Notes on Performance Handicap Systems in Yachting

R.E. Deakin[1] and R.A. Green[2]

[1] Dunsborough, WA, 6281, Australia; [2] Sandringham, VIC, 3191, Australia
email: randm.deakin@gmail.com

Abstract

Performance Handicap Systems in yachting yield a handicap (a number) that enables yachts of varying speed potential to compete in races where the yacht with the least corrected time wins – where corrected time is elapsed time multiplied by the handicap. After a race, *calculated handicaps* are obtained from a set of rules applied after the *standard corrected time* of the race has been determined. This paper demonstrates new methods of determining the standard corrected time and calculated handicaps as well as reviewing some current methods.

Keywords

Performance Handicap System, Standard Corrected Time, Back Calculated Handicap, Allocated Handicap, Calculated Handicap, mark boat, Performance Indicator Filter

Introduction

A Performance Handicap System (PHS) in yachting is a set of rules and mathematical calculations that enable yachts of varying speed potential to compete in races where the yacht with the least corrected time is the winner. The PHS produces a *handicap* which is a number, usually somewhere between 0.750 and 1.250. In local terminology (Australia) this handicap is known as the *allocated handicap AHC* and the yacht's *elapsed time ET* multiplied by the allocated handicap yields the *corrected time CT,* or

$$CT_k = ET_k \times AHC_k \tag{1}$$

where the subscript k denotes the kth boat in the fleet of n yachts and $k = 1, 2, 3, \ldots, n$ and

$$ET_k = \text{ finish time of } k^{th} \text{ yacht} - \text{start time} \tag{2}$$

So, the handicap is a numerical value that is a measure of the performance of both the yacht and the crew

> [A yacht's PHS derived handicap is different from a yacht's *rating* which is a numerical measure of potential speed based upon the yacht's parameters, e.g. waterline length, beam, displacement, sail area, etc. and a sequence of mathematical formula related to the physics of hydrodynamics and aerodynamics as applied to yachting force models (*World Sailing*[1] 2019). There are several measurement systems that give yacht ratings, e.g., The International Offshore Rule (IOR), the Chanel Rating System (CHS), the International Measurement System (IMS) and the International Rating Certificate (IRC). A popular measurement rating system in Australia and particularly in Victoria is the Australian Measurement System (AMS) administered by Yacht Racing Services Association Inc. (YRSA) for the Australian yachting community. This paper is not concerned with measurement systems or the ratings derived from them.]

The essence of a Performance Handicap System are the rules that enable the adjustment of handicaps after racing. Some of these rules may be arbitrary, some could be based on experience and some could be in place to achieve desired outcomes. Indeed, a PHS used in one yacht club could be different from that used in another club; or the PHS used for a regatta could be different from the usual club PHS. And, as we will discuss, a *calculated handicap* produced by the PHS after a race is related to the number of yachts in that race; their handicaps; the handicap of a mythical or real *mark boat;* and the allowable changes in handicaps.

---

[1] *World Sailing* is the world governing body for the sport of sailing formed in 1907 and then known as the International Yacht Racing Union (IYRU). The name was changed to the International Sailing Federation (ISAF) in 1996 before adopting the name *World Sailing* in 2015.

The sequence of steps in calculating a yacht's new PHS handicap are

1   A *Standard Corrected Time STC* is established for the race and this is the corrected time of the **mark boat** (which may be real or mythical).

2   A *Back Calculated Handicap BCH* is derived from the rule $ET_k \times BCH_k = SCT$ and

$$BCH_k = \frac{SCT}{ET_k} \qquad\qquad (3)$$

    There may be some screening of the *BCH* to detect anomalous results.

3   A *Calculated Handicap CHC* is evaluated from a function of the allocated and back-calculated handicaps. There may be some further screening of the calculated handicap before it becomes the yacht's allocated handicap for the next race.

This paper will outline some of the usual methods of determining the standard corrected time including a new method we are proposing. And, in addition, we will provide an explanation of some methods used to obtain the calculated handicap as well as a new method we call a **Performance Indicator Filter**. Examples of the calculations of a Performance Handicap System will be given for a series of yacht races that we now describe.

The Yacht Race Series

The Sandringham Yacht Club (SYC) 2018-19 Summer Aggregate, Division 1, Performance Handicapping System (PHS) was a 10-race series with published results using *TopYacht*[2] software showing race and handicapping details for races 1,2,3,4,6,7 and 10 (see Appendix A). For these **seven races** we have selected the 10 yachts that completed at least five of these races: *Bandit, Conquest, Dark and Stormy, Dream, Esprit, Joust, Niche, Scarlett Runner II, Sierra Chainsaw* and *Wicked*. Each of these yacht's elapsed times *ET* have been extracted from the published race results and a yacht's initial handicap is taken as the allocated handicap *AHC* for that race shown in the race results. These data form our **Yacht Race Series** and are shown in Appendix B. Also, Appendix B contains the results of the Yacht Race Series scored by two different methods – one method commonly used in Victorian yachting and the other by the new methods we are proposing.

Nomenclature

The following notation has been used

Symbol	Meaning	Definition
α	weight factor in *EWMA*	
AHC	allocated handicap	
BCH	back calculated handicap	$BCH = SCT\big/ET$
CHC	calculated handicap	
CT	corrected time	$CT = ET \times AHC$
ET	elapsed time	$ET = $ finish time $-$ start time
EWMA	exponentially weighted moving average	
H	true handicap	$H = AHC + \delta H$
SCT	standard corrected time	
δH	small unknown correction to a handicap	
j, k	integer counters usually seen as subscripts	
K	multiplier in Performance Indicator Filter	

Symbol	Meaning	Definition
M	corrected time position of mark boat	
MAD	Median Absolute Deviation	
n	number of yachts in the race	
N	number of races in a yacht race series or number of races in an averaging process	
p	mark boat place in a fleet of 100 yachts	
PI	performance indicator	$PI = BCH - AHC$
PIF	Performance Indicator Filter	
PM	performance multiplier	
R	range of corrected times (positive quantity)	
S	sum of corrected times	
x	handicap error	
z	a numeric value between zero and $\pm PI$	$CHC = AHC + z$

Many of the formula in the text have subscripts denoting either a yacht in a race fleet of n yachts, or a race in a series of N races. For example, $BCH_k = \dfrac{SCT}{ET_k}$ where the subscript k denotes the kth boat in a fleet of $k = 1,2,3,...,n$ yachts in a race. If we are discussing the kth yacht's back calculated handicap in the jth race of a series $j = 1,2,...,N$ then we might write $BCH_{k,j} = \dfrac{SCT_j}{ET_{k,j}}$ and if these were tabulated in an array of rows and columns then there would be $k = 1,2,3,...,n$ rows and $j = 1,2,...,N$ columns. For clarity, we have mostly avoided the use of double subscripts but care should be taken in relating the subscripts to the context, and we have used subscript k to denote a yacht in a particular race, and j to denote a race in a series of N races.

Standard Corrected Time *STC* and Back Calculated Handicap *BCH*: Usual Methods

There are several ways of determining the Standard Corrected Time STC – hence defining the **mark boat**, real or mythical – and then determining the Back Calculated Handicaps BCH. We describe two representative methods that are applied to the corrected times of Race 1a of our Yacht Race Series (Appendix B) and shown in Table 1.

| | | Allocated | | | Back Calculated | |
| | Elapsed Time | Handicap | Corrected Time | | Handicap *BCH* | |
Yacht	*ET*	*AHC*	*CT* (sec)	Place	*World Sailing*	*TopYacht*
Sierra Chainsaw	1:18:59 (4739 sec)	0.930	4407.270	**1**	0.977	0.955
Joust	1:18:40 (4720)	0.935	4413.200	**2**	0.980	0.959
Scarlett Runner II	1:19:28 (4768)	0.935	4458.080	**3**	0.971	0.949
Wicked	1:21:27 (4887)	0.926	4525.362	**4**	0.947	0.926
Bandit	1:23:33 (5013)	0.910	4561.830	**5**	0.923	0.903
Dream	1:21:47 (4907)	0.948	4651.836	**6**	0.943	0.922
Esprit	1:26:58 (5218)	0.895	4670.110	**7**	0.887	0.867
Conquest	1:26:30 (5190)	0.914	4743.660	**8**	0.892	0.872
Dark and Stormy	1:34:59 (5699)	0.885	5043.615	**9**	0.812	0.794
Niche	DNS					

Table 1. Race 1a of the Yacht Race Series (Appendix B). Yachts shown in corrected time order. Time in hour (h), minute (m) second (s) format as h:mm:ss

Method [1]   *World Sailing* in their document *International Empirical Handicap Scheme for Yachts* (*World Sailing* 2016), say that a *standard corrected time SCT* of the fleet be the average of the corrected times of the *reduced fleet* where the reduced fleet excludes the lowest 20% and highest 40% of the race fleet based on their corrected times. In the case of Race 1a, where $n = 9$ yachts, *Conquest* and *Dark and Stormy* are excluded, as are *Sierra Chainsaw, Joust, Scarlett Runner II* and *Wicked*. The reduced fleet is *Bandit, Dream* and *Esprit* and the standard corrected time is

$$SCT = \frac{4561.830 + 4651.836 + 4670.110}{3} = 4627.925 \text{ sec}$$

This is the corrected time of the **mark boat** (in this case, not a real yacht) and is used to calculate the new handicaps for the fleet from equation (3). The new handicaps are shown in Table 1 in the column headed *World Sailing* and are rounded to 3 decimal places.

Method [2]   *TopYacht* in the documentation of their Performance Handicap System suggest that a mark boat be selected by the race management and note that from experience "the 45% boat" gives acceptable results, where the 45% boat is the yacht finishing in 45th place on corrected time in a fleet of 100 yachts.

We define the **mark boat** as the yacht finishing in Mth position on corrected time where the finish order place M is given by

$$M = \text{round}\left(n \times \frac{p}{100} \right) \tag{4}$$

where $\text{round}(x)$ denotes the round function that rounds x to the nearest integer (if two integers are available the one furthest from zero is selected), n is the number of yachts in the race and p is the finish place of the mark boat in a 100 yacht fleet. For the 45% boat $p = 45$.

The *standard corrected time SCT* is now taken to be the corrected time of the mark boat (the yacht finishing in Mth position on corrected time).

For Race 1a in Table 1 above, *TopYacht's* 45% boat will have a finish order place $M = \text{round}(9 \times 0.45) = 4$ and the mark boat will be *Wicked,* the 4th boat on corrected time of 4525.362 sec. This becomes the *SCT* and new handicaps for the fleet are calculated from (3) and are shown in Table 1 in the column headed *TopYacht* (rounded to 3 decimal places).

Note that the *BCH* values in Table 1 in the *TopYacht* column will not be the same as those shown in Race 1 of Appendix A because the race above does not include the yacht *Cadibarra 8*. This demonstrates the dependence between the yachts in the race, the *SCT* and the *BCH's*.

Standard Corrected Time: A New Method

Of course, it would be extremely rare (and probably has never happened in practice) for all the corrected times in a particular race to be identical. This would imply that handicaps were derived from a perfect system of rules and yachts were sailed at their optimum in the prevailing conditions. But, nevertheless, we could consider the following two statements.

A If every yacht sailed to their *true handicap H* then their corrected times would be identical.

B The range of corrected times in an actual race is the sum of each yacht's *handicap error x.*

Considering statement A. If all the corrected times are the same, call this value the *standard corrected time SCT* and define a yacht's true handicap H as

$$H = AHC + \delta H \tag{5}$$

where AHC is the yacht's allocated handicap derived from a set of rules and δH is a small unknown correction. Now, we can write, for any individual yacht

$$ET \times \left(AHC + \delta H\right) = SCT$$
$$ET \times AHC + ET \times \delta H = SCT \tag{6}$$

and defining a yacht's handicap error x as

$$x = ET \times \delta H \tag{7}$$

Now using (1) we may write an *observation equation* for each of the $k = 1, 2, 3, \ldots, n$ yachts in the race as

$$x_k - SCT = -CT_k \tag{8}$$

And, from statement B we may write a single additional *constraint equation*

$$x_1 + x_2 + \cdots + x_n = R \tag{9}$$

where $R = \max\{CT\} - \min\{CT\}$ is a positive quantity equal to the range of the corrected times and $\{CT\}$ denotes the set of corrected times of the yachts in a race.

It turns out (see Appendix C) that the system of n observation equations (8) and the single constraint equation (9) can be solved to give the standard corrected time SCT as

$$SCT = \frac{1}{n}\left(S + R\right) \tag{10}$$

where S and R are the sum and range of the corrected times respectively.

The back calculated handicaps BCH_k are calculated using (3).

[Appendix C also shows that the true handicap $H_k = BCH_k$ and the small correction to the true handicap is $\delta H_k = BCH_k - AHC_k = PI_k$]

As an example, we will use Race 1a from our Yacht Race Series in Appendix B.

Yacht	Elapsed Time *ET*	Allocated Handicap *AHC*	Corrected Time *CT* (sec)	Place	Back Calculated Handicap *BCH*
Sierra Chainsaw	1:18:59 (4739 sec)	0.930	4407.270	**1**	0.987
Joust	1:18:40 (4720)	0.935	4413.200	**2**	0.991
Scarlett Runner II	1:19:28 (4768)	0.935	4458.080	**3**	0.981
Wicked	1:21:27 (4887)	0.926	4525.362	**4**	0.957
Bandit	1:23:33 (5013)	0.910	4561.830	**5**	0.933
Dream	1:21:47 (4907)	0.948	4651.836	**6**	0.954
Esprit	1:26:58 (5218)	0.895	4670.110	**7**	0.897
Conquest	1:26:30 (5190)	0.914	4743.660	**8**	0.902
Dark and Stormy	1:34:59 (5699)	0.885	5043.615	**9**	0.821
Niche	DNS				

Table 2. Race 1a of the Yacht Race Series (Appendix B). Yachts shown in corrected time order.

From Table 2, the range of corrected times $R = 5043.615 - 4407.270 = 636.345$ sec and the sum of the corrected times is $S = 41474.963$ sec and using equation (10)

$$SCT = \frac{1}{n}\left(S + R\right) = 4679.034 \text{ sec}$$

Back calculated handicaps are computed using (3) and are tabulated in Table 2. For example, the 6th yacht *Dream's BCH* is

$$BCH_6 = \frac{SCT}{ET_6} = \frac{4679.034}{4907} = 0.953543 \approx 0.954$$

The Calculated Handicap *CHC*

As we outlined in the Introduction, the third (and last) step in the production of a new handicap is the calculated handicap *CHC,* and if this doesn't appear to be an anomalous result then the *CHC* becomes the yacht's allocated handicap *AHC* for her next race.

There are many ways of determining the *CHC* and *TopYacht* (2018, 2019) describe or mention the following methods: Weighted Average; Exponential Average; Place Based Handicapping; Place Based Exponential Handicapping; Trend Biased Average and Boat Performance Ratio. *World Sailing* (2016) describes a system of Weighted Performance Indicators in calculating the *CHC*.

We will review three methods; **Weighted Average, Exponential Average,** and **Weighted Performance Indicators** that represent some of the possible methods listed above and Appendix D gives some useful background information. In addition, a fourth and new method of determining the *CHC* is described that we call a **Performance Indicator Filter** and Appendix E has some background information applicable to this topic. But first we set out the rationale for the general rule for calculating a yacht's new handicap.

The difference between a yacht's allocated handicap *AHC* (the handicap before the race) and the back calculated handicap *BCH* (the handicap it would have had if its corrected time equalled the standard corrected time of the race) is its *performance indicator PI* defined by *World Sailing* (2016) as

$$PI_k = BCH_k - AHC_k \tag{11}$$

noting here that $PI_k = \delta H_k$ in this paper [see Appendix C, equation (38)] and the subscript k denotes the kth boat in the fleet of n yachts.

If the conditions of the race were reasonable and unchanging and every yacht was sailed to its optimum then *PI's* simply reflect handicap errors, and in this perfect situation, each yacht's new handicap would simply be its back calculated handicap, i.e., $CHC_k = BCH_k$. If this perfect situation existed for the next race then all the *PI's* would be zero and each yacht's new handicap would be unchanged, i.e., $CHC_k = AHC_k$. So, we may establish a general rule for a yacht's calculated handicap

$$CHC_k = AHC_k + z_k \tag{12}$$

where AHC_k denotes the allocated handicap for the race and z_k is a numeric value between zero and $\pm PI_k$ that is a function of the yacht's *AHC* prior to the race and *BCH* after the race.

The Calculated Handicap: Weighted Performance Indicators (*World Sailing*)

World Sailing (2016) describe a system of weighted performance indicators that we summarize here.

- After a race, back calculated handicaps *BCH* are calculated using (3).
- Performance indicators *PI* are calculated using (11).
- Performance multipliers *PM* are selected from Table 3

Races completed	Portion of *PI*	Multiplier *PM*
1	All	1
2	Half	0.5
3	One third	0.333333
4	One quarter	0.25
5	One fifth	0.2
Greater than 5	One fifth	0.2

Table 3. Performance Multipliers, *World Sailing* (2016)

- Calculate new handicaps using

$$CHC_k = AHC_k + \left(PM \times PI_k \right) \tag{13}$$

We can see here that the 'weighted performance indicator' is $\dfrac{1}{\text{Race No.}} \times PI_k$ up to and including Race 5 and the weighted performance indicator is $PI_k/5$ for all races after Race 5.

The weighted performance indicator is the variable z_k in the general rule (12)

The Calculated Handicap: Weighted Average (*TopYacht*)

TopYacht (2019) describe a system they call 'weighted average' to obtain the calculated handicap from the handicaps of the race just completed and from races prior to that. We summarise their system as

- Select the number of races N that are to be used in the averaging process.
- Calculated handicaps *CHC* are given by

$$CHC_k = \begin{cases} \dfrac{1}{N}\left\{ \left(N - k \right) AHC_1 + \displaystyle\sum_{j=1}^{k} BCH_j \right\} & \text{for } k < N \\ \dfrac{1}{N} \displaystyle\sum_{j=k}^{k-N+1} BCH_j & \text{for } k \geq N \end{cases} \tag{14}$$

Where k is the race number, the subscript k denotes the handicap for the kth race, AHC_1 is the allocated handicap for the 1st race and BCH_j is the back calculated handicap for the jth race.

Suppose that $N = 4$ then the calculated handicaps for races 1,2,3,..., etc. are:

Race 1
$$CHC_1 = \frac{1}{4}\left\{ \left(4 - 1 \right) AHC_1 + \sum_{j=1}^{1} BCH_j \right\} = \frac{1}{4}\left\{ \left(3 \right) AHC_1 + BCH_1 \right\}$$

Race 2
$$CHC_2 = \frac{1}{4}\left\{ \left(4 - 2 \right) AHC_1 + \sum_{j=1}^{2} BCH_j \right\} = \frac{1}{4}\left\{ \left(2 \right) AHC_1 + BCH_1 + BCH_2 \right\}$$

Race 3
$$CHC_3 = \frac{1}{4}\left\{ \left(4 - 3 \right) AHC_1 + \sum_{j=1}^{3} BCH_j \right\} = \frac{1}{4}\left\{ AHC_1 + BCH_1 + BCH_2 + BCH_3 \right\}$$

Race 4
$$CHC_4 = \frac{1}{4}\sum_{j=4}^{1} BCH_j = \frac{1}{4}\left(BCH_4 + BCH_3 + BCH_2 + BCH_1 \right)$$

Race 5
$$CHC_5 = \frac{1}{4}\sum_{j=5}^{2} BCH_j = \frac{1}{4}\left(BCH_5 + BCH_4 + BCH_3 + BCH_2 \right)$$

etc.

