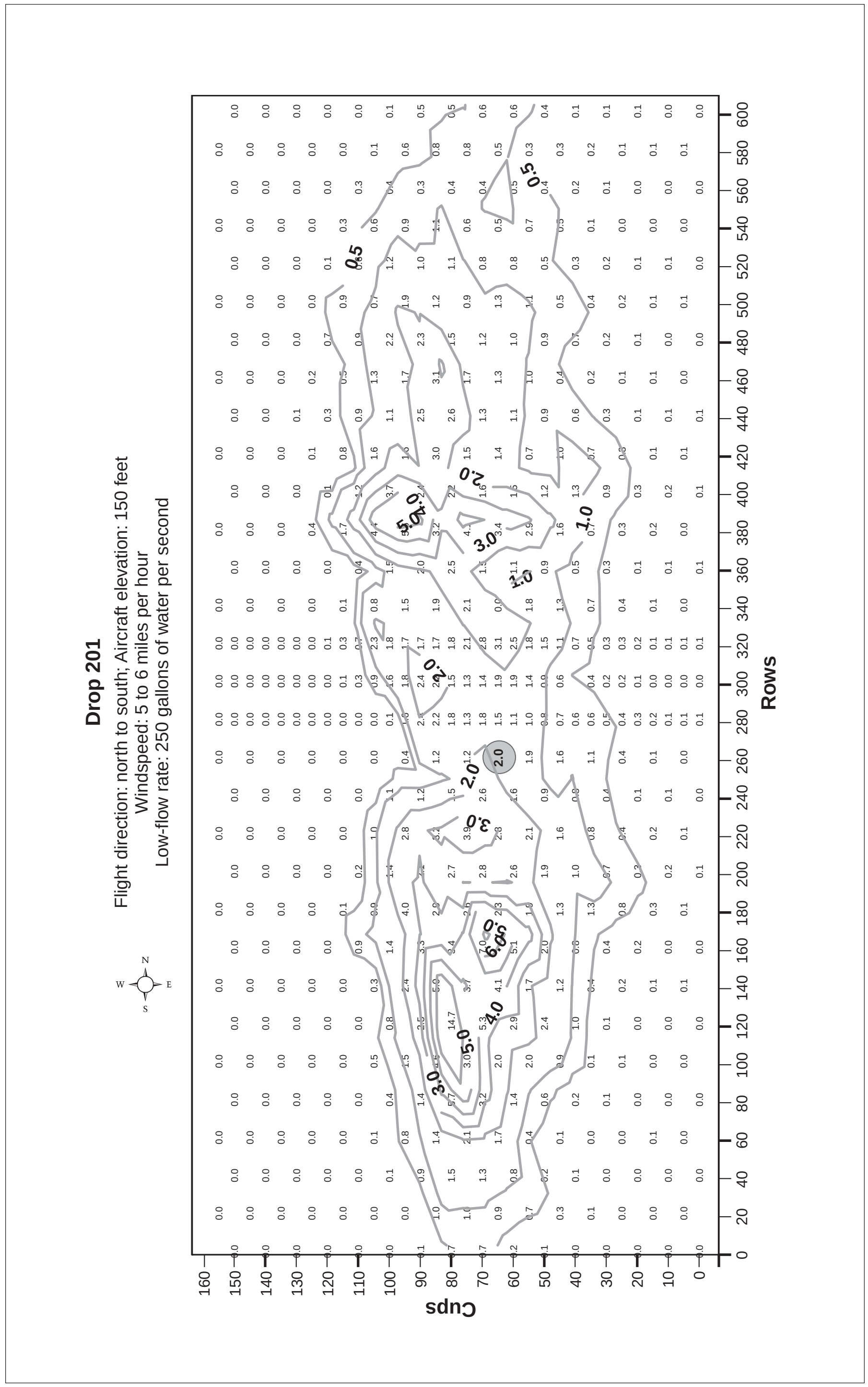

Figure 5—A contour plot showing observed gpc values.

A Comparison of Five Estimation Methods

Creating Contour Plots

The contour plot in figure 5 is generated by computer software using an internal interpolation method to estimate gpc values. The software has been found to be inconsistent. For example, if you look at the value in the small circle (2.0), you'll notice that the 2.0 contour line does not include this cup, as it should.

These plots help determine line lengths, usually expressed in feet, at different coverage levels within a ground pattern. In an attempt to improve the contour plots and line length estimates, five interpolation methods were examined and compared. The five methods are polygonal declustering, triangulation, inverse distance weighting, local sample mean, and ordinary kriging (Kaluzny and others 1998). These five methods are point estimators that use distance (and other factors) as a basis for estimation. When estimating points in space, it is generally assumed that points closer together are more alike than points farther apart. Under this assumption, more weight is given to points that are closer together.

Cross Validation

Cross validation was used to assess the performance of each of the five methods. Cross validation is a technique where the observed sample data are used to make estimations and the estimates are compared to the observed sample data. For example, 543 sample values make up the observed data set in drop 201. One observed value is removed and the remaining 542 values are used to predict a gpc value for the removed value. Once that calculation is complete, the observed value is put back and another observed value is removed. The remaining 542 values are used to predict gpc for the removed value. This process is repeated until a prediction has been made at each of the 543 locations. The result is 543 original observed sample values and 543 estimated values at the same locations. The estimates are compared to the observed data to determine how well the estimation method performed.

Table 1 shows the cross-validation results for drop 201. Triangulation depends on three points to make a prediction, so it cannot predict points in the corners of the grid. For this reason, cross validation produces fewer predicted values when triangulation is used. The observed gpc values at those sites were removed for comparison purposes.

The method that produces estimates that most closely resemble observed data is considered the best. Both triangulation and ordinary kriging have means identical to the observed data. The local sample mean has the least amount of variability, indicating that it smooths the most. Smoothing is similar to averaging. It provides an overview of underlying trends, but information can be lost with excessive smoothing. Examining the five-number summary (minimum, first quartile, median, third quartile, and maximum) gives an idea of the spread of the predicted values compared with the observed. Overall, the predictions have less spread than the true values except when polygonal declustering is used. All of the prediction methods, except for polygonal declustering, smooth data to some extent. Of the other four methods, local sample mean smooths the most and triangulation smooths the least. Triangulation has the highest correlation coefficient, while local sample mean has the lowest.

The second part of table 1 displays the summary statistics for the error of the five-point estimators. Error (also called residual) is the difference between the predicted value and the true value. The

Table 1—Summary statistics (gallons per 100 square feet, gpc) for five point-estimation methods for drop 201. MAE is mean absolute error and MSE is mean squared error. Triangulation cannot predict points in the corners of the grid, which is why the triangulation data only include 537 of the 543 points in the grid.

Summary statistics for five point-estimation methods for drop 201 (gpc)

	TRUE	Triangulation	TRUE	Ordinary kriging	Polygonal declustering	Inverse distance squared	Local sample mean
Mean	0.76	0.76	0.75	0.75	0.76	0.79	0.89
Standard deviation	1.21	1.12	1.20	1.05	1.20	0.59	0.30
Minimum	0.00	0.00	0.00	–0.40	0.00	0.07	0.28
1st quartile	0.01	0.02	0.01	0.04	0.02	0.34	0.69
Median	0.23	0.29	0.22	0.29	0.23	0.59	0.93
3rd quartile	1.15	1.12	1.13	1.19	1.13	1.13	1.06
Maximum	14.66	9.98	14.66	6.74	14.66	3.60	1.66
Correlation		0.92		0.84	0.70	0.80	0.09
n	**537**	**537**	**543**	**543**	**543**	**543**	**543**

Summary statistics for error distribution of point-estimation methods (gpc)

	Triangulation	Ordinary kriging	Polygonal declustering	Inverse distance squared	Local sample mean
Mean	–0.00016	0.00127	–0.00558	–0.034233	–0.13836
Standard deviation	0.465	0.660	0.924	0.813	1.213
Minimum	–6.070	–3.034	–12.140	–1.475	–1.614
1st quartile	–0.015	–0.126	–0.045	–0.387	–0.927
Median	0.000	–0.017	0.000	–0.186	–0.452
3rd quartile	0.080	0.102	0.160	0.080	0.357
Maximum	4.685	10.550	9.370	12.327	13.693
MAE	0.191	0.267	0.377	0.433	0.865
MSE	0.215	0.435	0.852	0.661	1.489
n	**537**	**543**	**543**	**543**	**543**

table of summary statistics for error shows extreme residuals as well as the mean absolute error (MAE) and the mean squared error (MSE). The MSE is the mean of the squared residuals. Residuals are squared to eliminate negative numbers. The MAE is the mean of the absolute value of the residuals. Taking the absolute value removes negative signs to provide a more meaningful statistic. A good prediction method would produce low MAE and MSE values (Isaaks and Srivastava 1989).

The residual means closest to zero were produced by triangulation and ordinary kriging. Triangulation produces the lowest MAE and MSE with ordinary kriging producing the second lowest.

After examining three drops (tables 1, 2, and 3), triangulation appears to perform the best as a prediction method, with ordinary kriging performing second best. These findings indicate that either triangulation or ordinary kriging could be used as a reliable estimator for drop-test data.

Table 2—Summary statistics (gallons per 100 square feet, gpc) for five point-estimation methods for drop 203. MAE is mean absolute error and MSE is mean squared error. Triangulation cannot predict points in the corners of the grid, which is why the triangulation data only include 538 of 544 points in the grid.

Summary statistics for five point-estimation methods for drop 203 (gpc)

	TRUE	Triangulation	TRUE	Ordinary kriging	Polygonal declustering	Inverse distance squared	Local sample mean
Mean	0.74	0.74	0.73	0.73	0.73	0.76	0.79
Standard deviation	1.24	1.13	1.24	1.07	1.24	0.53	0.14
Minimum	0.00	0.00	0.00	-0.52	0.00	0.10	0.39
1st quartile	0.01	0.01	0.00	0.02	0.01	0.36	0.70
Median	0.05	0.11	0.05	0.19	0.05	0.56	0.84
3rd quartile	1.06	1.14	1.04	1.20	1.04	1.08	0.89
Maximum	11.80	7.78	11.80	5.71	11.80	2.65	0.99
Correlation		0.91		0.83	0.66	0.78	0.03
n	**538**	**538**	**544**	**544**	**544**	**544**	**544**

Summary statistics for error distribution of point-estimation methods (gpc)

	Triangulation	Ordinary kriging	Polygonal declustering	Inverse distance squared	Local sample mean
Mean	–0.00002	0.00365	–0.00072	–0.03169	–0.06105
Standard deviation	0.512	0.695	1.018	0.884	1.240
Minimum	–5.040	–2.439	–10.090	–1.216	–0.992
1st quartile	–0.006	–0.155	–0.020	–0.443	–0.852
Median	0.000	–0.014	0.000	–0.271	–0.595
3rd quartile	0.070	0.140	0.140	–0.022	0.289
Maximum	4.025	8.160	8.050	10.083	11.089
MAE	0.225	0.345	0.445	0.537	0.898
MSE	0.262	0.482	1.035	0.780	1.539
n	**538**	**544**	**544**	**544**	**544**

Ordinary Kriging

Ordinary kriging is a weighted linear combination of the observed data. The weights are based on a model called the variogram. The variogram is the variance of the difference between two cups at the distance between the two cups. Through modeling, kriging attempts to minimize the prediction error variance to produce an unbiased estimate (Isaaks and Srivastava 1989). Because the findings showed that triangulation was the best prediction method, ordinary kriging was not used.

Table 3–Summary statistics (gallons per 100 square feet, gpc) for five point-estimation methods for drop 205. MAE is mean absolute error and MSE is mean squared error. Triangulation cannot predict points in the corners of the grid, which is why the triangulation data only include 538 of 544 points in the grid.

