Sailboat handicap racing

Some people like to race sailboats. Many of these people would like the best sailor to win. This is different from other types of sailboat racing. For example the America's Cup, where winning is also based on boat design, which can dominate the contest. One-design racing is the gold standard for a best-sailor-wins race. Everyone is sailing the same boat. Mostly. In reality, one-design classes vary in how exactly-the-same they require boats to be. Some one-design classes allow and favor the sailors with the newest best sails, the best equipment, the lightest boat, the best tuned boat, or the best kept boat. It's hard to keep money from being a factor. But still, one-design is as good as it gets. When sailors race one design boats, there are other factors that can keep the best sailor from winning, such as bad luck. One part of sailboat racing is predicting future wind and sea conditions across a course, which is not an exact skill. As courses get longer and involve land (islands, etc.) the luck factor is more significant. As courses become shorter, and simpler, like a windward-leeward buoy course, the luck factor is reduced.

While one design racing is the best format for best-sailor-wins racing, many sailors do not own one design class boats. Why not race the boat you do own? The first problem is that some boats are a lot faster than other boats and would simply win most of the time, unless we handicap them. Handicapping is any system that tries to compensate for differences in boat performance. Handicapping systems "model" a boat's performance so that different boats can be compared, the boat performance itself is factored out, and the best sailor can win. An accurate handicapping system should replicate the crew-skill-based-results of one-design racing. Unfortunately this is impossible to achieve.

Most popular today are single value handicap systems like PHRF. A single number is your "rating" for a race. They are simple to use. But they have significant inaccuracies in modeling boat performance across a range of conditions and across a range of different performing boats. As a result, race victories are heavily influenced by the actual conditions experienced in a race. In general, rating system accuracy gets worse, proportional to how much the actual conditions (TWS, TWA, sea state) deviate from the rated average, and proportional to the range of boat performance dissimilarity among the racers. An additional loss of accuracy for dissimilar boats is simply that dissimilar boats are on a course at different times and for different amounts of time. Winds can build or drop or shift. Sea states can change. Dissimilar boats do not sail in the same conditions.

All handicap systems in use today are either Time-On-Distance (TOD) or Time-On-Time (TOT). TOD is the most popular among Southern California PHRF monohull races. Its accuracy comes from the idea that on-average, all boats sail at one speed per boat. If you know a boat's one average speed (in all conditions) you can calculate how long it will take to sail 1 mile and compare this to how long it will take other boats to sail 1 mile, and represent this as a

seconds/mile difference. The rating system picks a scratch boat and rates it 0 seconds/mile. All other boats are compared to it and rated by how many seconds/mile faster or slower they are. So a boat who rates 25, is 25 seconds/mile slower than a 0 rated boat. Or 50 seconds per mile slower than a -25 rated boat. You might think that boats don't spend all their time sailing at one speed. But displacement boats have a hull speed upper limit. And one that is slow enough to be encountered frequently. Once they hit this speed, they do not go faster in more wind. As we imagine a race with a lot of different wind and course conditions, we can imagine that a fair amount of time is spent at hull speed, and therefore a rating system based on one average speed per boat will have a useful accuracy at modeling boat performance. TOD is most accurate for boats that spend the largest percentage of their race time at hull speed. CT = ET + (race distance * TOD rating)

But not all boats are displacement monohulls. Ultra Light Displacement Boats, sport boats, planning monohulls, sprit boats, are all types of monohulls who can get past a simple hull speed limit. They do go faster in more wind, especially for the wind ranges normally raced in. Multihulls fit in this group. For these boats, TOD is a poor choice from an accuracy standpoint. This leads us to TOT.

TOT does not assume boats will sail at one average speed in all conditions. TOT assumes that a boat will sail faster with more wind, along a linear curve. Your speed changes with wind strength. Again a scratch boat is chosen and is rated 1.0000. A boat that sails twice as fast as the scratch boat (on average in all conditions) would rate 2.000. A boat that sails half as fast would rate 0.5000. TOT is less accurate for boats limited by a hull speed, and more accurate for boats whose speed is linearly proportional to TWS over average racing conditions. But boats don't really have a simple linear performance curve. Every reef or sail change is a discontinuity and a change in its performance curve, as is a mode change like planning or flying a hull. Performance curves really change with every TWA change. TOT is less accurate when boats sail in very different TWS from each other. CT = ET * TOT rating.