7

When the race number is equal to or greater than N, the calculated handicap is just a simple moving average and there is no 'weighting' involved in these calculations. Appendix D has some useful information on moving averages and weighted averages. By studying the sequence of calculated handicaps above and understanding that the allocated handicap is the previous calculated handicap we may write

$$CHC_k = AHC_k + \frac{1}{N}\left(BCH_k - BCH_{k-N+1}\right) \quad \text{for } k > N \tag{15}$$

And the variable z_k in the general rule (12) is $z_k = \frac{1}{N}\left(BCH_k - BCH_{k-N+1}\right)$ for $k > N$.

The Calculated Handicap: Exponential Average (*TopYacht*)

TopYacht (2019) describe a method they call 'exponential average' to obtain the calculated handicap. We summarise their system as

- Select the gain G
- If $G \geq 1$ is an integer, say $1, 2, 3, \dots$ then the calculated handicap is

$$CHC = \frac{1}{G} \times BCH + \left(\frac{G-1}{G}\right)AHC \quad \text{for } G \geq 1 \tag{16}$$

- If $G \geq 0$ is a percentage, say 10% $(G = 10)$, 20% $(G = 20)$, ... then the calculated handicap is

$$CHC = \frac{G}{100} \times BCH + \left(\frac{100-G}{100}\right)AHC \quad \text{for } 0 \leq G \leq 100 \tag{17}$$

If the gain $G = 3$ then
$$CHC = \frac{1}{3}BCH + \frac{2}{3}AHC \tag{18}$$

If the gain $G = 33.33\%$ then
$$CHC = \frac{33.33}{100}BCH + \frac{66.67}{100}AHC$$

It is very common, in Victorian yachting, to use a gain $G = 3$.

> **Notes:** The derivation of (16) is shown in Appendix D – see equation (52) – and the use of the word 'gain' is probably a little misleading. A better expression might be
>
> $CHC = \left(\alpha\right)BCH + \left(1-\alpha\right)AHC$ for $0 < \alpha \leq 1$ – see Appendix D, equation (51) – and
>
> successive *CHC's* are terms in the Exponentially Weighted Moving Average (EWMA) defined by Hunter (1986). Equation (16) is really a recurrence formula for predicting or forecasting the next allocated handicap given the current back calculated handicap and allocated handicap. A weighting factor $\alpha = 1/G$ (see above) is used and since each new prediction is a function of prior allocated and back calculated handicaps – each with their own weighting factors – then the current prediction is in fact a summation of prior back calculated handicaps each multiplied by powers of $\left(1-\alpha\right)$ that get exponentially smaller as the number of prior values increases. This process is known as *exponential smoothing* and is often used in financial analysis and prediction. Appendix D has an example of exponential smoothing applied to the fluctuations of the exchange rate between US dollars and Australian dollars.

By a simple rearrangement of (16) we can write

$$CHC_k = AHC_k + \frac{1}{G}\left(BCH_k - AHC_k\right) = AHC_k + \frac{PI_k}{G} \tag{19}$$

And for the gain $G = 3$ (very common in Victorian yachting) then

$$CHC_k = AHC_k + \tfrac{1}{3}PI_k$$

And the variable z in the general rule (12) is $z_k = \frac{1}{3} PI_k$.

So, this means that after each race, the calculated handicap is equal to the allocated handicap plus one-third of the performance indicator $PI_k = BCH_k - AHC_k$.

The Calculated Handicap: A New Method

We propose a new method of determining the calculated handicap that we call a **Performance Indicator Filter (PIF)** that gives the z-value for yacht k in race j of a series as

$$z_{k,j} = z_{k,j-1} + K\left(PI_{k,j} - z_{k,j-1}\right) \text{ for } k, j > 0 \text{ and } z_{k,0} = 0 \tag{20}$$

Where $PI_{k,j} = BCH_{k,j} - AHC_{k,j}$ is the yacht's performance indicator for the jth race and K is a constant.

A yacht's calculated handicap CHC_j for race j of the series is then given by (12) and as this will be her allocated handicap for the next race we may write

$$AHC_{k,j+1} = AHC_{k,j} + z_{k,j} \tag{21}$$

From an analysis of the yacht *Dream's* performance in the Race Series (see Appendix B) we have determined that a value of $K = 2/5$ gives reasonable z-values. Appendix E has an explanation of our Performance Indicator Filter and how it is linked to a Kalman Filter which is a mathematical estimation process that is widely used in navigation.

As an example of the calculation of a yacht's z-values we use *Dream's PI* values in the Race Series scored using PIF and the new method for SCT (see Appendix B, Table B19)

$$PI = \begin{bmatrix} 0.006 & 0.082 & 0.074 & 0.006 & -0.012 & -0.027 & -0.044 \end{bmatrix} \tag{22}$$

Using (20) and (21) with $K = 2/5$ and the *PI* values (22) with *Dream's* allocated handicap in race $j = 1$ as 0.948 gives the following results

$$j = 1, \quad PI_{k,1} = 0.006, \quad z_{k,1} = 0 + 2/5\left(0.006 - 0\right) = 0.0024 \qquad AHC_{k,2} = 0.948 + 0.002 = 0.950$$
$$j = 2, \quad PI_{k,2} = 0.082, \quad z_{k,2} = 0.0024 + 2/5\left(0.082 - 0.0024\right) = 0.0342 \qquad AHC_{k,3} = 0.950 + 0.034 = 0.984$$
$$j = 3, \quad PI_{k,3} = 0.074, \quad z_{k,3} = 0.0342 + 2/5\left(0.074 - 0.0342\right) = 0.0501 \qquad AHC_{k,4} = 0.984 + 0.050 = 1.034$$

etc.

tabulated as

j	$PI_{k,j}$	$z_{k,j}$	$AHC_{k,j}$
1	0.006	0.002	0.948
2	0.082	0.034	0.984
3	0.074	0.050	1.034
4	0.006	0.032	1.066
5	-0.012	0.015	1.081
6	-0.027	-0.002	1.079
7	-0.044	-0.017	1.062

Table 4. *PI* values, z-values computed from a Performance Indicator Filter
and allocated handicaps for *Dream* for the seven-race series

A plot of *Dream's PI* values and the PIF z-values are shown in Figure 1, and since
$AHC_{k,j+1} = AHC_{k,j} + z_{k,j}$, *Dream's* handicap will be increasing when $z > 0$ and decreasing when $z < 0$,
i.e., increasing when the solid line is above zero and decreasing when it's below zero.
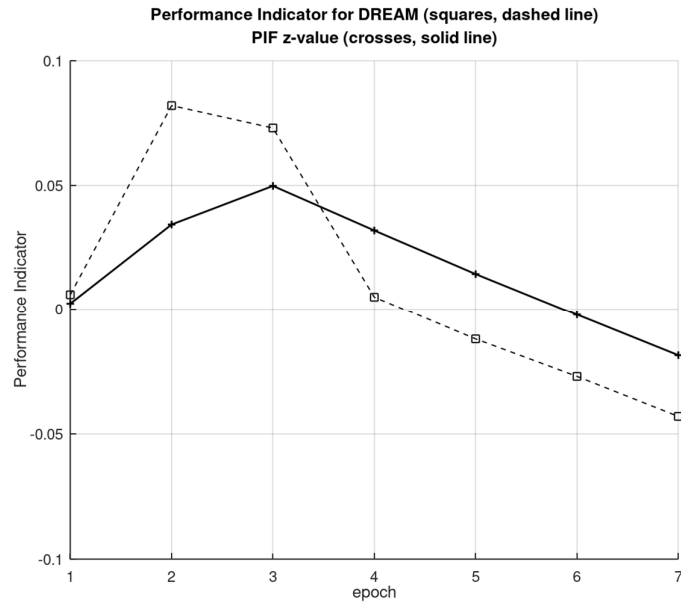
Figure 1. *Dream's PI* values (squares and dashed line) and PIF z-values (crosses and solid line). Note that $z_{k,1} = \frac{2}{5}PI_{k,1}$

Figure 2 shows *Dream's PI* values, PIF z-values and the values $\frac{1}{3}PI$. *TopYacht* scoring has $AHC_{k,j+1} = AHC_{k,j} + \frac{1}{3}PI_{k,j}$ and this figure shows the relationship between increasing and decreasing handicaps using *TopYacht* and PIF scoring.
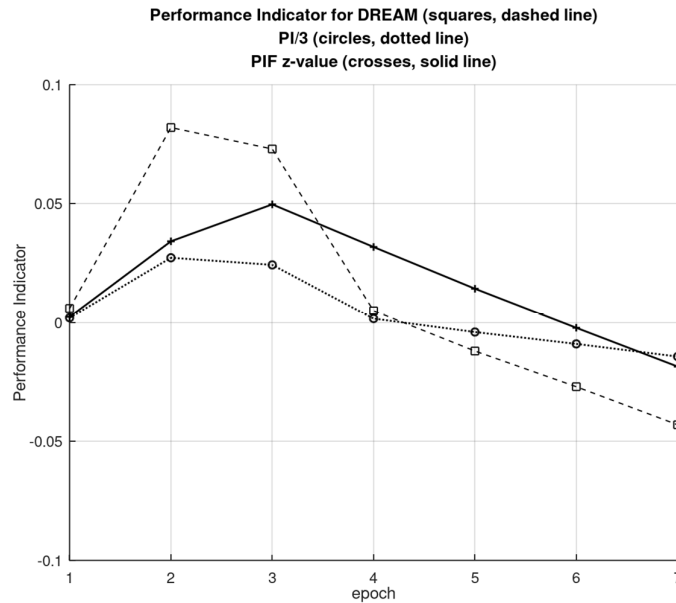

Figure 2. *Dream's PI* values (squares and dashed line), PIF z-values (crosses and solid line) and $\frac{1}{3}PI$ values (circles and dotted line). Note that

$$z_{k,1} = \frac{2}{5}PI_{k,1}$$

Discussion

The four basic components of a Performance Handicap System are

(i) Calculating the *corrected time CT* of each yacht in the race from $CT_k = ET_k \times AHC_k$ and determining the winner with the least corrected time.

(ii) Determining the *standard corrected time STC* (corrected time of the mark boat – real or mythical) for the race.

(iii) Calculating the *back calculated handicap* BCH_k and *performance indicator* $PI_k = BCH_k - AHC_k$ for each yacht in the race.

(iv) Determining the *allocated handicap* for yacht k for the next race j in the series from $AHC_{k,j+1} = AHC_{k,j} + z_{k,j}$ where $z_{k,j}$ is some function of the performance indicator for each yacht in race j.

We have reviewed several methods of determining the STC and have explained a new method we are proposing that is: $SCT_j = (S_j + R_j)/n$ where S_j and R_j are the sum and range of the corrected times respectively and n is the number of yachts in race j. The development of this simple formula is set out in Appendix C.

We have reviewed several methods of determining the *calculated handicap* CHC_k of yacht k and have shown that they can be described by a general formula: $CHC_k = AHC_k + z_k$ and since the calculated handicap of race j is the allocated handicap of the next race then $AHC_{k,j+1} = AHC_{k,j} + z_{k,j}$ where $z_{k,j}$ is a function of the performance indicator of yacht k in race j. And we have proposed a new method of calculating the z-value of yacht k in race j as: $z_{k,j} = z_{k,j-1} + K\left(PI_{k,j} - z_{k,j-1}\right)$ for $k, j > 0$ and $z_{k,0} = 0$ that we have called a Performance Indicator Filter (PIF). This new method of calculating the z-values is a simplification of a Kalman Filter which is a mathematical estimation process. A Kalman Filter is really a sequence of equations that processes measurements that are linked to the 'state' of a 'system'. In our case, the measurements are the performance indicators PI, the system is just a single entity, the PIF, and its state is the value z. As measurements are processed one after the other by the filter the state of the system changes and can be monitored. In Appendix E we have studied the output of a Kalman Filter to determine z-values given PI-values as measurements and reduced the sequence of operations to a single equation (20), the PIF, containing a constant K. We have taken $K = 2/5$ to be a reasonable value based on an analysis of the yacht *Dream* over the Race Series in Appendix B and we are confident that this value is suitable for the PIF applied to other yachts.

To test our two new methods the Race Series of Appendix B was scored using:

(i) *TopYacht* with a 45% yacht as the mark boat and allocated handicaps $AHC_{k,j+1} = AHC_{k,j} + \frac{1}{3}PI_{k,j}$ and

(ii) Performance Indicator Filter (PIF) $z_{k,j} = z_{k,j-1} + K\left(PI_{k,j} - z_{k,j-1}\right)$ for $k, j > 0$ and $z_{k,0} = 0$ with $K = 2/5$ and $SCT_j = (S_j + R_j)/n$, with allocated handicaps $AHC_{k,j+1} = AHC_{k,j} + z_{k,j}$.

The overall results for the series and series placings for the Best 5 of the 7-race series are shown in Tables B9 (TopYacht) and B18 (PIF) and are repeated here as Tables 5 and 6.

Yacht	Race Placings							Total (Best 5)	Series Place
	R1a	R2a	R3a	R4a	R6a	R7a	R10a		
Dream	6	1	1	2	1	3	8	8	**1**
Sierra Chainsaw	1	6	7	1	2	12	4	14	**2**
Esprit	7	3	5	6	3	2	9	19	**3**
Joust	2	2	8	10	10	7	1	20	**4**
Wicked	4	4	2	8	5	8	5	20	**5**
Niche	12	12	9	3	4	4	3	23	**6**
Bandit	5	5	3	7	6	5	6	24	**7**
Scarlett Runner II	3	12	10	4	7	12	2	26	**8**
Conquest	8	12	4	9	8	1	7	28	**9**
Dark and Stormy	9	12	6	5	9	6	11	35	**10**

Table 5. Yacht Race Series (*TopYacht* scoring): Race Placings; Total Points (Best 5); Overall Place
DNS = 12, RET = 11

Yacht	Race Placings							Total (Best 5)	Place
	R1a	R2a	R3a	R4a	R6a	R7a	R10a		
Sierra Chainsaw	1	6	8	3	2	12	3	15	**1**
Niche	12	12	6	2	1	4	5	18	**2**
Joust	2	2	9	10	10	6	1	20	**3**
Dream	6	1	1	5	9	8	8	21	**4**
Esprit	7	3	7	6	3	2	9	21	**5**
Dark and Stormy	9	12	5	1	4	3	11	22	**6**
Wicked	4	4	3	9	8	7	4	22	**7**
Bandit	5	5	2	8	7	5	6	23	**8**
Scarlett Runner II	3	12	10	4	5	12	2	24	**9**
Conquest	8	12	4	7	6	1	7	25	**10**

Table 6. Yacht Race Series (PIF scoring $K = 2/5$): Race Placings; Total Points (Best 5); Overall Place
DNS = 12, RET = 11

For *TopYact* scoring (Table 5) the point score range for the 10 entrants is $35 - 8 = 27$. The first 3 yachts have points 8, 14 and 19; the middle 4 yachts 20, 20, 23 and 24; and the bottom 3 have points 26, 28 and 35. *Dream* wins 3 races and *Sierra Chainsaw* 2 races. The other two race wins are by *Joust* and *Conquest*.

For PIF scoring (Table 6) the point score range for the 10 entrants is $25 - 15 = 10$. The first 3 yachts have points 15, 18 and 20; the middle 4 yachts 21, 21, 22 and 22; and the bottom 3 have points 23, 24 and 25. *Dream* wins 2 races and the other 5 races are wins by *Sierra Chainsaw, Niche, Joust, Dark and Stormy* and *Conquest*.

It would appear that PIF scoring [with $SCT_j = \left(S_j + R_j \right)/n$] gives series point scores with a much smaller range between first and last and much tighter groupings in the first-3, the middle-4 and the bottom-3 groups. Indeed, the points range in the middle-4 group is one point, and in the bottom-3 group it is 2 points. Three yachts; *Niche, Dark and Stormy,* and *Dream* are the most affected by the two scoring methods. *Dark and Stormy* and *Niche* both rising 4 places and *Dream* falling 3 when scored using PIF. *Esprit* and *Wicked* fall by 2 places; *Bandit, Scarlett Runner II* and *Conquest* fall by 1 place; and *Joust* and *Sierra Chainsaw* rise by 1 place.

The scoring of the Race Series using the two methods was done using a program written in GNU Octave[3] and the required information for each yacht is:

TopYacht – Elapsed time and allocated handicap,

PIF – Elapsed time, allocated handicap and z-value from previous race ($z_{k,j-1} = 0$ for race $j = 1$)

Our new method of scoring requires an additional piece of information (the prior z-value) but this shouldn't be a trouble if the spreadsheets are linked to a race series master sheet containing handicap information that can be automatically updated.

Conclusion

This paper provides a useful summary of the principles and operation of a *Performance Handicap System* in yachting and reviews some current methods of scoring as well as introducing new methods of calculating the standard corrected time STC and allocated handicaps AHC using a Performance Indicator Filter. In addition, examples of the calculations required have been given and the development of formula set down and supported with technical information in several appendices. References have been provided for the interested reader.

We believe our new method of scoring deserves further investigation and study and we hope that it might be taken up by a yacht club or several yacht clubs and compared with existing methods.

---

[3] GNU Octave is a high-level language, primarily intended for numerical computations. It provides a convenient command line interface for solving linear and nonlinear problems numerically, and for performing other numerical experiments using a language that is mostly compatible with Matlab. GNU Octave is freely redistributable software from the Free Software Foundation

References

Brown, R.G. and Hwang, P.Y.C, 1992, *Introduction to Random Signals and Applied Kalman Filtering,* 2nd ed., John Wiley & Sons, New York.

Cross, P.A. 1992, *Advanced Least Squares Applied to Position Fixing,* Working Paper No. 6, Department of Land Information, University of East London. https://seabedhabitats.files.wordpress.com/2011/10/cross_1994.pdf (accessed 10-Aug-2015)

Douguet, R., Diguet, J-P., Laurent, J. and Riou, Y., 2013, 'A new real-time method for sailboat performance estimation based on leeway modelling', *The 21st Chesapeake Sailing Yacht Symposium,* Annapolis, Maryland, March 2013, 13 pages. http://vm2330.sgvps.net/~syrftest/images/library/20150805142512.pdf [accessed 03-Apr-2020]

Gauss, C.F. 1809, *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections,* a translation of *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium* by C.H. Davis, Dover, New York, 1963.

Hunter, J. Stuart, 1986, 'The exponentiall weighted moving average', *Journal of Quality Technology,* Vol. 18, No. 4 (Oct., 1986), pp. 203-210

Kalman, R.E. 1960, 'A new approach to linear filtering and prediction', *Journal of Basic Engineering,* American Society of Mechanical Engineers (ASME), Series 82D, pp. 35-45, March 1960.

