Comparison of various handicapping systems on a 15 race Series.

(A TopYacht Discussion Paper. May 2001.)

Objective

To compare a number of automatically updating handicap systems for in shore keel boat races. [I refer to these as "Measured Performance Handicap" systems.]

Methodology

The same 15 race Series was re-scored / re-handicapped under a number of different handicap (HC) systems and different parameter set-ups.

Systems compared

1. Running average
2. Trend biased running average
3. Recent race biased running average
4. Exponential average
5. Percentage performance

Parameters explored

Number of races used in running average

Lower clamp

a percentage (of the race AHC) at which a BCH will be clamped if the BCH falls below that value.

Upper clamp

a percentage (of the race AHC) at which a BCH will be clamped if the BCH exceeds that value.

Lower Limit

a percentage (of the race AHC) below which a BCH will be totally ignored for handicap calculations.

Upper Limit

a percentage (of the race AHC) above which a second or later BCH will have its total (vs upper clamped) value included in the calculations of the new handicap

Cross Series (vs In Series)

Cross Series handicapping uses data from other Series. In Series only uses data from within this Series and uses a special algorithm to calculate handicaps for the first few races until a sufficient number of BCHs are available.

Minimum step size

The smallest percentage that a new handicap must alter before the new value is stored for future use. New handicaps below this value are ignored and the current HC carried forward.

Maximum step size

The largest percentage that a new HC can alter from the current HC.

Fast Start

Provides a mechanism for adjusting HC more quickly at the start of a Series than will be done when sufficient BCHs are available for normal calculation. Primarily used for In Series HCing.

First place Correction. [1P]

If activated this calculates the first place getter's new HC as being the HC needed for that competitor to have be between the second and third places getters in the current race.

First place Correction extended. [2P]

As per P1 but the BCH for this race is also altered by the same scale factor as the AHC.

The Series analysed

Summer 2001 Series from SYC. Only Divisions 1 and 2 were considered as both had in excess of 20 boats. Some races contained just SYC boats other races contained many casual entries.

Comparison of outcomes

Four criteria were used as a measure of the effectiveness of the HC system under test.

Spread of scores of the first 10 place getters in the Aggregate results

For place getters below 10th place the DNCs etc were a larger influence than the HC corrected times.

Initially we used….

Spread of scores = average score of the places getters 8,9,10 minus the average score of the first 3 place getters.

This was later refined to the standard deviation of the first 10 place getters. The two approaches provided the similar answers and the later was easier to calculate so it has been exclusively used.

So for an *approximation* to the actual spread (1st to 10th), multiply the figures provided by 2.5.

Average number of different competitors who received a first through to a 10th.

The number of *different* competitors with a first place was added to the number of *different* comps with a second place etc then the resultant averaged.

Spread of normalised HC corrected times.

This is the standard deviation of the HC corrected times for each race where the average race time was normalised to 100 minutes. These are then averaged for the 15 races in the Series.

Drift of the average HC over the Series

The average AHC for each race was listed and any trend manually noted.

Outcome

Over 40 different HC systems or system configurations were trialled.

The test results were graphed.

Notes

- The HC system used for the testing does provide the facilities to allow individual minimum HCs for each competitor. So when it is was appropriate for the test, the poor performances were ignored by setting a Low Limit of 5% *of* the current HC or put another way, 95% *below* the current HC.

- It soon became apparent that the average spread of normalised (to 100 minutes) HC corrected times for the Series under any HC system was rarely even 10 seconds different to any other system tested so these results are not commented on further.

- While there was some variations of average AHC *per race* across the HC systems, not one system provided a significant drift and so this will not be further commented on.

But what is a "good" handicap system?

- Before evaluating any HC system it is necessary to determine the criteria for a "good" HC system. This definition will vary from club to club.
- Some clubs want *all* boats to have an even chance of winning a race.

- Others say they want a HC system where *all competently sailed boats* have an even change of winning a race.
- Still other clubs say the Aggregate Series results are what matters, so they are looking for a system that gives the closest bunching of Aggregate scores.
- Some clubs argue that the best crew(s) should win even if this means that only 5 or so competitors ever win a race.
- Other clubs hold the view that the aim is share the prizes around.
- Two further criteria might come from the two questions often asked of handicappers..
  "I sail consistent and well so how come I never win a race, what's wrong with this HC system?"

  "I came half way down the fleet in this race but my HC for next week has gone up %%$#$#@!!! What sort of dumb Handicapping system is this?"

Results

All the systems explored were designed to automatically provide a new HC for all boats that performed reasonably to their HC. The basic aim being to establish an "average performance" indicator i.e. the handicap. The outcomes was purely computer generated with no manual intervention by a HCer.

Overview.

- Virtually all systems produced "sensible" handicaps.
- The boats in the first 5 places in the Aggs were very similar but with some place changes. Likewise the places 6 to 10 was mainly filled with the same set of boats under all systems.
- For Div 1 the spread of standard deviations of aggregate scores was from 10 to 19 while for Div 2 this was 17 to 27. [Remember multiply this by approx. 2.5 for the range of scores from 1st to 10th.]
- The number of *different* competitors who were placed first was summed with the number of *different* competitors who came 2nd etc. These figures were averaged. They varied from 10 to just 11.5 for Div 1 and from 9.75 to 11.4 for Div 2 over the 15 race Series.
- Systems that allowed all competitors to have their HC adjusted [even if some competitors consistently performed well below their HC] moved more competitors into contention thereby spreading the number of different place getters and tightening the Aggs.
- In general terms systems that allowed the new HC to vary significantly on a race by race basis tended to be the ones that provided the tighter spread of Aggs and the larger number of different first place getters etc. and conversely.

Details

All tests out use the 40% boat on HC Corrected time as a reference for the generations of the BCHs. Changing the value of this parameter does not alter the relative performances but rather leads to HC drift.
As I think of competitors as performing "within a +/ 3% range" the following tables use the terminology of "Low Limit of 20%" meaning a limit set at 20% *below* the HC. This is the same as saying at 80% *of* the HC.

As the bulk of the HC systems tested were variants of the running average, the basic running average was investigated first.

The number of BCHs to average

Division 1

Number BCHs Averaged	Stand. Dev 1~10 Agg Places	No Diff Place Getters

4BCH, 20%Clmps, 20%LL, 20%UL	11.1	11.2
6BCH, 20%Clmps, 20%LL, 20%UL	11.0	11.2
8BCH, 20%Clmps, 20%LL, 20%UL	12.4	10.9
4BCH, 4%Clmps, 5%LL, 5%UL	14.2	10.7
6BCH, 4%Clmps, 5%LL, 5%UL	15.5	11
8BCH, 4%Clmps, 5%LL, 5%UL	18.4	10.2

Division 2

Number BCHs Averaged	Stand. Dev 1~10 Agg Places	No Diff Place Getters
4BCH, 20%Clmps, 20%LL, 20%UL	17.3	11.0
6BCH, 20%Clmps, 20%LL, 20%UL	18.3	10.3
8BCH, 20%Clmps, 20%LL, 20%UL	22.0	10.1
4BCH, 4%Clmps, 5%LL, 5%UL	20.5	11.1
6BCH, 4%Clmps, 5%LL, 5%UL	23.8	10.6
8BCH, 4%Clmps, 5%LL, 5%UL	26.8	10.6

The differences are small. But **if** the aim was small spread of Aggs **and** more different place getters then 4 BCHs appears the best choice. For the rest of this discussion "better" will be understood to mean "the aim was a small spread of Aggs **and** a large number of different winners, second place getters etc.".

Reducing the 4 BCHs to 3 was not evaluated for the following [personal] reasons.

Observations of many HC performance graphs shows that the vast majority of competitors' consecutive performance oscillates around a mean value.

- There is a certain randomness about this process. Less than 4 BCHs can provide an average that poorly reflects the "normal" performance of the competitor.
- It is most unusual to find consecutive performances moving away from this average for more than 4 races before heading back towards the average value.
- Even averaging 4 races can provide rather large steps from one HC to the next. With 3 or less races these steps can become very large.

Below is a "classic" performance curve in that the performances neatly oscillates about the HC. [This is a real competitor.] The dots are the performance and the squares are the HC for each race. You can also see the "+/- 4% clamps" and "+/- 5% limits" as discussed below.

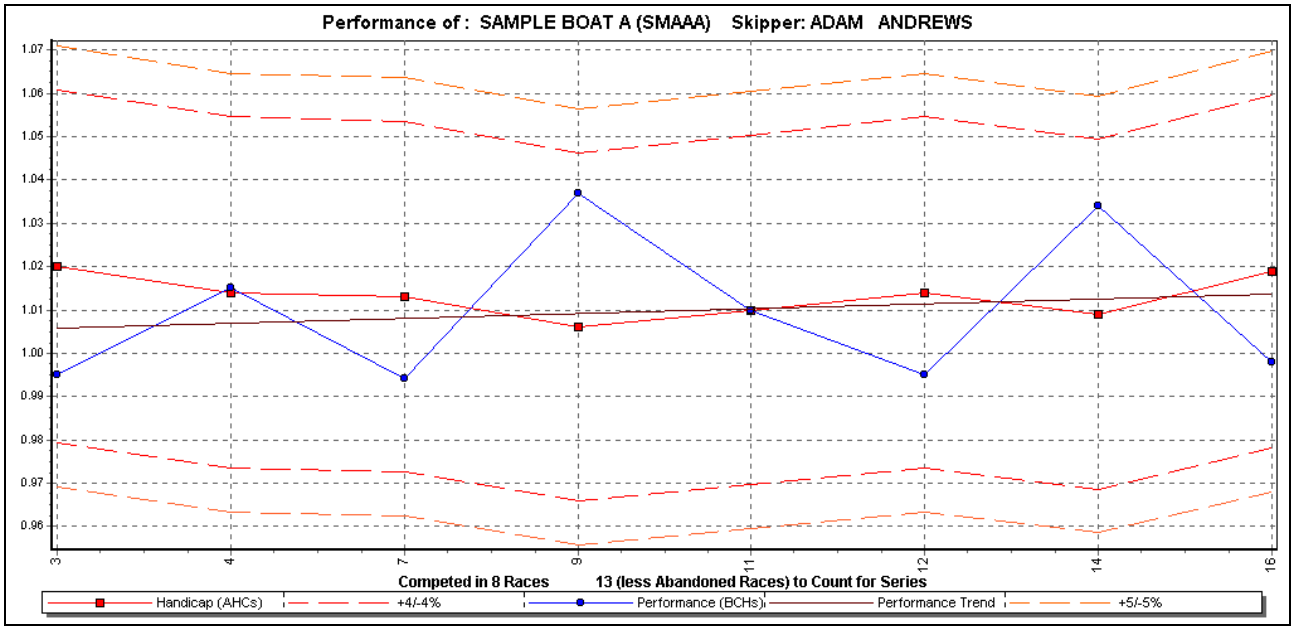

Figure 1 "Classic" Performance Curve

Next is more a “normal” performance graph in that it moves up and down and has a certain randomness. Again this is a graph of a real competitor.

