Hertzler Systems Inc.®
Turning Data into Knowledge

Statistical Qualifications Testing

Product QA/S GainSeeker® SPC
Version 5.200 and above
Date of origin September 30, 1997
Date of last edit October 14, 2021

This Statistical Qualifications test document is part of Hertzler Systems Inc. Master Validation and Verification Plan. Specifically it is the essential component of qualification testing for this product. QA/S GainSeeker® SPC software is a collection of programs which allow users to collect, manage and analyze statistical process control data. At the heart of these programs is a collection of statistical calculations. According to ANSI/IEEE 1012-1986, qualification testing is formal testing designed to "demonstrate to the customer that the software meets its specified requirements." This Qualification Plan, therefore, will demonstrate to the customer that the software meets its specified requirements in the area of *statistical calculations.*

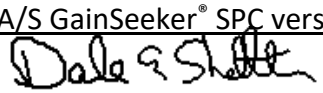
Hertzler established a baseline for statistical verification during the release for GainSeeker® SPC version 5.2. All statistical values produced by the software were identified and defined, and divided into five groups, as follows:

1. Seven foundational values that form the core for most calculations performed by the program. These statistics are verified by an internal study conducted by Hertzler Systems Inc. staff. The statistics are then confirmed by using another statistical software package. The software package we use to confirm these seven values is MINITAB.
2. Twenty advanced statistical calculations that are very difficult to calculate. Most of these are used in calculations of normality. These are based on complex equations and they would be very difficult for the average user to validate. An independent statistical consultant validates these values.
3. Six advanced statistical calculations that are somewhat difficult to calculate. Although it may be somewhat difficult for the average user to validate, no other calculations are based on these statistical values. Comparing values with those derived from another statistical software package validates these values. The software package we use to validate these six values is MINITAB.
4. Statistical values that are derived from the seven foundational values, or the advanced calculations. These values are simplistic calculations. In most cases, error in these calculations would be obvious to both Hertzler staff and the customer. Because they are simple to verify, we have made no attempt to further verify these calculations. If customers wish to verify these for their own satisfaction, they will find the task time consuming but simple. With each revision, we compare these values with those calculated in previous versions to verify that the values have not changed.
5. Non-statistical values reported by the QA/S GainSeeker® SPC statistical engine. These values can be simply reporting settings or pass/fail conditions for other statistical values.

All five of these groups are identified in Appendix A.

This statistical list and groupings are referenced during the Requirements Phase of each GainSeeker® SPC development cycle. If product requirements/specifications are introduced with potential effect to the calculations portion of the code, the category of change is measured against which categories of statistics are affected in order to design a Statistical Qualifications Test Plan, which is then executed for that product release.

Product and Version: QA/S GainSeeker® SPC version 9.4.1

Signed: 

Title: Vice President of Product Development

Date: October 14, 2021

Appendix A

Group 1: Seven foundational values validated by internal Hertzler Systems Inc. study and confirmed by another statistical software package

Internal Reference Number	Value	Description	Values used
23	Mean	Summation, Division	Count
24	R-Bar	Summation, Division	Count
100	LCLx	Subtraction, Multiplication	Mean, SD
101	UCLx	Addition, Multiplication	Mean, SD
103	UCLr	Addition, Multiplication	Mean, SD
121	Cpk	Division	Zl, Zu
125	Cp	Subtraction, Division	6SD

Group 2: Twenty advanced statistical values tested by outside contractor

Internal Reference Number	Value	Values used
79	% below spec predicted	Derived from distribution equation
80	% above spec predicted	Derived from distribution equation
81	% out of spec predicted	Derived from distribution equation
82	% in spec predicted	Derived from distribution equation
110	A1	Result of distribution equation
111	A2	Result of distribution equation
112	M1	Result of distribution equation
113	M2	Result of distribution equation
114	Non-normal area	Derived from distribution equation
117	Distribution	Derived from distribution equation
118	Spread used	Derived from distribution equation
119	Skewness	Sum of Differences Squared, Sum of Differences Cubed
120	Kurtosis	Sum of Differences to the 4th, Sum of Differences Squared
134	Zl non-normal	Derived from distribution equation
135	Zu non-normal	Derived from distribution equation
139	Chi-squared	Sum of Differences, Sort
140	Trend correlation coefficient	Sums of Squares, Sums of Differences
142	Trend slope	Sums of Squares, Sums of Differences
143	Trend Y-intercept	Sums, Slope Sub, Count
172	Kruskal-Wallis value	Counts, Sort, Sums