Krakiwsky, E.J. 1975, *A Synthesis of Recent Advances in the Method of Least Squares,* Lecture Notes No. 42, 1992 reprint, Department of Surveying Engineering, University of New Brunswick, Fredericton, Canada

NIST/SEMATECH *e-Handbook of Statistical Methods,* National Institute of Standards and Technology, U.S. Department of Commerce, and Semiconductor Manufacturing Technology (University of Albany - State University of New York) http://www.itl.nist.gov/div898/handbook [accessed 20-Dec-2019]

Sorenson, Harold W. 1985, *Kalman Filtering: Theory and Application,* The Institute of Electrical and Electronics Engineers (IEEE) Press selected reprint series, Editor Harold W. Sorenson, IEEE Press, New York.

TopYacht, 2018, 'Measured Performance Handicapping', *TopYacht* Technical Documents and Discussion Papers, 7 pages, 10-Oct-2018. https://topyacht.net.au/results/shared/technical/Measured%20Performance%20Handicapping.pdf?ty= 1631 [accessed 25-Feb-2020]

TopYacht, 2019, 'How the Next Handicap is Calculated', *TopYacht* Technical Documents and Discussion Papers, 14 pages, 28-Oct-2019. https://topyacht.net.au/results/shared/technical/How%20the%20Next%20Handicap%20is%20Calculat ed.pdf?ty=0729 [accessed 25-Feb-2020]

World Sailing, 2019, 'Ratings and Handicap Systems', *World Sailing,* Classes & Equipment, Offshore, 1 page. https://www.sailing.org/classesandequipment/offshore/ratings_and_handicap_systems.php [accessed 25-Feb-2020]

World Sailing, 2016, 'International Empirical Handicap Scheme for Yachts', *World Sailing,* 3 pages. https://www.sailing.org/tools/documents/TurnkeytextVer2-[7780].pdf [accessed 25-Feb-2020]

APPENDIX A

2018-19 SUMMER AGGREGATE SANDRINGHAM YACHT CLUB

Division 1, Performance Handicap System (PHS)

https://www.syc.com.au/raceresults/2018/klbt/agg/series.htm [accessed 18-Feb-2020]

Race Results produced by *TopYacht*

Race 1 (20/10/2018) PROVISIONAL RESULTS

Updated: 20/10/2018 5:17:56 PM

D1 PHS results Start : 14:40

Place	Sail Num	Boat Name	Skipper	Fin Tim	Elapsd	AHC	Cor'd T	BCH	CHC	Score
1	SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	15:58:59	01:18:59	0.930	01:13:27	0.955	0.938	1.0
2	SM1110	JOUST	Rod Warren	15:58:40	01:18:40	0.935	01:13:33	0.959	0.943	2.0
3	SM11	SCARLET RUNNER-11	Rob Date	15:59:28	01:19:28	0.935	01:14:18	0.949	0.940	3.0
4	SM420	CADIBARRA 8	Paul Roberts	15:57:47	01:17:47	0.965	01:15:04	0.970	0.967	4.0
5	SM4	WICKED	Mike And Mark Welsh	16:01:27	01:21:27	0.926	01:15:25	0.926	0.926	5.0
6	SM4040	BANDIT	Ashley Trebilcock	16:03:33	01:23:33	0.910	01:16:02	0.903	0.908	6.0
7	SM8	DREAM	Robert Green	16:01:47	01:21:47	0.948	01:17:32	0.922	0.939	7.0
8	SM477	ESPRIT	Garry Anderson	16:06:58	01:26:58	0.895	01:17:50	0.867	0.886	8.0
9	SM45	CONQUEST	Paul Jacka	16:06:30	01:26:30	0.914	01:19:04	0.872	0.902	9.0
10	SM69	DARK AND STORMY	S.Mellington T.Kourtis	16:14:59	01:34:59	0.885	01:24:04	0.794	0.885	10.0
RET	SM424	BARNSTORMER	Brian Barnes			0.885				13.0
RET	SM1100	SMOOTH CRIMINAL	Helen Cameron			0.975				13.0

Race 2 (24/11/2018) PROVISIONAL RESULTS

Updated: 24/11/2018 4:52:41 PM

D1 PHS results Start : 12:00

Place	Sail No	Boat Name	Skipper	Fin Tim	Elapsd	AHC	Cor'd T	BCH	CHC	Score	Vis
1	SM8	DREAM	Robert Green	14:33:06	02:33:06	0.939	02:23:46	1.002	0.952	1.0	
2	SM420	CADIBARRA 8	Paul Roberts	14:32:44	02:32:44	0.967	02:27:42	1.004	0.979	2.0	
3	SM1110	JOUST	Rod Warren	14:39:23	02:39:23	0.943	02:30:18	0.963	0.950	3.0	
4	SM477	ESPRIT	Garry Anderson	14:52:06	02:52:06	0.886	02:32:29	0.891	0.888	4.0	
5	SM4	WICKED	Mike And Mark Welsh	14:45:41	02:45:41	0.926	02:33:25	0.926	0.926	5.0	
6	SM9535	AMBITION	Chris Dare	14:15:18	02:15:18	1.150	02:35:36	1.134	1.145	6.0	
7	B2	BKT JAMHU	Phil Bedlington	14:44:52	02:44:52	0.944	02:35:38	0.931	0.940		Yes
8	SM1113	PLAYLIST	Matt/Glenn/John Powell/Chesser/Cox	14:50:20	02:50:20	0.920	02:36:42	0.901	0.914	7.0	
9	SM4040	BANDIT	Ashley Trebilcock	14:53:45	02:53:45	0.908	02:37:46	0.883	0.900	8.0	
10	SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	14:51:11	02:51:11	0.938	02:40:34	0.896	0.925	9.0	
11	B666	RECKLESS	Rory Thomas	15:01:13	03:01:13	0.910	02:44:54	0.847	0.898		Yes

Race 3 (1/12/2018) PROVISIONAL RESULTS

Updated: 19/12/2018 9:50:45 AM

D1 PHS results Start : 14:00

Place	Sail Num	Boat Name	Skipper	Fin Tim	Elapsd	AHC	Cor'd T	BCH	CHC	Score	DidNot
1	SM9535	AMBITION	Chris Dare	15:21:17	01:21:17	1.145	01:33:04	1.252	1.160	1.0	
2	SM8	DREAM	Robert Green	15:38:13	01:38:13	0.952	01:33:30	1.036	0.965	2.0	
3	SM4	WICKED	Mike And Mark Welsh	15:45:44	01:45:44	0.926	01:37:54	0.962	0.938	3.0	
4	SM4040	BANDIT	Ashley Trebilcock	15:49:01	01:49:01	0.900	01:38:07	0.933	0.911	4.0	
5	SM45	CONQUEST	Paul Jacka	15:51:07	01:51:07	0.902	01:40:14	0.916	0.907	5.0	
6	SM477	ESPRIT	Garry Anderson	15:54:36	01:54:36	0.888	01:41:46	0.888	0.888	6.0	
7	SM1110	JOUST	Rod Warren	15:48:14	01:48:14	0.950	01:42:49	0.940	0.947	7.0	
8	SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	15:51:12	01:51:12	0.925	01:42:52	0.915	0.922	8.0	
9	SM4000	NICHE	Peter Jackson	15:54:53	01:54:53	0.900	01:43:24	0.886	0.895	9.0	
10	SM007	VELOCE	Phil Simpfendorfer	15:49:37	01:49:37	0.945	01:43:35	0.928	0.939	10.5	
10	SM1113	PLAYLIST	Matt/Glenn/John Powell/Chesser/Cox	15:53:20	01:53:20	0.914	01:43:35	0.898	0.909	10.5	
12	SM424	BARNSTORMER	Brian Barnes	15:57:45	01:57:45	0.885	01:44:12	0.864	0.878	12.0	
13	SM69	DARK AND STORMY	S.Mellington T.Kourtis	15:58:50	01:58:50	0.885	01:45:10	0.856	0.875	13.0	
14	SM11	SCARLET RUNNER-11	Rob Date	15:52:24	01:52:24	0.940	01:45:39	0.905	0.928	14.0	

Race 4 (2/02/2019) PROVISIONAL RESULTS

Updated: 2/02/2019 4:02:58 PM

D1 PHS results Start : 12:00

Place	Sail Num	Boat Name	Skipper	Fin Tim	Elapsd	AHC	Cor'd T	BCH	CHC	Score
1	SM69	DARK AND STORMY	S.Mellington T.Kourtis	13:17:19	01:17:19	0.833	01:04:24	0.859	0.842	1.5
1	SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	13:10:51	01:10:51	0.909	01:04:24	0.938	0.919	1.5
3	SM4000	NICHE	Peter Jackson	13:13:00	01:13:00	0.888	01:04:49	0.910	0.895	3.0
4	SM11	SCARLET RUNNER-11	Rob Date	13:11:07	01:11:07	0.933	01:06:21	0.934	0.933	4.0
5	SM8	DREAM	Robert Green	13:06:23	01:06:23	1.001	01:06:27	1.001	1.001	5.0
6	SM1110	JOUST	Rod Warren	13:12:20	01:12:20	0.923	01:06:46	0.919	0.922	6.0
7	SM45	CONQUEST	Paul Jacka	13:15:19	01:15:19	0.887	01:06:48	0.882	0.885	7.5
7	SM4040	BANDIT	Ashley Trebilcock	13:13:59	01:13:59	0.903	01:06:48	0.898	0.901	7.5
9	SM4	WICKED	Mike And Mark Welsh	13:12:16	01:12:16	0.927	01:07:00	0.920	0.925	9.0
10	SM477	ESPRIT	Garry Anderson	13:15:30	01:15:30	0.888	01:07:03	0.880	0.885	10.0

Race 5 (19/01/2019) PROVISIONAL RESULTS

Updated: 23/02/2019 5:36:24 PM

D1 PHS results Start : 12:00

Place	Sail Num	Boat Name	Skipper	Score	DidNot
1	SM4	WICKED	Mike And Mark Welsh	6.0	
1	SM8	DREAM	Robert Green	1.0	
1	SM45	CONQUEST	Paul Jacka	1.0	
1	SM69	DARK AND STORMY	S.Mellington T.Kourtis	12.0	
1	SM424	BARNSTORMER	Brian Barnes	13.0	
1	SM477	ESPRIT	Garry Anderson	3.0	
1	SM1100	SMOOTH CRIMINAL	Helen Cameron	14.0	
1	SM1110	JOUST	Rod Warren	7.0	
1	SM1113	PLAYLIST	Matt/Glenn/John Powell/Chesser/Cox	7.0	
1	SM4000	NICHE	Peter Jackson	2.0	
1	SM4040	BANDIT	Ashley Trebilcock	5.0	
1	SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	5.0	
1	SM9535	AMBITION	Chris Dare	2.0	
14	SM11	SCARLET RUNNER-11	Rob Date	3.0	

Race 6 (2/02/2019) PROVISIONAL RESULTS

Updated: 2/02/2019 4:49:45 PM

D1 PHS results Start : 14:00

Place	Sail Num	Boat Name	Skipper	Fin Tim	Elapsd	AHC	Cor'd T	BCH	CHC	Score
1	SM4000	NICHE	Peter Jackson	15:38:20	01:38:20	0.888	01:27:19	0.906	0.894	1.0
2	SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	15:35:17	01:35:17	0.924	01:28:02	0.935	0.928	2.0
3	SM477	ESPRIT	Garry Anderson	15:39:49	01:39:49	0.888	01:28:38	0.893	0.890	3.0
4	SM8	DREAM	Robert Green	15:28:47	01:28:47	1.001	01:28:52	1.004	1.002	4.0
5	SM69	DARK AND STORMY	S.Mellington T.Kourtis	15:46:58	01:46:58	0.833	01:29:06	0.833	0.833	5.0
6	SM4	WICKED	Mike And Mark Welsh	15:36:24	01:36:24	0.927	01:29:22	0.924	0.926	6.0
7	SM4040	BANDIT	Ashley Trebilcock	15:39:24	01:39:24	0.903	01:29:46	0.896	0.901	7.0
8	SM45	CONQUEST	Paul Jacka	15:41:19	01:41:19	0.887	01:29:52	0.879	0.884	8.0
9	SM11	SCARLET RUNNER-11	Rob Date	15:36:45	01:36:45	0.933	01:30:16	0.921	0.929	9.0
10	SM1110	JOUST	Rod Warren	15:38:30	01:38:30	0.923	01:30:55	0.905	0.917	10.0

Race 7 (23/02/2019) PROVISIONAL RESULTS

Updated: 2/04/2019 10:04:28 PM

D1 PHS results Start : 14:00

Place	Sail Num	Boat Name	Skipper	Fin Tim	Elapsd	AHC	Cor'd T	BCH	CHC	Score
1	SM11	SCARLET RUNNER-11	Rob Date			0.929				8.9
1	SM9535	AMBITION	Chris Dare			1.196				11.2
3	SM45	CONQUEST	Paul Jacka	15:41:26	01:41:26	0.889	01:30:10	0.924	0.901	1.0
4	SM477	ESPRIT	Garry Anderson	15:43:30	01:43:30	0.890	01:32:07	0.905	0.895	2.0
5	SM4000	NICHE	Peter Jackson	15:44:08	01:44:08	0.898	01:33:31	0.900	0.899	3.0
6	SM69	DARK AND STORMY	S.Mellington T.Kourtis	15:52:29	01:52:29	0.833	01:33:42	0.833	0.833	4.0
7	SM4040	BANDIT	Ashley Trebilcock	15:44:25	01:44:25	0.900	01:33:58	0.897	0.899	5.0
8	SM1110	JOUST	Rod Warren	15:43:09	01:43:09	0.912	01:34:04	0.908	0.911	6.0
9	SM1113	PLAYLIST	Matt/Glenn/John Powell/Chesser/Cox	15:44:31	01:44:31	0.903	01:34:23	0.897	0.901	7.0
10	SM8	DREAM	Robert Green	15:34:20	01:34:20	1.001	01:34:26	0.993	0.998	8.0
11	SM4	WICKED	Mike And Mark Welsh	15:43:15	01:43:15	0.923	01:35:18	0.908	0.918	9.0

Race 8 (2/03/2019) PROVISIONAL RESULTS

Updated: 2/03/2019 5:12:59 PM

D1 PHS results Start : 10:30

Sail Num	Boat Name	Skipper	Score
SM4	WICKED	Mike And Mark Welsh	8.0
SM8	DREAM	Robert Green	2.0
SM11	SCARLET RUNNER-11	Rob Date	11.0
SM45	CONQUEST	Paul Jacka	6.0
SM69	DARK AND STORMY	S.Mellington T.Kourtis	3.0
SM1110	JOUST	Rod Warren	5.0
SM1113	PLAYLIST	Matt/Glenn/John Powell/Chesser/Cox	9.0
SM4000	NICHE	Peter Jackson	4.0
SM4040	BANDIT	Ashley Trebilcock	7.0
SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	10.0
SM477	ESPRIT	Garry Anderson	1.0

Race 9 (2/03/2019) PROVISIONAL RESULTS

Updated: 2/03/2019 5:26:03 PM

D1 PHS results Start : 13:30

Sail Num	Boat Name	Skipper	Score
SM4	WICKED	Mike And Mark Welsh	5.0
SM11	SCARLET RUNNER-11	Rob Date	9.0
SM45	CONQUEST	Paul Jacka	8.0
SM69	DARK AND STORMY	S.Mellington T.Kourtis	6.0
SM477	ESPRIT	Garry Anderson	10.0
SM1110	JOUST	Rod Warren	11.0
SM1113	PLAYLIST	Matt/Glenn/John Powell/Chesser/Cox	4.0
SM4000	NICHE	Peter Jackson	2.0
SM4040	BANDIT	Ashley Trebilcock	3.0
SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	7.0
SM8	DREAM	Robert Green	1.0

Race 10   (23/03/2019) PROVISIONAL RESULTS

Updated: 24/03/2019 1:13:36 PM

D1 PHS results Start : 14:00

Place	Sail Num	Boat Name	Skipper	Fin Tim	Elapsd	AHC	Cor'd T	BCH	CHC	Score
1	SM1110	JOUST	Rod Warren	15:24:33	01:24:33	0.899	01:16:01	0.965	0.911	1.0
2	SM11	SCARLET RUNNER-11	Rob Date	15:25:01	01:25:01	0.897	01:16:16	0.960	0.909	2.0
3	SM007	VELOCE	Phil Simpfendorfer	15:25:54	01:25:54	0.939	01:20:40	0.950	0.943	3.0
4	SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Mollison	15:28:52	01:28:52	0.909	01:20:47	0.918	0.912	4.0
5	SM4000	NICHE	Peter Jackson	15:30:35	01:30:35	0.901	01:21:37	0.901	0.901	5.0
6	SM4	WICKED	Mike And Mark Welsh	15:29:33	01:29:33	0.914	01:21:51	0.911	0.913	6.0
7	SM45	CONQUEST	Paul Jacka	15:32:40	01:32:40	0.888	01:22:17	0.881	0.886	7.0
8	SM4040	BANDIT	Ashley Trebilcock	15:31:52	01:31:52	0.896	01:22:19	0.888	0.893	8.0
9	SM8	DREAM	Robert Green	15:25:34	01:25:34	1.014	01:26:46	0.954	1.000	9.0
10	SM477	ESPRIT	Garry Anderson	15:38:32	01:38:32	0.886	01:27:18	0.828	0.874	10.0
RET	SM69	DARK AND STORMY	S.Mellington T.Kourtis			0.945				13.0
RET	SM424	BARNSTORMER	Brian Barnes			0.866				13.0

Series Results [PHS] for D1 up to Race 10 (Drops = 2)