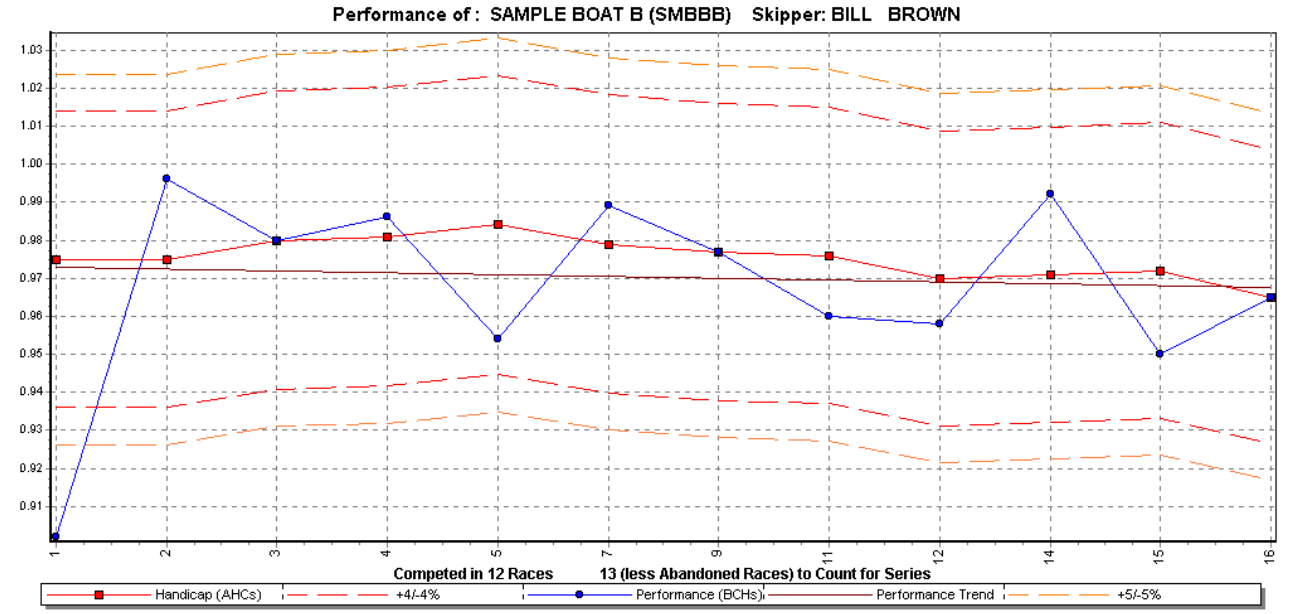

Figure 2 "Normal" Performance Curve

> **Note**: Rather than carry out all the following evaluations on a various numbers of BCHs in the averaging process, many of the following tests were carried out with just 4 BCHS as this appeared the best option from the first test above.

L

Limit values.

While the upper Limit had minimal effect, the lower Limit had significant effect. If this was lowered to –20% then virtually all competitors had every BCH included in the averaging process with the result mentioned earlier.

Division 1

Number BCHs Averaged	Low Limit	Stand. Dev 1~10 Agg Places	No Diff Place Getters
4	20%	11.1	11.2
4	10%	11.2	11.2
4	5%	17	11.2

Division 2

Number BCHs Averaged	Low Limit	Stand. Dev 1~10 Agg Places	No Diff Place Getters
4	20%	17.3	11.1
4	10%	18.2	10.7
4	5%	21.8	10.7

Clearly the 20% lower limit tightened the Aggs **BUT**, if the lower limit was the mechanism used to exclude the performances that were consider to be too far below the HC then you may opt for a lower limit of say 95% even though it produced poorer outcomes in this test.

Clamps

Observation of many performance graphs suggests that the performance of the vast majority of competitors oscillate around their average performance and that the range is normally +/- 3% to +/-4%. This is so clear from the many performance graphs observed that it was not considered useful to expend too much time investigating this so only a few comparisons were undertaken.

Both Clamp Comparisons

In both cases 4BCHs were averaged. The first test was +/-3% the second +/-4%.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4% Lower 4% High	14.2	10.7
3% Lower 3% High	14.9	10.9

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4% Lower 4% High	20.5	11.1
3% Lower 3% High	21.7	11.2

Small differences, but the choice depends on whether you are after tighter Aggs or more different place getters.

Lower Clamp comparisons

The VYC standard is a 3% lower clamp, a 10% lower limit and all BCHs in the range 5% to 3% be clamped at 1.5%.

This was compared with a simple lower clamp at 97% and a lower limit of 95%.

In both cases 8 races were considered with 2 low discards and 1 high discard.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
VYC	13.8	10.8
3% Lower 3% High	13.4	10.8

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
VYC	24.6	10.1
3% Lower 3% High	25.2	10.4

As can be seen the additional step in the VYC system had only a small impact and inconclusive effect on the results for the race Series that was studied.

Upper Clamp comparisons.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
3% Lower 5% High	13.8	10.8
3% Lower 3% High	13.4	10.8

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
3% Lower 5% High	24.6	10.1
3% Lower 3% High	25.2	10.4

[I know these results look the same as the previous set. The results *do* reflect the tests.] The difference are small and in terms of the Stand Dev Aggs they are at odds between the divisions.

An interesting variant on altering Clamps AND limits is to deliberately uses the clamps and limits to "push up" the competitors HC towards their better performances. In the following table the 'push up" variant used an upper clamp of 3% but a lower clamp of just 1.5% and different lower limits. A non "push up" reference is also provided. In all cases below the 4 BCHs were averaged.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4%Clamps, 5%Low Lim, 5% Hi L	14.2	10.7
1.5%L Clamp, 3% Up Clamp, 4%Low Limit, 5% Hi Limit	14.1	10.8
1.5%L Clamp, 3% Up Clamp, 5%Low Limit, 5% Hi L	14.8	10.6

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4%Clamps, 5%Low Lim, 5% Hi L	20.5	11.1
1.5%L Clamp, 3% Up Clamp, 4%Low L, 5% Hi L	21.1	10.9
1.5%L Clamp, 3% Up Clamp, 5%Low L, 5% Hi L	23.2	10.9

Different outcomes for the two divisions. Suggesting that there is not much in it??

Win correction for 1st Place [referred to in test data as 1P]

To win a race a competitor must have performed the highest above his HC when compared to others. This "highest" is usually in the order or 3 to 6 percent. Observations of many performance graphs shows that competitors rarely stray outside their 3 to 4 percent oscillations. But on a rare occasion they venture somewhat above their normal upper range. On these occasions they usually win the race or at least come well up the fleet. Because such occasions are rare I am personally reluctant to adjust a competitor's HC on the basis of a win. BUT some clubs believe that it is necessary to attempt to stop the same competitor winning consecutive races. So some tests were run. The algorithm implemented increased a first place getters HC for the next race to the HC that the winner needed to have come half way between 2nd and 3rd place in the race he won.

For a second test [referred to a 2P] the first place getter also had his BCH for this race increased by the same scale factor as applied to their HC.

The "In Series" noted below is where no BCHs from previous Series are used in the calculations but a "fiddle factor on the Initial HC" is used until 4 BCHs are available for averaging.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	14.2	10.7
4BCH, 4%Clmps, 5%LL, 5%UL,1P	13.7	11.6
4BCH, 4%Clmps, 5%LL, 5%UL,2P	13.3	11.4
4BCH, 4%Clmps, 5%LL, 5%UL, In Sers	15.3	10.8
4BCH, 4%Clmps, 5%LL, 5%UL, In Sers, 1P	16.0	10.7
4BCH, 4%Clmps, 5%LL, 5%UL, In Sers, 2P	15.3	10.6
VYC	13.8	10.8
VYC, 1P	13.9	11.1
8BCH,3%Clmps, 5%LL, 7%HL	14.6	11.2
8BCH,3%Clmps, 5%LL, 7%HL, 1P	15.3	11.3

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps,95%LL,105%UL	20.5	11.1
4BCH, 4%Clmps,95%LL,105%UL,1P	20.7	11.1
4BCH, 4%Clmps,95%LL,105%UL,2P	20.7	11.0
4BCH, 4%Clmps,95%LL,105%UL, In Sers	20.3	11.4
4BCH, 4%Clmps,95%LL,105%UL, In Sers, 1P	20.5	11.3
4BCH, 4%Clmps,95%LL,105%UL, In Sers, 2P	20.5	11.1
VYC	10.1	10.1
VYC, 1P	11.1	10.4
8BCH,3%Clmps, 95%LL,107%HL	10.1	10.1
8BCH,3%Clmps, 95%LL,107%HL, 1P	10.6	10.6

The win corrections do not seem to have much *overall* effect but they may well provide the intended effect of lowering the chances of a competitor winning two consecutive races.

Limiting the maximum handicap step size.

Some clubs apply a limit to how much a HC can change from one race to the next. This dampening effect slows the rate of change of the HC. The only tests carried out with this concept were with the Exponential Averaging system shown later.

Leaving the HC unchanged unless the new calculated HC exceeds a specified Step Size.