Summary statistics for five point-estimation methods for drop 205 (gpc)

	TRUE	Triangulation	TRUE	Ordinary kriging	Polygonal declustering	Inverse distance squared	Local sample mean
Mean	0.78	0.78	0.77	0.77	0.80	0.81	0.81
Standard deviation	1.47	1.38	1.47	1.26	1.46	0.81	0.35
Minimum	0.00	0.00	0.00	–0.64	0.00	0.01	0.08
1st quartile	0.00	0.00	0.00	0.02	0.00	0.21	0.58
Median	0.02	0.06	0.02	0.13	0.08	0.47	0.96
3rd quartile	0.92	0.96	0.91	1.07	0.97	1.25	1.07
Maximum	9.38	7.62	9.38	5.91	9.38	4.11	1.32
Correlation		0.94		0.89	0.76	0.82	0.29
n	**538**	**538**	**544**	**544**	**544**	**544**	**544**

(Continued —>)

Table 3–Continued.

Summary statistics for error distribution of point-estimation methods (gpc)					
	Triangulation	Ordinary kriging	Polygonal declustering	Inverse distance squared	Local sample mean
Mean	–0.00005	0.00631	–0.03029	–0.04323	–0.04037
Standard deviation	0.507	0.679	1.008	0.923	1.408
Minimum	–3.955	–2.950	–7.910	–1.520	–1.174
1st quartile	–0.005	–0.138	–0.063	–0.398	–0.938
Median	0.000	–0.022	0.000	–0.205	–0.372
3rd quartile	0.050	0.052	0.053	–0.025	0.078
Maximum	2.365	6.601	4.730	7.489	8.308
MAE	0.233	0.303	0.463	0.512	0.916
MSE	0.256	0.460	1.015	0.853	1.979
n	**538**	**544**	**544**	**544**	**544**

Triangulation

The Delaunay triangulation method that was used is a weighted linear combination. The result is that closer points receive more weight. Delaunay triangulation uses polygons to determine triangles. In figure 6, the known points are points 1, 2, and 3. The unknown point is V. Point 1 is weighted from area 1, which is the area of the largest triangle. This gives point 1 the most weight, because it is the closest point. Figure 7 illustrates the triangles generated from drop 201.

Triangulation was used to estimate gpc values between observed points. Plotting the estimated points with the observed points created the 10- by 5-foot grid in figure 8.

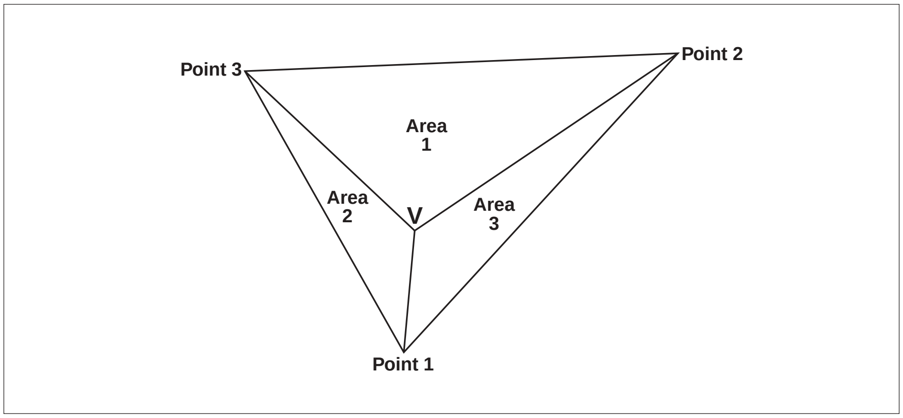

Figure 6—Three triangles constructed to estimate an unknown point.

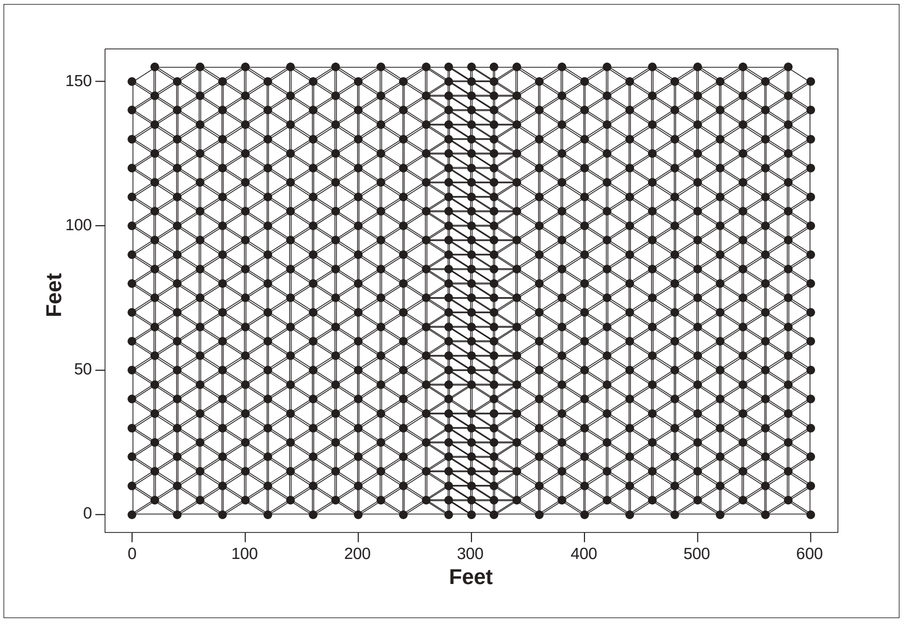

Figure 7—Triangles constructed from sample points.

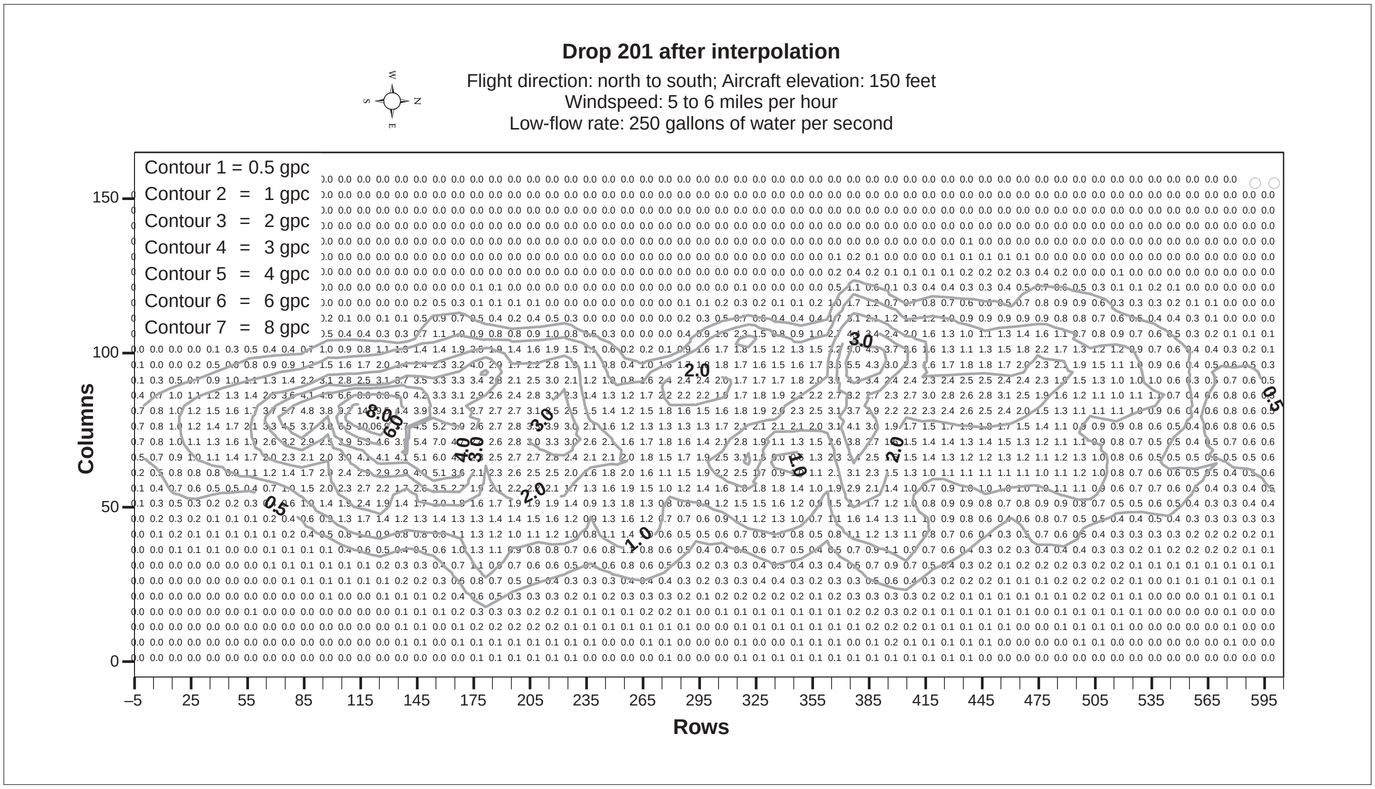

Figure 8—Contour plot redrawn after triangulated gpc values were added to observed gpc values.

The contour plot generated by computer software was overlaid as in figure 8. Once the triangulation was complete, an algorithm was used to calculate line length.

Lengths of retardant line at different coverage levels are calculated by searching crossrange rows for values above a threshold. Line segments begin at the point of the first downrange value above the threshold and end at the point of the last value. The points immediately uprange and downrange of the starting and ending points are used to perform a linear interpolation between the two. This technique allows reporting lengths with accuracy greater than the grid spacing. Lengths for each coverage level of interest are reported as both longest continuous segment and total length. This provides an indication of overall continuity of the line. Uncertainty in coverage level is applied as a single estimated value to all points when checking for the threshold condition. A coverage level value of 3.98 will be at the threshold of 4.00 if the estimated uncertainty is 0.02.

Results

Tabular

Six drop tests numbered 201 to 206 were made over the grid and the results collected on July 12, 2001. Table 4 shows the data collected and the calculated line lengths (longest continuous segment). The actual volumes for drops 201 and 202 were not measured. The measured volumes were obtained using a flow meter when the tank was being filled. There was no way to measure volume dropped or the actual flow rate. For this reason, percent recovery was not included in this table.

The flow rates shown are also a target. No usable flow-rate data were recovered from this drop test. The two numbers for height and speed were provided by the military. For height, one number is from a radar altimeter and the other is from the self-contained navigation system. For speed, one number is from a pitot tube and the other is from the self-contained navigation system. The Forest Service measured height and speed using video analysis.

Table 4—Results from six drop tests over the grid. Wind from 0° would be a head wind. GS is ground speed and RH is relative humidity. GTS-R is a fire retardant.

Drop	Volume in tank (gallons)	Time	Target flow rate (gallons per second)	Height[1] (feet, Forest Service)	Height[2] (feet, military)	GS[1] (knots, Forest Service)	GS[3] (knots, military)	Wind (miles per hour)	Wind direction (0 degrees headwind)	Temp (°F)	Relative humidity (percent)	Fire-fighting chemical	Density (grams per cubic centimeter)
201	900[4]	10:31 a.m.	250	159	170/160	129	133/132	4	45	70	56	H_2O	1.000
202	900[4]	11:08 a.m.	250	149	150/150	128	136/131	4	320	73	52	H_2O	1.000
203	871	11:41 a.m.	250	167	160/160	132	130/138	2	95	74	50	GTS-R	1.094
204	877	12:17 p.m.	250	144	140/150	134	132/124	5	130	78	45	GTS-R	1.097
205	958	2:56 p.m.	500	Missed	160/150	Missed	133/133	0 to 5	Missed	Missed	Missed	GTS-R	1.096
206	821	3:37 p.m.	500	157	160/160	143	138/131	4	130	89	29	GTS-R	1.095

[1] The Forest Service measured height and ground speed using video analysis.
[2] The military measured height using a radar altimeter (first value) and a self-contained navigation system (second value).
[3] The military measured ground speed using a pitot tube (first value) and a self-contained navigation system (second value).
[4] The actual volumes for drops 201 and 202 were not measured. The target volume was 900 gallons.