Place	Ties	Sail Num	Boat Name	Skipper	From	Sers Score	Race 10	Race 9	Race 8	Race 7	Race 6	Race 5	Race 4	Race 3	Race 2	Race 1
1		SM8	DREAM	Robert Green	SYC	23.0	[9.0]	1.0	2.0	[8.0]	4.0	1.0	5.0	2.0	1.0	7.0
2		SM4000	NICHE	Peter Jackson	SYC	29.0	5.0	2.0	4.0	3.0	1.0	2.0	3.0	9.0	[20.0C]	[18.0C]
3	3.05	SM477	ESPRIT	Garry Anderson	SYC	37.0	10.0	[10.0]	1.0	2.0	3.0	3.0	[10.0]	6.0	4.0	8.0
4		SM1110	JOUST	Rod Warren	SYC	37.0	1.0	[11.0]	5.0	6.0	[10.0]	7.0	6.0	7.0	3.0	2.0
5		SM6115	SIERRA CHAINSAW	Mel/Andy/Pat Molliso	SYC	37.5	4.0	7.0	[10.0]	[20.0C]	2.0	5.0	1.5	8.0	9.0	1.0
6		SM45	CONQUEST	Paul Jacka	SYC	43.5	7.0	8.0	6.0	1.0	8.0	1.0	7.5	5.0	[20.0C]	[9.0]
7		SM4	WICKED	Mike And Mark Welsh	SYC	44.0	6.0	5.0	8.0	[9.0]	6.0	6.0	[9.0]	3.0	5.0	5.0
8		SM4040	BANDIT	Ashley Trebilcock	SYC	44.5	[8.0]	3.0	7.0	5.0	7.0	5.0	7.5	4.0	[8.0]	6.0
9		SM11	SCARLET RUNNER-11	Rob Date	SYC	49.9	2.0	9.0	11.0	8.9	9.0	3.0	4.0	[14.0]	[20.0C]	3.0
10		SM69	DARK AND STORMY	S.Mellington T.Kourt	SYC	54.5	13.0R	6.0	3.0	4.0	5.0	12.0	1.5	[13.0]	[20.0C]	10.0
11		SM1113	PLAYLIST	Matt/Glenn/John Powe	SYC	82.5	20.0C	4.0	9.0	7.0	[20.0C]	7.0	[20.0C]	10.5	7.0	18.0C
12		SM9535	AMBITION	Chris Dare	SYC/SSCBC	98.2	20.0C	20.0C	20.0C	11.2	[20.0C]	2.0	[20.0C]	1.0	6.0	18.0C
13		SM420	CADIBARRA 8	Paul Roberts	SYC	126.0	20.0C	20.0C	20.0C	20.0C	20.0C	20.0C	[20.0C]	[20.0C]	2.0	4.0
14		SM424	BARNSTORMER	Brian Barnes	SYC	131.0	13.0R	20.0C	20.0C	20.0C	20.0C	20.0C	[20.0C]	12.0	[20.0C]	13.0R
15		SM007	VELOCE	Phil Simpfendorfer	SYC	131.5	3.0	20.0C	20.0C	20.0C	20.0C	20.0C	[20.0C]	10.5	[20.0C]	18.0C
16		SM1100	SMOOTH CRIMINAL	Helen Cameron	SYC	147.0	20.0C	20.0C	20.0C	20.0C	20.0C	14.0	20.0C	[20.0C]	[20.0C]	13.0R
17		SM300	THE MUPPET SHOW	Peter Williams	SYC	158.0	20.0C	20.0C	20.0C	20.0C	20.0C	20.0C	20.0C	[20.0C]	[20.0C]	18.0C
17		SM1750	CARPE DIEM	Paul Commins	SYC	158.0	20.0C	20.0C	20.0C	20.0C	20.0C	20.0C	20.0C	[20.0C]	[20.0C]	18.0C
17		SM1112	JAKE	Stuart Lyon	SYC	158.0	20.0C	20.0C	20.0C	20.0C	20.0C	20.0C	20.0C	[20.0C]	[20.0C]	18.0#

APPENDIX B

The Yacht Race Series

The Sandringham Yacht Club (SYC) 2018-19 Summer Aggregate, Division 1, Performance Handicapping System (PHS) was a 10-race series with published results showing race and handicapping details for races 1,2,3,4,6,7 and 10 (see Appendix A). For these **seven races** we have selected the 10 yachts that completed at least five of these races: *Bandit, Conquest, Dark and Stormy, Dream, Esprit, Joust, Niche, Scarlett Runner II, Sierra Chainsaw* and *Wicked*. Each yacht's elapsed times *ET* have been extracted from the published race results and a yacht's initial handicap is taken as the allocated handicap *AHC* for that race shown in the race results; these data form our **yacht race series** and are shown below.

Yacht	Initial Handicap	R1a	R2a	R3a	R4a	R6a	R7a	R10a
				Elapsed Time *ET*				
Bandit	0.910	1:23:33	2:53:45	1:49:01	1:13:59	1:39:24	1:44:25	1:31:52
Conquest	0.914	1:26:30	DNS	1:51:07	1:15:19	1:41:19	1:41:26	1:32:40
Dark and Stormy	0.885	1:34:59	DNS	1:58:50	1:17:19	1:46:58	1:52:29	RET
Dream	0.948	1:21:47	2:33:06	1:38:13	1:06:23	1:28:47	1:34:20	1:25:34
Esprit	0.895	1:26:58	2:52:06	1:54:36	1:15:30	1:39:49	1:43:30	1:38:32
Joust	0.935	1:18:40	2:39:23	1:48:14	1:12:20	1:38:30	1:43:09	1:24:33
Niche	0.900	DNS	DNS	1:54:53	1:13:00	1:38:20	1:44:08	1:30:35
Scarlett Runner II	0.935	1:19:28	DNS	1:52:24	1:11:07	1:36:45	DNS	1:25:01
Sierra Chainsaw	0.930	1:18:59	2:51:11	1:51:12	1:10:51	1:35:17	DNS	1:28:52
Wicked	0.926	1:21:27	2:45:41	1:45:44	1:12:16	1:36:24	1:43:15	1:29:33

Table B1. Yacht Race Series elapsed times and initial Allocated Handicaps
Elapsed Times are shown in hour (h), minute (m), second (s) format as h:mm:ss
DNS = Did Not Start, RET = Retired

Race Series Scored using *TopYacht*

The following tables show the results for each of the seven races in the series using *TopYacht* scoring with 45% mark boat [see equation (4)] and $CHC = \frac{1}{3}BCH + \frac{2}{3}AHC = AHC + \frac{1}{3}PI$ where $PI = BCH - AHC$

Yacht	*ET*	*AHC*	*CT* (sec)	Place	*BCH*	*PI*	*CHC*
Sierra Chainsaw	1:18:59	0.930	4407.270	**1**	0.955	0.025	0.938
Joust	1:18:40	0.935	4413.200	**2**	0.959	0.024	0.943
Scarlett Runner II	1:19:28	0.935	4458.080	**3**	0.949	0.014	0.940
Wicked	1:21:27	0.926	4525.362	**4**	0.926	0.000	0.926
Bandit	1:23:33	0.910	4561.830	**5**	0.903	-0.007	0.908
Dream	1:21:47	0.948	4651.836	**6**	0.922	-0.026	0.939
Esprit	1:26:58	0.895	4670.110	**7**	0.867	-0.028	0.886
Conquest	1:26:30	0.914	4743.660	**8**	0.872	-0.042	0.900
Dark and Stormy	1:34:59	0.885	5043.615	**9**	0.794	-0.091	0.855
Niche	DNS			**12**			

Table B2. Race 1a: Mark Boat is *Wicked.*

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	CHC
Dream	2:33:06	0.939	8625.654	**1**	0.996	0.057	0.958
Joust	2:39:23	0.943	9017.909	**2**	0.957	0.014	0.948
Esprit	2:52:06	0.886	9148.836	**3**	0.886	0.000	0.886
Wicked	2:45:41	0.926	9205.366	**4**	0.920	-0.006	0.924
Bandit	2:53:45	0.908	9465.900	**5**	0.878	-0.030	0.898
Sierra Chainsaw	2:51:11	0.938	9634.198	**6**	0.891	-0.047	0.922
Scarlett Runner II	DNS			**12**			
Conquest	DNS			**12**			
Dark and Stormy	DNS			**12**			
Niche	DNS			**12**			

Table B3. Race 2a: Mark Boat is *Esprit*.

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	CHC
Dream	1:38:13	0.958	5645.494	**1**	1.034	0.076	0.983
Wicked	1:45:44	0.924	5861.856	**2**	0.960	0.036	0.936
Bandit	1:49:01	0.898	5873.818	**3**	0.931	0.033	0.909
Conquest	1:51:07	0.900	6000.300	**4**	0.914	0.014	0.905
Esprit	1:54:36	0.886	6092.136	**5**	0.886	0.000	0.886
Dark and Stormy	1:58:50	0.855	6096.150	**6**	0.854	-0.001	0.855
Sierra Chainsaw	1:51:12	0.922	6151.584	**7**	0.913	-0.009	0.919
Joust	1:48:14	0.948	6156.312	**8**	0.938	-0.010	0.945
Niche	1:54:53	0.900	6203.700	**9**	0.884	-0.016	0.895
Scarlett Runner II	1:52:24	0.940	6339.360	**10**	0.903	-0.037	0.928

Table B4. Race 3a: Mark Boat is *Esprit*

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	CHC
Sierra Chainsaw	1:10:51	0.919	3906.669	**1**	0.933	0.014	0.924
Dream	1:06:23	0.983	3915.289	**2**	0.996	0.013	0.987
Niche	1:13:00	0.895	3920.100	**3**	0.906	0.011	0.899
Scarlett Runner II	1:11:07	0.928	3959.776	**4**	0.930	0.002	0.929
Dark and Stormy	1:17:19	0.855	3966.345	**5**	0.855	0.000	0.855
Esprit	1:15:30	0.886	4013.580	**6**	0.876	-0.010	0.883
Bandit	1:13:59	0.909	4035.051	**7**	0.894	-0.015	0.904
Wicked	1:12:16	0.936	4058.496	**8**	0.915	-0.021	0.929
Conquest	1:15:19	0.905	4089.695	**9**	0.878	-0.027	0.896
Joust	1:12:20	0.945	4101.300	**10**	0.914	-0.031	0.935

Table B5. Race 4a: Mark Boat is *Dark and Stormy*.

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	CHC
Dream	1:28:47	0.987	5257.749	1	1.009	0.022	0.994
Sierra Chainsaw	1:35:17	0.924	5282.508	2	0.940	0.016	0.929
Esprit	1:39:49	0.883	5288.287	3	0.897	0.014	0.888
Niche	1:38:20	0.899	5304.100	4	0.911	0.012	0.903
Wicked	1:36:24	0.929	5373.336	5	0.929	0.000	0.929
Bandit	1:39:24	0.904	5391.456	6	0.901	-0.003	0.903
Scarlett Runner II	1:36:45	0.929	5392.845	7	0.926	-0.003	0.928
Conquest	1:41:19	0.896	5446.784	8	0.884	-0.012	0.892
Dark and Stormy	1:46:58	0.855	5487.390	9	0.837	-0.018	0.849
Joust	1:38:30	0.935	5525.850	10	0.909	-0.026	0.926

Table B6. Race 6a: Mark Boat is *Wicked.*

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	CHC
Conquest	1:41:26	0.892	5428.712	1	0.927	0.035	0.904
Esprit	1:43:30	0.888	5514.480	2	0.909	0.021	0.895
Dream	1:34:20	0.994	5626.040	3	0.997	0.003	0.995
Niche	1:44:08	0.903	5641.944	4	0.903	0.000	0.903
Bandit	1:44:25	0.903	5657.295	5	0.901	-0.002	0.902
Dark and Stormy	1:52:29	0.849	5729.901	6	0.836	-0.013	0.845
Joust	1:43:09	0.926	5731.014	7	0.912	-0.014	0.921
Wicked	1:43:15	0.929	5755.155	8	0.911	-0.018	0.923
Sierra Chainsaw	DNS			12			
Scarlett Runner II	DNS			12			

Table B7. Race 7a: Mark Boat is *Niche.*

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	CHC
Joust	1:24:33	0.921	4672.233	1	0.976	0.055	0.939
Scarlett Runner II	1:25:01	0.928	4733.728	2	0.971	0.043	0.942
Niche	1:30:35	0.903	4907.805	3	0.911	0.008	0.906
Sierra Chainsaw	1:28:52	0.929	4953.428	4	0.929	0.000	0.929
Wicked	1:29:33	0.923	4959.279	5	0.922	-0.001	0.923
Bandit	1:31:52	0.902	4971.824	6	0.899	-0.003	0.901
Conquest	1:32:40	0.904	5026.240	7	0.891	-0.013	0.900
Dream	1:25:34	0.995	5108.330	8	0.965	-0.030	0.985
Esprit	1:38:32	0.895	5291.240	9	0.838	-0.057	0.876
Dark and Stormy	RET			11			

Table B8. Race 10a: Mark Boat is *Sierra Chainsaw.*

The race finish placings for the yachts in the race series are aggregated in Table B9 and the total of the best five placings are used to decide the Series Place of each yacht. Yachts that did not start (DNS) a race are given a race place of 12 (two more than the number of entrants in the series) and yachts that retired (RET) from a race a given a race place of 11 (one more than the number of entrants in the series).

| | Race Placings | | | | | | | Total | Series |
Yacht	R1a	R2a	R3a	R4a	R6a	R7a	R10a	(Best 5)	Place
Dream	6	1	1	2	1	3	8	8	**1**
Sierra Chainsaw	1	6	7	1	2	12	4	14	**2**
Esprit	7	3	5	6	3	2	9	19	**3**
Joust	2	2	8	10	10	7	1	20	**4**
Wicked	4	4	2	8	5	8	5	20	**5**
Niche	12	12	9	3	4	4	3	23	**6**
Bandit	5	5	3	7	6	5	6	24	**7**
Scarlett Runner II	3	12	10	4	7	12	2	26	**8**
Conquest	8	12	4	9	8	1	7	28	**9**
Dark and Stormy	9	12	6	5	9	6	11	35	**10**

Table B9. Yacht Race Series (*TopYacht* scoring): Race Placings; Total Points (Best 5); Overall Place
DNS = 12, RET = 11

The Performance Indicators $PI = AHC - BCH$ for each race of the series are shown in Table B10 together with two other values denoted MAD and $\hat{\sigma}$. These are tabulated for each yacht in the series (right-most columns) and for each race (last two rows). MAD is Median Absolute Deviation and $\hat{\sigma} = 1.4826(\text{MAD})$ is an estimate of the standard deviation of the *PI*'s (see Appendix E). Both are measures of variation of the data and are used in the development of a Performance Indicator Filter (see Appendix D)

| | Performance Indicator *PI* | | | | | | | *Series* | |
Yacht	R1a	R2a	R3a	R4a	R6a	R7a	R10a	MAD	$\hat{\sigma}$
Bandit	-0.007	-0.030	0.033	-0.015	-0.003	-0.002	-0.003	0.004	0.006
Conquest	-0.042	DNS	0.014	-0.027	-0.012	0.035	-0.013	0.020	0.030
Dark and Stormy	-0.091	DNS	-0.001	0.000	-0.018	-0.013	RET	0.010	0.015
Dream	-0.026	0.057	0.076	0.013	0.022	0.003	-0.030	0.025	0.037
Esprit	-0.028	0.000	0.000	-0.010	0.014	0.021	-0.057	0.017	0.025
Joust	0.024	0.014	-0.010	-0.031	-0.026	-0.014	0.055	0.023	0.034
Niche	DNS	DNS	-0.016	0.011	0.012	0.000	0.008	0.013	0.019
Scarlett Runner II	0.014	DNS	-0.037	0.002	-0.003	DNS	0.043	0.017	0.025
Sierra Chainsaw	0.025	-0.047	-0.009	0.014	0.016	DNS	0.000	0.018	0.027
Wicked	0.000	-0.006	0.036	-0.021	0.000	-0.018	-0.001	0.003	0.004
MAD for race	0.025	0.022	0.015	0.015	0.015	0.010	0.011		
$\hat{\sigma}$ for race	0.037	0.033	0.022	0.022	0.022	0.016	0.016		

Table B10. Yacht Race Series (*TopYacht* scoring): Performance Indicators; MAD; $\hat{\sigma}$.
DNS = Did Not Start, RET = Retired

Race Series Scored using *Performance Indicator Filter*

The following tables show the results for each of the seven races in the series scored using our new method of Standard Corrected Time $SCT = \dfrac{1}{n}\big(S + R\big)$ where S and R are the sum and the range of the corrected times respectively and n is the number of entrants in the race. And with $CHC_k = AHC_k + z_k$ where $z_k = z_{k-1} + K\big(PI_k - z_{k-1}\big)$ for $k > 0$ and $z_0 = 0$ is a Performance Indicator Filter (PIF) with $K = 2/5$

Yacht	*ET*	*AHC*	*CT* (sec)	Place	*BCH*	*PI*	old z	new z	*CHC*
Sierra Chainsaw	1:18:59	0.930	4407.270	**1**	0.987	0.057	0	0.023	0.953
Joust	1:18:40	0.935	4413.200	**2**	0.991	0.056	0	0.023	0.958
Scarlett Runner II	1:19:28	0.935	4458.080	**3**	0.981	0.046	0	0.019	0.954
Wicked	1:21:27	0.926	4525.362	**4**	0.957	0.031	0	0.013	0.939
Bandit	1:23:33	0.910	4561.830	**5**	0.933	0.023	0	0.009	0.919
Dream	1:21:47	0.948	4651.836	**6**	0.954	0.006	0	0.002	0.950
Esprit	1:26:58	0.895	4670.110	**7**	0.897	0.002	0	0.001	0.896
Conquest	1:26:30	0.914	4743.660	**8**	0.902	-0.012	0	-0.005	0.909
Dark and Stormy	1:34:59	0.885	5043.615	**9**	0.821	-0.064	0	-0.026	0.859
Niche	DNS			**12**					

Table B11. Race 1a: *SCT* is 4679.034 sec

Yacht	*ET*	*AHC*	*CT* (sec)	Place	*BCH*	*PI*	old z	new z	*CHC*
Dream	2:33:06	0.950	8726.700	**1**	1.032	0.082	0.002	0.034	0.984
Joust	2:39:23	0.958	9161.354	**2**	0.992	0.034	0.022	0.027	0.985
Esprit	2:52:06	0.896	9252.096	**3**	0.918	0.022	0.001	0.010	0.906
Wicked	2:45:41	0.939	9334.599	**4**	0.954	0.015	0.012	0.014	0.953
Bandit	2:53:45	0.919	9580.575	**5**	0.910	-0.009	0.009	0.002	0.921
Sierra Chainsaw	2:51:11	0.953	9788.263	**6**	0.923	-0.030	0.023	0.002	0.955
Scarlett Runner II	DNS			**12**					
Conquest	DNS			**12**					
Dark and Stormy	DNS			**12**					
Niche	DNS			**12**					

Table B12. Race 2a: *SCT* is 9484.192 sec

Yacht	*ET*	*AHC*	*CT* (sec)	Place	*BCH*	*PI*	old *z*	new *z*	*CHC*
Dream	1:38:13	0.984	5798.712	**1**	1.058	0.074	0.034	0.050	1.034
Bandit	1:49:01	0.921	6024.261	**2**	0.953	0.032	0.002	0.014	0.935
Wicked	1:45:44	0.953	6045.832	**3**	0.982	0.029	0.014	0.020	0.973
Conquest	1:51:07	0.909	6060.303	**4**	0.935	0.026	-0.005	0.007	0.916
Dark and Stormy	1:58:50	0.859	6124.670	**5**	0.874	0.015	-0.026	-0.010	0.849
Niche	1:54:53	0.900	6203.700	**6**	0.904	0.004	0.000	0.002	0.902
Esprit	1:54:36	0.906	6229.656	**7**	0.906	0.000	0.010	0.006	0.912
Sierra Chainsaw	1:51:12	0.955	6371.760	**8**	0.934	-0.021	0.002	-0.007	0.948
Joust	1:48:14	0.985	6396.590	**9**	0.960	-0.025	0.027	0.006	0.991
Scarlett Runner II	1:52:24	0.954	6433.776	**10**	0.924	-0.030	0.019	-0.001	0.953