Group 3: Advanced statistical values validated by another statistical software package

Internal Reference Number	Value	Values used
141	Trend line	Sums, Sums of Squares, Sums of Products
150	Trend UCLx	Sums, Sums of Squares, Sums of Products
151	Trend LCLx	Sums, Sums of Squares, Sums of Products

Group 4: Simplistic statistical values that do not normally undergo specific testing

Internal Reference Number	Value
3	Total values
4	Count included values
5	Count excluded values
6	Total subgroups
7	Count included subgroups
8	Count excluded subgroups
51	Maximum subgroup
52	Minimum subgroup
53	Maximum range
54	Minimum range
55	Maximum value
56	Minimum value
59	Maximum subgroup included
60	Minimum subgroup included
61	Median subgroup included
62	Maximum range included
63	Minimum range included
64	Median range included
65	Maximum value included
66	Minimum value included
67	Median
68	% above spec
69	% below spec
70	% in spec
71	% out of spec
160	Mode
194	Count out of spec
195	Count in spec
210	Count anchor points
211	Count above spec
212	Count below spec
213	Count above target
214	% above target
215	Count below target
216	% below target
229	First Quartile
230	Third Quartile
233	Bypassed values
234	Bypassed subgroups
235	Count subgroups not bypassed
236	Count values not bypassed
237	Maximum subgroup not bypassed
238	Minimum subgroup not bypassed
239	Maximum range not bypassed
240	Minimum range not bypassed
241	Maximum value not bypassed
242	Minimum value not bypassed
249	% above Ind Limit
250	% below Ind Limit
251	% out of Ind Limit
252	% in Ind Limit

253	% above lower Ind Limit
254	% below upper Ind Limit
255	% above lower spec
256	% below upper spec
257	% above lower gate
258	% below upper gate
264	Mean – nominal
283	Count subgroups with all missing values
284	Count incomplete subgroup values
285	No data found
287	Count above spec predicted
288	Count below spec predicted
289	Count in spec predicted
290	Count out of spec predicted
291	Count above ind. limit
292	Count below ind. Limit
293	Count in ind. Limit
294	Count out of ind. limit
295	Count above gate
296	Count below gate
297	Count in gate
298	Count out of gate
299	Count subgroups within 1 sigma
300	Count subgroups within 3 sigma
302	Insufficient non-normal resolution
314	% above range gate
315	Count above range gate
319	Trend crosses in

Internal Reference Number	Value	Description	Values used
26	SD (factors)	Sum of Squares, Square Root	Mean
27	SD (sample)	Sum of Squares, Square Root	Mean
29	SD (Range)	Sum of Squares, Square Root	R-Bar sub
32	Nominal sub gate	Addition, Division	
33	Tolerance sub gate	Addition, Division	
36	Nominal Ind Limit	Addition, Division	
37	Tolerance Ind Limit	Addition, Division	
40	Offset percentage		
44	Nominal spec	Addition, Division	
45	Tolerance spec	Addition, Division	
48	Nominal range gate	Addition, Division	
49	Tolerance range gate	Addition, Division	
72	PPM above spec	Division, Multiplication	Count, Sort
73	PPM below spec	Division, Multiplication	Count, Sort
74	PPM in spec	Division, Multiplication	Count, Sort
75	PPM out of spec	Division, Multiplication	Count, Sort
83	PPM below spec predicted	Multiplication	Out of spec predicted
84	PPM above spec predicted	Multiplication	Below spec predicted
85	PPM out of spec predicted	Multiplication	Above spec predicted
86	PPM in spec predicted	Multiplication	In spec predicted
87	Mean + 3SD	Addition, Multiplication	SD, Mean
88	Mean - 3SD	Addition, Multiplication	SD, Mean
89	Six SD	Multiplication	SD

