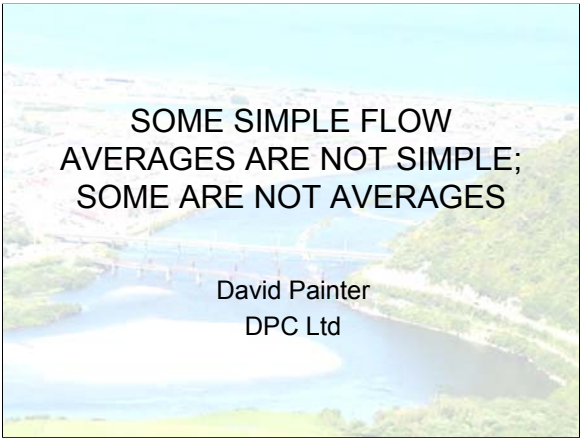
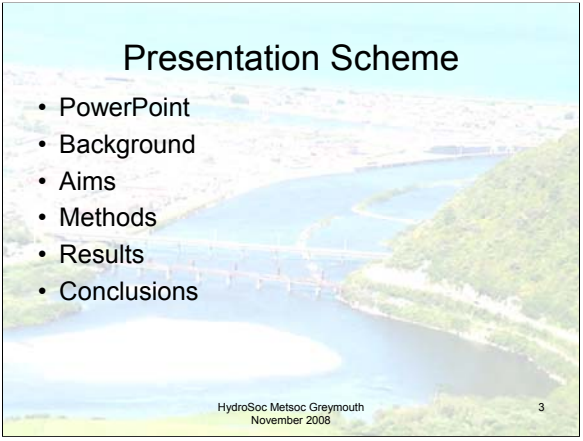
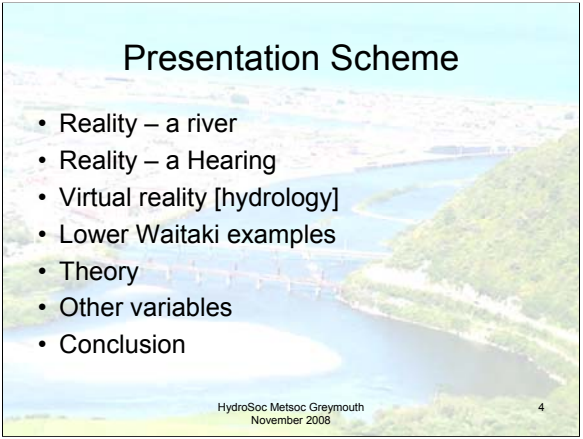


This material in this presentation started life as written evidence to an RMA Hearing [Lower Waitaki] in August 2008. I made a minor error in my evidence related to terminology about averages. Correcting it, I noticed that other submitters and expert witnesses besides myself had used ambiguous terminology related to average values of river flow data. That led to an addendum to my evidence and a request from the Hearing Commissioners for further information. It became a presentation for the NZ Hydrological and Meteorological Societies Conference at Shanty Town [Greymouth] in November 2008, then in slightly modified form as

I put this slide [Greymouth from chopper] in as a background so that people would keep a ‘real river’ in mind while I talked about arithmetic!



This is what the conference organisers specified.



This is what I did instead.

Real river, to level logger, to wireless network, to database manager office, via rating [level to discharge], to user.

Reality: RMA Hearing

Acknowledgement: ME. From an acted Hearing for a training course

My interest in this aspect of averages actually arose from involvement in the Lower Waitaki Hearing. A Hearing is also ‘reality’.

Project Aqua hydro 03, North Bank hydro 06, Hunter Downs Irrigation Scheme, existing irrigators, new applicants – all relevant background to the Lower Waitaki River RMA Hearing of resource consent applications.