Table B13. Race 3a: *SCT* is 6232.432 sec

Yacht	*ET*	*AHC*	*CT* (sec)	Place	*BCH*	*PI*	old *z*	new *z*	*CHC*
Dark and Stormy	1:17:19	0.849	3938.511	**1**	0.893	0.044	-0.010	0.011	0.860
Niche	1:13:00	0.902	3950.760	**2**	0.945	0.043	0.002	0.019	0.921
Sierra Chainsaw	1:10:51	0.948	4029.948	**3**	0.974	0.026	-0.007	0.006	0.954
Scarlett Runner II	1:11:07	0.953	4066.451	**4**	0.970	0.017	-0.001	0.006	0.959
Dream	1:06:23	1.034	4118.422	**5**	1.040	0.006	0.050	0.032	1.066
Esprit	1:15:30	0.912	4131.360	**6**	0.914	0.002	0.006	0.004	0.916
Conquest	1:15:19	0.916	4139.404	**7**	0.916	0.000	0.007	0.004	0.920
Bandit	1:13:59	0.935	4150.465	**8**	0.933	-0.002	0.014	0.008	0.943
Wicked	1:12:16	0.973	4218.928	**9**	0.955	-0.018	0.020	0.005	0.978
Joust	1:12:20	0.991	4300.940	**10**	0.954	-0.037	0.006	-0.011	0.980

Table B14. Race 4a: *SCT* is 4140.762 sec

Yacht	*ET*	*AHC*	*CT* (sec)	Place	*BCH*	*PI*	old *z*	new *z*	*CHC*
Niche	1:38:20	0.921	5433.900	**1**	0.952	0.031	0.019	0.024	0.945
Sierra Chainsaw	1:35:17	0.954	5454.018	**2**	0.982	0.028	0.006	0.015	0.969
Esprit	1:39:49	0.916	5485.924	**3**	0.938	0.022	0.004	0.011	0.927
Dark and Stormy	1:46:58	0.860	5519.480	**4**	0.875	0.015	0.011	0.013	0.873
Scarlett Runner II	1:36:45	0.959	5566.995	**5**	0.967	0.008	0.006	0.007	0.966
Conquest	1:41:19	0.920	5592.680	**6**	0.924	0.004	0.004	0.004	0.924
Bandit	1:39:24	0.943	5624.052	**7**	0.942	-0.001	0.008	0.004	0.947
Wicked	1:36:24	0.978	5656.752	**8**	0.971	-0.007	0.005	0.000	0.978
Dream	1:28:47	1.066	5678.582	**9**	1.054	-0.012	0.032	0.015	1.081
Joust	1:38:30	0.980	5791.800	**10**	0.950	-0.030	-0.011	-0.018	0.962

Table B15. Race 6a: *SCT* is 5616.208 sec

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	old z	new z	CHC
Conquest	1:41:26	0.924	5623.464	**1**	0.980	0.056	0.005	0.025	0.949
Esprit	1:43:30	0.927	5756.670	**2**	0.961	0.034	0.012	0.020	0.947
Dark and Stormy	1:52:29	0.873	5891.877	**3**	0.884	0.011	0.013	0.012	0.885
Niche	1:44:08	0.945	5904.360	**4**	0.955	0.010	0.024	0.018	0.963
Bandit	1:44:25	0.947	5932.955	**5**	0.952	0.005	0.006	0.005	0.952
Joust	1:43:09	0.962	5953.818	**6**	0.964	0.002	-0.020	-0.010	0.952
Wicked	1:43:15	0.978	6058.710	**7**	0.963	-0.015	0.001	-0.006	0.972
Dream	1:34:20	1.081	6118.460	**8**	1.054	-0.027	0.014	-0.002	1.079
Sierra Chainsaw	*DNS*			**12**					
Scarlett Runner II	*DNS*			**12**					

Table B16. Race 7a: *SCT* is 5966.914 sec

Yacht	ET	AHC	CT (sec)	Place	BCH	PI	old z	new z	CHC
Joust	1:24:33	0.952	4829.496	**1**	1.047	0.095	-0.010	0.032	0.984
Scarlett Runner II	1:25:01	0.966	4927.566	**2**	1.041	0.075	0.007	0.034	1.000
Sierra Chainsaw	1:28:52	0.969	5166.708	**3**	0.996	0.027	0.015	0.020	0.989
Wicked	1:29:33	0.972	5222.556	**4**	0.989	0.017	-0.006	0.003	0.975
Niche	1:30:35	0.963	5233.905	**5**	0.977	0.014	0.018	0.017	0.980
Bandit	1:31:52	0.952	5247.424	**6**	0.964	0.012	0.005	0.008	0.960
Conquest	1:32:40	0.949	5276.440	**7**	0.955	0.007	0.025	0.018	0.967
Dream	1:25:34	1.079	5539.586	**8**	1.035	-0.044	0.002	-0.017	1.062
Esprit	1:38:32	0.947	5598.664	**9**	0.899	-0.048	0.020	-0.007	0.940
Dark and Stormy	*RET*			**11**					

Table B17. Race 10a: *SCT* is 5312.390 sec

The race finish placings for the yachts in the race series are aggregated in Table B18 and the total of the best five placings are used to decide the Series Place of each yacht. Yachts that did not start (DNS) a race are given a race place of 12 (two more than the number of entrants in the series) and yachts that retired (RET) from a race a given a race place of 11 (one more than the number of entrants in the series).

Yacht	R1a	R2a	R3a	R4a	R6a	R7a	R10a	Total (Best 5)	Place
				Race Placings				Total	
Sierra Chainsaw	1	6	8	3	2	12	3	15	**1**
Niche	12	12	6	2	1	4	5	18	**2**
Joust	2	2	9	10	10	6	1	20	**3**
Dream	6	1	1	5	9	8	8	21	**4**
Esprit	7	3	7	6	3	2	9	21	**5**
Dark and Stormy	9	12	5	1	4	3	11	22	**6**
Wicked	4	4	3	9	8	7	4	22	**7**
Bandit	5	5	2	8	7	5	6	23	**8**
Scarlett Runner II	3	12	10	4	5	12	2	24	**9**
Conquest	8	12	4	7	6	1	7	25	**10**

Table B18. Yacht Race Series (PIF scoring $K = 2/5$): Race Placings; Total Points (Best 5); Overall Place
DNS $= 12$, RET $= 11$

The Performance Indicators $PI = AHC - BCH$ for each race of the series are shown in Table B19 together with two other values denoted MAD and $\hat{\sigma}$. These are tabulated for each yacht in the series (right-most columns) and for each race (last two rows). MAD is Median Absolute Deviation and $\hat{\sigma} = 1.4826(\text{MAD})$ is an estimate of the standard deviation of the PI's (see Appendix F). Both are measures of variation of the data and are used in the development of a Performance Indicator Filter (see Appendix E)

Yacht	R1a	R2a	R3a	R4a	R6a	R7a	R10a	MAD	$\hat{\sigma}$
			Performance Indicator *PI*					*Series*	
Bandit	0.023	-0.009	0.032	-0.002	-0.001	0.005	0.012	0.007	0.010
Conquest	-0.012	DNS	0.026	0.000	0.004	0.056	0.006	0.010	0.016
Dark and Stormy	-0.064	DNS	0.015	0.044	0.015	0.011	RET	0.010	0.015
Dream	0.006	0.082	0.074	0.006	-0.012	-0.027	-0.044	0.032	0.047
Esprit	0.002	0.022	0.000	0.002	0.022	0.034	-0.048	0.017	0.025
Joust	0.056	0.034	-0.025	-0.037	-0.030	0.002	0.095	0.035	0.052
Niche	DNS	DNS	0.004	0.043	0.031	0.010	0.014	0.009	0.013
Scarlett Runner II	0.046	DNS	-0.030	0.017	0.008	DNS	0.075	0.035	0.052
Sierra Chainsaw	0.057	-0.030	-0.021	0.026	0.028	DNS	0.027	0.024	0.036
Wicked	0.031	0.015	0.029	-0.018	-0.007	-0.015	0.017	0.020	0.030
MAD for race	0.023	0.023	0.025	0.016	0.014	0.013	0.022		
$\hat{\sigma}$ for race	0.034	0.034	0.037	0.024	0.021	0.019	0.033		

Table B19. Yacht Race Series (PIF scoring): Performance Indicators; MAD; $\hat{\sigma}$.
DNS = Did Not Start, RET = Retired

APPENDIX C

A New Method of Determining the Standard Corrected Time

A method of determining the *Standard Corrected Time* $\left(SCT \right)$ of a yacht race can be deduced from the following two statements:

    A If every yacht sailed to their *true handicap H* then their corrected times would be identical.

    B The range of corrected times in an actual race is the sum of each yacht's *handicap error x*.

Considering statement A. If all the corrected times are the same, call this value the *standard corrected time SCT* and define a yacht's true handicap H as

$$H = AHC + \delta H \tag{23}$$

where AHC is the yacht's allocated handicap derived from a set of rules and δH is a small unknown correction. Now, we can write, for any individual yacht

$$
\begin{aligned}
ET \times \left(AHC + \delta H \right) &= SCT \\
ET \times AHC + ET \times \delta H &= SCT
\end{aligned}
\tag{24}
$$

and defining a yacht's handicap error x as

$$x = ET \times \delta H \tag{25}$$

Now using (1) we may write an *observation equation* for each of the $k = 1, 2, 3, \ldots, n$ yachts in the race as

$$x_k - SCT = -CT_k \tag{26}$$

And, from statement B we may write a single additional *constraint equation*

$$x_1 + x_2 + \cdots + x_n = R \tag{27}$$

where $R = \max \left\{ CT \right\} - \min \left\{ CT \right\}$ is a positive quantity equal to the range of the corrected times and $\left\{ CT \right\}$ denotes the set of corrected times of the yachts in a race.

Suppose there were $n = 5$ yachts in a race, then the set of observation equations and the constraint equation would have the following form

$$
\begin{aligned}
x_1 - SCT &= -CT_1 \\
x_2 - SCT &= -CT_2 \\
x_3 - SCT &= -CT_3 \\
x_4 - SCT &= -CT_4 \\
x_5 - SCT &= -CT_5 \\
x_1 + x_2 + x_3 + x_4 + x_5 &= R
\end{aligned}
\tag{28}
$$

We may write these equations in matrix form $\mathbf{B}\mathbf{x} = \mathbf{f}$ as

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & -1 \\
0 & 1 & 0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & -1 \\
1 & 1 & 1 & 1 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ SCT
\end{bmatrix}
=
\begin{bmatrix}
-CT_1 \\ -CT_2 \\ -CT_3 \\ -CT_4 \\ -CT_5 \\ R
\end{bmatrix}
\tag{29}
$$

where $\mathbf{B}$ is a coefficient matrix having $n + 1$ rows and $n + 1$ columns, $\mathbf{x}$ is a column vector having $n + 1$ rows and containing the unknown handicap errors $x_1, x_2, \ldots, x_n$ and the unknown standard corrected time

SCT. The column vector $\mathbf{f}$ on the right-hand side of the equals sign has $n+1$ rows containing the negative corrected times $-CT_1, -CT_2, \ldots, -CT_n$ and the range of the corrected times R.

The solution for the unknown vector $\mathbf{x}$ is given by matrix algebra as

$$\mathbf{x} = \mathbf{B}^{-1}\mathbf{f} \tag{30}$$

where the superscript $^{-1}$ denotes matrix inverse defined as $\mathbf{B}\mathbf{B}^{-1} = \mathbf{I}$ and $\mathbf{I}$ is the Identity matrix.

Because of the form of $\mathbf{B}$ its inverse $\mathbf{B}^{-1}$ has a simple structure

$$\mathbf{B}^{-1} = \left[\begin{array}{ccccc|c}
\left(1-\dfrac{1}{n}\right) & -\dfrac{1}{n} & -\dfrac{1}{n} & -\dfrac{1}{n} & \cdots & -\dfrac{1}{n} & \dfrac{1}{n} \\
-\dfrac{1}{n} & \left(1-\dfrac{1}{n}\right) & -\dfrac{1}{n} & -\dfrac{1}{n} & \cdots & -\dfrac{1}{n} & \dfrac{1}{n} \\
-\dfrac{1}{n} & -\dfrac{1}{n} & \left(1-\dfrac{1}{n}\right) & -\dfrac{1}{n} & \cdots & -\dfrac{1}{n} & \dfrac{1}{n} \\
-\dfrac{1}{n} & -\dfrac{1}{n} & -\dfrac{1}{n} & \left(1-\dfrac{1}{n}\right) & \cdots & -\dfrac{1}{n} & \dfrac{1}{n} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
-\dfrac{1}{n} & -\dfrac{1}{n} & -\dfrac{1}{n} & -\dfrac{1}{n} & \cdots & \left(1-\dfrac{1}{n}\right) & \dfrac{1}{n} \\
\hline
-\dfrac{1}{n} & -\dfrac{1}{n} & -\dfrac{1}{n} & -\dfrac{1}{n} & \cdots & -\dfrac{1}{n} & \dfrac{1}{n}
\end{array}\right] \tag{31}$$

$\mathbf{B}^{-1}$ is square, having $n+1$ rows and columns that can be partitioned into 4 parts. The upper-left part being a square and symmetric submatrix of size n rows by n columns where the leading-diagonal elements are identical and equal to $1 - 1/n$ and the off-diagonal elements are all identical and equal to $-1/n$. The upper-right submatrix is a column vector having n rows with all elements equal to $1/n$ and the lower-left submatrix is a row vector having n columns with all elements equal to $-1/n$. The lower-right submatrix contains a single element equal to $1/n$.

For the case where $n = 5$

$$\mathbf{B}^{-1} = \left[\begin{array}{ccccc|c}
0.8 & -0.2 & -0.2 & -0.2 & -0.2 & 0.2 \\
-0.2 & 0.8 & -0.2 & -0.2 & -0.2 & 0.2 \\
-0.2 & -0.2 & 0.8 & -0.2 & -0.2 & 0.2 \\
-0.2 & -0.2 & -0.2 & 0.8 & -0.2 & 0.2 \\
-0.2 & -0.2 & -0.2 & -0.2 & 0.8 & 0.2 \\
\hline
-0.2 & -0.2 & -0.2 & -0.2 & -0.2 & 0.2
\end{array}\right] \tag{32}$$

The solution for $\mathbf{x}$ is given by (30) and using the general form of $\mathbf{B}^{-1}$ given by (31) we may give the solutions for the elements of $\mathbf{x}$ as

$$x_1 = -\left(1 - \frac{1}{n}\right)CT_1 + \frac{1}{n}CT_2 + \frac{1}{n}CT_3 + \cdots + \frac{1}{n}CT_n + \frac{1}{n}R$$

$$= \frac{1}{n}\left\{\left(\sum_{k=1}^{n}CT_k\right) + R\right\} - CT_1$$

$$x_2 = \frac{1}{n}\left\{\left(\sum_{k=1}^{n}CT_k\right) + R\right\} - CT_2$$

$$x_3 = \frac{1}{n}\left\{\left(\sum_{k=1}^{n}CT_k\right) + R\right\} - CT_3 \tag{33}$$

$$\vdots$$

$$x_n = \frac{1}{n}\left\{\left(\sum_{k=1}^{n}CT_k\right) + R\right\} - CT_n$$

$$SCT = \frac{1}{n}\left\{\left(\sum_{k=1}^{n}CT_k\right) + R\right\}$$

If the standard corrected time is calculated first by defining the sum of the corrected times as S where

$$S = \sum_{k=1}^{n}CT_k \tag{34}$$

then

$$SCT = \frac{1}{n}\left(S + R\right) \tag{35}$$

and the handicap errors $x_1, x_2, x_3, \ldots, x_n$ are given by

$$x_k = SCT - CT_k \quad \text{for} \quad k = 1, 2, 3, \ldots, n \tag{36}$$

Finally, the corrections to each yacht's handicap are given by

$$\delta H_k = \frac{x_k}{ET_k} \tag{37}$$

and then, according to (5) each yacht's δH is added to their allocated handicap AHC to give their true handicap H.

We now show that the true handicap H is in fact the back calculated handicap BCH defined by (3).

Using (7) and (8) we may write $SCT - CT_k = ET_k + \delta H_k$ which can be rearranged as $\delta H_k = \dfrac{SCT}{ET_k} - \dfrac{CT_k}{ET_k}$

where the first term on the right-hand side of the equals sign is the back calculated handicap [see (3)] and the second term is the allocated handicap [see (1)], hence

$$\delta H_k = BCH_k - AHC_k \tag{38}$$

and from (5)

$$H_k = BCH_k \tag{39}$$

As an example, we will use Race 1a from our Yacht Race Series in Appendix B.

| | | Allocated | | | Back Calculated |
Yacht	Elapsed Time ET	Handicap AHC	Corrected Time CT (sec)	Place	Handicap BCH
Sierra Chainsaw	1:18:59 (4739 sec)	0.930	4407.270	1	0.987
Joust	1:18:40 (4720)	0.935	4413.200	2	0.991
Scarlett Runner II	1:19:28 (4768)	0.935	4458.080	3	0.981
Wicked	1:21:27 (4887)	0.926	4525.362	4	0.957
Bandit	1:23:33 (5013)	0.910	4561.830	5	0.933
Dream	1:21:47 (4907)	0.948	4651.836	6	0.954
Esprit	1:26:58 (5218)	0.895	4670.110	7	0.897
Conquest	1:26:30 (5190)	0.914	4743.660	8	0.902
Dark and Stormy	1:34:59 (5699)	0.885	5043.615	9	0.821
Niche	DNS				

Table C1. Race 1a of the Yacht Race Series (Appendix B). Yachts shown in corrected time order. Time in hour (h), minute (m) second (s) format as h:mm:ss

From Table C1, the range of corrected times $R = 5043.615 - 4407.270 = 636.345$ sec and the sum of the corrected times is $S = 41474.963$ sec and using equation (35)

$$SCT = \frac{1}{n}\left(S + R\right) = 4679.034 \text{ sec}$$

and

$$x_k = SCT - CT_k$$

giving

$$x_1 = 271.764 \quad x_2 = 265.834 \quad x_3 = 220.954 \quad x_4 = 153.672 \quad x_5 = 117.204$$
$$x_6 = 27.198 \quad x_7 = 8.924 \quad x_8 = -64.626 \quad x_9 = -364.581$$

where the values are seconds and the subscripts are the finish place order in the race (see Table C1).