94	Mean + 4SD	Addition, Multiplication	SD, Mean
95	Mean - 4SD	Addition, Multiplication	SD, Mean
102	LCLr	Subtraction, Multiplication	Mean, SD
108	LCLx for zones		LCLx sub
122	Ppk	Division	Zl, Zu
123	CR	Subtraction, Division	6SD
124	Pr	Subtraction, Division	6SD
126	Pp	Subtraction, Division	6SD
127	PC	Subtraction, Division	6SD
128	PP	Subtraction, Division	6SD
129	Cpm	Square Root, Squares, Division	SD, Nominal, Mean
130	Zl factors	Subtraction, Division	SD
131	Zu factors	Subtraction, Division	SD
132	Zl SSD	Subtraction, Division	SD
133	Zu SSD	Subtraction, Division	SD
144	Trend crosses at	Subtraction, Division	Slope Sub, Y-Intercept Sub, SD
152	Trend Y-intercept UCLx	Addition	SD, Trend line sub
153	Trend Y-intercept LCLx	Subtraction	SD, Trend line sub
159	Variance	Square	SD
163	% above control	Division, Multiplication	Count, Sort
164	% below control	Division, Multiplication	Count, Sort
165	% out of control	Division, Multiplication	Count, Sort
166	% in control	Division, Multiplication	Count, Sort
167	% above gate	Division, Multiplication	Count, Sort
168	% below gate	Division, Multiplication	Count, Sort
169	% out of gate	Division, Multiplication	Count, Sort
170	% in gate	Division, Multiplication	Count, Sort
176	EWMA UCLx	Division, Multiplication, Subtraction, Addition, Square Root	Mean, R-Bar
177	EWMA LCLx	Division, Multiplication, Subtraction, Square Root	Mean, R-Bar
180	Moving average UCLx	Addition, Multiplication	Moving average mean, SD
181	Moving average mean	Summation, Division	Count
182	Moving average LCLx	Subtraction, Multiplication	Moving average mean, SD
183	Moving average UCLr	Addition, Multiplication	Moving average R-Bar, SD
184	Moving average R-Bar	Summation, Division	Count
185	Moving average LCLr	Subtraction, Multiplication	Moving average R-Bar, SD
187	Overall Range	Subtraction	Minimum, Maximum
188	% subgroups within 1 sigma	Subtraction, Division, Multiplication	Control Limits
189	% subgroups within 2 sigma	Subtraction, Division, Multiplication	Control Limits
198	Sum values	Addition	
199	Sum values squared	Addition, Multiplication	
205	CuSum start value (S)	Multiplication	k, L
206	CuSum low sample unit tolerance (C1)	Division, Multiplication, Subtraction, Addition	k, h, Target X
207	CuSum high sample unit tolerance (C2)	Division, Multiplication, Subtraction, Addition	k, h, Target X
245	Count above control	Addition	Count, Sort
246	Count below control	Addition	Count, Sort
247	Count out of control	Addition	Count, Sort
248	Count in control	Addition	Count, Sort
260	Degrees of freedom	Addition	Count, Subgroup size
286	SD (subgroup)	Sum of Squares, Square Root	Mean
304	Sigma Level	Subtraction, Division	SD