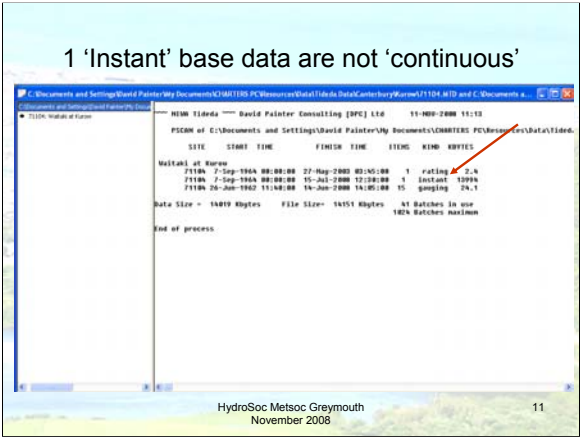
It is little use trying to measure flow in the braided lower river by conventional means. The recorders are back up above Kurow in a single channel and at the Waitaki Dam.

5 Important Points about Averages

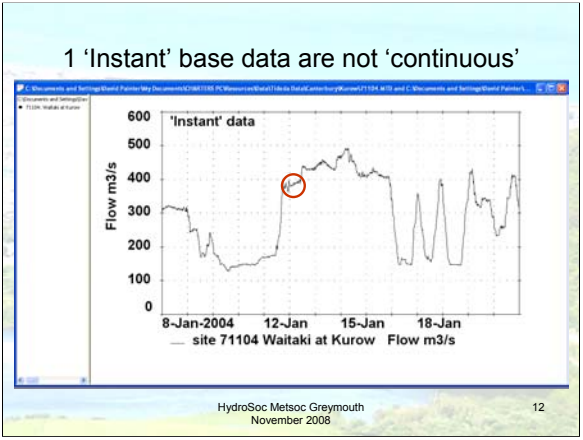
1. ‘Instant’ base data are not ‘continuous’
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need base data interval, window length, step length and time-label position all specified

5 Important Points about Averages

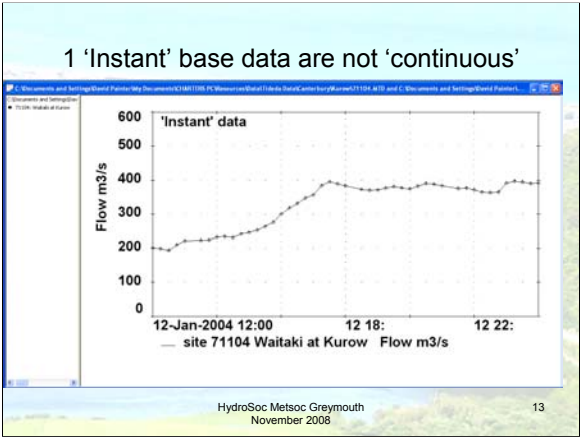
1. ‘Instant’ base data are not ‘continuous’
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need window length, step length and time-label position all specified



These are typical ‘instant’ data recorded by regional government and research agencies in New Zealand.



The data LOOK to be continuous.

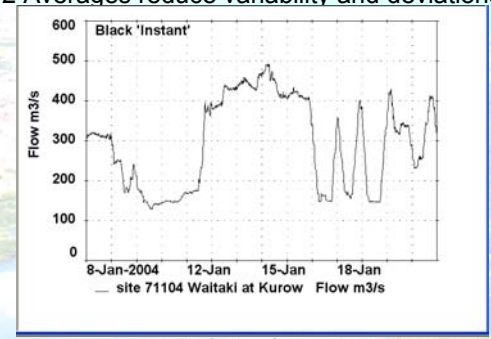


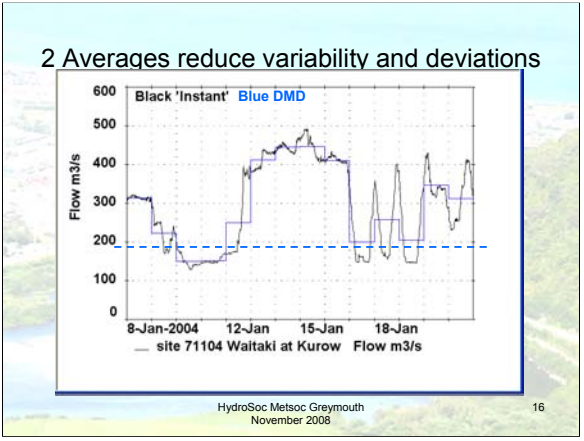
But they are actually sampled each 15 minutes. Most level data now are sampled each 15-minutes.