The correction to each yacht's handicap is given by (37) and as an example we take the yacht *Dream* who was the 6th boat over the line with an elapsed time $ET_6 = 1:21:47 = 4907$ sec and with handicap error $x_6 = 27.198$ sec. The handicap correction δH for *Dream* is

$$\delta H_6 = \frac{x_6}{ET_6} = \frac{27.198}{4907} = 0.005543 \approx 0.006$$

and *Dream's* true handicap H is

$$H_6 = AHC_6 + \delta H_6 = 0.948 + 0.006 = 0.954$$

Also, as we have shown above, this is *Dream's* back calculated handicap computed using (3)

$$BCH_6 = \frac{SCT}{ET_6} = \frac{4679.034}{4907} = 0.953543 \approx 0.954$$

The back calculated handicaps for Race 1a using this new method of computing the standard corrected time are tabulated in Table C1

APPENDIX D

Moving Averages, Recurrence Relations, Weighting and Exponential Smoothing

Moving Averages

In time series[4] analysis, moving averages are often used to produce trendlines that provide a visual summary of the data. Suppose a time series of length n has y-variables in an ordered list $y_1, y_2, y_3, ..., y_n$ where the subscripts denotes times $t = 1, 2, 3, ..., n$, and a 'window' of period (or length) p, where $p = 2k + 1$ is an odd number, is superimposed over this list and the average of the values in the window is calculated and denoted A_t. And then the window is moved one place to the right over the ordered list and a new average calculated and denoted A_{t+1} and this process repeated for averages A_{t+2}, A_{t+3}, etc. If the initial position of the window has the first element of the window coinciding with the first element of the ordered list then three simple **moving averages** are

Right
$$A_t = \frac{1}{2k+1} \sum_{j=-2k}^{0} y_{t+j} \quad \text{for} \quad t = 2k+1, 2k+2, 2k+3, ..., n \tag{40}$$

Left
$$A_t = \frac{1}{2k+1} \sum_{j=0}^{2k} y_{t+j} \quad \text{for} \quad t = 1, 2, 3, ..., n - 2k \tag{41}$$

Centre
$$A_t = \frac{1}{2k+1} \sum_{j=-k}^{k} y_{t+j} \quad \text{for} \quad t = k+1, k+2, ..., n - k \tag{42}$$

Each equation will yield $n - p + 1 = n - 2k$ identical values but they will be denoted in differing sequences. For example, say $n = 10$, $p = 2k + 1 = 5$ and $k = 2$ there will be $n - 2k = 6$ averages and equation (40) will yield the sequence $\{A_5, A_6, A_7, A_8, A_9, A_{10}\}$ where the initial average is associated with the right-hand end of the window, (41) will yield $\{A_1, A_2, A_3, A_4, A_5, A_6\}$ where the initial average is associated with the left-hand end of the window and (42) yields $\{A_3, A_4, A_5, A_6, A_7, A_8\}$ where the initial average is associated with the centre of the window.

An example of moving average trendlines is shown in Figure D1 where the data (shown Table D1) is the exchange rate between Australian Dollars (AUD) and US Dollars (USD) $[USD \times Rate = AUD]$ for a 27-day period beginning on Tuesday 28th January 2020 and ending on Thursday 27th February 2020. The exchange rate for each day is shown as a small square and three moving average trendlines give a visual summary of the exchange rate fluctuations. The right moving average trendline (40) lags the data and the left moving average trendline (41) leads the data.

---

[4] A time series is often shown in graphical form as a scatter plot of x-y data where the x-values are a sequence of times (e.g. days, months, years, etc.) or some other (regular) intervals and the y-values are the variables of interest.
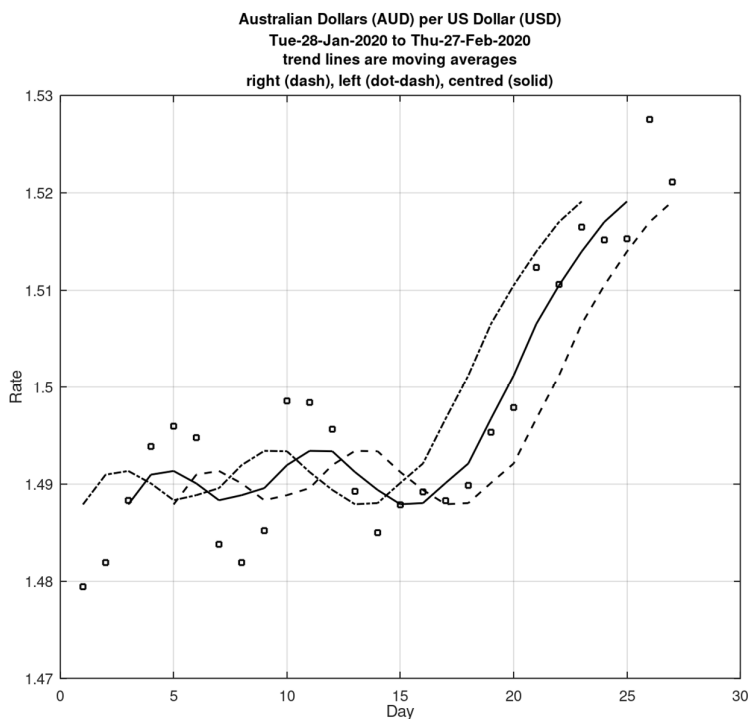
Figure D1. Trendlines showing the fluctuation in the exchange rate between Australian Dollars (AUD) and US Dollars (USD). $USD \times Rate = AUD$; day 1 = 28-Jan-2020, day 27 = 27-Feb-2020.

Day	Rate	Day	Rate	Day	Rate	Day	Rate
1	1.47943	8	1.48192	15	1.48787	22	1.51058
2	1.48192	9	1.48521	16	1.48920	23	1.51647
3	1.48832	10	1.49858	17	1.48830	24	1.51515
4	1.49388	11	1.49842	18	1.48987	25	1.51527
5	1.49597	12	1.49566	19	1.49535	26	1.52753
6	1.49479	13	1.48926	20	1.49790	27	1.52110
7	1.48380	14	1.48500	21	1.51232		

Table D1. Exchange rate between Australian Dollars (AUD) and US Dollars (USD)
$USD \times Rate = AUD$. Day 1 = 28-Jan-2020, Day 27 = 27-Feb-2020
Data source: Exchange-Rates.org (https://www.exchange-rates.org/history/AUD/USD/G/30)

Recurrence relations

The calculation of moving averages may be made more efficient by expressing them as recurrence relations. For example, consider the $n = 10$ data elements $\left\{ y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6 \quad y_7 \quad y_8 \quad y_9 \quad y_{10} \right\}$ and the calculation of the right moving average (40) for a window of period $p = 3$

$$A_3 = \tfrac{1}{3}\left(y_1 + y_2 + y_3\right)$$
$$A_4 = \tfrac{1}{3}\left(y_2 + y_3 + y_4\right) = A_3 + \tfrac{1}{3}\left(y_4 - y_1\right)$$
$$A_5 = \tfrac{1}{3}\left(y_3 + y_4 + y_5\right) = A_4 + \tfrac{1}{3}\left(y_5 - y_2\right)$$
$$\vdots$$
$$A_9 = \tfrac{1}{3}\left(y_7 + y_8 + y_9\right) = A_8 + \tfrac{1}{3}\left(y_9 - y_6\right)$$
$$A_{10} = \tfrac{1}{3}\left(y_8 + y_9 + y_{10}\right) = A_9 + \tfrac{1}{3}\left(y_{10} - y_7\right)$$

and a recurrence relation for the right moving average is

$$A_t = A_{t-1} + \frac{1}{p}\left(y_t - y_{t-p}\right) \quad \text{for} \quad t = p+1, p+2, \ldots, n \tag{43}$$

This requires only the first average A_{t-1} to be calculated.

Weighted Moving Averages

Weighted moving averages can be developed by considering the weighted average q of a set of values $y_1, y_2, y_3, \ldots y_n$ each having weight $w_1, w_2, w_3, \ldots, w_n$ where

$$q = \frac{w_1 y_1 + w_2 y_2 + \cdots + w_n y_n}{w_1 + w_2 + \cdots + w_n} = \frac{\displaystyle\sum_{k=1}^{n} w_k y_k}{\displaystyle\sum_{k=1}^{n} w_k} \tag{44}$$

and weights are numbers that reflect the degree of importance attached to the y-value, i.e., larger weights reflect values of higher importance.

Suppose the window has $p = 2k+1$ elements and p is odd, and that weights are just the location of window elements. In the right moving average, the weight of the right-most element is p and the weights of elements to the left are $p-1, p-2, \ldots, 3, 2, 1$. The **weighted right moving average** is

$$WA_t = \frac{\displaystyle\sum_{j=-2k}^{0}\left(p+j\right)y_{t+j}}{p+\left(p-1\right)+\cdots+2+1} = \frac{2}{p\left(p+1\right)}\sum_{j=-2k}^{0}\left(p+j\right)y_{t+j} \quad \text{for} \quad t = p, p+1, \ldots, n \tag{45}$$

Note here that $p+\left(p-1\right)+\left(p-2\right)+\cdots+3+2+1 = \dfrac{p\left(p+1\right)}{2}$ is a triangular number[5] and the weights $w_1 = p, w_2 = p-1, \ldots, w_p = 1$ form an arithmetic progression with a common difference of 1.

---

[5] Triangular numbers are given by the formula $T_n = n+\left(n-1\right)+\left(n-2\right)+\cdots+3+2+1 = \dfrac{n\left(n+1\right)}{2}$ which can be verified by considering an $n \times n$ square array of elements (n rows by n columns). T_n is the sum of the elements in each row beginning at the leading diagonal element of each row and will be equal to the number of elements on the leading diagonal plus one half of the remaining elements in the array, i.e., $T_n = n + \dfrac{1}{2}\left(n^2 - n\right) = \dfrac{n\left(n+1\right)}{2}$

Exponential Smoothing

We follow here the work of Hunter[6] (1986) and the NIST/SEMATECH *e-Handbook.*

Suppose we have observations y up to and including time $t-1$, that is, our observations form the set $\{y_1, y_2, y_3, \ldots, y_{t-2}, y_{t-1}\}$ and we wish to forecast or predict the next observation y_t. We denote our prediction as $\hat{y}_t$ and when the actual observation y_t becomes available the prediction error is

$$e_t = y_t - \hat{y}_t \tag{46}$$

The **method of exponential smoothing** takes the prediction for the previous period and adds to it a proportion of the prediction error at that previous time period to give the next prediction or the update as

$$\hat{y}_t = \begin{cases} \mu & \text{for } t = 1 \\ \hat{y}_{t-1} + \alpha\left(y_{t-1} - \hat{y}_{t-1}\right) & \text{for } t > 1 \end{cases} \tag{47}$$

where $0 < \alpha \le 1$ is a constant known as the weighting factor and μ is an apriori[7] value of the data. We can rearrange the recurrence relationship (47) and write the update as

$$\hat{y}_t = \begin{cases} \mu & \text{for } t = 1 \\ \alpha y_{t-1} + \left(1 - \alpha\right)\hat{y}_{t-1} & \text{for } t > 1 \end{cases} \tag{48}$$

[Note here that Hunter (1986) describes the Exponentially Weighted Moving Average (EWMA) as (in his notation)

$$\begin{align} \text{EWMA} = \hat{y}_{t+1} = \hat{y}_t + \lambda e_t \qquad & \left(3\right) \\ = \hat{y}_t + \lambda\left(y_t - \hat{y}_t\right) \qquad & \left(4\right) \end{align}$$

And then re-writes his equation (4) as

$$\hat{y}_{t+1} = \lambda y_t + \left(1 - \lambda\right)\hat{y}_t \qquad \left(5\right)$$

The only difference in his notation and ours is that λ replaces α and the subscript t is advanced one unit.]

Using (48) in a sequence gives

$$\begin{align} \hat{y}_2 &= \alpha y_1 + \left(1 - \alpha\right)\mu \\ \hat{y}_3 &= \alpha y_2 + \left(1 - \alpha\right)\hat{y}_2 \\ &= \alpha y_2 + \left(1 - \alpha\right)\left\{\alpha y_1 + \left(1 - \alpha\right)\mu\right\} \\ &= \alpha y_2 + \alpha\left(1 - \alpha\right)y_1 + \left(1 - \alpha\right)^2\mu \\ \hat{y}_4 &= \alpha y_3 + \left(1 - \alpha\right)\hat{y}_3 \\ &= \alpha y_3 + \left(1 - \alpha\right)\left\{\alpha y_2 + \alpha\left(1 - \alpha\right)y_1 + \left(1 - \alpha\right)^2\mu\right\} \\ &= \alpha y_3 + \alpha\left(1 - \alpha\right)y_2 + \alpha\left(1 - \alpha\right)^2 y_1 + \left(1 - \alpha\right)^3\mu \end{align}$$

and from this sequence we can write a general form as

$$\hat{y}_t = \left[\sum_{k=1}^{t-1}\alpha\left(1 - \alpha\right)^{k-1}y_{t-k}\right] + \left(1 - \alpha\right)^{t-1}\mu \quad \text{for } t > 1 \tag{49}$$

---

[6] J. Stuart Hunter Professor Emeritus at Princeton University.

[7] *apriori* [Latin *a priori* what is known before] often taken to mean something known beforehand, or known from prior knowledge

For $t = 5$ equation (49) gives

$$\hat{y}_5 = \left[\sum_{k=1}^{4} \alpha \left(1 - \alpha \right)^{k-1} y_{5-k} \right] + \left(1 - \alpha \right)^{5-1} \mu$$

$$= \left[\alpha \left(1 - \alpha \right)^0 y_4 + \alpha \left(1 - \alpha \right)^1 y_3 + \alpha \left(1 - \alpha \right)^2 y_2 + \alpha \left(1 - \alpha \right)^3 y_1 \right] + \left(1 - \alpha \right)^4 \mu$$

$$= \alpha y_4 + \alpha \left(1 - \alpha \right) y_3 + \alpha \left(1 - \alpha \right)^2 y_2 + \alpha \left(1 - \alpha \right)^3 y_1 + \left(1 - \alpha \right)^4 \mu$$

Letting weights $w_4 = \alpha, w_3 = \alpha \left(1 - \alpha \right), w_2 = \alpha \left(1 - \alpha \right)^2, w_1 = \alpha \left(1 - \alpha \right)^3$ we have

$$\hat{y}_5 = w_4 y_4 + w_3 y_3 + w_2 y_2 + w_1 y_1 + \left(1 - \alpha \right)^4 \mu$$

and for $\alpha = 2/3$ and $1 - \alpha = 1/3$ the weights are

$$
\begin{aligned}
w_4 &= \alpha & &= 0.666667 \\
w_3 &= \alpha \left(1 - \alpha \right) & &= 0.222222 \\
w_2 &= \alpha \left(1 - \alpha \right)^2 & &= 0.074074 \\
w_1 &= \alpha \left(1 - \alpha \right)^3 & &= 0.024691
\end{aligned}
$$

This demonstrates the exponential nature of the weights in the calculation of $\hat{y}_t$ where the weights are tending to zero and the difference between successive weights is also tending to zero. It should be noted here that the last term in the summation for $\hat{y}_5$ is $\left(1 - \alpha \right)^4 \mu$ and for $\alpha = 2/3$ then $\left(1 - \alpha \right)^4 = 0.012346$ and for any t this coefficient is $\left(1 - \alpha \right)^{t-1}$ and for the calculation of $\hat{y}_{10}$ then the coefficient $\left(1 - \alpha \right)^9 = 0.000051$.

An example of exponential smoothing is shown in Figure D2 that is a scatter plot of the exchange rate between Australian Dollars (AUD) and US Dollars (USD) where $USD \times Rate = AUD$ for a 27-day period beginning on Tuesday 28th January 2020 and ending on Thursday 27th February 2020 [see data in Table D1]. The predicted exchange rate $\hat{y}$ is given by $\hat{y}_{\text{today}} = \alpha \times Rate_{\text{yesterday}} + \left(1 - \alpha \right) \times \hat{y}_{\text{yesterday}}$ where the weighting factor $\alpha = 2/3$ and is shown as a solid line.

Australian Dollars (AUD) per US Dollar (USD)
Tue-28-Jan-2020 to Thu-27-Feb-2020
centred 5-point Moving Average dashed line
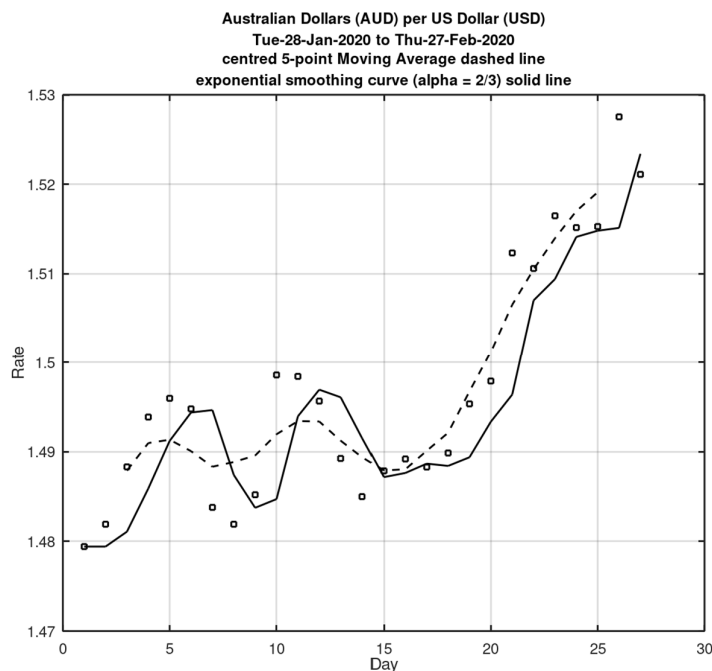**exponential smoothing curve (alpha = 2/3) solid line**

Figure D2. Trendlines showing the fluctuation in the exchange rate between Australian Dollars (AUD) and US Dollars (USD). $USD \times Rate = AUD$; day 1 = 28-Jan-2020, day 27 = 27-Feb-2020.

Performance Handicap Systems in yachting sometimes use moving averages or exponential smoothing in calculating new handicaps after each yacht race. A rule that is sometimes used is:

$$\text{Calculated Handicap} = 1\big/3 \ \text{Back Calculated Handicap} + 2\big/3 \ \text{Allocated Handicap}$$

or
$$CHC = \frac{1}{3} BCH + \frac{2}{3} AHC \tag{50}$$

If we assume that allocated handicaps are values predicted from past performances in yacht races then we can denote these as $\hat{y}_{t-1}, \hat{y}_{t-2}, \hat{y}_{t-3}, \ldots$ and at each of these prior races the measurements or observations related to performance were the back calculated handicaps denoted as $y_{t-1}, y_{t-2}, y_{t-3}, \ldots$ then we can use (48) and write

$$AHC_t = \alpha \times BCH_{t-1} + \left(1 - \alpha\right) AHC_{t-1} \quad \text{for } 0 < \alpha \leq 1 \text{ and } t > 1 \tag{51}$$

and if the weighting factor $\alpha = 1\big/3$ then we have the rule (50).

Handicaps based on this recurrence relationship could be called exponentially weighted handicaps and they are sometimes called exponentially averaged handicaps.

The prediction formula for handicaps is sometimes written as

$$AHC_t = \frac{1}{G} \times BCH_{t-1} + \left(\frac{G-1}{G}\right) AHC_{t-1} \quad \text{for } G > 1 \text{ and } t > 1 \tag{52}$$

where $G = 1\big/\alpha$ is known as the *gain*. (If $\alpha = 1\big/3$ then the gain is 3)

APPENDIX E

The Kalman Filter (KF) and Performance Indicator Filter (PIF)

The Kalman Filter

The Kalman filter is a set of equations that are applied recursively to estimate the <u>state</u> of a <u>system</u> from a sequence of <u>noisy</u> measurements at times $t_1, t_2, t_3, \ldots$ etc. The state of the system is its value or values at times $t_1, t_2, t_3, \ldots$ etc. and a system may have a single value or multiple values. Say, for instance, the system is a ship steaming on a particular heading in a shipping channel and the state of the system (the ship) is its east and north coordinates $\left(E_k, N_k \right)$ and its velocity $\left(\dot{E}_k, \dot{N}_k \right)$ We say that this system (the ship) has a <u>state</u> <u>vector</u> $\mathbf{x}_k = \left[E_k, N_k, \dot{E}_k, \dot{N}_k \right]^T$ containing four elements and the subscript k indicates a value at time t_k.