Some HCers like the concept of the HCs remaining reasonably fixed unless a competitor's performance exceeds a predetermined "step size". Others prefer to have the HCs move after each race so the competitors understand that the HC are under constant revision. In the following tests the Step size was set to 0.05% in one instance and to 0.5% in the other.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps,5%LL, 5%UL 0.05Step Size	14.2	10.7
4BCH, 4%Clmps,5%LL, 5%UL 0.5Step Size	15.9	11

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps,95%LL,105%UL 0.05Step Size	20.5	11.1
4BCH, 4%Clmps,95%LL,105%UL 0.5Step Size	20.3	10.6

Like so many previous tests the results are inconclusive.
Here are two graphs. The first has a step size of 0.05% the second a step size of 0.5%.
Notice that the second shows less race by race fluctuations in the HC.

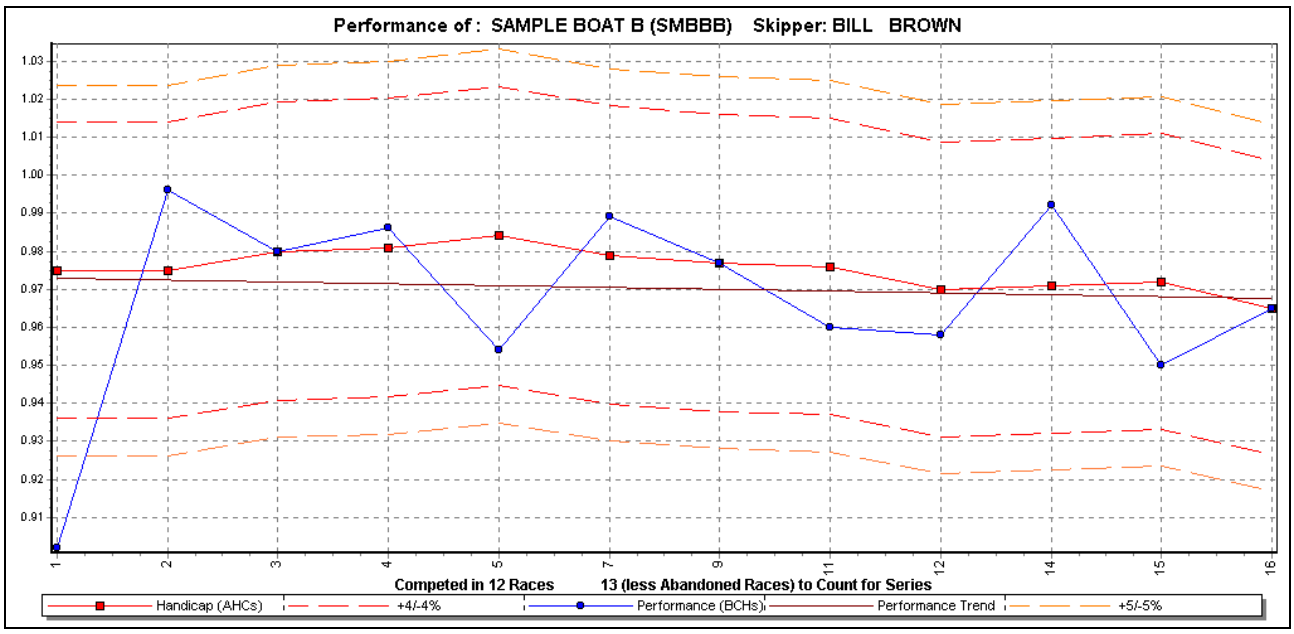

Figure 3 Performance Curve with 0.05% minimum step size

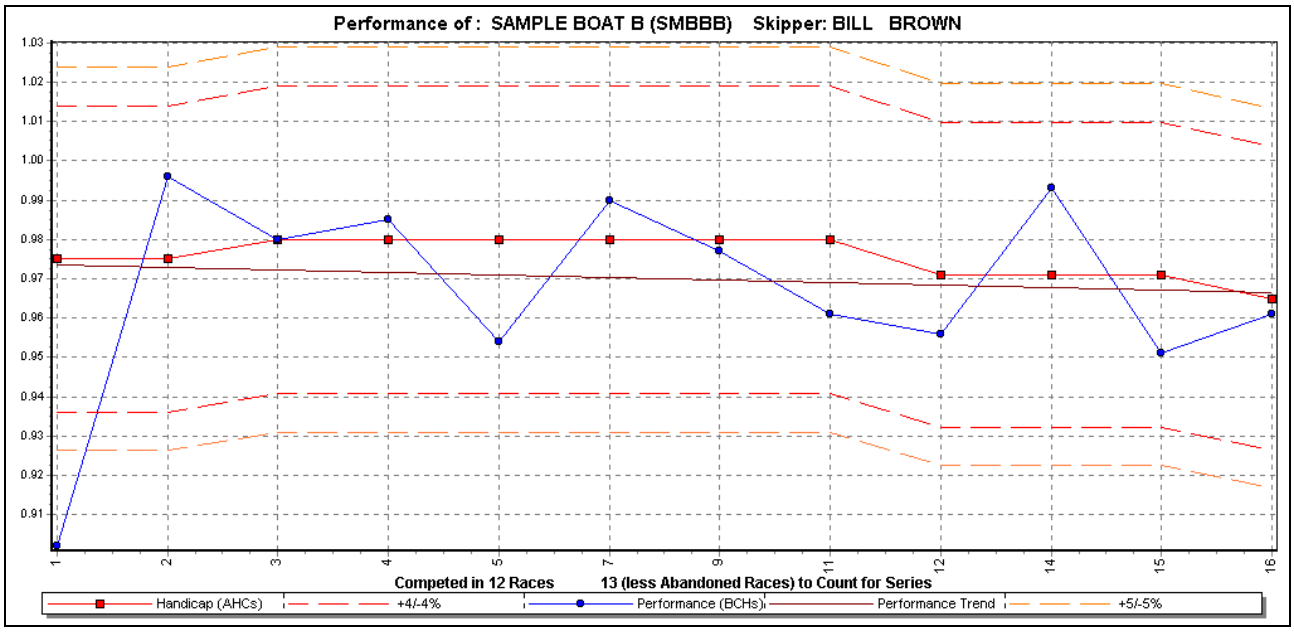

Figure 4 Curve with 0.5% minimum step size

Variants on the Running Average Handicapping system

One thing has become clear from all the above tests. Provided you use sensible values for the various HC parameters then the HC results produced by any of the above systems will be reasonable with surprisingly little differences evident in the many tests carried out.

Next to be explored were a few Alternate Handicapping systems that are based around the running average concepts.

Recent Biased Averaging.

This system uses a running average of BCHs but provides a greater weighting to the more recent BHCs. Some HCers believe that this provides a more 'up to date' HC. Others believe it may help as a mechanism to counter those competitors whose performance is slowly rising and as such are often in the winners circles. As mentioned above, observation shows that competitors' performances are generally random with a basic oscillation around an average value. For this reason I am personally wary of this form of biasing as it can provide emphasis on a particular BCH(s) which, as luck may have it, is rather uncharacteristic of the competitor's average performance.
The algorithm implemented for these tests provides a different weighting dependant on the order of the BCHs. The latest BCH = "BCH1".

"Sum of Weightings" [SoW] is the sum of each individual weight so if the weights were 7,6,5,4 the SoW = 7+6+5+4 = 22 [So the overall average is unity based.]
When averaging 6 BCHs.
New HC = (7/SoW.) * BCH1 + (6/SoW) * BCH2 + (5/SoW) * BCH3 + (4/SoW) * BCH4. + (3/SoW) * BCH5 + (2/SoW) * BCH6.

Notes:
The weightings do not decrease below 2. So for 8 BCHs averaged the last two have the same weighting.

The weighting always starts at 7 irrespective of the number of BCHs being averaged. The concept here was that if the number of BCHs was small then the relative weights of just the first few will not be too dominant for the reason stated in the previous paragraph.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	14.2	10.7
4BCH, 4%Clmps, 5%LL, 5%UL, Recent Bias	13.6	10.7
4BCH, 4%Clmps, 5%LL, 5%UL, In Sers	15.3	10.8
4BCH, 4%Clmps, 5%LL, 5%UL, IS, Recent Bs	15.2	10.8

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	20.5	11.1
4BCH, 4%Clmps, 5%LL, 5%UL, Recent Bias	20.2	10.8
4BCH, 4%Clmps, 5%LL, 5%UL, In Sers	20.3	11.4
4BCH, 4%Clmps, 5%LL, 5%UL, IS, Recent Bs	20.5	10.9

The above tables do not provide any consistent indication of a substantial improvement under such a system.

Trend Biased

The situation can arise where a competitor is slowly improving and therefore often performs above his HC. Consequently the competitor wins several consecutive races Some clubs believe that it is not appropriate for one competitor to win a number of consecutive race / trophies.
By considering the gradient of the trend of consecutive BCHs a "Trend Biased" factor can be applied to calculation of the new HC.

Below the performance graphs of a competitor under 4BCH, 4%Clamps, 5%LL, 5%UL then with the same parameters but with Trend Bias. Notice how much quicker the HCs "catch up" particularly for races 10 and 11 and 14.