The final accuracy issue for TOD & TOT systems, is how is the rating determined in the first place? VPP stands for Velocity Prediction Program, software used by boat designers to produce polar diagrams for a specific boat design. As you may know, polar diagrams are a common way for showing predicted boat speed at a variety of TWS, TWA, and sail combinations. VPP software is not perfect to begin with. Worse, if the input data is incorrect, so are the calculated results. Boats can evolve beyond a VPP's capabilities, with lifting foils and canting masts. Design VPPs do not produce a handicap rating. We can imagine picking our average TWS, and from typical course estimates coming up with time spent at different TWAs, then averaging the VPP data for these conditions into a single value rating. If the TWS and TWA assumptions are perfectly matched for all boats in a specific race, then the only error would be the VPP

calculation error. While boat designer VPPs are not used for ratings as far as I know, there are rating organizations that have VPP based software for calculating ratings. BAMA (Bay Area Multihull Association) & ORCA (Ocean Racing Catamaran Association, an odd name considering the number of trimarans rated), OMR (Australia), and Texcel (Dutch) are some examples. All calculate ratings from raw boat data. The most important factors being sail area, hull length and boat weight in racing condition.

The opposite method for rating a boat is to use a human estimate. A "rater" looks at a boat and says "it looks like a -10". He is basing his rating on his knowledge of a lot of other boats and their actual race performance. He may have VPP type software to guide his decision. He may have performance and rating data for specific models and brands of boats. In considering actual race performances, he must judge how much of the actual race results are due to boat design, boat condition, crew performance, wind conditions, and random factors. He wants the rating to be on boat performance only. It is easier for him when he has a lot of racing data for the same model boat, and harder for him when he is rating one-off designs. Perhaps he works with a committee of raters. Perhaps a sailor can appeal his rating. The rater(s) will review the sailor's boat and race performance. He/they will judge his sailing ability and boat condition so as to try and determine the actual boat's design-based performance.

How can we improve the accuracy of a rating system?

I think a lot of people have no idea that their system isn't already "perfectly" accurate. People who are passionate about handicap race results being accurate, haven't really considered the realities described above. Handicap racing is not and cannot be very accurate, by one-design standards. You have to get your head wrapped around all the inaccuracy issues, get a feel for which are biggest, and go after them first. Improving trivial errors won't have much effect on the end result.

PHRF has 3 ratings per boat; Offwind, Random Leg, & around Buoys. By matching these to a race course they are able to better predict the TWA average for rating that type of course. This should improve displacement boat rating accuracy using TOD, when you guess the race type correctly. There are always exceptions. Like the 2014 N2E race which used Offwind ratings, but it was actually an upwind, then reaching race, in atypical winds.

Portsmouth has wind ranges to assign to a race, and different ratings per wind range. Which is useful for 20 minute long windward/leeward courses in stable winds. In a typical multi-hour coastal race the wind will vary thru many TWAs and TWSs and be different for fast versus slow boats.

What else can we do?
Wherever possible, set up classes with the smallest rating ranges and the most similar boats.
Separate non-displacement boats from displacement boats, into different classes.
Use TOT for non-displacement boats and multihulls
It never hurts to improve the accuracy of the initial rating. Better data for example.

Other variations and issues.
There are a number of topics in handicapping revolving around "why does my rating suck".
The first possibility is that it doesn't really suck. Accuracy is not achievable. People can enjoy poker even if they don't get great cards every hand. If the conditions fall against you, someone else will have the advantage.

Objective versus Subjective systems –Objective ratings means the numbers are plugged into formula, and the chips fall where they may. Subjective ratings mean there are people judging what your rating should be. With Subjective ratings there are always issues with human error, human bias, and lack of consistency. This is usually where "fair" is raised as an issue. "Fair" means applying a system equally to everyone. Sometimes Subjective systems have the perception of favoring some people over others. They must judge how well you sail, in order to adjust your rating from actual results. People who appeal and negotiate better may have advantages. There can be the perception of politics. We are all familiar with tales of boats saddled with a bad rating and what it does to resale value, or vice versa. On the other hand, ORCA ratings are Objective only. Some systems may be a mix.

Golf handicaps – The purpose of a golf handicap sailboat rating is to let everyone have an equal chance of winning, regardless of boat, boat condition or sailing skill. This is typically accomplished by adjusting ratings after each race, by a fairly significant amount. A large, frequent, adjustment to or towards the fleet average, so the participants have a better chance in the next race.

Hybrid systems. – Rating systems that use more than one technique. For example, ORCA's creators were aerospace scientists concerned with 2 problems; improving VPP based calculated rating accuracy using actual race data, and avoiding the problems of subjectively guessed ratings and adjustments. VPP errors arise not only because the software did not calculate the correct rating, but also because the "conditions" assumptions do not match what is actually raced in. ORCA creators thought that if a lot of race data was used, it would refine both the condition assumption errors and any calculation/data errors, converging on more accurate local ratings. So they developed an adjustment methodology that adjusts a boat's rating once a quarter, from all races in that period. Maximum adjustment is about 1% per race, with most

adjustments much smaller (new boats adjust heavier initially). Is this also adjusting for poor boat condition and poor sailing skills? Unfortunately yes, because there is no objective way to interpret race data for boat performance only. In practice most boats converge on a small range of ratings. An improving skipper can watch his rating go up. This system favors whoever out-sails their own average performance by the most in a race. And any rating complaint can be addressed by doing more races.

Bill Gibbs
ORCA rating chairman 6/23/14