Length of the retardant line at specific coverage levels (in gallons per 100 square feet, gpc).

Drop	0.5	1	2	3	4	6	8
201	600	521	218	122	105	26	25
202	600	485	375	100	73	14	0
203	600	570	457	185	80	38	21
204	593	524	275	202	118	39	28
205	513	479	373	174	97	68	12
206	511	477	356	190	73	28	0

Graphical

Figures 9 through 14 show contour plots of the six drops after triangulation. The outer, dark contour represents the lightest coverage level of 0.5 gpc. The inner hachured contour indicates one of the heaviest coverage levels, 8 gpc.

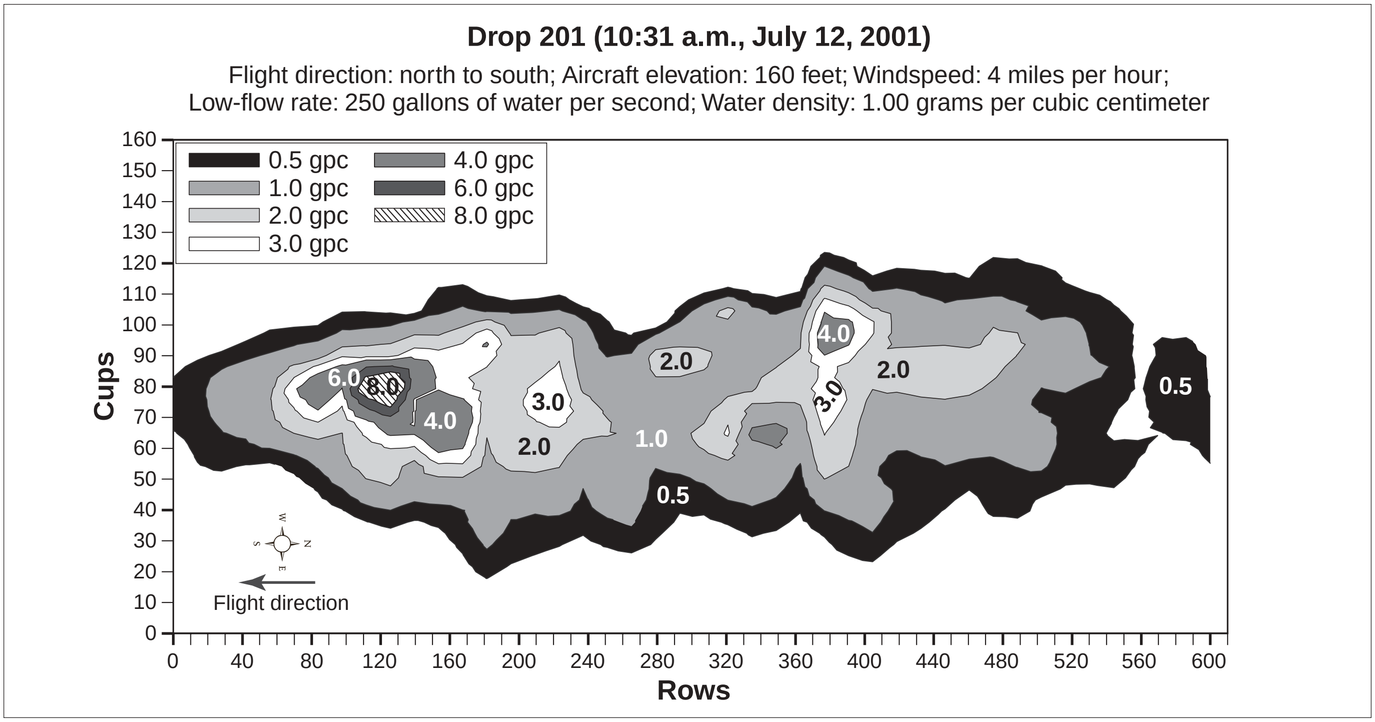

Figure 9—Contour plot of drop 201 after triangulation.

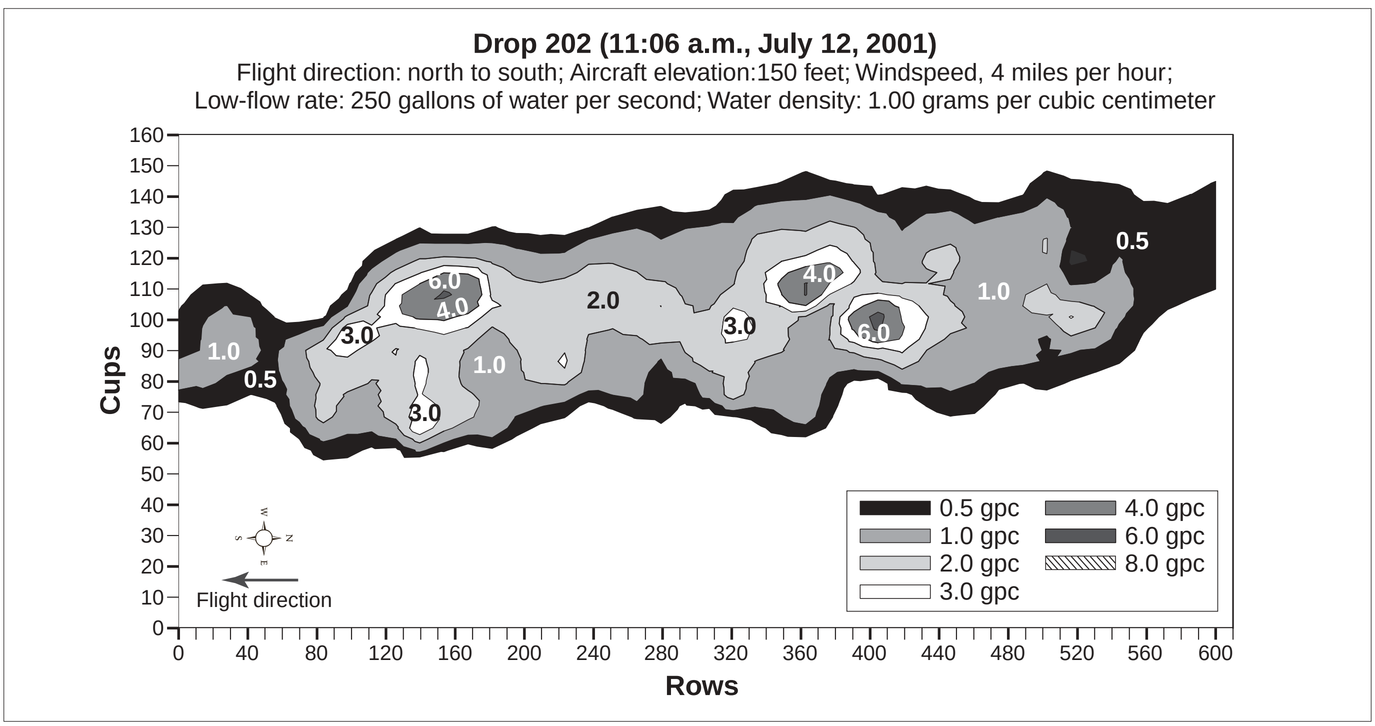

Figure 10—Contour plot of drop 202 after triangulation.

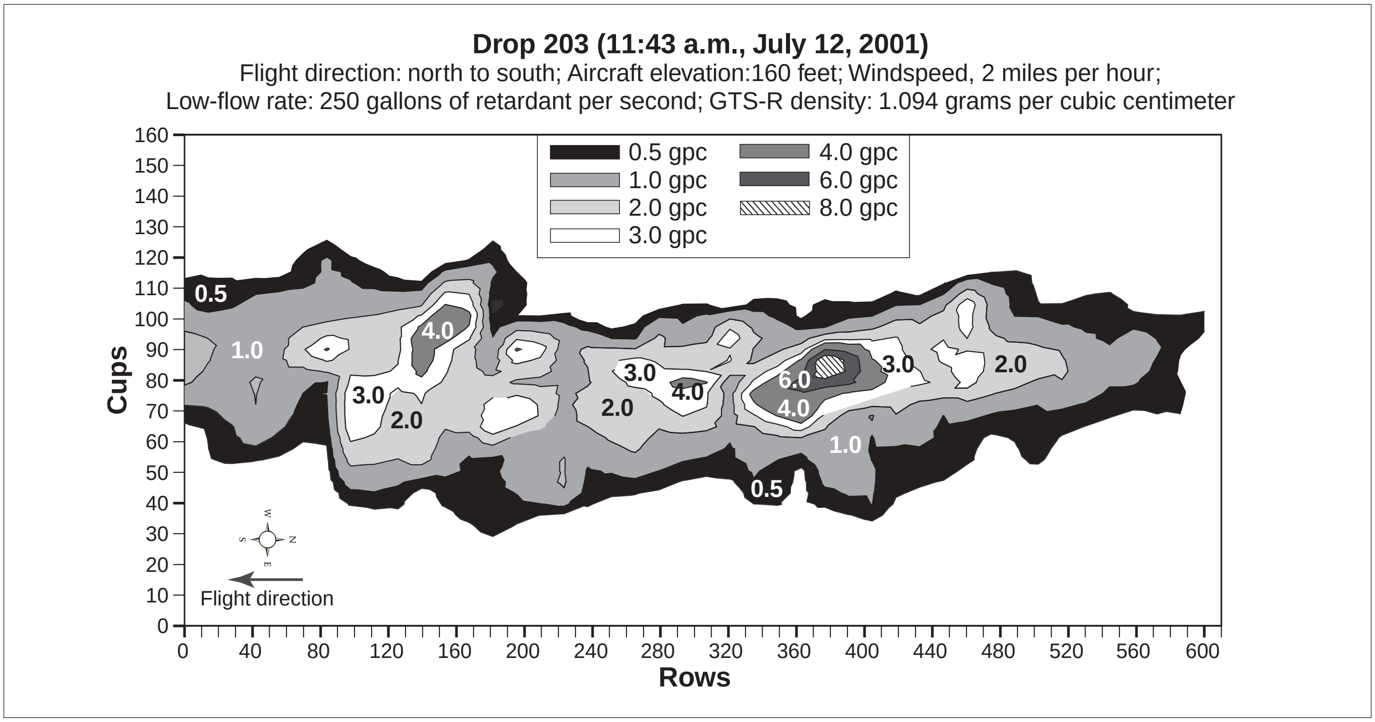

Figure 11—Contour plot of drop 203 after triangulation.