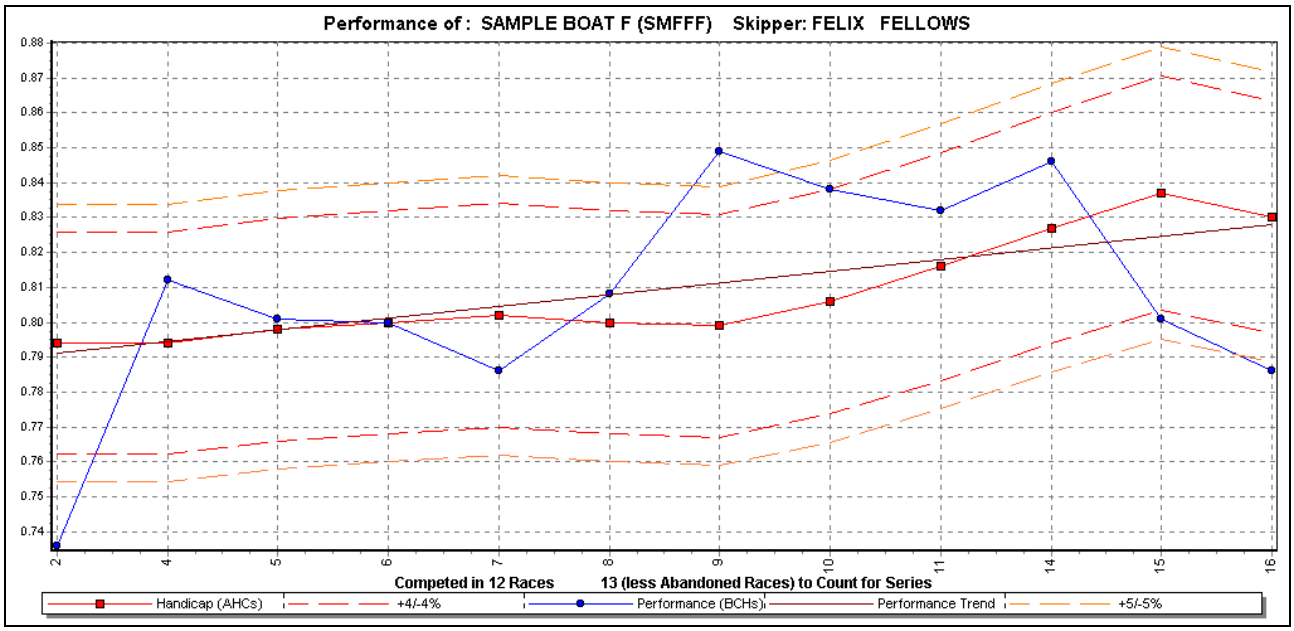

Figure 5 Performance Curve for standard running average of BHCs

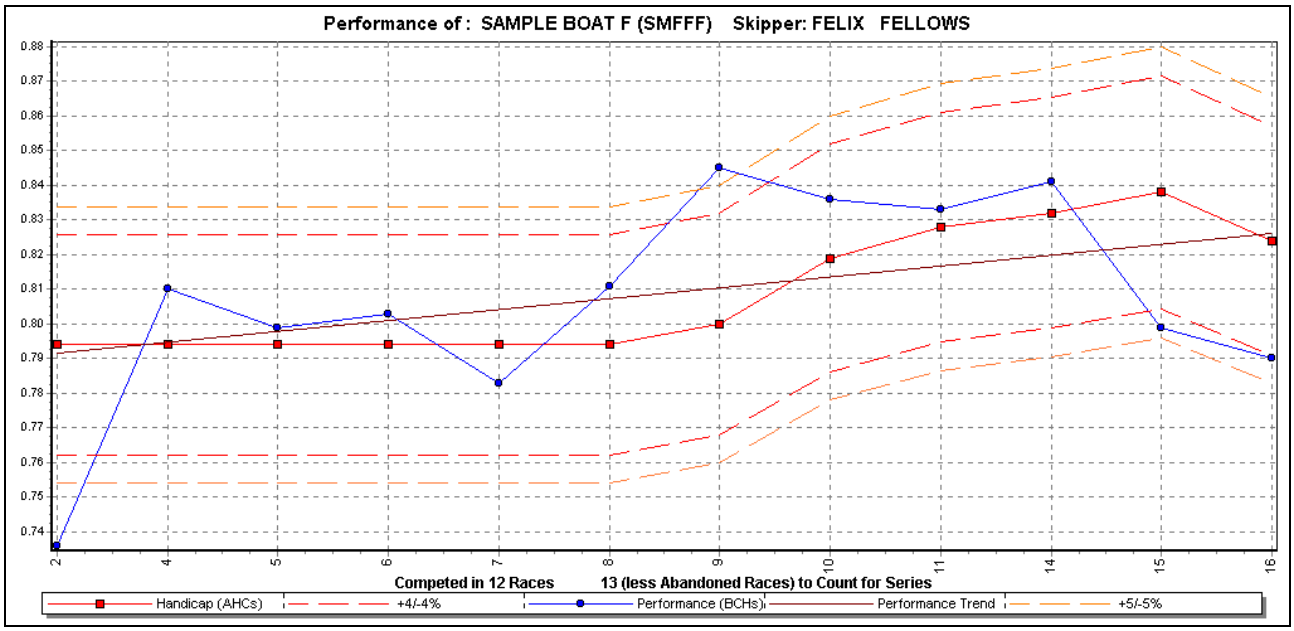

Figure 6 Performance Curve for Trend Biased running average of BHCs

What of the two criteria?

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	14.2	10.7
4BCH, 4%Clmps, 5%LL, 5%UL, Trend Bias	14.6	11.3
4BCH, 4%Clmps, 5%LL, 5%UL, In Sers	15.3	10.8
4BCH, 4%Clmps, 5%LL, 5%UL, IS, Trend Bs	16.5	10.6

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	20.5	11.1
4BCH, 4%Clmps, 5%LL, 5%UL, Recent Bias	19.1	11.1
4BCH, 4%Clmps, 5%LL, 5%UL, In Sers	20.3	11.4
4BCH, 4%Clmps, 5%LL, 5%UL, IS, Trend Bs	20.4	11.1

While some "benefit" may be seen if there is a reduction of the number of times an "improver" is awarded a race trophy, but overall this concept shows no apparent improvement to the HCing.

Exponential Averaging

New HC = ((Gain-1)/Gain) * HC) + (1/Gain) * BCH for this race.
This system is popular in a number of clubs because of its simplicity and because if a competitor performs poorly in this race then his HC will go down for the next race and conversely. [This is *not* necessarily the case for running average systems]. Thus this system solves one of the questions posed by sailors earlier in this document.
Observations of real data suggest that this system has dual flaws in that it seems slow to react to changes and yet can appear to provide largish steps from one new HC to the next.

One solution is to use a largish gain but to also apply a limit to the step size to the next HC.

Division 1

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	14.2	10.7
Expo G=4, 4% Clamps, 5%LL, 5% UL	16.9	10.2
Expo G=4, 4% Clamps, 5%LL, 5% UL, 1% Step Limit	16.2	10.1
Expo G=5, 4% Clamps, 5%LL, 5% UL	18.6	10.1

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	14.2	10.7
Expo G=4, 4% Clamps, 5%LL, 5% UL	23.8	10.5
Expo G=4, 4% Clamps, 5%LL, 5% UL, 1% Step Limit	23.3	10.1
Expo G=5, 4% Clamps, 5%LL, 5% UL	25.0	10.8

These system appear to expand the Aggs rather than compress them and offers no advantage in the spread of scores area.

Percentage Performance

This is a non conventional system and so will need a full explanation.
An often heard complaints from experienced sailors is ".I am a reasonably a good and a consistent sailor. I have many years experience but I have *never* yet won a single race. What's wrong with the HC system ??"

Under a conventional HC system the winner is the competitor with the lowest HC corrected time.

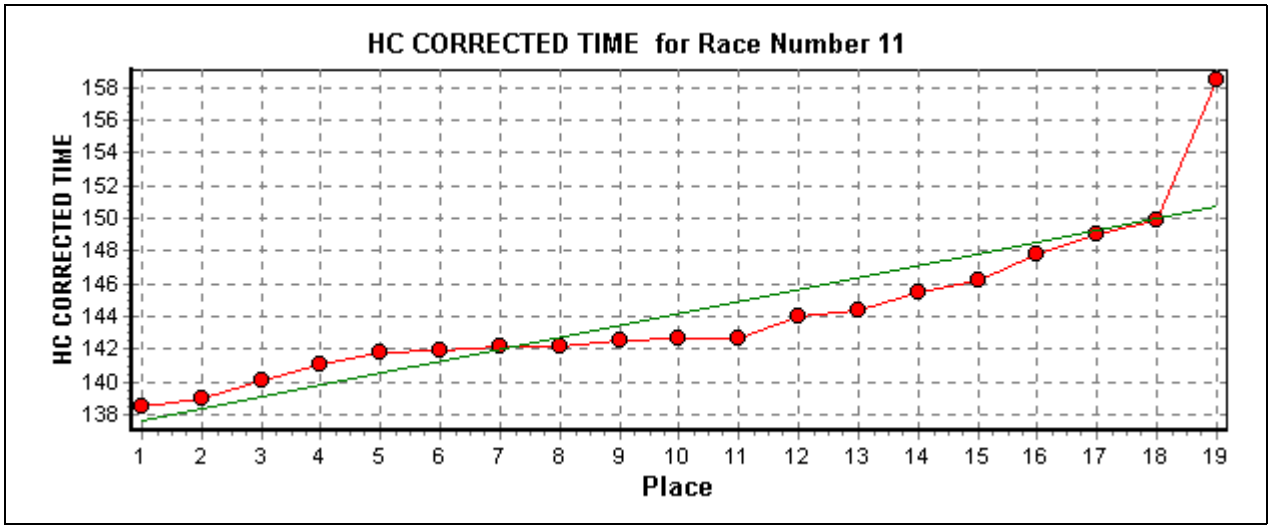

Figure 7 Handicap Corrected Time for Race 11 against place in the race.

If reality what that means is the winner is the competitor who sails most above his HC. The following graphs show the percentage above their HC to which each competitor performed in a particular race. As for the previous graph the order is by HC corrected time i.e. by "place".