On the other hand, a system may be a process such as a Performance Indicator Filter (PIF) used in scoring a yacht race series. The state of this system is a single value $\left(z_k \right)$ that is a function of the performance indicator $\left(PI_k \right)$ determined at times $t_1, t_2, t_3, \ldots$ etc., and this system (the PIF) has a state vector $\mathbf{x}_k = \left[z_k \right]$ containing a single element and the subscript k indicates a value at time t_k.

Noisy measurements are measurements that contain small <u>random errors</u> assumed to be <u>normally distributed</u>, i.e., the aggregation of errors in size groupings would form the familiar symmetric bell-shaped histogram with positive and negative errors equally likely and small errors more frequent than large errors. It is often convenient to talk of residuals (or corrections) rather than errors, where a residual is the same magnitude as an error but of opposite sign.

A Kalman filter gives the best estimates of the state of a <u>dynamic system</u> at a particular instant of time. And a dynamic system can be one whose values are changing with time, due to the motion of the system and measurement errors, or one whose values are measured at various instants of time and appear to change due to measurement errors. Dynamic systems do not have a single state (consisting of one or many values) that can be determined from a finite set of measurements but instead have a continuously changing state that has values sampled at different instants of time.

The Kalman filter equations were published in 1960 by Dr. R.E. Kalman in his famous paper describing a new approach to the solution of linear filtering and prediction (Kalman 1960). Since that time, papers on the application of the technique have been filling numerous scientific journals and it is regarded as one of the most important algorithmic techniques ever devised. It has been used in applications ranging from navigating the Ranger and Apollo spacecraft in their lunar missions to predicting short-term fluctuations in the stock market. Sorenson (1985) shows Kalman's technique to be an extension of C.F. Gauss' original method of least squares developed in 1795 and provides an historical commentary on its practical solution of linear filtering problems studied by 20th century mathematicians.

The derivation of the Kalman filter equations can be found in many texts related to signal processing that is the usual domain of Electrical Engineers, e.g., Brown and Hwang (1992). These notes use notation and terminology that would be familiar to surveyors (one of the authors has a surveying background) and the derivation of the equations of the Kalman Filter used here follow two authors, Krakiwsky (1975) and Cross (1992) both with surveying backgrounds. The Kalman filter equations and the associated measurement and dynamic models are given below with a brief explanation of the terms.

In the explanation that follows the 'hat' symbol (^) above a vector $\mathbf{x}$ indicates that it is an estimate of the true (but unknown) state of the system derived from the Kalman Filter (a least squares process). This is also known as the filtered state. The 'prime' symbol (′) indicates a predicted quantity.

Primary and Secondary (or Dynamic) Measurement Models

Suppose that $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{k-1}, \mathbf{x}_k$ are vectors of parameters or <u>state vectors</u> of a system at times $t_1, t_2, t_3, \ldots, t_{k-1}, t_k$ and that $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3, \ldots, \mathbf{l}_{k-1}, \mathbf{l}_k$ are the corresponding vectors of measurements associated with the parameters. We may write three equations as follows:

$$
\begin{aligned}
\mathbf{v}_{k-1} + \mathbf{B}_{k-1}\mathbf{x}_{k-1} &= \mathbf{f}_{k-1} && \text{primary } t_{k-1} \\
\mathbf{v}_k + \mathbf{B}_k\mathbf{x}_k &= \mathbf{f}_k && \text{primary } t_k \\
\mathbf{x}_k &= \mathbf{T}\mathbf{x}_{k-1} + \mathbf{v}_m && \text{secondary or dynamic}
\end{aligned}
\tag{53}
$$

where

  $\mathbf{x}$ is the state vector containing the parameters of the system
  $\mathbf{v}$ is the vector of residuals associated with the measurements $\mathbf{l}$ where $\hat{\mathbf{l}} = \mathbf{l} + \mathbf{v}$
  $\mathbf{B}$ is a coefficient matrix
  $\mathbf{f}$ is a vector of numeric terms derived from the measurements $\mathbf{l}$
  $\mathbf{T}$ is the <u>transition matrix</u>
  $\mathbf{v}_m$ is a vector of residuals associated with the <u>dynamic</u> model

The <u>primary measurement models</u> in (53) link measurements $\mathbf{l}$ (contained in the vector of numeric terms $\mathbf{f}$) with parameters in the state vector $\mathbf{x}$ at times t_{k-1} and t_k.

The <u>secondary or dynamic model</u> in (53) links the state vectors $\mathbf{x}$ at times t_{k-1} and t_k. The transition matrix $\mathbf{T}$ is an attempt to model temporal changes between the state vectors (the dynamics of the system) and $\mathbf{v}_m$ is a vector of corrections reflecting the fact that the transition matrix $\mathbf{T}$ is an approximation of the true dynamics. The elements of $\mathbf{v}_m$ are assumed to be small, random and normally distributed with a mean of zero.

The measurements $\mathbf{l}_{k-1}$ and $\mathbf{l}_k$ and the model corrections $\mathbf{v}_m$ have associated weight matrices $\mathbf{W}_{k-1}, \mathbf{W}_k$ and $\mathbf{W}_m$ and cofactor matrices $\mathbf{Q}_{k-1}, \mathbf{Q}_k$ and $\mathbf{Q}_m$ where in general $\mathbf{Q} = \mathbf{W}^{-1}$.

System Driving Noise of the Secondary Model

For the solution of many practical problems, it is useful to assume the vector $\mathbf{v}_m$ as being the product of two matrices

$$
\mathbf{v}_m = \mathbf{H}\mathbf{w}
\tag{54}
$$

where $\mathbf{w}$ is a vector of quantities known as the <u>system driving noise</u> which cause the secondary model to be incorrect and $\mathbf{H}$ is a coefficient matrix chosen so that the product $\mathbf{H}\mathbf{w}$ represents the effect of these quantities on the parameters. Note that in general $\mathbf{H}$ will not be a square matrix as the number of error sources causing the system noise in the secondary model is not necessarily equal to the number of parameters in $\mathbf{x}$.

The Cofactor Matrix of the Secondary Model $\mathbf{Q}_m$

The system driving noise $\mathbf{w}$ in (54) is assumed to be a vector of random quantities with zero mean and variance matrix estimated by the cofactor matrix $\mathbf{Q}_w$. The cofactor matrix $\mathbf{Q}_m$ can be obtained using *Propagation of Variances* (or propagation of cofactors) that can be summarized as follows:

If linear (or linearized) equations can be expressed in a matrix form

$$
\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}
$$

where $\mathbf{y}$ is a vector of variables, $\mathbf{A}$ is a coefficient matrix, $\mathbf{x}$ is a vector of variables having an associated cofactor matrix $\mathbf{Q}_x$ and $\mathbf{b}$ is a vector of constants then the cofactor matrix of the variables $\mathbf{y}$ is given by

$$\mathbf{Q}_y = \mathbf{A}\mathbf{Q}_x\mathbf{A}^T$$

Cofactor propagation using (54) gives

$$\mathbf{Q}_m = \mathbf{H}\mathbf{Q}_w\mathbf{H}^T \tag{55}$$

and the weight matrix of the secondary model is given by

$$\mathbf{W}_m = \left(\mathbf{Q}_m\right)^{-1} = \left(\mathbf{H}\mathbf{Q}_w\mathbf{H}^T\right)^{-1} \tag{56}$$

The Kalman Filter Equations

With initial estimates of the state vector $\hat{\mathbf{x}}_{k-1}$ and the state cofactor matrix $\mathbf{Q}_{\hat{x}_{k-1}}$; and with the cofactor matrix of the dynamic model $\mathbf{Q}_m$ a Kalman Filter has the following five general steps

(1) Compute the predicted state vector at t_k

$$\mathbf{x}'_k = \mathbf{T}\hat{\mathbf{x}}_{k-1}$$

(2) Compute the predicted state cofactor matrix at t_k

$$\mathbf{Q}_{x'_k} = \mathbf{T}\mathbf{Q}_{\hat{x}_{k-1}}\mathbf{T} + \mathbf{Q}_m$$

(3) Compute the Kalman *Gain* matrix

$$\mathbf{K} = \mathbf{Q}_{x'_k}\mathbf{B}_k^T\left(\mathbf{Q}_k + \mathbf{B}_k\mathbf{Q}_{x'_k}\mathbf{B}_k^T\right)^{-1}$$

(4) Compute the filtered state vector by updating the predicted state with the measurements at t_k

$$\mathbf{x}_k = \mathbf{x}'_k + \mathbf{K}\left(\mathbf{f}_k - \mathbf{B}_k\mathbf{x}'_k\right)$$

(5) Compute the filtered state cofactor matrix

$$\mathbf{Q}_{\hat{x}_k} = \left(\mathbf{I} - \mathbf{K}\mathbf{B}_k\right)\mathbf{Q}_{x'_k}$$

Go to step (1) and repeat the process for the next measurement epoch t_{k+1}

The Kalman filter equations are relatively easy to implement on modern computers (a reason for its popularity) and the example studied below will be supplemented by Octave computer code.

A Kalman Filter in Calculated Handicaps for Yachting.

Kalman Filter's are widely used in navigation and systems to assess the performance of yachts. For example, Douguet et al. (2013) use Kalman Filters to model the leeway of a sailboat as part of a method for evaluating sailboat performance and there are numerous other published applications. Here we wish to use a Kalman Filter to model the performance of the yacht *Dream* in the Race Series shown in Appendix B and scored using *TopYacht* scoring with 45% mark boat [see equation (4)] and calculated handicaps

$CHC = \frac{1}{3}BCH + \frac{2}{3}AHC = AHC + \frac{1}{3}PI$ where the performance indicator $PI = BCH - AHC$ and AHC and BCH denote allocated and back calculated handicaps respectively.

In particular, *Dream's* performance indicators for the seven races are shown in Table B10 and repeated here as

$$PI = \begin{bmatrix} -0.026 & 0.057 & 0.076 & 0.013 & 0.022 & 0.003 & -0.030 \end{bmatrix} \tag{57}$$

A plot of *Dream's* performance indicators is shown in Figure E1

Performance Indicator (PI) for DREAM
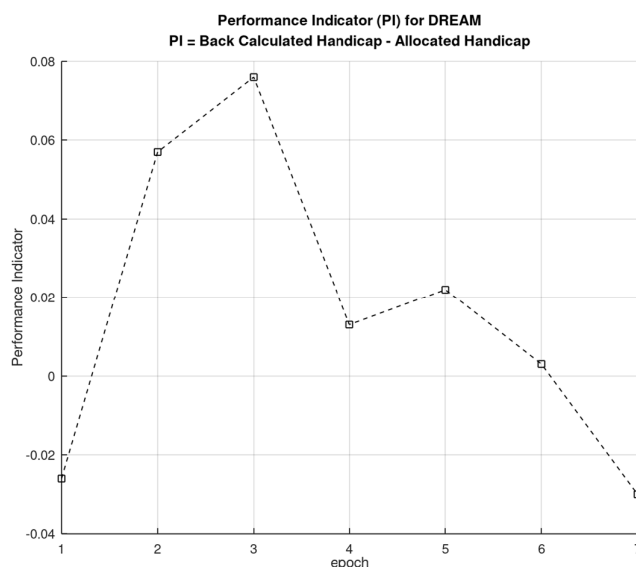**PI = Back Calculated Handicap - Allocated Handicap**

Figure E1. *Dream's* performance indicators for the seven-race series

The general form of the calculated handicap is $CHC_k = AHC_k + z_k$ [see equation (12)] where z is a numeric value between zero and $\pm PI_k$ that is a function of the yacht's AHC prior to the race and BCH after the race. And, since the CHC for one race is the AHC for the next race we may write

$$AHC_{k+1} = AHC_k + z_k \tag{58}$$

We now wish to use a Kalman Filter to estimate z_k given measurements PI_k

The measurements PI_k (affected by random errors), and residuals can be expressed as

$$\mathbf{l}_k + \mathbf{v}_k = \hat{\mathbf{l}}_k \tag{59}$$

where $\mathbf{l}_k$ is the *n,1* vector of measurements, $\mathbf{v}_k$ is the *n,1* vector of residuals (small unknown corrections to the measurements) and $\hat{\mathbf{l}}_k$ are estimates of the true (but unknown) value of the measurements. n is the number of measurements at each epoch, that in this case is one. The primary measurement model can be expressed in terms of the filtered state vector $\hat{\mathbf{x}}_k$ at time t_k as

$$\mathbf{v}_k + \mathbf{B}_k\hat{\mathbf{x}}_k = \mathbf{f}_k \tag{60}$$

40

In this case $\hat{\mathbf{x}}_k$ contains the elements of $\hat{\mathbf{l}}_k$, both vectors containing only single quantities. And $\mathbf{f}_k = -\mathbf{l}_k$ also both containing single quantities (the *PI* at t_k). The matrix $\mathbf{B}$ will contain a single quantity, $\mathbf{B} = \begin{bmatrix} -1 \end{bmatrix}$.

The dynamic model linking the elements of the state vector at times t_{k-1} and t_k is

$$\mathbf{x}_k = \mathbf{T}\mathbf{x}_{k-1} + \mathbf{v}_m \tag{61}$$

The state vector contains a single element that may change between t_{k-1} and t_k but the dynamics are unknown and changes will be due to measurement error and model error. The transition matrix $\mathbf{T}$ will contain a single element $\mathbf{T} = \begin{bmatrix} 1 \end{bmatrix}$ and the model correction $\mathbf{v}_m = \mathbf{H}\mathbf{w}$ is a single element vector containing a random quantity. By setting $\mathbf{H} = \begin{bmatrix} 1 \end{bmatrix}$ the model correction is equal to the system driving noise, i.e., $\mathbf{v}_m = \mathbf{w}$. We can "assign" an appropriate estimate of the standard deviation s_w and the cofactor matrix $\mathbf{Q}_w = \begin{bmatrix} s_w^2 \end{bmatrix}$ is now defined, and by cofactor propagation [see (55)], the cofactor matrix of the dynamic model $\mathbf{Q}_m = \mathbf{Q}_w$. We choose $s_w = 0.033$ which gives $\mathbf{Q}_w = \begin{bmatrix} s_w^2 \end{bmatrix} = \begin{bmatrix} (0.033)^2 \end{bmatrix}$. The value $s_w = 0.033$ chosen here is representative of the values of $\hat{\sigma} \approx 1.4826 \left(\mathrm{MAD} \right)$ shown in Table B10 in Appendix B.

Lastly, an estimate of the cofactor matrix of the measurements $\mathbf{Q}$ and the filtered state cofactor matrix $\mathbf{Q}_{x_k}$ must be made. Let us assume (guess) that the measurements have a standard deviation of $s_l = 0.075$ and $\mathbf{Q} = \begin{bmatrix} s_l^2 \end{bmatrix} = \begin{bmatrix} (0.075)^2 \end{bmatrix}$. Also, since our primary measurement model has a state vector containing a single value $\left(z_k \right)$, then $\mathbf{Q}_x$ will only contain a single value, and we have as a starting estimate $\mathbf{Q}_{x_1} = \begin{bmatrix} (0.075)^2 \end{bmatrix}$, the same as $\mathbf{Q}$. Note here that the value $\mathbf{Q} = \begin{bmatrix} s_l^2 \end{bmatrix} = \begin{bmatrix} (0.075)^2 \end{bmatrix} = \begin{bmatrix} 0.005625 \end{bmatrix}$ chosen as a 'guess' is not too far removed from the final filtered estimate $\mathbf{Q}_{x_7} = \begin{bmatrix} 0.002001 \end{bmatrix} \approx \begin{bmatrix} (0.045)^2 \end{bmatrix}$ shown below in the output of Octave function *yacht1.m*

Now we can now start the Kalman filter at epoch t_2 using the value at t_1 as filtered estimates noting here that we have decided that the filtered state $\hat{\mathbf{x}}_1 = z_1 = \frac{2}{5}l_1 = \frac{2}{5}PI_1$.

---

(1) Compute the predicted state vector at epoch t_2 using $\frac{2}{5}$ of the measurement -0.026 at t_1 as the filtered state $\hat{\mathbf{x}}_1$

$$\begin{aligned} \mathbf{x}_2' &= \mathbf{T}\hat{\mathbf{x}}_1 \\ &= \begin{bmatrix} 1 \end{bmatrix}\begin{bmatrix} -0.0104 \end{bmatrix} \\ &= -0.0104 \end{aligned}$$

(2) Compute the predicted state cofactor matrix at t_2 using $\mathbf{Q}_{x_1} = \begin{bmatrix} (0.075)^2 \end{bmatrix}$ as the filtered estimate

$$\begin{aligned} \mathbf{Q}_{x_2'} &= \mathbf{T}\mathbf{Q}_{x_1}\mathbf{T}^T + \mathbf{Q}_m \\ &= \begin{bmatrix} 1 \end{bmatrix}\begin{bmatrix} (0.075)^2 \end{bmatrix}\begin{bmatrix} 1 \end{bmatrix} + \begin{bmatrix} (0.033)^2 \end{bmatrix} \\ &= \begin{bmatrix} 0.006714 \end{bmatrix} \end{aligned}$$

(3) Compute the Kalman *Gain matrix* noting that $\mathbf{Q} = \begin{bmatrix} (0.075)^2 \end{bmatrix}$

$$\begin{aligned} \mathbf{K} &= \mathbf{Q}_{x_2'}\mathbf{B}_2^T \left(\mathbf{Q} + \mathbf{B}_2\mathbf{Q}_{x_2'}\mathbf{B}_2^T \right)^{-1} \\ &= \begin{bmatrix} 0.006714 \end{bmatrix}\begin{bmatrix} -1 \end{bmatrix} \\ &\quad \times \left(\begin{bmatrix} (0.075)^2 \end{bmatrix} + \begin{bmatrix} -1 \end{bmatrix}\begin{bmatrix} 0.006714 \end{bmatrix}\begin{bmatrix} -1 \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} -0.006714 \end{bmatrix}\begin{bmatrix} 0.012339 \end{bmatrix}^{-1} \\ &= -0.544128 \end{aligned}$$

(4) Compute the filtered state vector $\hat{\mathbf{x}}_2$ by updating the predicted state with the measurements at t_2

$$\begin{aligned}
\hat{\mathbf{x}}_2 &= \mathbf{x}_2' + \mathbf{K}\left(\mathbf{f}_2 - \mathbf{B}_2\mathbf{x}_2'\right) \\
&= [-0.0104] + \left\{[-0.544128] \right. \\
&\quad \left. \times\left([-0.057] - [-1][-0.0104]\right)\right\} \\
&= 0.026274
\end{aligned}$$

(5) Compute the filtered state cofactor matrix at t_2

$$\begin{aligned}
\mathbf{Q}_{x_2} &= \left(\mathbf{I} - \mathbf{K}_2\mathbf{B}_2\right)\mathbf{Q}_{x_2'} \\
&= \left([1] - [-0.544128][-1]\right) \\
&\quad \times[0.006714] \\
&= 0.003061
\end{aligned}$$

Go to step **(1)** and repeat the process for the next measurement epoch t_3.