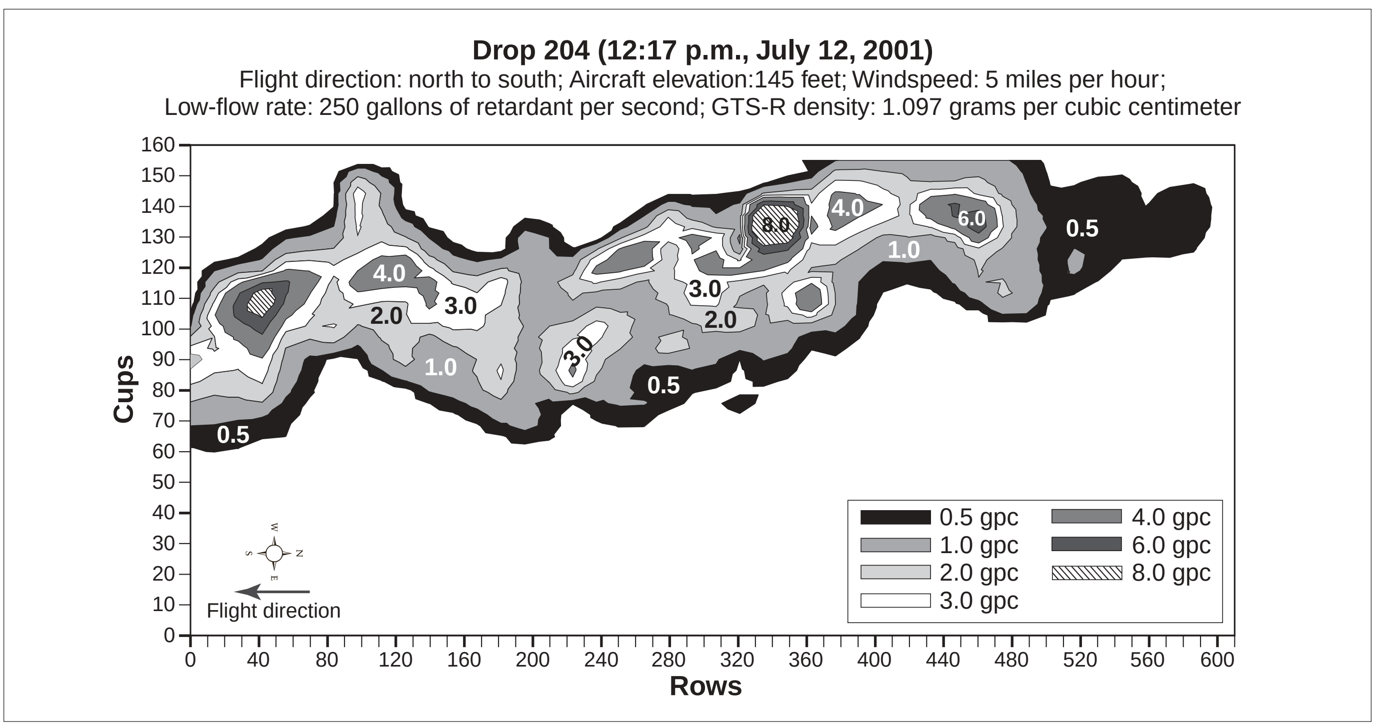

Figure 12—Contour plot of drop 204 after triangulation.

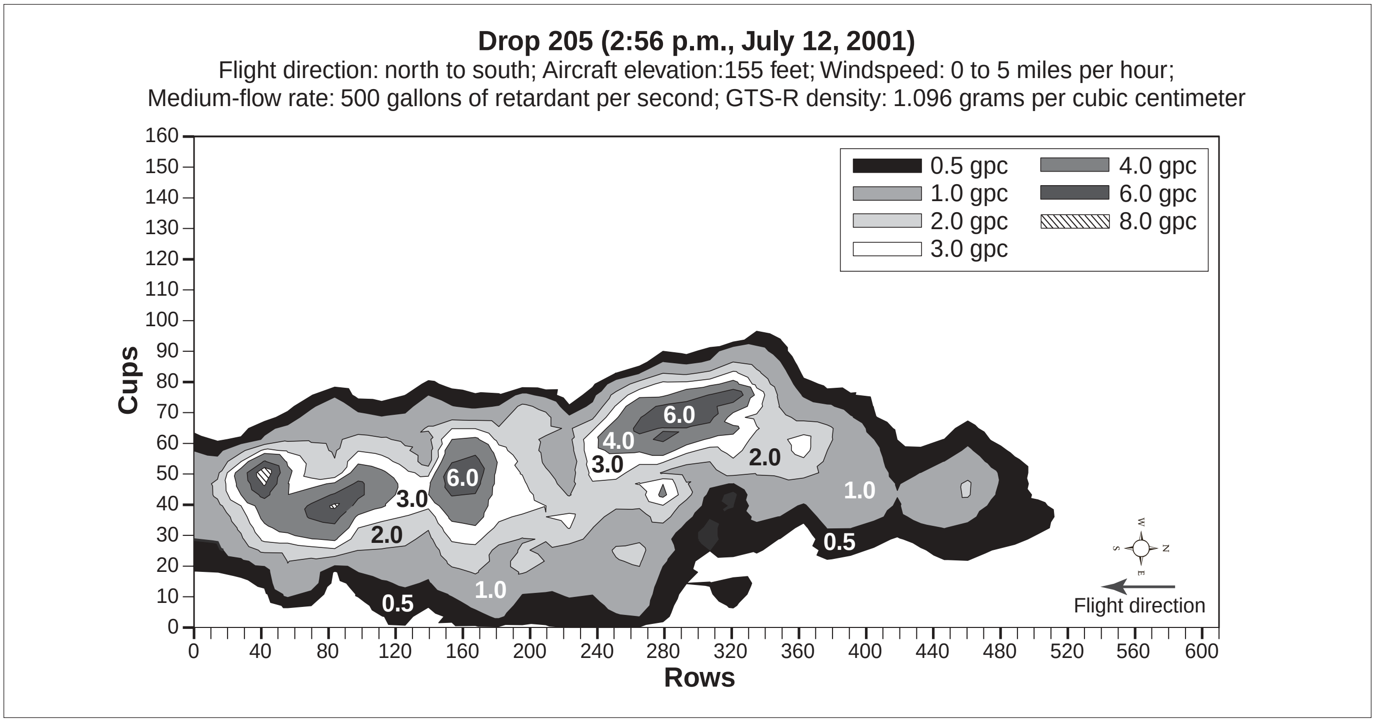

Figure 13—Contour plot of drop 205 after triangulation.

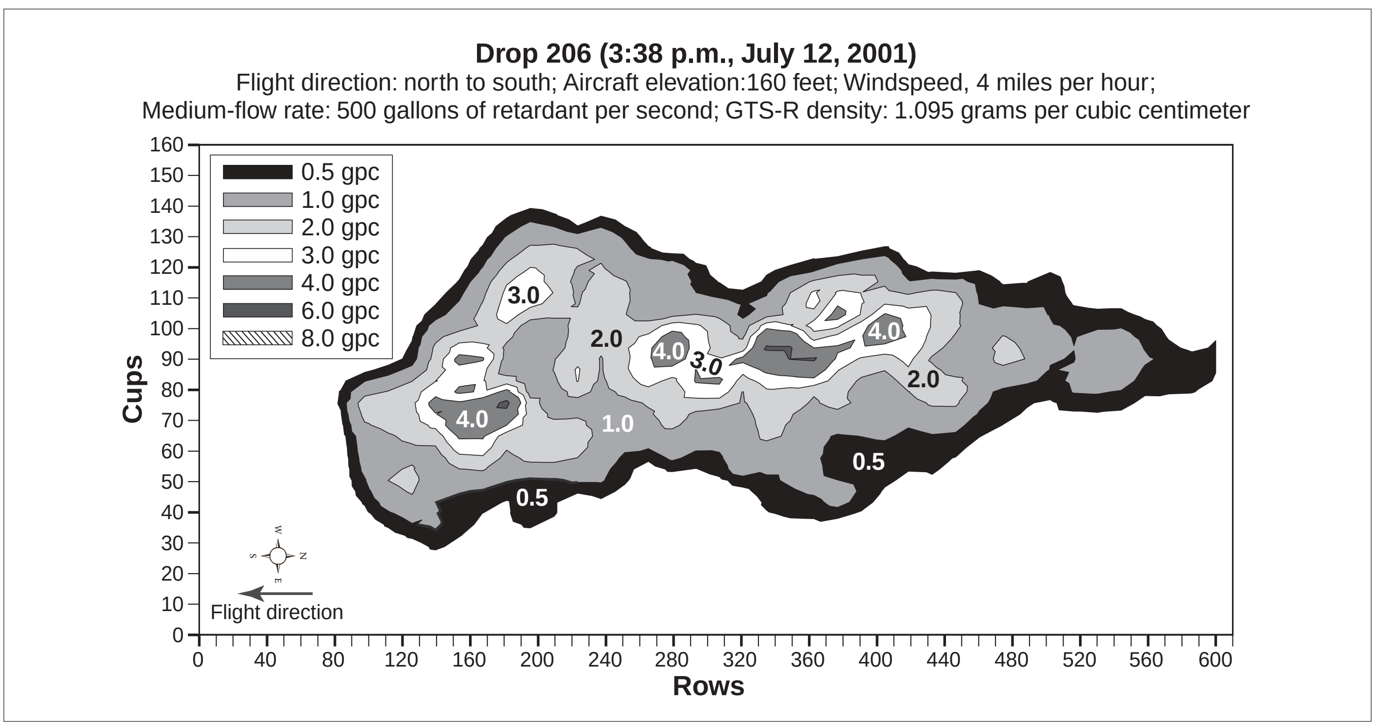

Figure 14—Contour plot of drop 206 after triangulation.

Variability

Replicate Drops

To understand the variability between drops and within the experiment, replicate drops were made where the height, flow rate, speed, volume, and material dropped were constant. The effects of humidity, wind, and temperature were low enough to be assumed to be negligible. Basically, three drop types were tested with two replicates each (table 5).

The winds were equal to or less than 5 miles per hour, the temperature was between 70 and 90 °F, and the relative humidity was between 29 and 56 percent. Lids were placed on all the cups within 10 minutes, minimizing the liquid lost to evaporation.

Analysis of Variance (ANOVA) Results

The ANOVA results (appendix A) demonstrate how differences between factors can be compared. For instance, the mean continuous line length at 0.5 gpc for water is greater than the mean, continuous line length for GTS-R retardant (*p*-value of 0.000388). It also shows that the line length associated with the low flow rate of 250 gallons per second is longer at 0.5 gpc than the length associated with the high flow rate of 500 gallons per second (*p*-value of 0.0000955).

Graphical Results

Figure 15 illustrates some differences within the replicates. The scatterplots show gpc levels by row, presenting a horizontal profile of the drop. Drops 201 and 202 are fairly similar. They have two distinct peaks, tapering off on either end. The line length charts indicate similar line lengths except for 2 and 8 gpc.

Table 5—Summary of the three drop types. The retardant used in the drop tests was GTS-R.