Group 5: Non-statistical values reported by QA/S GainSeeker® SPC

Internal Reference Number	Value
0	Subgroup size
1	Decimal places
2	Maximum count for retrieval
9	Run length
10	Part Number/Process
11	High date/time queried
12	Low date/time queried
13	Filter name
14	Filter contents
15	Short Run method
16	Description
17	Meas. System
18	Meas. unit
19	Range chart
20	Standard deviation method
21	Method used to calculate control limits
22	Exclude outliers
30	Lower gate
31	Upper gate
34	Lower Individual gate
35	Upper Individual gate
38	Lower reasonable limit
39	Upper reasonable limit
41	Column
42	Lower spec
43	Upper spec
46	Lower Range gate
47	Upper Range gate
50	Non-normal method
90	Mean + 3SD within spec
91	Mean – 3SD within spec
92	Mean + 4SD within spec
93	Mean – 4SD within spec
96	Standard variable 1
97	Standard variable 2
98	Standard variable 3
99	Standard variable 4
115	Skewness test
116	Kurtosis test
136	Zl used
137	Zu used
138	Chi-squared hypothesis
158	Data entry constant
161	High date/time retrieved
162	Low date/time retrieved
171	Kruskal-Wallis test
173	Kruskal-Wallis 0.10 limit
174	Kruskal-Wallis 0.01 limit
175	X-bar chart all inside control
186	Moving average span
190	Scale high Range

191	Scale high X-Bar
192	Scale low X-Bar
193	Exponent
197	Limit non-normal slices
200	Lower control cannot be less than zero
201	CuSum acceptance limit (L)
202	CuSum start value constant (k)
203	CuSum decision interval (h)
208	Date period
209	SQL query statement
217	Data in control
218	Current date/time
219	Current date
220	Current time
221	High date queried
222	High time queried
223	High date retrieved
224	High time retrieved
225	Low date queried
226	Low time queried
227	Low date retrieved
228	Low time retrieved
231	Count data groups
232	Data grouped by
259	Varying subgroup size found
261	Standard subgroup size
262	Target X
263	Target R
265	DMS Part Number
266	DMS Process
267	Non-symmetrical control limits
268	Add groups for varing subgroup sizes
269	Default data grouping
270	Runs based on
271	Exclude varying subgroup size (percent)
272	Exclude varying subgroup size (degrees of freedom)
273	Standardize data for varying subgroup size
274	Amount of time to display
275	Force distribution
276	Evaluate more data sets as non-normal
277	Trend chart limits
278	EWMA weight
279	Real-time checks
280	Standard flags
281	Short Run index
282	Use exponent
301	CuSum acceptance limit type
303	Exclude varying subgroup size (all)
305	Conversion multiplier
306	Drill-down conditions
307	External data
308	Last X-Bar
309	Last Event
310	Last Cause
311	Last Action Taken

312	Last Note
313	Last Part Number
316	Window login name
317	Retrieval name
318	Paired Sample Grouping

Statistical changes made between SPC version 6 and SPC version 7.1

1. We changed our table for the constant d4. d4 is used to calculate upper control limits on the r chart. Therefore all range upper control limits may change slightly. A.T.A. Holden of the Rochester Institute of Technology recomputed these values in 1987.

2. The calculation of the Y-Intercept for control limit lines (not best-fit regression lines) has changed to more accurately calculate the error term. This change will be noticed on both the upper and lower trending control limit line equations as well as the Y-Intercept for control limit statistical values.

3. The EWMA upper and lower control limits were reported switched in version 6 of SPC. We now report the limits correctly.

4. We now report PP, Pp, Ppk and PR differently if non-normal data. We now set these values to their corresponding normal values, e.g. Ppk = Cpk.

Statistical changes made between SPC version 7.1 and SPC version 7.2

None.

Statistical changes made between SPC version 7.2 and SPC version 7.3

1. The following statistics are now calculated for data with unilateral specifications: Cp, CR, PC, PP, Pp, and PR.

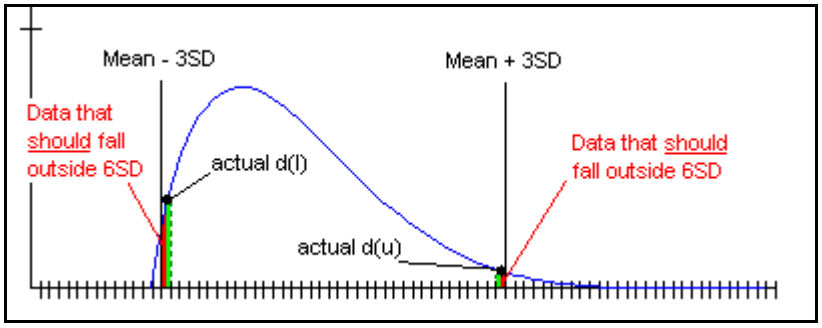
Statistical changes made between SPC version 7.3 and SPC version 7.4