The values from the Kalman Filter for epochs t_3, t_4 and t_5 are

epoch t_3	epoch t_4	epoch t_5
$\mathbf{x}_3' = 0.026274$	$\mathbf{x}_4' = 0.047385$	$\mathbf{x}_5' = 0.034250$
$\mathbf{Q}_{x_3'} = 0.004150$	$\mathbf{Q}_{x_4'} = 0.003477$	$\mathbf{Q}_{x_5'} = 0.003238$
$\mathbf{K}_3 = -0.424536$	$\mathbf{K}_4 = -0.382005$	$\mathbf{K}_5 = -0.365323$
$\hat{\mathbf{x}}_3 = 0.047385$	$\hat{\mathbf{x}}_4 = 0.034250$	$\hat{\mathbf{x}}_5 = 0.029774$
$\mathbf{Q}_{\hat{x}_3} = 0.002388$	$\mathbf{Q}_{\hat{x}_4} = 0.002149$	$\mathbf{Q}_{\hat{x}_5} = 0.002055$

An Octave function *yacht1.m* using the performance indicator data given above has the following output

```
>> yacht1

Epoch =    2, measurement PI =   0.057
Gain K =   -0.544128 1/K = -1.8378
Filtered State  Corrn        Filtered State cofactor matrix Qxx
 0.026274        0.036674     0.003061

Epoch =    3, measurement PI =   0.076
Gain K =   -0.424536 1/K = -2.35551
Filtered State  Corrn        Filtered State cofactor matrix Qxx
 0.047385        0.021110     0.002388

Epoch =    4, measurement PI =   0.013
Gain K =   -0.382005 1/K = -2.61777
Filtered State  Corrn        Filtered State cofactor matrix Qxx
 0.034250        -0.013135    0.002149
:
:
:
Epoch =    7, measurement PI = -0.030
Gain K =   -0.355725 1/K = -2.81116
Filtered State  Corrn        Filtered State cofactor matrix Qxx
 0.002327        -0.017848    0.002001

xhat = -0.0104000    0.0262743    0.0473846    0.0342495    0.0297745
        0.0201750    0.0023265

>>
```

The variable 'xhat' above is the filtered state vector $\hat{\mathbf{x}}$ and it contains the z-values noting that $z_1 = \frac{2}{5}PI_1$

$$z = \begin{bmatrix} -0.010 & 0.026 & 0.047 & 0.034 & 0.030 & 0.020 & 0.002 \end{bmatrix} \tag{62}$$
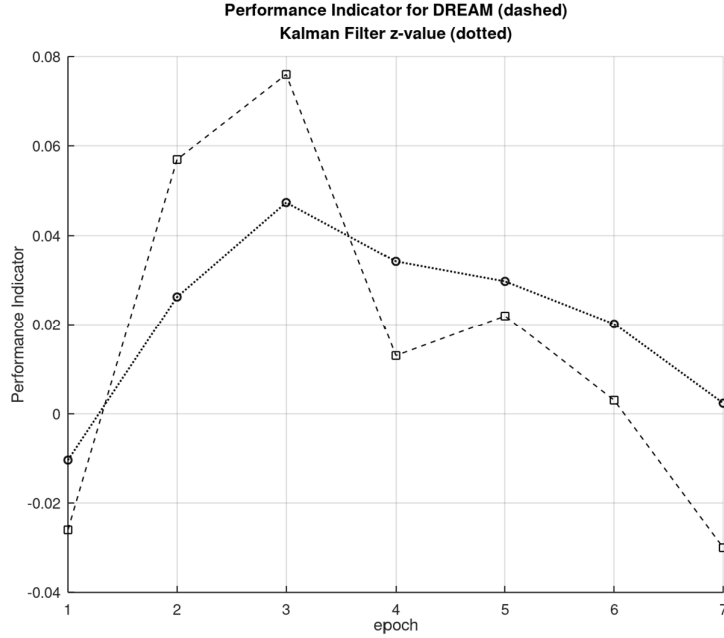

Figure E2. *Dream's* filtered z-values (circles and dotted line) derived from
PI values (squares and dashed line). Note that $z_1 = \frac{2}{5}PI_1$

A Performance Indicator Filter in Yachting.

By inspection of the process outlined above we write a simplified recurrence formula for computing the values z_k and call this a Performance Indicator Filter (PIF)

$$z_k = z_{k-1} + K\left(PI_k - z_{k-1}\right) \text{ for } k > 0 \text{ and } z_0 = 0 \tag{63}$$

Using (63) with $K = 2/5$ and $PI = \begin{bmatrix} -0.026 & 0.057 & 0.076 & 0.013 & 0.022 & 0.003 & -0.030 \end{bmatrix}$ gives the following tabulated results

k	PI	z
1	-0.026	-0.010
2	0.057	0.017
3	0.076	0.041
4	0.013	0.030
5	0.022	0.027
6	0.003	0.017
7	-0.030	-0.002

Table E1. *PI* values and z-values computed from a Performance Indicator
Filter for *Dream* for the seven-race series

These z-values are a reasonable approximation of those from the Kalman Filter (62) and Figure E3 shows the relationships between the Performance Indicators, the z-values from the Kalman Filter and the z-values from the Performance Indicator Filter (63)
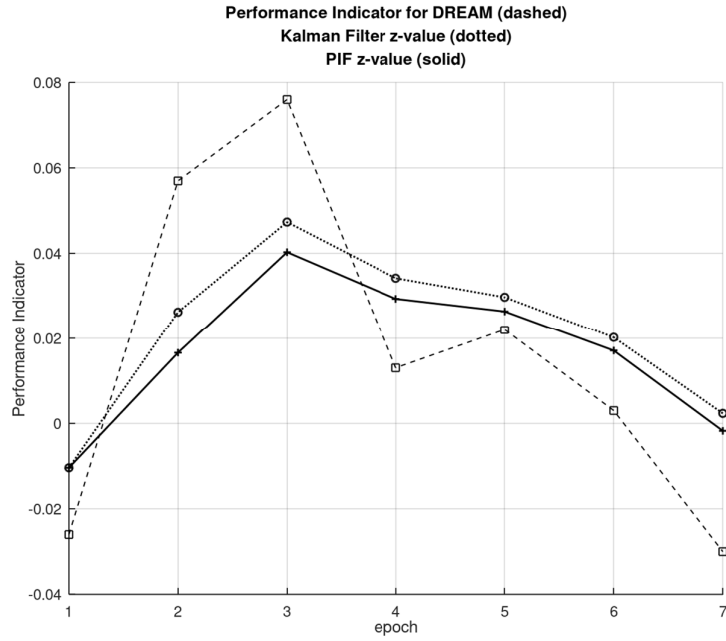

Performance Indicator for DREAM (dashed)
Kalman Filter z-value (dotted)
PIF z-value (solid)

Figure E3. *Dream's PI* values (squares and dashed line), Kalman Filter z-values (circles and dotted line) and PIF z-values (crosses and solid line).

Note that $z_1 = \frac{2}{5} PI_1$

Appendix F

Mean, Variance, Standard Deviation and Median Absolute Deviation (MAD)

Sample Mean, Standard Deviation and Variance

In the estimation of the precision of data $x_1, x_2, x_3, \ldots, x_n$ we often use the *sample variance* s_x^2, and its positive square-root the *sample standard deviation* s_x as measures of the variability of the data computed from the data itself. And we also assume that the data is a finite sample of size n drawn from an infinite population N having a mean μ_x, variance σ_x^2 and standard deviation σ_x being the positive square-root of the variance. The population values μ_x, σ_x^2 are often unknown and estimated by the sample values $\overline{x}, s_x^2$ where $\overline{x} = \dfrac{1}{n} \sum_{k=1}^{n} x_k$ is the sample mean. The sample standard deviation and its squared value the sample variance, are also measures of the scale of the data and the mean of a set of data is also a measure of the location of the data sample.

For the data $x_1, x_2, x_3, \ldots, x_n$ that are a sample of size n, the sample standard deviation s_x computed from the sample variance $s_x^2 = \dfrac{1}{n-1} \sum_{k=1}^{n} \left(x_k - \overline{x} \right)^2$ where s_x is the positive square root of s_x^2

But the sample mean and variance (and hence the sample standard deviation) are known to suffer from the effects of outliers, since large values affect the mean $\overline{x}$ and also the squared differences $\left(x_k - \overline{x} \right)^2$ in the calculation of the variance.

Median Absolute Deviation (MAD)

A more robust measure of the location of a sample is the *median M* and a more robust measure of the scale is the *Median Absolute Deviation* (MAD) that is defined as the median of the absolute deviations from the sample's median M, i.e.,

$$\text{MAD} = \text{median}\left(\left| x_k - M \right| \right) \quad \text{where } M = \text{median}\left(x_k \right) \tag{64}$$

The median M of a sample of n values is obtained by first ordering the values from smallest to largest and then choosing the middle value if n is an odd number, or the average of the two middle values if n is even. In either case there will be the same number values that are larger than or equal to the median, and smaller than or equal to the median M,

$$M = \begin{cases} x_{j+1}^{\text{ord}} & n = 2j+1 \\ \frac{1}{2}\left(x_j^{\text{ord}} + x_{j+1}^{\text{ord}} \right) & n = 2j \end{cases} \tag{65}$$

For example, suppose x_k is a set of $k = 1, 2, \ldots, n$ values and for $n = 7$, $x_k = \left\{ -2 \ \ 7 \ \ 4 \ \ 16 \ \ 1 \ \ 0 \ \ 8 \right\}$. The set is ordered from smallest to largest as $x_k^{\text{ord}} = \left\{ -2 \ \ 0 \ \ 1 \ \ \underset{\uparrow}{4} \ \ 7 \ \ 8 \ \ 16 \right\}$ and since n is odd, the median M is the middle value indicated with $\uparrow$ and $j = \left(n - 1 \right)/2 = 3$ and using (65)

$M = v_{j+1}^{\text{ord}} = v_4^{\text{ord}} = 4$. There are 3 values less than M (the values to the left of the 4th value) and 3 values greater than M (the values to the right of the 4th value).

Now suppose the set $x_k = \left\{ -2 \quad 7 \quad 2 \quad 16 \quad 1 \quad 0 \quad 8 \quad 4 \quad 4 \quad -5 \right\}$ has $n = 10$ values that is ordered from smallest to largest as $x_k^{\mathrm{ord}} = \left\{ -5 \quad -2 \quad 0 \quad 1 \quad \underset{\uparrow}{2} \quad \underset{\uparrow}{4} \quad 4 \quad 7 \quad 8 \quad 16 \right\}$, and since n is even, the median M is the average of the two middle values and $j = n/2 = 5$ and using (65) $M = \frac{1}{2}\left(x_j^{\mathrm{ord}} + x_{j+1}^{\mathrm{ord}} \right) = \frac{1}{2}\left(2 + 4 \right) = 3$. There are 5 values less than the median (the first 5 values) and 5 values greater than the median (the last 5 values).

It should be noted here that if X is a random variable that can take values n values $x_1, x_2, ..., x_n$ having a median M then the probability that any X is less than or equal to the median is exactly $\frac{1}{2}$ or $\Pr\left(X \leq M \right) = \frac{1}{2}$ and if X is a continuous random variable with probability density function $f_X\left(x \right)$ and cumulative distribution function $F_X\left(x \right)$, so that $F_X\left(x \right) = \int\limits_{-\infty}^{x} f_X\left(y \right) dy$, or $\dfrac{d}{dx} F_X\left(x \right) = f_X\left(x \right)$ then the median M is defined by the solution of the integral equation $\Pr\left(X \leq M \right) = F_X\left(M \right) = \int\limits_{-\infty}^{M} f_X\left(x \right) dx = \dfrac{1}{2}$.

Appendix G shows how this result can be used to determine the value of a scale factor b that enables the MAD to be used as a consistent estimator of the standard deviation σ of normally distributed data where

$$\hat{\sigma} = b \times \mathrm{MAD} \approx 1.4826\left(\mathrm{MAD} \right) \tag{66}$$

Appendix G

Population Median and Median Absolute Deviation (MAD)

The derivation of the probability statements $\Pr\left(\left|X - \mu\right| \leq \mathrm{MAD}\right) = \dfrac{1}{2}$ and $\Pr\left(Z \leq \dfrac{\mathrm{MAD}}{\sigma}\right) = \dfrac{3}{4}$ [see equations (70) and (71) below] are the work of Dr Max Hunter, who turned his keen eye and talent for rigour to a topic not often treated in the statistical literature. It's a joy. [Max Hunter is a mathematician, friend and former colleague of one of the authors].

Let X be a random variable with density function $f_X\left(x\right)$ and distribution function $F_X\left(x\right)$, so that

$$F_X\left(x\right) = \int_{-\infty}^{x} f_X\left(y\right)dy \,, \text{ or } \frac{d}{dx}F_X\left(x\right) = f_X\left(x\right).$$

The population median m is defined by the solution of the integral equation

$$\Pr\left(X \leq m\right) = F_X\left(m\right) = \int_{-\infty}^{m} f_X\left(x\right)dx = \frac{1}{2} \tag{67}$$

The alternative equation

$$\int_{m}^{\infty} f_X\left(x\right)dx = \frac{1}{2} \tag{68}$$

Can also be used to define m.

Let the random variable Y be defined by

$$0 \leq Y = \left|X - m\right| = \begin{cases} X - m, \text{ if } X \geq m \\ m - X, \text{ if } X < m \end{cases}$$

And suppose its density function is $g_Y\left(y\right)$ with distribution function $G_Y\left(y\right)$. Then for $y \geq 0$,

$$\begin{aligned} G_Y\left(y\right) &= \Pr\left(Y \leq y\right) \\ &= \Pr\left(\left|X - m\right| \leq y\right) \\ &= \Pr\left(-y \leq X - m \leq y\right) \\ &= \Pr\left(m - y \leq X \leq m + y\right) \\ &= \Pr\left(X \leq m + y\right) - \Pr\left(X \leq m - y\right) \\ &= F_X\left(m + y\right) - F_X\left(m - y\right) \end{aligned}$$

And for $y < 0$, $G_Y\left(y\right) = 0$.

Hence for $y \geq 0$

$$g_Y\left(y\right) = \frac{d}{dy}G_Y\left(y\right) = \frac{d}{dy}\left\{F_X\left(m + y\right) - F_X\left(m - y\right)\right\} = f_X\left(m + y\right) - f_X\left(m - y\right)$$

and $g_Y\left(y\right) = 0$ for $y < 0$.

The population median M of the random variable Y satisfies the equation

$$\int_{-\infty}^{M} g_Y\left(y\right)dy = \frac{1}{2}$$

And

47

$$\int\limits_{-\infty}^{M} g_Y(y)\,dy = \int\limits_{-\infty}^{0} g_Y(y)\,dy + \int\limits_{0}^{M} g_Y(y)\,dy = \int\limits_{0}^{M} g_Y(y)\,dy$$

$$= \int\limits_{0}^{M} \left\{ f_X(m+y) + f_X(m-y) \right\} dy$$

$$= \int\limits_{m}^{m+M} f_X(s)\,ds - \int\limits_{m}^{m-M} f_X(t)\,dt \quad \text{(with substitutions } s = m + y, \ t = m - y)$$

$$= \int\limits_{m-M}^{m+M} f_X(s)\,ds$$

and therefore

$$\int\limits_{m-M}^{m+M} f_X(s)\,ds = \frac{1}{2} \tag{69}$$

Suppose now that $f_X(x)$ is symmetric about the origin then $m = 0$ from (67). So by (69)

$$\frac{1}{2} = \int\limits_{-M}^{M} f_X(x)\,dx = 2\int\limits_{0}^{M} f_X(x)\,dx$$

and therefore

$$\int\limits_{0}^{M} f_X(x)\,dx = \frac{1}{4}$$

Thus the interval $\left[-M, M\right]$ encloses an area of 0.5 under the density function for X, or since

$$\int\limits_{-\infty}^{M} f_X(x)\,dx = \frac{3}{4},$$

M is the 75 percentile of X.

But M is just the definition of MAD, so for any random variable X with population mean $E\{X\} = \mu$ and a symmetric density function about $E\{X\} = \mu$

$$\Pr\left(\left|X - \mu\right| \le \text{MAD}\right) = \frac{1}{2} \tag{70}$$

Now

$$\Pr\left(\left|X - \mu\right| \le \text{MAD}\right) = \Pr\left(\frac{\left|X - \mu\right|}{\sigma} \le \frac{\text{MAD}}{\sigma}\right)$$

$$= \Pr\left(\left|Z\right| \le \frac{\text{MAD}}{\sigma}\right) \qquad Z \text{ is a standard normal random variable}$$

$$= \Pr\left(-\frac{\text{MAD}}{\sigma} \le Z \le \frac{\text{MAD}}{\sigma}\right) \qquad \text{by definition of modulus}$$

$$= \Pr\left(Z \le \frac{\text{MAD}}{\sigma}\right) - \Pr\left(Z \le -\frac{\text{MAD}}{\sigma}\right) \qquad \text{by symmetry}$$

$$= \Pr\left(Z \le \frac{\text{MAD}}{\sigma}\right) - \left[1 - \Pr\left(Z \le \frac{\text{MAD}}{\sigma}\right)\right] \qquad \text{by definition}$$

$$= 2\Pr\left(Z \le \frac{\text{MAD}}{\sigma}\right) - 1$$

and, using (70)

$$\Pr\left(Z \le \frac{\text{MAD}}{\sigma}\right) = \frac{3}{4} \tag{71}$$

If $f_Z(z)$ is the density function of the standard normal distribution and $F_Z(z)$ is the distribution function then

$$\frac{\text{MAD}}{\sigma} = F_Z^{-1}\left(\frac{3}{4}\right) \tag{72}$$

Where F_Z^{-1} denotes the standard normal inverse cumulative distribution function. Most mathematical software packages (Maple, Mathematica, Matlab, R, etc.) have functions to compute inverse cumulative distribution functions and for the standard normal distribution GNU Octave has a function `norminv()` that computes the value of $F_Z^{-1}(x)$ and for $\text{MAD}/\sigma = F_Z^{-1}(3/4)$ can be computed from the following instructions in the Octave Command Window.

```
>> format long g
>> MAD_on_sigma = norminv(3/4)
MAD_on_sigma = 0.6744897501960818
>>
```

And

$$\frac{\text{MAD}}{\sigma} = F_Z^{-1}\left(\frac{3}{4}\right) \approx 0.6744897501960818 \tag{73}$$

Inspection of (66) leads to

$$b = \frac{\sigma}{\text{MAD}} \approx 1.482602218505602 \tag{74}$$

We can use this relationship to estimate the standard deviation from

$$\hat{\sigma} = b \times \text{MAD} \approx 1.4826\left(\text{MAD}\right) \tag{75}$$