Summary of three drop types

Drop tests	Height (feet)	Speed (knots)	Flow rate (gallons per second)	Volume (gallons)	Material
201 and 202	150 to 160	131	250	900	H_2O
203 and 204	145 to 160	130 to 133	250	871 to 877	GTS-R
205 and 206	155 to 160	133 to 137	500	821 to 958	GTS-R

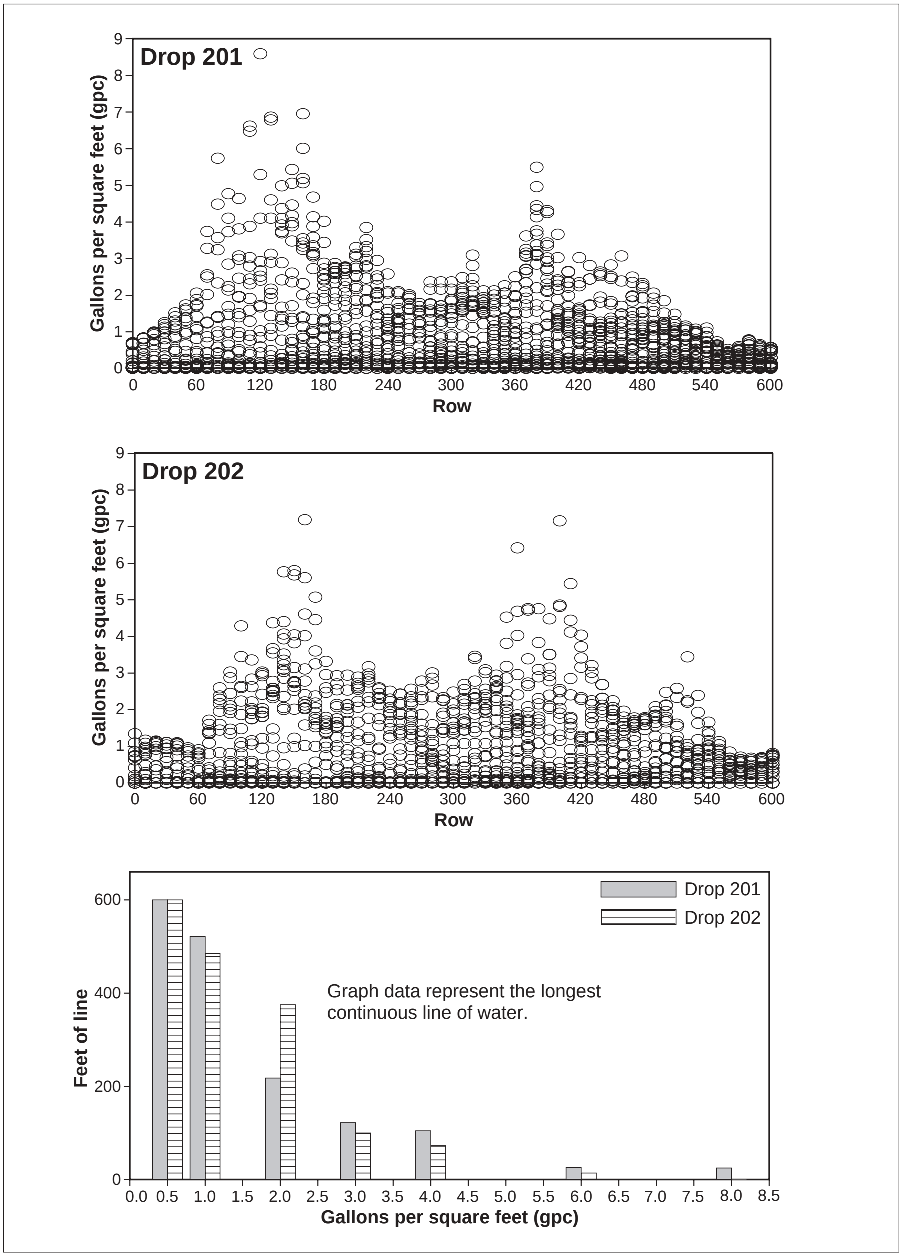

Figure 15—Comparison of line lengths at different gpc levels for drops 201 and 202.

The scatterplots for drops 203 and 204 (figure 16) are not quite as similar as the scatterplots for drops 201 and 202, but the line length chart for drop 203 is similar to that for 204, except for 2 gpc.

Figure 17 shows similar profiles for drops 205 and 206. Drop 205 has more points at coverage levels higher than 4 gpc. The line lengths are similar with discrepancies increasing as the gpc values increase.

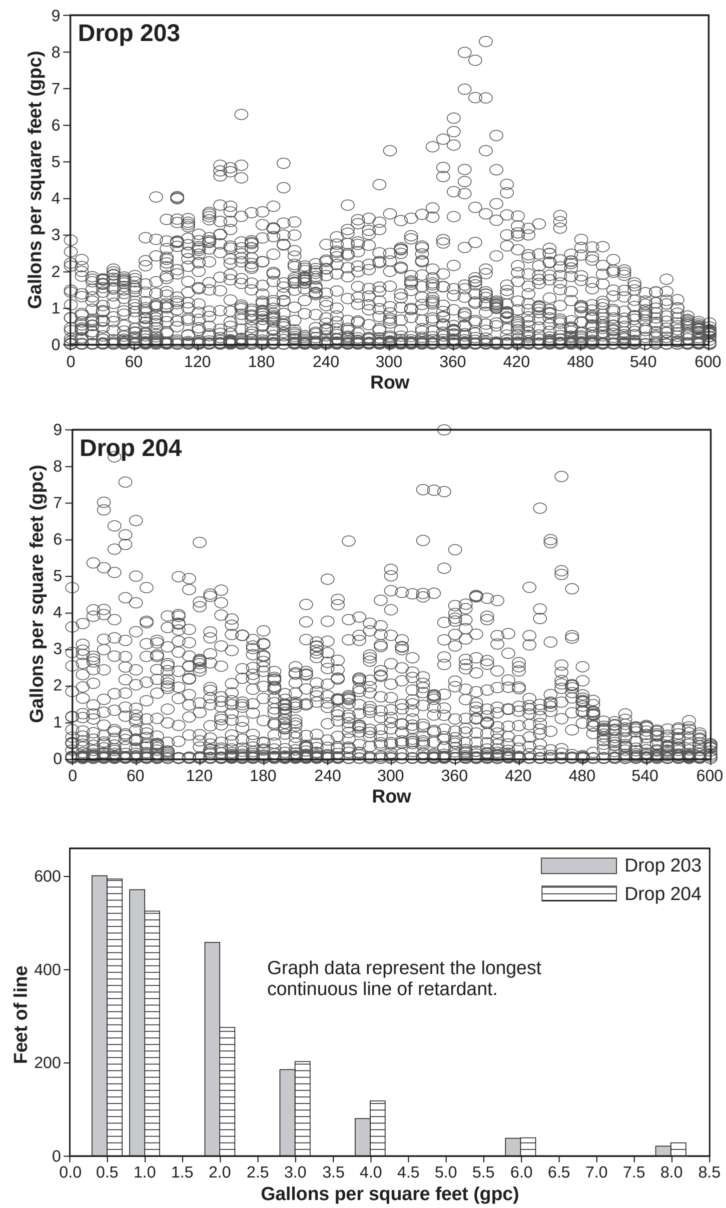

Figure 16—Comparison of line lengths at different gpc levels for drops 203 and 204.

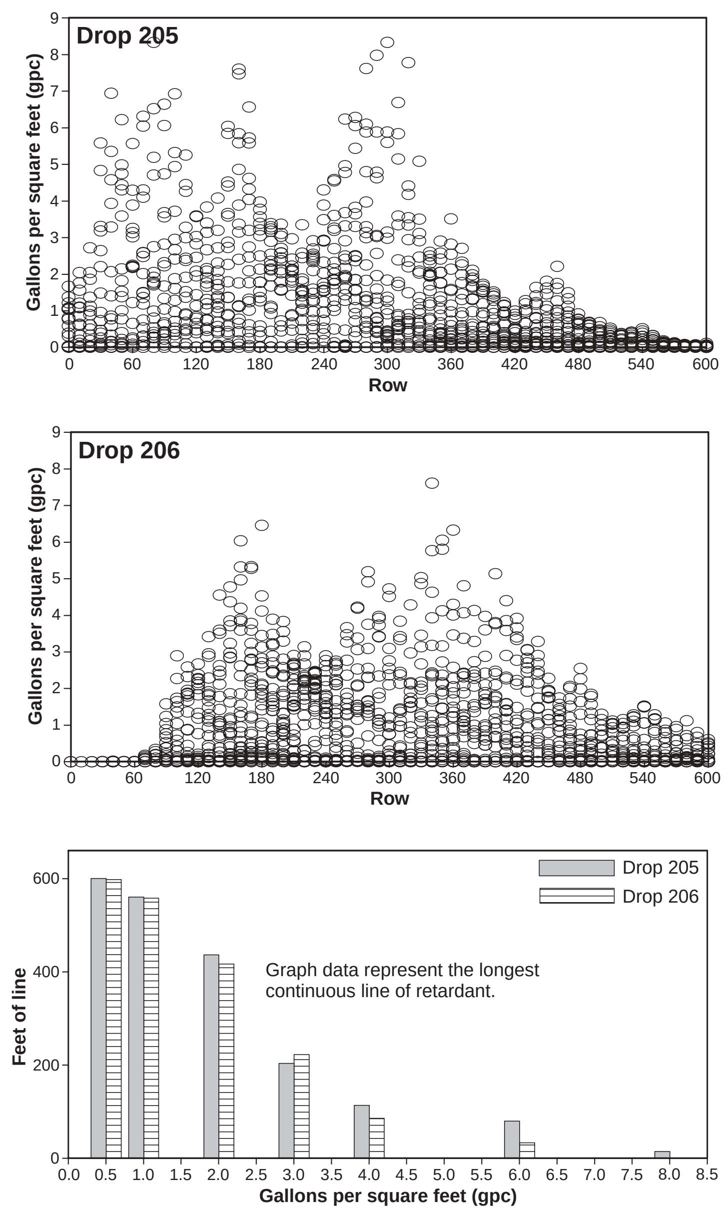

Figure 17—Comparison of line lengths at different gpc levels for drops 205 and 206.

Sampling

Spacing in the Downrange Direction

The original grid design shown in figure 2 was compared to a design with wider spacing through cross validation. Every other row was removed in the observed data set to obtain the 40- by 10-foot spacing (figure 18). This subset of 256 values was used to predict, by triangulation, values at the 288 sites that were removed.

A comparison (table 6) of the predicted gpc values from the 40-foot spacings to the observed values shows a correlation of 0.78, a mean error of –0.01195 gpc and a median error of zero.

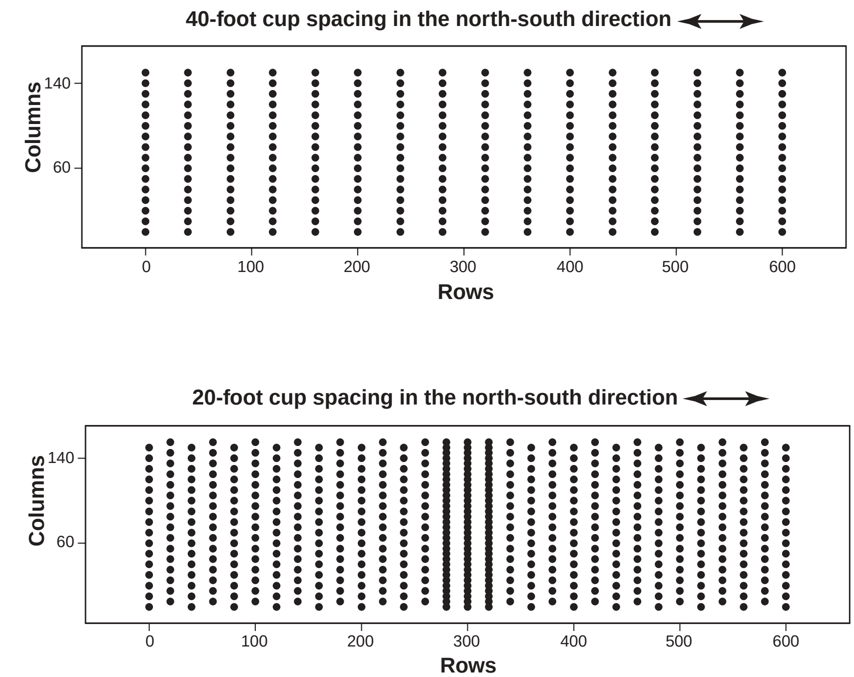

Figure 18—The original 20-foot cup spacing was increased to 40 feet in the north-south direction to evaluate sampling density.

Table 6—Comparison of observed gpc values with predicted gpc values from a 40- by 10-foot spacing. MAE is mean absolute error and MSE is mean squared error.