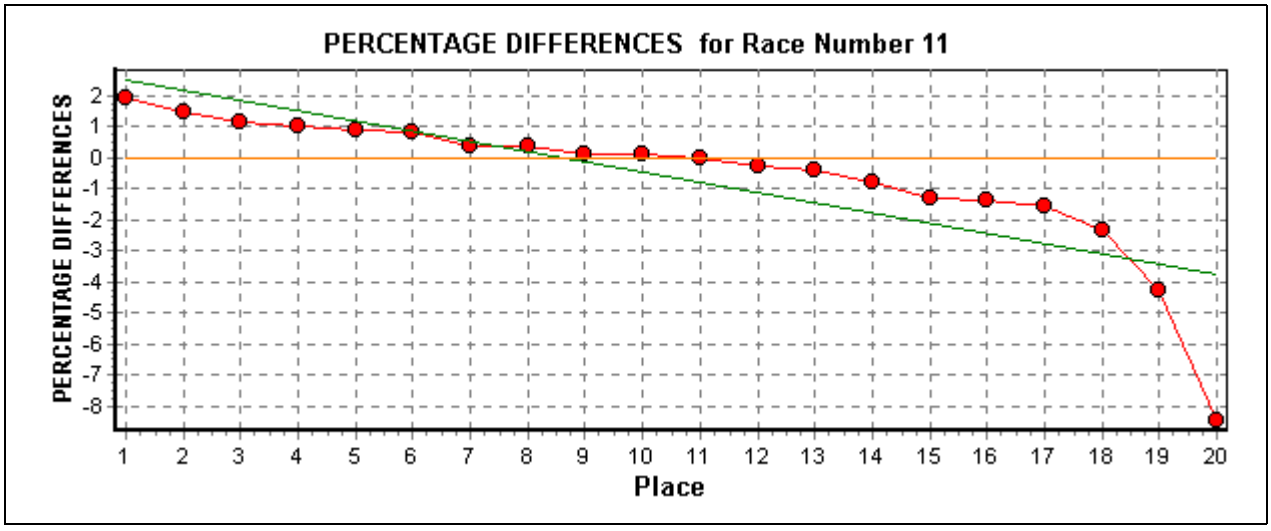

Figure 8 Percentage Difference for all place getters in Race 11

Percentage Difference =100 * (HC – BCH)/HC.

For this particular race the first place getter was 2% above his HC and the second place getter 1.5% above his HC. Not surprisingly the middle of the fleet sailed to their HC and so recorded a very small Percentage Difference.
For most races the winning percentage is approximately 3 to 4 percent i.e. the winner would performed 3 to 4 percent higher than his HC.

This Percentage Difference is a relative term in that it relates to each HC. For example, in this race the HC values for each competitor are show below and clearly have minimal influence on their placement in the race.

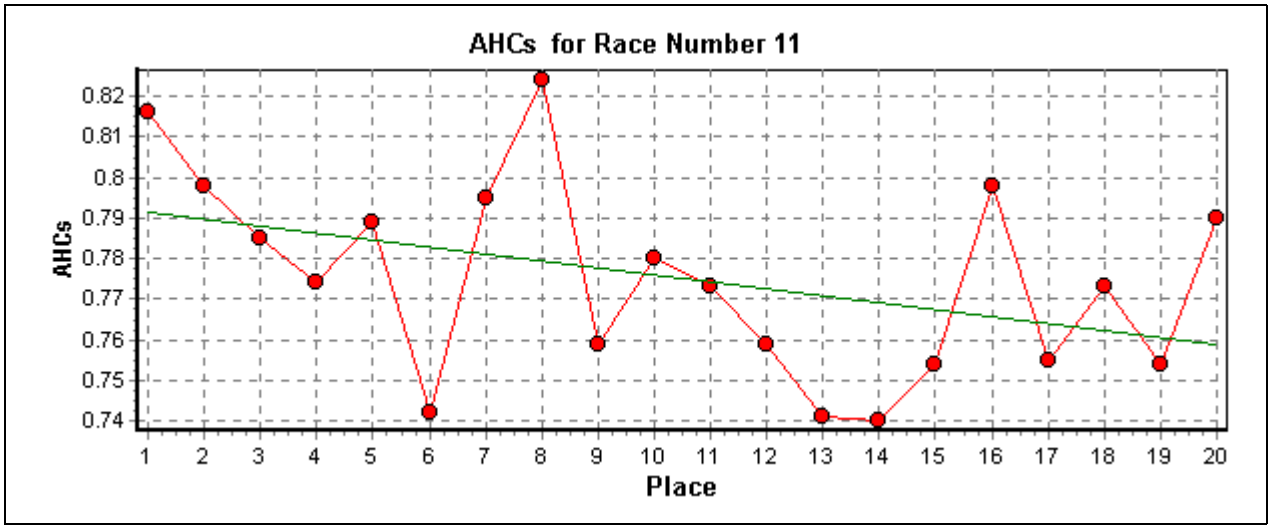

Figure 9 Handicaps for competitors in Race 11

This system of "he who sails most above his HC wins" works well ***but*** it is fair ***if and only if*** all competitors have the ***same*** normal spread of performances. But do they???

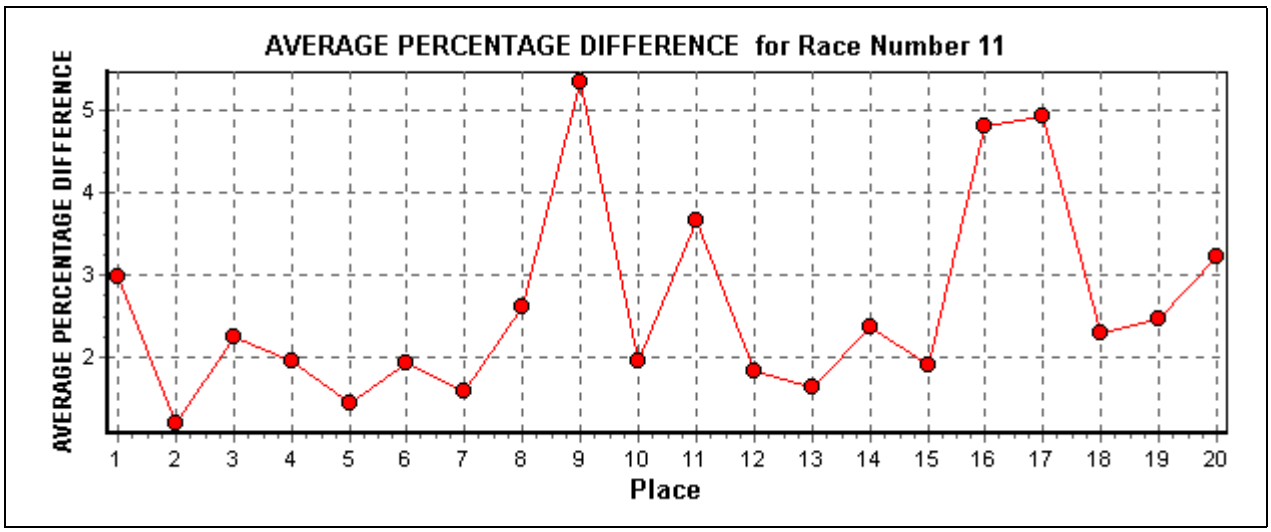

Figure 10 Average difference for competitors in Race 11

The Average Percentage Performance/Difference figures are established over many races and are ***not*** something derived from this single race.

Unfortunately, as the above graph clearly shows the competitors do ***not*** all have the same Average Percentage Performance. This fact I believe, ***totally*** undermines the premise on which conventional HCing is based.

The graph shows that the Average Percentage Performance Differences for competitors in this Series ranges from just over 1% to greater than 5%. This is independent of the competitor's actual HC. You can have two competitors both with the same HC (i.e. average performance) of say 0.900 but one achieves this by alternating between 0.945 and 0.855 i.e. plus and minus 5% where as another may only have a normal spread of performances of +/- 2% or 0.918 to 0.882.

Below are the performance graphs for two competitors, Felix Fellows and Charles Chase. [Real data from the test Series but names etc are changed.]