5 Important Points about Averages

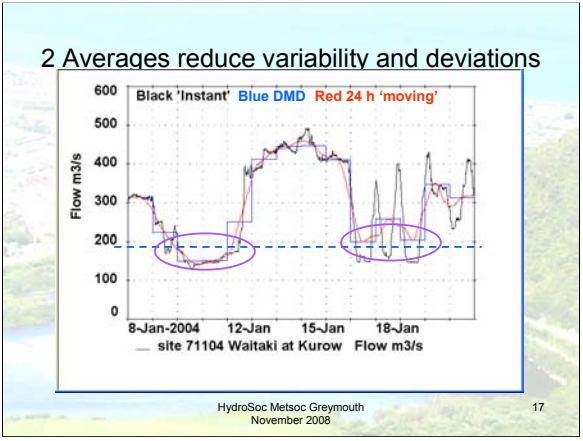
1. ‘Instant’ base data are not ‘continuous’
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need base data interval, window length, step length and time-label position all specified

2 Averages reduce variability and deviations





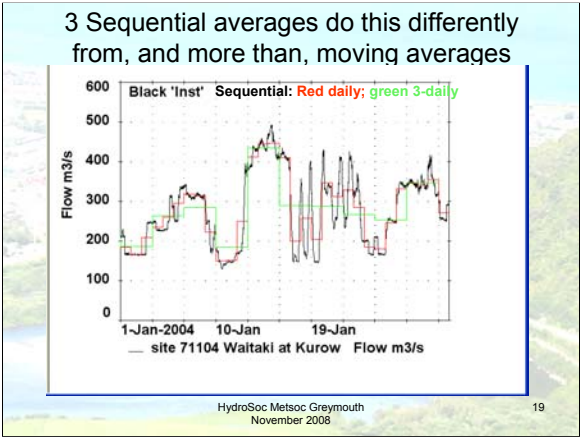
Daily Mean Discharge. Often there is a ‘threshold’ e.g. a ’minimum flow’ [190] for environmental or other reasons.



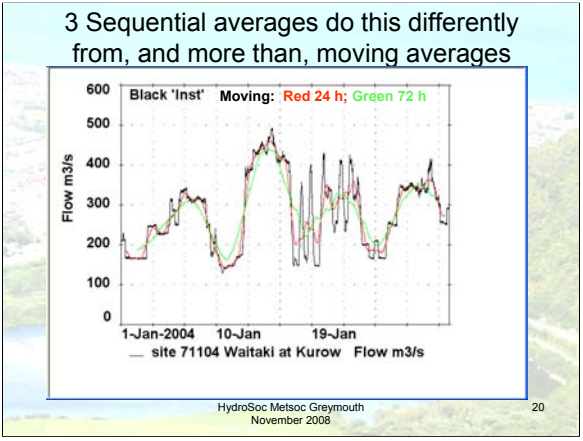
The right-hand oval shows both DMD and 24-hr moving average flows above minimum while the ‘instant’ flows go below it. It is less marked, but the left-hand oval also shows average flows above minimum for different times when the ‘instant’ flows are below it.

5 Important Points about Averages

1. ‘Instant’ base data are not ‘continuous’
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need base data interval, window length, step length and time-label position all specified

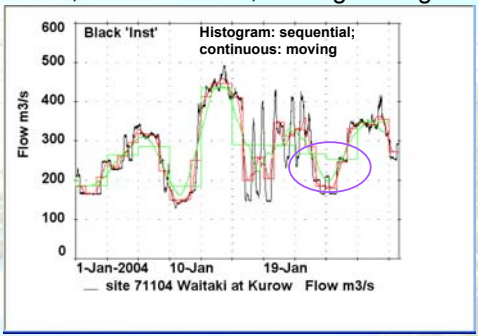


The histograms show sequential averages. The greater the ‘window length’, the greater the reduction in variation from the mean.