Drop 201	TRUE	Triangulation (gpc)	Summary statistics for error distribution: Drop 201	Triangulation (gpc)
Mean	0.81	0.82	Mean	–0.01195
Standard deviation	1.08	1.24	Standard deviation	0.793
Minimum	0.00	0.00	Minimum	–7.065
1st quartile	0.02	0.03	1st quartile	–0.110
Median	0.33	0.39	Median	0.000
3rd quartile	1.27	1.16	3rd quartile	0.085
Maximum	5.49	10.80	Maximum	3.535
Correlation		0.78	MAE	0.341
n	**241**	**241**	MSE	0.627
			n	**241**

Spacing in the Crossrange Direction

Spacing in the crossrange direction was also examined (figure 19). The 20- by 5-foot spacing was compared to a 20- by 10-foot spacing.

Figure 20 shows quantile-quantile (QQ) plots. These plots are used to compare the distribution of estimated to observed values. The distributions are equal when x = y or when the data fall on the 45-degree line. In both plots the distributions are similar. The 40-foot QQ plot shows that the wider spacing is not going to pick up unusually high cup weights. In effect, widening the spacing smooths the results. Going from a 20-foot spacing to a 40-foot spacing is probably too great a jump.

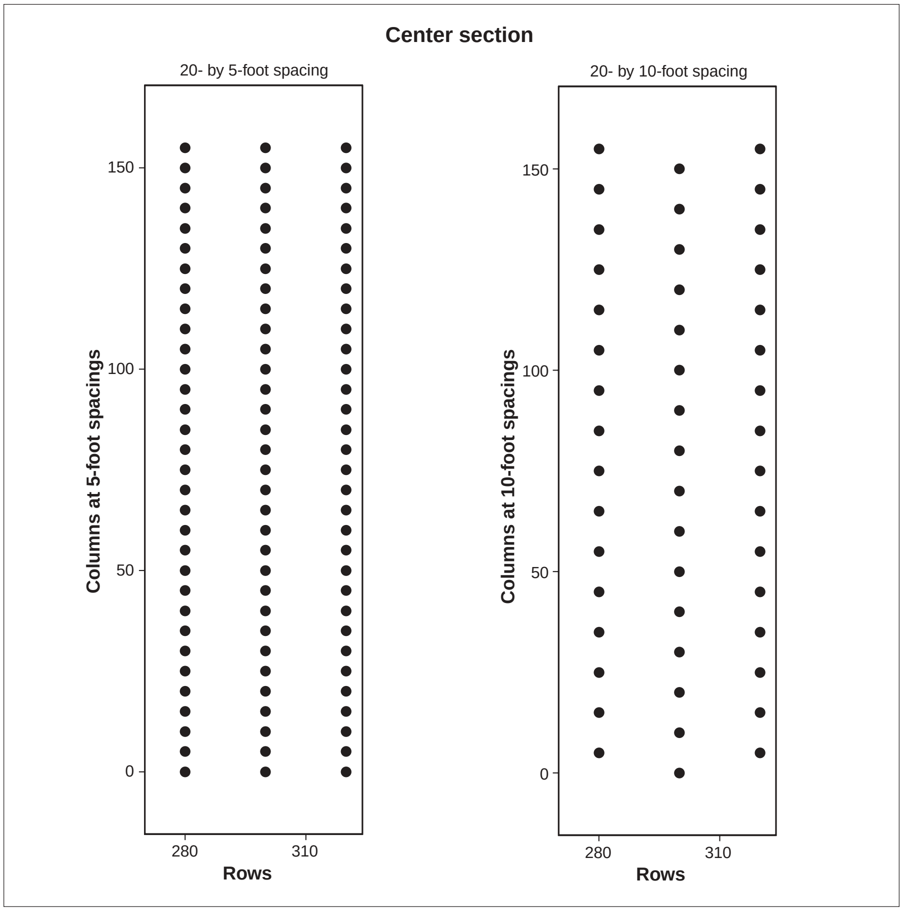

Figure 19—The original 5-foot spacing was increased to 10 feet in the crossrange direction to evaluate sampling density.

However, a 25-foot spacing may be appropriate. More information on the comparison of spacing in the direction of flight can be found in the appendix.

The 10- versus 5-foot comparison shows almost identical distributions. The correlation is high (0.97, table 7) meaning the change in spacing is not producing a big change in the results. Considering the time and costs, the 5-foot spacing would probably not be necessary.

Staggered Spacing

If drop-test data are viewed as spatial data, it is assumed that two cups close together are more likely to have similar values than two that are far apart (Isaaks and Srivastava 1989). To reduce the distance between cups, the stakes can be staggered. Without staggering, the greatest distance between two stakes is 22.36 feet. This distance can be reduced to 20.62 feet with staggering. Even though the difference in distance is less than 10 percent, this small step can help improve accuracy.

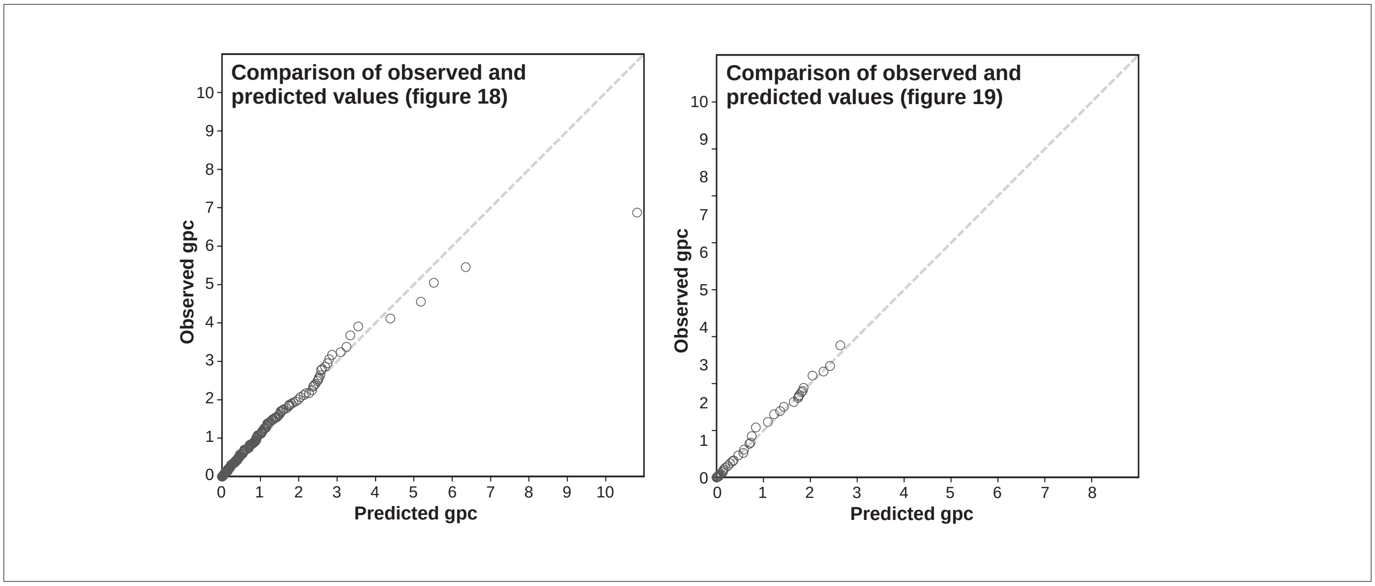

Figure 20—Quantile-quantile (QQ) plots comparing distributions.

Table 7—Comparison of observed gpc (gallons per 100 square feet) values from the center section of cups with 5-foot spacing with predicted gpc values from the center section of cups with 10-foot spacing. MAE is mean absolute error and MSE is mean squared error.

Drop 201	TRUE	Triangulation (gpc)	Summary statistics for error distribution Drop 201	Triangulation (gpc)
Mean	0.74	0.72	Mean	–0.02375
Standard deviation	0.85	0.79	Standard deviation	0.204
Minimum	0.00	0.00	Minimum	–0.965
1st quartile	0.01	0.03	1st quartile	–0.043
Median	0.33	0.39	Median	0.003
3rd quartile	1.50	1.37	3rd quartile	0.048
Maximum	3.09	2.64	Maximum	0.430
Correlation		0.97	MAE	0.110
n	**48**	**48**	MSE	0.041
			n	**48**

Conclusions

Cross validation showed that triangulation and ordinary kriging were the two best estimation methods for drop-test data. Because replicate drops were made, an analysis of variance was performed to determine whether differences in line lengths due to the firefighting chemical and flow rate were significant. Also, cross validation helped to determine whether changing the grid spacing improved the accuracy of the results. Either triangulation or ordinary kriging are the recommended interpolation methods. If the grid spacing is changed, cross validation can be performed again to see which of the two methods is superior.

Replicate drops should be made whenever investigators need to know whether differences in line length are due to changes in factor levels or whether they are just a reflection of the inherent variability in the test. An analysis of variance can determine how much variability is due to changes in factor levels versus the variability inherent in the experiment. Many sources of variability are associated with drop testing. For instance, variability exists in how we measure wind, height, speed, flow rate, and volume. There may also be unknown variation in retardant cloud formation and deposition. The variance associated with predicting gpc values must also be considered. For more information on calculating the prediction variance of a triangulated gpc value, see appendix A.

The investigation into the sampling scheme reveals that increasing the spacing reduces the accuracy of the estimates. This fact must be weighed against the added time and cost of tighter spacing. While going from a 20-foot spacing to a 40-foot spacing is probably too large an increase, the cups could be spaced a little farther apart in the downrange direction without losing much information. In the crossrange direction, the present 10-foot spacing is recommended. A 5-foot spacing wouldn't give that much more accuracy, but it would cost much more in time and money. The appendix examines the predictive capabilities of 20- and 30-foot spacings.

Overall, drop testing gives us a relatively good idea of the performance of an airtanker in a controlled setting. Drop tests would be even more accurate if a permanent grid could be set up. This would allow greater consistency in the experiment.

Because gpc values from a drop test are used to calculate line lengths, it is important to remember that the gpc values are simply estimates. Specifications based on these estimates should probably be expressed in a range that reflects the variability around the estimate.

References

Isaaks, E.H.; Srivastava, R.M. 1989. An introduction to applied geostatistics. New York: Oxford University Press. 561 p.

Kaluzny, Stephen P.; Vega, Silvia C.; Cardoso, Tamre P.; Shelly, Alice A. 1998. S+ spatial stats: user's manual for Windows and UNIX. New York: Springer-Verlag. 327 p.

Ott, Lyman. 1993. An introduction to statistical methods and data analysis. California: Wadsworth Publishing Co. 1,051 p.

Suter, Ann. 2000. Drop testing airtankers: a discussion of the cup-and-grid method. Tech. Rep. 0057-2868-MTDC. Missoula, MT: U.S. Department of Agriculture, Forest Service, Missoula Technology and Development Center. 14 p.

Glossary

Aerial delivery system—A fixed- or rotary-winged aircraft capable of delivering firefighting chemicals.

Aerial drop—A release of firefighting chemical from an aerial delivery system in flight.

Algorithm—A rule for solving a certain type of problem.

Analysis of variance—A statistical technique by which the total variation in a set of data may be reduced to components associated with the possible sources of variation, allowing the relative importance of each source to be assessed.

Contour plot—A graphical picture on which the characteristics of a surface are shown by contour lines. In drop testing, the isopleths join points of equal coverage level on a surface.

Correlation coefficient—A number between –1 and 1 that measures the degree to which two variables are linearly related.

Coverage level—A recommended amount in gallons of retardant applied to a specific area (100 square feet) of surface. Coverage level 2 represents 2 gallons per 100 square feet (gpc).

Crossrange—Perpendicular to the direction of flight.

Cross validation—A method of comparing predicted and observed values.