1. Data with a subgroup size of one which is non-normal can now use non-symmetrical control limits. Your system can be configured to use non-symmetrical or symmetrical limits (i.e. the limits reported in previous versions). The new limits affect the following statistics: “Chart in control”, “Data in control”, “LCLx ind”, “LCLx sub”, “UCLx ind”, “UCLx sub”, “Percent above control”, “Percent below control”, “Percent in control”, “Percent out of control”, “Percent within 1 Sigma”, “Percent within 2 sigma”.

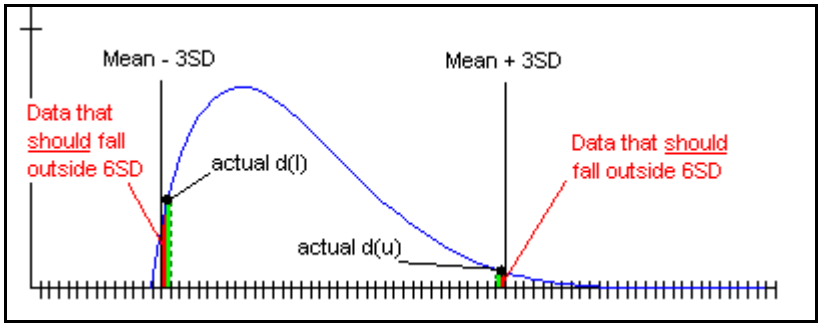
2. GainSeeker 7.4 now calculates **Mean +/- 3SD** more accurately for data that is non-normal and modeled.

   For data that is non-normal and modeled, GainSeeker integrates the area under the curve to approximate the points on the curve beyond which only 0.135% of the data rests (**d(l)** and **d(u)**).

In previous versions, GainSeeker used a minimum of 16 slices for the integration. After isolating which slices contained d(l) and d(u), GainSeeker set the statistics **Mean +/- 3SD** to the value of the "innermost" side of each slice (closest to the median). Using such large slices for the integration and basing process spread on the inside edge of the slices provided a somewhat optimistic interpretation of the actual spread of the process.



Beginning with version 7.4, GainSeeker now uses a minimum of 1024 slices for integration of the data curve. It also sets **Mean +/- 3SD** to the value of the outside edge of the slices containing d(l) and d(u). This provides greater accuracy and a more realistic picture of the true spread of the data. Notice how the red area (error) has been greatly reduced in the following example:



Because this tends to slightly widen the reported spread of the data, you may find that Cpk calculations for non-normal data are lower than those reported by previous versions of GainSeeker for the same data set. Other statistics whose calculations are based on the Mean +/- 3SD will also be affected. They are Mean + 3SD, Mean – 3SD, Mean + 4SD, Mean – 4SD, Mean + 3SD within Spec, Mean – 3SD within Spec, Mean + 4SD within Spec, Mean – 4SD within Spec, Six SD, Cpk, CR, Ppk, PR, Cp, PC, Pp, PP, PPM out/in/below/above Spec, In/out of Spec predicted, Non normal Area, Below Spec actual, zl/zu non normal, zl/zu used

If you must use the earlier, less accurate approximation of the data spread, add nn_sd_old=1 to CMS.ini.

Statistical changes made between SPC version 7.4 and SPC version 7.5

1. The statistical label 'Values cannot be less than zero' was changed to 'Lower control cannot be less than zero'.

2. The following statistical values are new in version 7.5: 'Percent above Ind. Gate', 'Percent above lower Gate', 'Percent above lower Ind. Gate', 'Percent above lower Spec', 'Percent below Ind. Gate', 'Percent below upper Gate', 'Percent below upper Ind. Gate', 'Percent below upper Spec', 'Percent inside Ind. Gate', 'Percent outside Ind. Gate', 'Varying subgroup sizes found', 'Degrees of Freedom', and 'Standard subgroup size'.