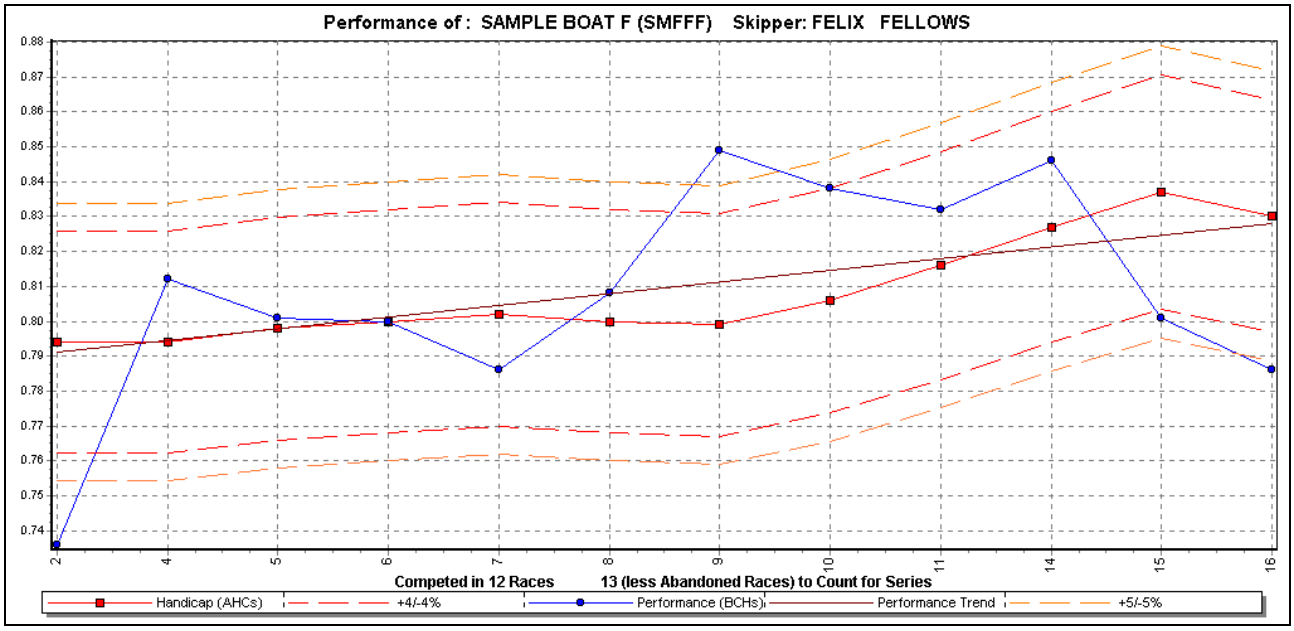

Figure 11 Performances of Felix Fellows for Test Series

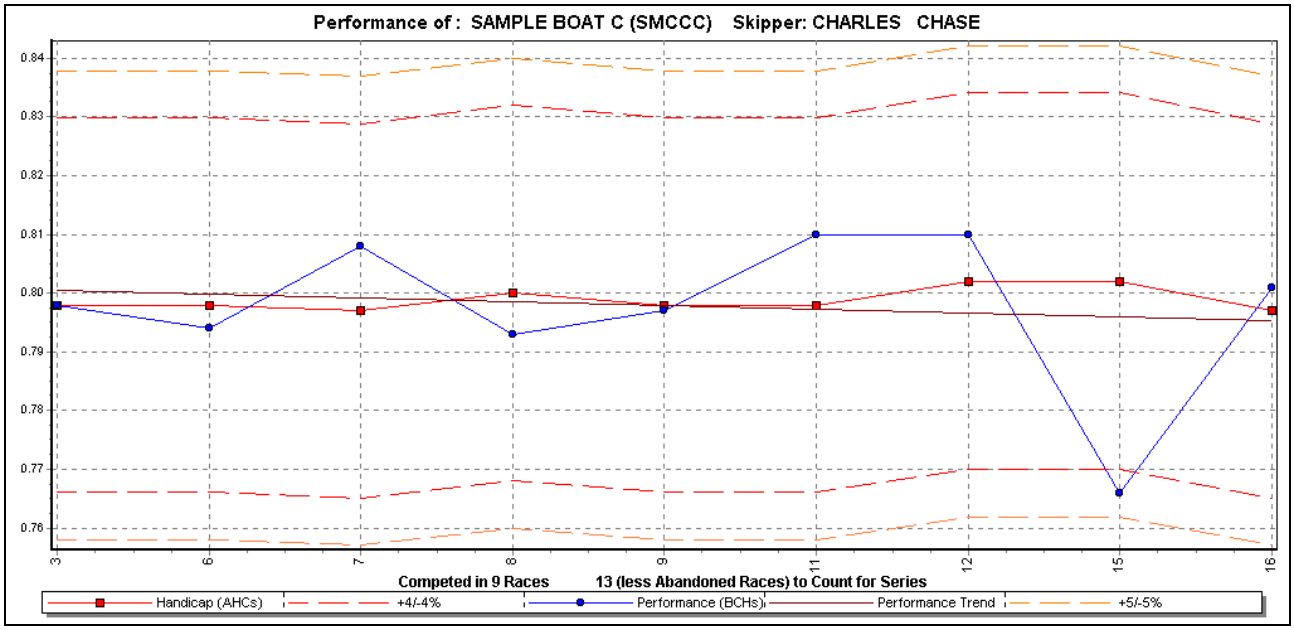

Figure 12 Performaces of Charles Chase for Test Series

Notice the relative consistency of Charles' graph.

Presented another way, here is the same data but just showing the actual Percentage Difference for the two competitors for the races in this Series.

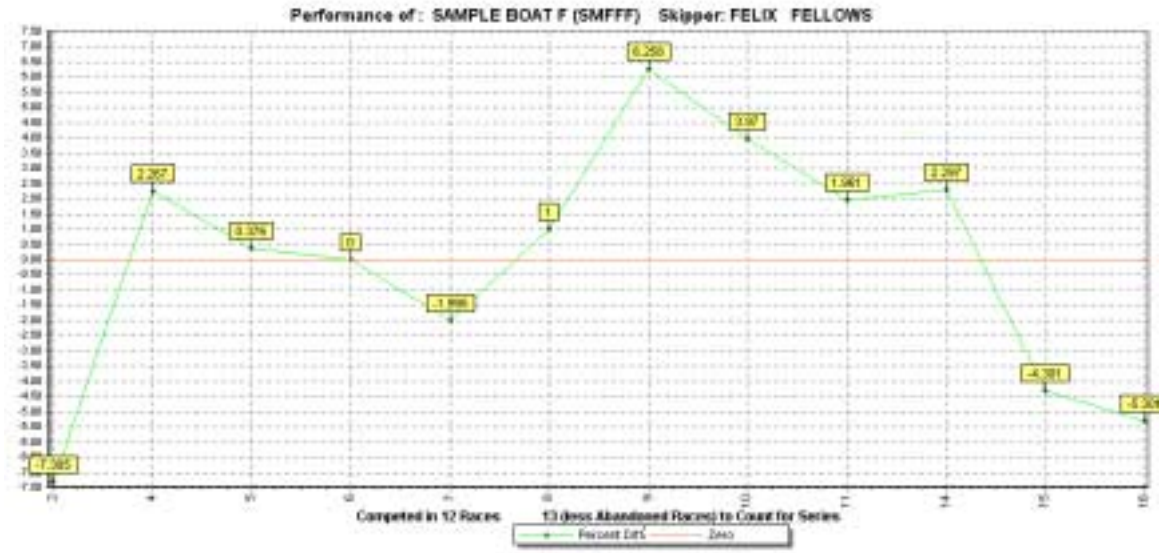

Figure 13 Performance Differences per race for Felix Fellows

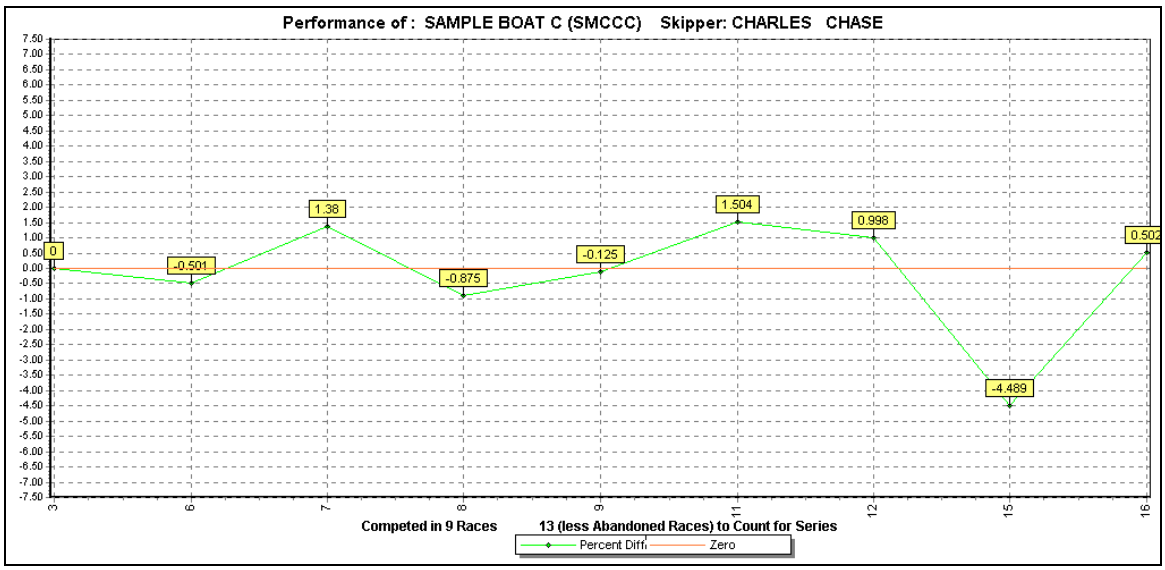

Figure 14 Performance Differences per race for Charles Chase

Felix Fellows is a fairly typical competitor with an Average Percentage Performance of 3.1%. That is he normally sails within a window of +/- 3.1% of his HC. Charles Chase has an average percentage performance spread of just 1.15%
This differences in their APP may be due to different boat types, different skill levels etc, etc. Whatever what the reason for the difference in APP, both sailors sail well and sail at around their HC. ***But*** under conventional Hcing, for the reason explained above, Charles is most *unlikely* to ever win a race.

In race 11 [the one used for these graphs] the first place getter was Felix Fellows, the second place was taken out by Charles Chase.

In this race Felix, with a performance of 1.96% above his HC, beat Charles who was 1.5% above his. Yet for Charles that was an exception performance i.e 1.5/1.15 = 1.3 times above his Average Percentage Performance . While for Felix his 1.97% is1.96/3.1 this is just 0.63 of his Average Percentage Performance.
This is real data and illustrates why skippers like Charles get frustrated and think that "normal" HC systems are biased against them.

By using Percentage Performance Handicapping a different place order can be calculated for each race. It is based on each competitor's performance against his normal spread of performances i.e. his Average Percentage Performance.
The normal HCing is run and then each competitor's Percentage Performance is calculated. PP = Percent Difference/Av Percentage Difference. Under this system the competitor with the largest (positive) PP wins.

Those who perform very well will have a PP of approx 1.4 , the middle of the fleet will be approximately 0 while those who performed poorly will have a negative PP value.
Under this system the competitor is competing against him self as much as against the rest of the fleet.

For those who prefer to think of winning on HC corrected time then the PPs can be converted to a time by the formula PPCorrected Time = (1 - 3 * PP/100) * (Base Corrected Time). The "3" comes from the observation that most competitors normal spread of performance is +/- 3% and is just a scale factor that does not affect who wins but rather it produces corrected times that are in the range that competitors expect.

The system has two down sides.

1. Each competitor now carries two HC parameters i.e. a HC and an Average Percentage Performance figure.
2. The average sailor may have trouble understanding this "less conventional" system?

Having said that, the Charles Chases of this world would probably welcome the system. Maybe it could be run as an alternate HC system in parallel with the standard HC system that it needs to run with anyway???

Q/ What effect does it have on the boats that perform more erratically whether by design or by helmsmanship?
A/ They can still win but they have to perform above their normal performance spread as does Charles.

Q/ What effects does it have on the Series Aggs compared to conventional HCing?
A/ Surprisingly little!

Q/ What effect does it have on most races?
A/ The place getter order will vary for a number of competitors while many will stay as they were.