Data array—Data arranged in a matrix with columns and rows.

Distribution (frequency)—A frequency distribution shows the number of observations falling into each of several ranges of values. Frequency distributions are sometimes displayed as histograms.

Downrange—Parallel to the direction of flight.

Error (residual)—The difference between the predicted value and the observed value.

Firefighting chemicals—Chemical products such as long-term retardants and water enhancers (chemicals containing ingredients designed to alter the physical behavior of water) used in firefighting.

Fire retardant—Any substance, except plain water, that reduces the flammability of fuels or slows their rate of combustion.

Fire suppressant—An agent that extinguishes the flaming and glowing phases of combustion when applied directly to the burning fuel.

First quartile—The 25th percentile. After a set of values has been arranged in order of magnitude, the first quartile is the value that has 75 percent of the values below it.

Flow rate—The rate at which retardant exits a tank or bucket, usually expressed in gallons per second.

GPC—A unit for measuring coverage expressed in gallons per 100 square feet.

Grid—A physical array incorporating containers set in a regular, defined pattern to measure deposition patterns created by the aerial release of fire chemicals.

Ground pattern—The characteristics of ground deposition from aerially delivered liquid.

Histogram—A graph of a frequency distribution table in which rectangles with bases on the horizontal axis are given widths equal to the class intervals. The heights of the rectangles are equal to the corresponding frequencies.

Isopleth—A line drawn on a map through all points having the same numerical value.

Line length—The length, usually expressed in feet, of a ground pattern. Line length is used to relate the length of different coverage levels within a ground pattern.

Linear interpolation—Estimation of a value of a variable between two known values when it is assumed there is uniform change between the two known values.

Mean absolute error (MAE)—The average of the absolute value of a set of residuals.

Mean square error (MSE)—The average of a set of residuals after each one has been squared.

Median—The 50th percentile. After a set of values has been arranged in order of magnitude, the median is the value that has 50 percent of the values below it.

P-value—In a hypothesis test, the probability of observing an outcome “more contradictory to the null hypothesis than the observed sample result” is called the *p*-value (Ott 1993).

QQ plots—Quantile-quantile plots. A graph comparing the distributions of two variables.

Replicates—Duplicates. A replicate drop or a duplicate drop is one that has the same factor levels, specifically, the same height, speed, volume, flow rate, and so forth.

Residual (error)—The difference between the predicted value and the observed value.

Sampling—The process of selecting a sample for testing.

Sampling density—The number of samples in a fixed area.

Tare—The weight of the empty container.

Third quartile—The 75th percentile. After a set of values has been arranged in order of magnitude, the third quartile is the value that has 25 percent of the values below it.

Triangulation—A weighted linear combination used for estimating values at specific locations. The weights depend on the distance and location.

Variability—Data variability refers to the spread of values along the scale of measurement and the extent to which the data are grouped.

Appendix—Details on Cups, Error Variance, and Grid Spacing

Cups

The following cups (table 8) and lids (table 9) were used and their weight recorded.

Table 8—Weight of cups used in the six drop tests.

Weight range (grams)	Average (grams)	Total cups
26.85 to 26.95	26.90	5,000
26.75 to 26.85	26.80	4,000
26.65 to 26.75	26.70	3,000

Table 9—Weight of the lids used in the six drop tests.

Weight range (grams)	Average (grams)	Total lids
16.45 to 16.55	16.50	2,500
16.35 to 16.45	16.40	8,500

- ❏ Average cup weight = [(26.9*5) + (26.8*4) + (26.7*3)]/12 = 26.816667 grams
- ❏ Standard deviation = 0.07993 grams. Variance = 0.0063888049 grams
- ❏ Average lid weight = [(16.4*8.5) + (16.5*2.5)]/11 = 16.422727 grams
- ❏ Standard deviation = 0.04191 grams. Variance = 0.0017564481 grams
- ❏ Tare (average weight of cup and lid) = 43.23939 grams

Combined standard deviation:

$$\sqrt{(0.07993)^2 + (0.04191)^2} = 0.09025$$

The lowest possible cup and lid weight was 43.00 grams and the highest was 43.50. If a cup with retardant in it weighed less than 43.23939 grams, the computer program automatically switched to a tare weight of 43.00 to avoid negative gpc.

At a 99-percent confidence level (CI), the margin of error for the tare weight of
43.2393 grams is ± 0.23249
(2.576*0.09025 = 0.23249)

At a 95 percent CI, the margin of error for the tare weight is ± 0.17689 grams. (1.960*0.09025 = 0.17689 grams)

Error Variance Estimate for GPC

The error variance estimator for triangulated gpc values is:

$$\hat{V}\,(tgpc) = \left[\hat{V}\,(triangulation) + \hat{V}\,(cups)/n_c + \hat{V}\,(lids)/n_l\right] * 0.124087^2$$

Where $\hat{V}$ (triangulation) is the triangulation variance. $\hat{V}$ (cups) is the variance for empty cups, and $\hat{V}$ (lids) is the variance for empty lids. n_c *and* n_l are the number of cups and the number of lids, respectively. 0.124087 is a constant that converts grams of retardant with density 1.095 grams per milliliter into gpc.

Mean square error (MSE) is an estimate of the triangulation variance. The three MSEs are 0.215, 0.262, and 0.256, which is an average MSE of 0.244.

$$0.003804 = [0.244 + 0.000000532 + 0.0000001597] * 0.124087^2$$

Variance around triangulated gpc = 0.0038. Standard deviation around triangulated gpc = 0.0616.

Analysis of Variance

An example of an analysis of variance (ANOVA) model (figure 21).

```
***  Analysis of Variance Model ***

Short Output:
Cell:
        aov(formula – Continuous  ~ Ret + FlowRate, data – LineLengths05, qr – T,
n.action – na.exclude)

Terms:
                      Ret   FlowRate  Residuals
Sum of Squares    2790.75    7140.25      26.50
Deg. of Freedom         1          1          3

Residual standard error: 2.972092
Estimated effects may be unbalanced

              Df   Sum of Sq    Mean Sq    F Value          Pr(F)
      Ret      1     2790.75   2790.750   315.9340   0.0003882841
 FlowRate      1     7140.25   7140.250   808.3302   0.0000955337
Residuals      3       26.50      8.833

Tables of means
Grand mean
 569.5

 Ret
        GTSR      Water
      554.25     600.00
rep     4.00       2.00

FlowRate
        High        Low
      527.25     590.63
rep     2.00       4.00
```

Figure 21—Analysis of variance results.

Grid Spacing

Gpc data collected from a previous drop test, which used a grid with cups in a 10- by 10-foot spacing, were used for the following comparisons. The 10- by 10-foot spacing provided a data set that could be divided into subsets for cross validation. Two subsets were created with the points in a 20- by 10-foot spacing, and three subsets were created with the points in a 30- by 10-foot spacing. Figures 22 and 23 show examples of these subsets. Tables 10 to 14 display the cross validation tabular results. Figures 24 and 25 display the QQ-plots comparing distributions.

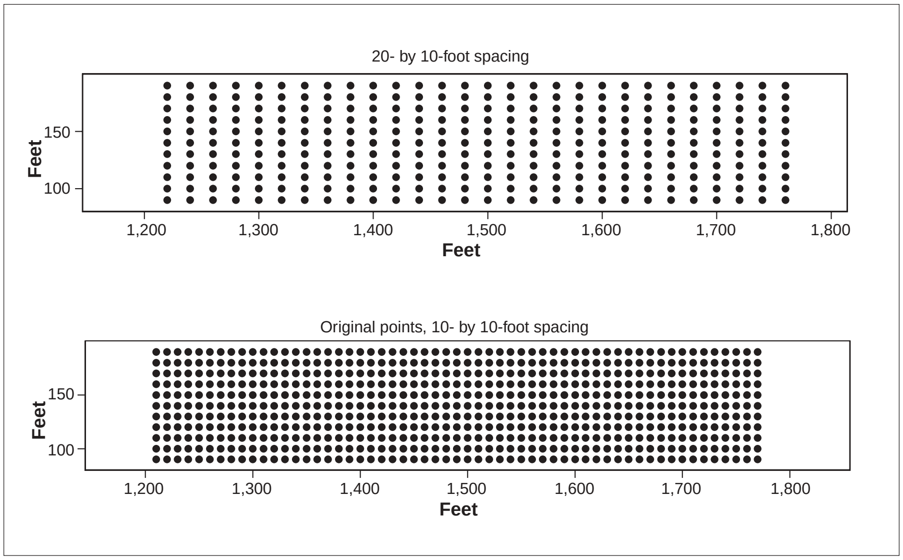

Figure 22—The original 10-foot spacing was increased to 20 feet to evaluate sampling density.

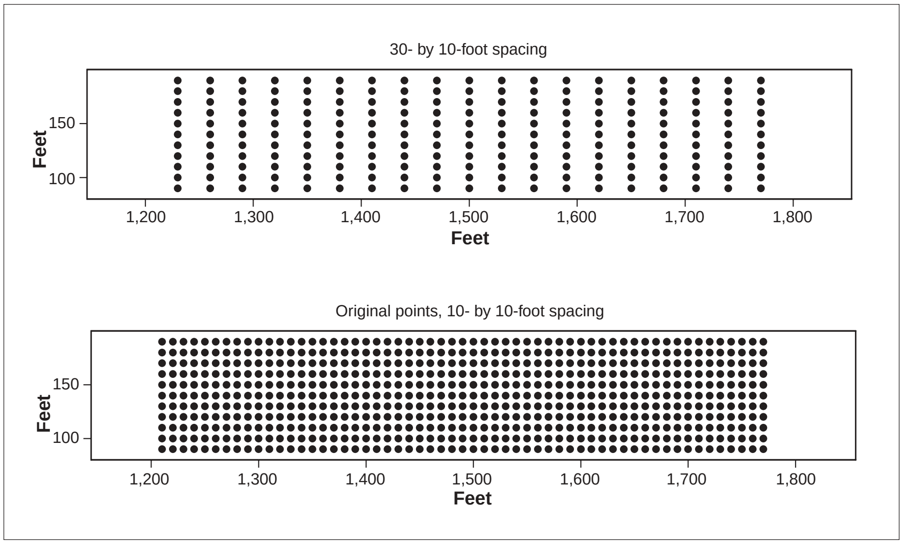

Figure 23—The original 10-foot spacing was increased to 30 feet to evaluate sampling density.

Table 10—First comparison of observed gpc (gallons per hundred square feet) values from a 10- by 10-foot spacing with predicted values from a 20- by 10-foot spacing. MAE is mean absolute error and MSE is mean squared error.

	TRUE	Triangulation (gpc)	Summary statistics for error distribution	Triangulation (gpc)
Mean	1.534	1.509	Mean	0.02483
Standard deviation	0.905	0.819	Standard deviation	0.386
Minimum	0.037	0.068	Minimum	–1.856
1st quartile	0.950	0.962	1st quartile	–0.144
Median	1.380	1.419	Median	–0.021
3rd quartile	2.008	2.052	3rd quartile	0.150
Maximum	5.081	5.428	Maximum	2.206
Correlation		0.905	MAE	0.246
n	**319**	**319**	MSE	0.149
			n	**319**

Table 11—Second comparison of observed gpc (gallons per 100 square feet) values from a 10- by 10-foot spacing with predicted values from a 20- by 10-foot spacing. MAE is mean absolute error and MSE is mean squared error.