3. The following statistical values would report a zero when no specifications were set. These statistics now report a 'NA' instead of zero: 'CR', 'PC', 'PP', 'Pp', and 'PR'.

4. The following statistical values would report a 'None' when no specifications were set. These statistics now report a 'NA' instead of 'None': 'Cp', 'zl non-normal', and 'zu non-normal'.

5. If the data set contains subgroups of varying sizes, the statistical value for 'Subgroup size' will return a value of zero. To get the subgroup size stored in the standard, use 'Standard subgroup size'.

Statistical changes made between SPC version 7.5 and SPC version 7.6

None.

Statistical changes made between SPC version 7.6 and SPC version 7.7

None.

Statistical changes made between SPC version 7.7 and SPC version 8.0

1. Bypassed values (233) and Bypassed subgroups (234) changed from 'NA' to '0' when no data is found.

2. If reporting performance indexes as capability indexes, we no longer set Ppk, Pp and PR to NA.

3. The following statistics were incorrect in some rare cases when excluding outliers: Cp, Cpk, CR, Mean, Mean – 3SD, Mean -4SD, Mean + 3SD, Mean + 4SD, PC, Pp, PP, Ppk, PR, Six SD, ChiSqr Value, ChiSqr Hypothesis, Zl used, Zl non-normal, Zu used, Zu non-normal, Maximum good subgroup, and Percent within 2 sigma.

4. Changed the Nominal used for zone checks when using Gates and multiple data groups are found.

5. The default number of slices to use for non-normal calculations has changed from 2048 to 4096.

6. Several statistics have been removed. Individual and Subgroup statistics are no longer both calculated; just one statistic is now reported. For example, instead of 'LCLx ind' and 'LCLx sub'; there is just 'LCLx' now.

7. The statistics that report percentages changed to report the value to two decimal places.

8. The Maximum Slices statistic (#197) changed to Limit non-normal slices and is reported differently in Dynamic Reports. The Charts and Reports module returns the number of slices and Dynamic Reports returns a Yes or No.

9. Several statistic labels where changed for capitalization or to make the label more clear.

10. Several statistic values changed from True/False to Yes/No and from Not Set to Not set.

Statistical changes made between SPC version 8.0 and SPC version 8.4

None.

Statistical changes made between SPC version 8.4 and SPC version 8.5

Statistical differences between SPC Data Entry (SDE) and PC Collect (PCC) noted in version 8.5

For the real-time checks 2 of 3 above/below 2SD and 4 of 5 above/below 1SD, SDE does an extra check and ignores the failure if the point not a part of the run is on the opposite side of the mean. PCC follows Wikipedia documentation and considers the run a failure regardless of the position of the point not above/below nSD. Also, PCC tends to report a run twice. For example, if there is a point above the mean, then four below 1SD, then another point above the mean, the four points below 1SD will be reported with the initial point and with the trailing point.

If the data set is non-normal, PCC calculates control limits for a subgroup size of one based on the model of the curve at the +/- 3SD positions. SDE calculates the control limits based on Factors or the Sample Standard Deviation (SSD) for a subgroup size of one for non-normal data.

Note: References to PCC are meant to include both PC Collect and the Python script interface.

Statistical changes made between SPC version 8.5 and SPC version 9.3.2

None.

Statistical changes made in SPC version 9.4

Sigma Level (304) should return the smaller of the two Z values – Zl (136) or Zu (137) – used to calculate process capability. In previous versions, Sigma Level returned ‘NA’ if Zl was ‘NA’, even if Zu was not ‘NA’. Sigma Level now returns ‘NA’ only if both Zl and Zu are ‘NA’.

Statistical changes made in SPC version 9.4.1

None.