Bear in mind that this system does not produce a new HC for the next race but rather Offers an alternative system of using the race data to determine who is deem the winner etc.

How does this form of HCing fair in the 2 criteria used above for comparing HC systems?
Division 1

Description	**Stan Dev scores 1~10 Agg Places**	**No Diff Place getters**
4%Clamps, 5%Low Lim, 5% Hi L	14.2	10.7
4%Clamps, 5%Low Lim, 5% Hi L, PerPerf	12.9	10.9
VYC	13.8	10.8

Division 2

Description	Stan Dev scores 1~10 Agg Places	No Diff Place getters
4BCH, 4%Clmps, 5%LL, 5%UL	20.5	11.1
4%Clamps, 5%Low Lim, 5% Hi L, PerPerf	21.2	10.8
VYC	24.6	10.1

Reasonably well against the standard definition of "better" as above, particularly for Div 1.

Over all Conclusions.

See also the "Overview" section earlier in this document.
Note: all the HC systems tested used a single HC parameter [except the Percent Performance system].

I set out learn which HC system provided the best results and/or how to tweak the parameters for the "best" results.
Rather a lot of work later ……
Two thing stands out above all else.

1. With a single HC parameter there is a definite limit beyond which no amount of adjusting of the various HC parameters can "improve" the HCing.
2. Provided that sensible parameters are set, then most of the systems tested provided *reasonable* HC results. Depending on your club's HC philosophy different systems/tweaks offer different advantages..

Comparisons.

To judge the relative performances of the various systems under test, it is now necessary to select the criteria by which to judge them.
Let us define the aim as….
"*to give all competent competitors an even chance to win a race and to win the Series if they sail at their best*"
I believe this translates to…
A small spread of Aggs
and a large number of *different* winners, *different* second place getters etc
and the use of a low limit of 5% to stop performances below this from altering a competitor's HC downwards to a unrealistic value.
These criteria having been selected, this reduced the test results to be compared to 40ish alternatives. Even this is was a daunting task so further simplification was needed. Relatively similar systems were ignored and the total reduced to 25.

The results are presented as the last 3 pages of this document.
One page sorts the systems by "least spread of Aggs" [remember this is standard deviation so you need to multiple this by approx 2.5 to get the likely spread from 1st to 10th in the test data.]
A second page is sorted by largest number of different competitors who received a first/2nd/3rd etc through to 10^{th} . There is not much in this with many tied values!!
The final page is the combined order.
After the initial comparisons where printed it suggested that Percentage Performance was worth further exploration. To this end the "Percentage Performance Corrected Time" method mentioned above was implemented and further tests run. The earlier Tests with a fixed Average Percentage Performance are labelled "PercPerf"; the later tests with an automatically updating Average Percentage Performance are labelled "AutoPP" or "AutoPercPerf" in the results table below.

Warning: If you try a new HC system then there is *every* possibility that competitors' HCs will change from "accepted" values. This may be necessary to produce different outcomes to your club's current system. Competitors must understand that it is unlikely that they can "have their old HC" ***and*** have "improved" HCing as well. The possible exception is the Percentage Performance system which is an "add on" to other HC systems ***but*** it has two parameters a HC and an Average Percentage Performance figure.

Rods Comments/Conclusions…

Two things are very obvious.

1. The exponential systems cleared faired the worst under both criteria. So unless the major consideration is competitors wanting a system that raises their HC after a good performance and conversely, then maybe give this one a miss.
2. Systems that used more than 4 races in the averaging system did not compare well against those that used 4. But, as mentioned previously, the HCs did change significantly race by race when only 4 races are averaged.

Before further explanation it must be remembered that some tests used data (BCHs) from races in a *previous* Series while many tests just used the data available from "In (the) Series" to calculate new HCs. One real concern with any new Series is the validity of the Initial HCs of all the competitors. The "Fast Start " system which more quickly adjusts HCs in the first few races seemed to improve that situation particularly in the "In Series" calculations. Alternately, or as well, the "Trend Bias" system also seem to improve the results under both criteria used by the tests.

For a simple system the system number 11 [a running average over the last 4 races, 4% clamps 5% limits] is easy to run and relatively easy to explain to the competitors.

If you aim to reward those "who deserve to win" then it appears that the Percentage Performance system is well worth considering as it not only appears to meet that aim but also faired very well on the two test criteria. [Remember it is an "add on" to any HC system. It is ***not*** stand alone.]

The outstanding successes are clearly the Percentage Performance, the Trend Bias and the Fast Start. Individually they tightened the Aggs. Collectively they tightened the Aggs and provided a larger number of different winners, second placegetters etc.

From a mathematical point of view I personally have reservations about both the Fast start and the Trend Biases for reasons noted above, but they certainly did well in the tests.

Well there you have it. Vastly more work then I ever intended; but the more I did, the more intriguing it became!. Because the volume of data grew to way beyond my expectations, there must be a real possibility of a occasional translation error as I manipulated and tabulated all the data. I tried to double check everything, but there may be an occasion(s) in which I failed.

Where to from here??

To improve measured performance HCing systems two things (at least) must happen.

1. Clubs must define what they expect the HC system to do for them.
2. Maybe the addition of a second HC parameter is required whether that be the Percent Performance concept or maybe an average wind speed factor or something else. [I suspect that an "Average Wind Speed Factor " may have value for shortish i.e. 3hr or less races around the same type of set course with a reasonably consistent breeze??? Yeah, I know, we all wish….]

The graphs / tables on which this report is based are freely available upon request but I believe that the summary sheets provide the bulk of the more interesting information
.

All feedback and/or any further HCing ideas would be warmly welcomed and explored where possible.

Rod McCubbin
TopYacht software
2001-05-22
Melbourne

Sorted by Spread of Aggs

Sys No	Description	Div 1 Sprd Aggs	Div 2 Sprd Aggs	Sum Sprd Aggs	Order Sprd Aggs	Div 1 No. of Dif Place Getters	Div 2 No. of Dif Place Getters	Sum No. of Dif Place Getters	Order Dif Place Getters	Sum Orders	O/A Order
50	4B,4Clmps,5Lims,InSers,FS,AutoPP	13.0	18.6	31.6	**1**	11.8	10.8	22.6	**3**	4	**1**
13	4B,4Clmps,5Lims.InSers,FstStrt	14.3	18.5	32.8	**2**	11.0	9.8	20.8	**24**	26	**14**
18	4B,4Clmps,5Lims,TrendBias	14.6	19.1	33.7	**3**	11.3	11.1	22.4	**6**	9	**3**
15	4B,4Clmps,5Lims,RecBias	13.6	20.2	33.8	**4**	10.7	10.5	21.2	**20**	24	**11**
20	4B,4Clmps,5Lims,2P	13.2	20.7	33.9	**5**	11.4	11.0	22.4	**6**	11	**6**
49	4B,4Clmps,5Lims,AutoPercPerf	13.2	20.8	34.0	**6**	11.1	11.5	22.6	**3**	9	**3**
43	4B,4Clmps,5Lims,PercPerf	12.9	21.2	34.1	**7**	10.9	10.8	21.7	**13**	20	**8**
51	4B,4Clmps,5Lims,InSers,FS,AutoPP,F	16.7	17.4	34.1	**7**	11.7	11.0	22.7	**1**	8	**2**
19	4B,4Clmps,5Lims,1P	13.7	20.7	34.4	**9**	11.6	11.1	22.7	**1**	10	**5**
11	4B,4Clmps,5Lims	14.1	20.5	34.6	**10**	10.7	11.1	21.8	**12**	22	**10**
48	4B,4Clmps,5Lims,AutoPercPerf	13.5	21.6	35.1	**11**	11	11.2	22.2	**8**	19	**7**
37	4B,1.5LC,3UC,4LL,5UL,PushUp	14.1	21.1	35.2	**12**	10.8	10.9	21.7	**13**	25	**13**
12	4B,4Clmps,5Lims,InSers	15.3	20.3	35.6	**13**	10.8	11.4	22.2	**8**	21	**9**
16	4B,4Clmps,5Lims,RecBias.InSers	15.2	20.5	35.7	**14**	10.8	10.9	21.7	**13**	27	**15**
22	4B,4Clmps,5Lims,InSers,2P	15.3	20.5	35.8	**15**	10.6	11.1	21.7	**13**	28	**16**
14	4B,4Clmps,5Lims,InSers,BnL	15.3	20.6	35.9	**16**	10.9	10.8	21.7	**13**	29	**19**
21	4B,4Clmps,5Lims,InSers,1P	16.0	20.5	36.5	**17**	10.9	11.1	22.0	**11**	28	**16**
29	4B,4Clmps,5Lims,0.5MinStep	14.9	21.7	36.6	**18**	10.9	11.2	22.1	**10**	28	**16**
52	4B,4Clmps,5Lims,InSers.TrendB	15.9	21.2	37.1	**19**	11.2	11.3	22.5	**5**	24	**11**
35	VYC	13.8	24.6	38.4	**20**	10.8	10.1	20.9	**23**	43	**22**
32	VYC, BUT 3%Clamps	13.4	25.2	38.6	**21**	10.8	10.4	21.2	**20**	41	**21**
36	VYC,1P	13.9	24.9	38.8	**22**	11.1	10.4	21.5	**18**	40	**20**
40	Expo,G=4,4Clmps,5Lim	16.2	23.3	39.5	**23**	10.1	10.1	20.2	**25**	48	**25**
2	8B,3Clpms,2LD,1HD,5LL,7UL	14.6	26.7	41.3	**24**	11.2	10.1	21.3	**19**	43	**22**
41	Expo,G=5,4Clmps,5Lim	18.6	25.0	43.6	**25**	10.1	10.8	20.9	**22**	47	**24**