	TRUE	Triangulation (gpc)	Summary statistics for error distribution	Triangulation (gpc)
Mean	1.524	1.550	Mean	–0.02628
Standard deviation	0.922	0.805	Standard deviation	0.416
Minimum	0.040	0.093	Minimum	–1.044
1st quartile	0.909	0.994	1st quartile	–0.201
Median	1.412	1.411	Median	–0.035
3rd quartile	2.012	2.071	3rd quartile	0.096
Maximum	8.183	4.060	Maximum	4.122
Correlation		0.893	MAE	0.241
n	**308**	**308**	MSE	0.173
			n	**308**

Table 12—First comparison of observed gpc (gallons per 100 square feet) values from a 10- by 10-foot spacing with predicted values from a 30- by 10-foot spacing. MAE is mean absolute error and MSE is mean squared error.

	TRUE	Triangulation (gpc)	Summary statistics for error distribution	Triangulation (gpc)
Mean	1.553	1.534	Mean	0.01857
Standard deviation	0.962	0.736	Standard deviation	0.574
Minimum	0.051	0.053	Minimum	–1.337
1st quartile	0.947	1.044	1st quartile	–0.217
Median	1.400	1.481	Median	–0.020
3rd quartile	1.978	2.069	3rd quartile	0.147
Maximum	8.183	3.676	Maximum	5.265
Correlation		0.803	MAE	0.334
n	**396**	**490**	MSE	0.329
			n	**396**

Table 13–Second comparison of observed gpc (gallons per 100 square feet) values from a 10- by 10-foot spacing with predicted values from a 30- by 10-foot spacing. MAE is mean absolute error and MSE is mean squared error.

	TRUE	Triangulation (gpc)	Summary statistics for error distribution	Triangulation (gpc)
Mean	1.520	1.534	Mean	–0.01385
Standard deviation	0.904	0.787	Standard deviation	0.450
Minimum	0.037	0.081	Minimum	–1.426
1st quartile	0.914	0.980	1st quartile	–0.218
Median	1.374	1.468	Median	–0.017
3rd quartile	2.046	2.018	3rd quartile	0.137
Maximum	8.183	3.956	Maximum	4.259
Correlation		0.867	MAE	0.279
n	**418**	**418**	MSE	0.202
			n	**418**

Table 14–Third comparison of observed gpc (gallons per 100 square feet) values from a 10- by 10-foot spacing with predicted values from a 30- by 10-foot spacing. MAE is mean absolute error and MSE is mean squared error.

	TRUE	Triangulation (gpc)	Summary statistics for error distribution	Triangulation (gpc)
Mean	1.531	1.538	Mean	–0.00609
Standard deviation	0.881	0.846	Standard deviation	0.480
Minimum	0.037	0.102	Minimum	–3.092
1st quartile	0.916	0.998	1st quartile	–0.191
Median	1.400	1.387	Median	–0.028
3rd quartile	2.027	2.060	3rd quartile	0.171
Maximum	5.081	6.132	Maximum	2.545
Correlation		0.846	MAE	1.531
n	**396**	**396**	MSE	0.230
			n	**396**

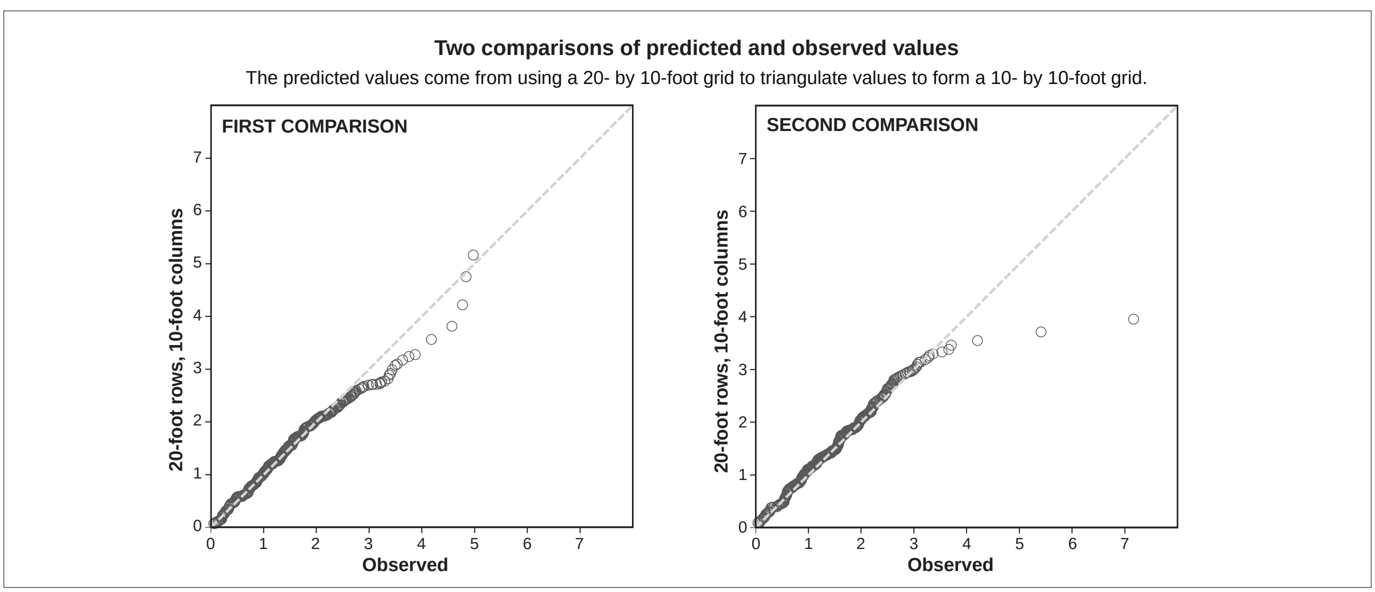

Figure 24—Quantile-quantile (QQ) plots comparing distributions.

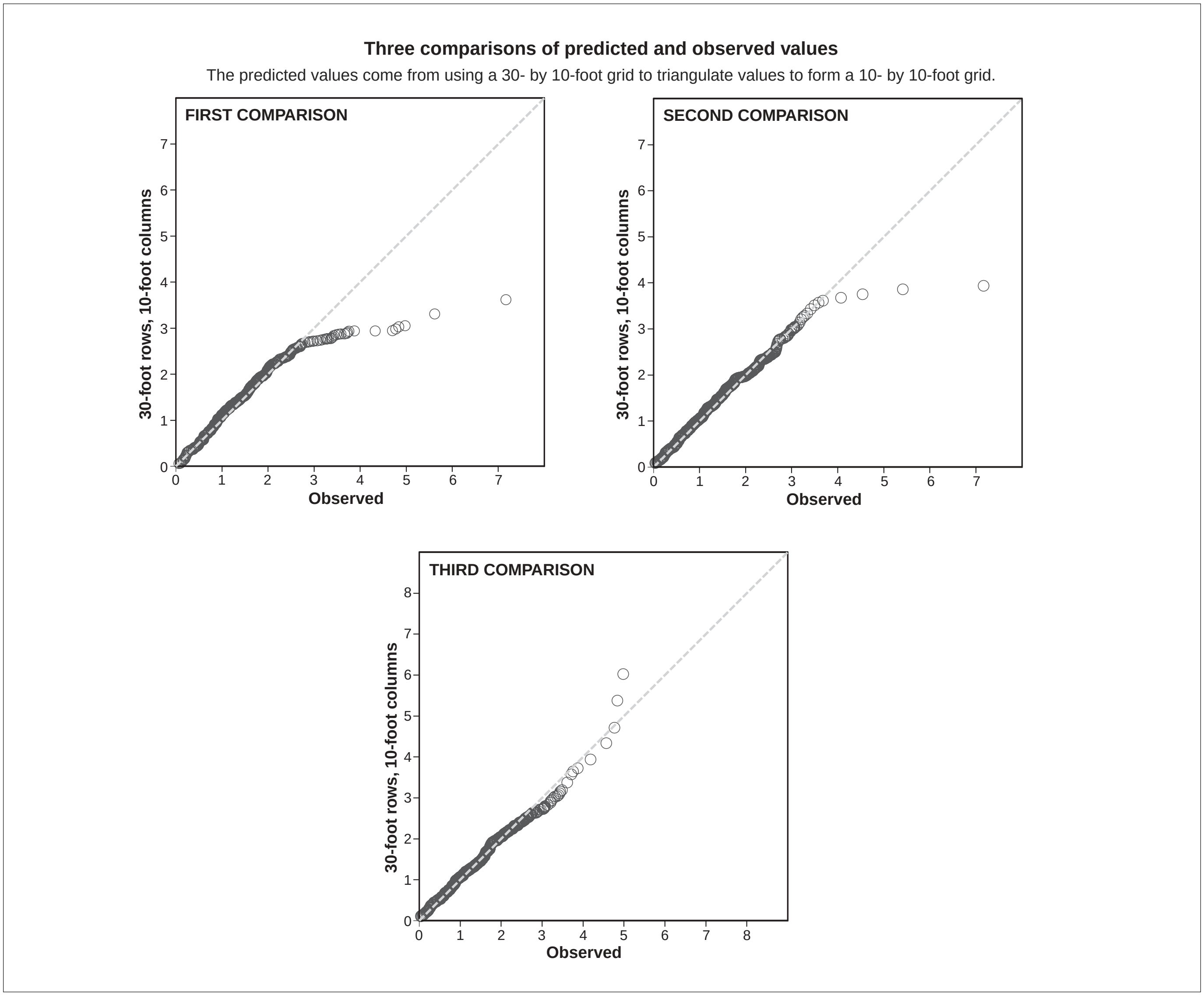

Figure 25—Quantile-quantile (QQ) plots comparing distributions.

Notes

Notes

Notes

About the Author

Ann Suter is a statistician for the Missoula Technology and Development Center's Wildland Fire Chemical Systems group. She joined the Forest Service in 1997 after serving 2 years as a Peace Corps volunteer in Jamaica, where she worked on reforestation and soil erosion control. She holds a master's degree in international development from the American University.

Library Card

Suter, Ann. 2002. Estimating methods, variability, and sampling for drop-test data. 0257-2826-MTDC. Missoula, MT: U.S. Department of Agriculture, Forest Service, Missoula Technology and Development Center. 30 p.

Discusses the testing process the Forest Service has used for the past six decades to analyze the ground patterns made by aerial drops of fire retardants or suppressants. The process involves dropping firefighting chemicals from an airtanker flying over open cups arranged in a regularly spaced grid. This report uses data collected from six airtanker drops to investigate estimation methods, variability, and sampling. Five estimation methods were compared: triangulation, ordinary kriging, polygonal declustering, inverse distance squared, and local sample mean. Cross validation showed that triangulation and ordinary kriging were the two best estimation methods for drop-test data. Replicate drops should be made whenever investigators need to know whether differences in line length are due to changes in factor levels or whether they are just a reflection of the inherent variability in the test. Investigation of the sampling scheme shows that increasing the spacing of the cups reduces the accuracy of the estimates. In the crossrange direction (perpendicular to the flight path), a 10-foot spacing is recommended. In the downrange direction (in the direction of the flight path), spacing could be increased slightly from the present 20 feet without seriously affecting the accuracy of the estimates.

Keywords: airtankers, coverage levels, cross validation, ground pattern, history, kriging, line length, sampling, spatial statistics, triangulation, variance

---

Additional single copies of this document may be ordered from:
USDA Forest Service, MTDC
5786 Hwy. 10 West
Missoula, MT 59808-9361
Phone: 406–329–3978
Fax: 406–329–3719
E-mail: *wo_mtdc_pubs@fs.fed.us*

An electronic copy of this report will be available on the Internet at:
http://fsweb.mtdc.wo.fs.fed.us/cgi-bin/enter.pl?link=pubs/htmlpubs/htm02572826/

For further technical information, contact Ann Suter at MTDC.
Phone: 406–329–4815
Fax: 406–329–4763
E-mail: *asuter@fs.fed.us*