KEY
4B = 4 Races of BCHs
4Clamps = Upper & Lower clamps at 4
5Lims = Upper & Lower Limits = 5%
0.5MinStep = Min Step size set to 0.5%
1.5LC = 1.5% Lower Clamp
3UC = 3% Upper Clamp
2LD = 2 Lowest BCHS Discarded
1HD = 1 Highest BCH discarded
G=5 = Gain of Exponential Average = 5
Trend = Trend Biased
RecBias = Recent Biased
FS = Fast Start
FstStrt = Fast Start
InSers = HCing only used data from In this Series
1P First Place Correction
AutoPercPerf = Auto update Percentage Performance
PerPerf = Fixed Av Percent Performance
BnL = Biggest & Littlest BCHs halved in average
Expo = Exponential Averaging

Sorted by number of *Different* Competitors with a first etc

Sys No	Description	Div 1 Sprd Aggs	Div 2 Sprd Aggs	Sum Sprd Aggs	Order Sprd Aggs	Div 1 No. of Dif Place Getters	Div 2 No. of Dif Place Getters	Sum No. of Dif Place Getters	Order Dif Place Getters	Sum Orders	O/A Order
51	4B,4Clmps,5Lims,InSers,FS,AutoPP,F	16.7	17.4	34.1	**7**	11.7	11.0	22.7	**1**	8	**2**
19	4B,4Clmps,5Lims,1P	13.7	20.7	34.4	**9**	11.6	11.1	22.7	**1**	10	**5**
50	4B,4Clmps,5Lims,InSers,FS,AutoPP	13.0	18.6	31.6	**1**	11.8	10.8	22.6	**3**	4	**1**
49	4B,4Clmps,5Lims,AutoPercPerf	13.2	20.8	34.0	**6**	11.1	11.5	22.6	**3**	9	**3**
52	4B,4Clmps,5Lims,InSers.TrendB	15.9	21.2	37.1	**19**	11.2	11.3	22.5	**5**	24	**11**
18	4B,4Clmps,5Lims,TrendBias	14.6	19.1	33.7	**3**	11.3	11.1	22.4	**6**	9	**3**
20	4B,4Clmps,5Lims,2P	13.2	20.7	33.9	**5**	11.4	11.0	22.4	**6**	11	**6**
12	4B,4Clmps,5Lims,InSers	15.3	20.3	35.6	**13**	10.8	11.4	22.2	**8**	21	**9**
48	4B,4Clmps,5Lims,AutoPercPerf	13.5	21.6	35.1	**11**	11	11.2	22.2	**8**	19	**7**
29	4B,4Clmps,5Lims,0.5MinStep	14.9	21.7	36.6	**18**	10.9	11.2	22.1	**10**	28	**16**
21	4B,4Clmps,5Lims,InSers,1P	16.0	20.5	36.5	**17**	10.9	11.1	22.0	**11**	28	**16**
11	4B,4Clmps,5Lims	14.1	20.5	34.6	**10**	10.7	11.1	21.8	**12**	22	**10**
43	4B,4Clmps,5Lims,PercPerf	12.9	21.2	34.1	**7**	10.9	10.8	21.7	**13**	20	**8**
37	4B,1.5LC,3UC,4LL,5UL,PushUp	14.1	21.1	35.2	**12**	10.8	10.9	21.7	**13**	25	**13**
16	4B,4Clmps,5Lims,RecBias.InSers	15.2	20.5	35.7	**14**	10.8	10.9	21.7	**13**	27	**15**
14	4B,4Clmps,5Lims,InSers,BnL	15.3	20.6	35.9	**16**	10.9	10.8	21.7	**13**	29	**19**
22	4B,4Clmps,5Lims,InSers,2P	15.3	20.5	35.8	**15**	10.6	11.1	21.7	**13**	28	**16**
36	VYC,1P	13.9	24.9	38.8	**22**	11.1	10.4	21.5	**18**	40	**20**
2	8B,3Clpms,2LD,1HD,5LL,7UL	14.6	26.7	41.3	**24**	11.2	10.1	21.3	**19**	43	**22**
32	VYC, BUT 3%Clamps	13.4	25.2	38.6	**21**	10.8	10.4	21.2	**20**	41	**21**
15	4B,4Clmps,5Lims,RecBias	13.6	20.2	33.8	**4**	10.7	10.5	21.2	**20**	24	**11**
35	VYC	13.8	24.6	38.4	**20**	10.8	10.1	20.9	**23**	43	**22**
41	Expo,G=5,4Clmps,5Lim	18.6	25.0	43.6	**25**	10.1	10.8	20.9	**22**	47	**24**
13	4B,4Clmps,5Lims.InSers,FstStrt	14.3	18.5	32.8	**2**	11.0	9.8	20.8	**24**	26	**14**
40	Expo,G=4,4Clmps,5Lim	16.2	23.3	39.5	**23**	10.1	10.1	20.2	**25**	48	**25**

KEY
4B = 4 Races of BCHs
4Clamps = Upper & Lower clamps at 4
5Lims = Upper & Lower Limits = 5%
0.5MinStep = Min Step size set to 0.5%
1.5LC = 1.5% Lower Clamp
3UC = 3% Upper Clamp
2LD = 2 Lowest BCHS Discarded
1HD = 1 Highest BCH discarded
G=5 = Gain of Exponential Average = 5
Trend = Trend Biased
RecBias = Recent Biased
FS = Fast Start
InSers = HCing only used data from In this Series
1P First Place Correction
AutoPercPerf = Auto update Percentage Performance
PerPerf = Fixed Av Percent Performance
BnL = Biggest & Littlest BCHs halved in average
Expo = Exponential Averaging

Sorted by Overall Order

Sys No	Description	Div 1 Sprd Aggs	Div 2 Sprd Aggs	Sum Sprd Aggs	Order Sprd Aggs	Div 1 No. of Dif Place Getters	Div 2 No. of Dif Place Getters	Sum No. of Dif Place Getters	Order Dif Place Getters	Sum Orders	O/A Order
50	4B,4Clmps,5Lims,InSers,FS,AutoPP	13.0	18.6	31.6	**1**	11.8	10.8	22.6	**3**	4	**1**
51	4B,4Clmps,5Lims,InSers,FS,AutoPP,F	16.7	17.4	34.1	**7**	11.7	11.0	22.7	**1**	8	**2**
49	4B,4Clmps,5Lims,AutoPercPerf	13.2	20.8	34.0	**6**	11.1	11.5	22.6	**3**	9	**3**
18	4B,4Clmps,5Lims,TrendBias	14.6	19.1	33.7	**3**	11.3	11.1	22.4	**6**	9	**3**
19	4B,4Clmps,5Lims,1P	13.7	20.7	34.4	**9**	11.6	11.1	22.7	**1**	10	**5**
20	4B,4Clmps,5Lims,2P	13.2	20.7	33.9	**5**	11.4	11.0	22.4	**6**	11	**6**
48	4B,4Clmps,5Lims,AutoPercPerf	13.5	21.6	35.1	**11**	11	11.2	22.2	**8**	19	**7**
43	4B,4Clmps,5Lims,PercPerf	12.9	21.2	34.1	**7**	10.9	10.8	21.7	**13**	20	**8**
12	4B,4Clmps,5Lims,InSers	15.3	20.3	35.6	**13**	10.8	11.4	22.2	**8**	21	**9**
11	4B,4Clmps,5Lims	14.1	20.5	34.6	**10**	10.7	11.1	21.8	**12**	22	**10**
52	4B,4Clmps,5Lims,InSers.TrendB	15.9	21.2	37.1	**19**	11.2	11.3	22.5	**5**	24	**11**
15	4B,4Clmps,5Lims,RecBias	13.6	20.2	33.8	**4**	10.7	10.5	21.2	**20**	24	**11**
37	4B,1.5LC,3UC,4LL,5UL,PushUp	14.1	21.1	35.2	**12**	10.8	10.9	21.7	**13**	25	**13**
13	4B,4Clmps,5Lims.InSers,FstStrt	14.3	18.5	32.8	**2**	11.0	9.8	20.8	**24**	26	**14**
16	4B,4Clmps,5Lims,RecBias.InSers	15.2	20.5	35.7	**14**	10.8	10.9	21.7	**13**	27	**15**
29	4B,4Clmps,5Lims,0.5MinStep	14.9	21.7	36.6	**18**	10.9	11.2	22.1	**10**	28	**16**
21	4B,4Clmps,5Lims,InSers,1P	16.0	20.5	36.5	**17**	10.9	11.1	22.0	**11**	28	**16**
22	4B,4Clmps,5Lims,InSers,2P	15.3	20.5	35.8	**15**	10.6	11.1	21.7	**13**	28	**16**
14	4B,4Clmps,5Lims,InSers,BnL	15.3	20.6	35.9	**16**	10.9	10.8	21.7	**13**	29	**19**
36	VYC,1P	13.9	24.9	38.8	**22**	11.1	10.4	21.5	**18**	40	**20**
32	VYC, BUT 3%Clamps	13.4	25.2	38.6	**21**	10.8	10.4	21.2	**20**	41	**21**
2	8B,3Clpms,2LD,1HD,5LL,7UL	14.6	26.7	41.3	**24**	11.2	10.1	21.3	**19**	43	**22**
35	VYC	13.8	24.6	38.4	**20**	10.8	10.1	20.9	**23**	43	**22**
41	Expo,G=5,4Clmps,5Lim	18.6	25.0	43.6	**25**	10.1	10.8	20.9	**22**	47	**24**
40	Expo,G=4,4Clmps,5Lim	16.2	23.3	39.5	**23**	10.1	10.1	20.2	**25**	48	**25**

KEY
4B = 4 Races of BCHs
4Clamps = Upper & Lower clamps at 4
5Lims = Upper & Lower Limits = 5%
0.5MinStep = Min Step size set to 0.5%
1.5LC = 1.5% Lower Clamp
3UC = 3% Upper Clamp
2LD = 2 Lowest BCHS Discarded
1HD = 1 Highest BCH discarded
G=5 = Gain of Exponential Average = 5
Trend = Trend Biased
RecBias = Recent Biased
FS = Fast Start
FstStrt = Fast Start
InSers = HCing only used data from In this Series
1P First Place Correction
AutoPercPerf = Auto update Percentage Performance
PerPerf = Fixed Av Percent Performance
BnL = Biggest & Littlest BCHs halved in average
Expo = Exponential Averaging