The continuous squiggly* lines are moving averages. There is a similar reduction in variation as the window increases in length. [*Strictly, there are a whole lot of over-lapping histograms – but it is too messy to draw like that.]

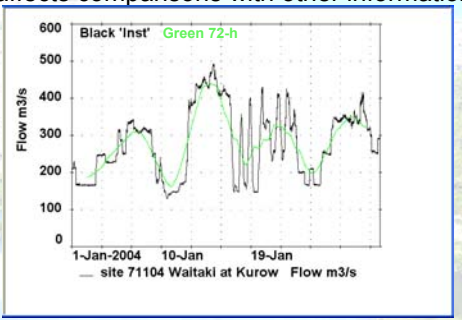
3 Sequential averages do this differently from, and more than, moving averages



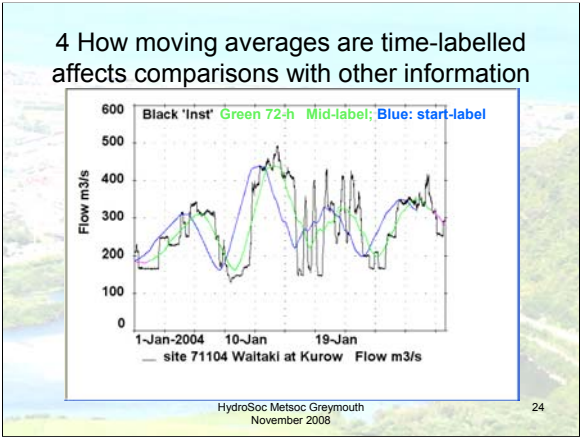
5 Important Points about Averages

1. ‘Instant’ base data are not ‘continuous’
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need base data interval, window length, step length and time-label position all specified

4 How moving averages are time-labelled affects comparisons with other information



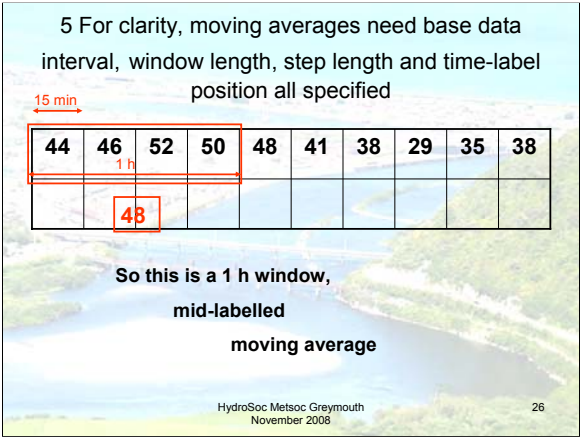
Note the coincidence of peaks and troughs.



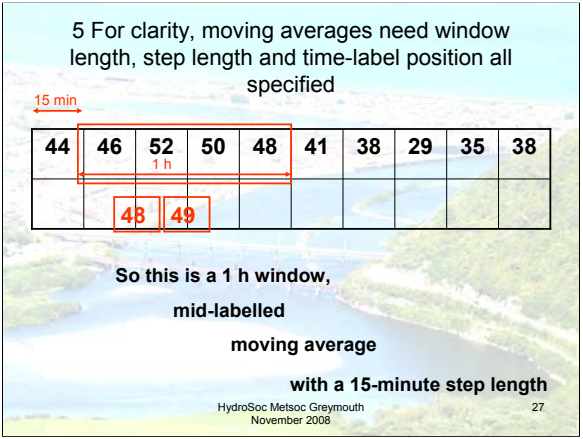
These are the same data but blue are ‘labelled’ and plotted at the start of the window. Green is centre-labelled. Some people prefer ‘end-labelled’.

5 Important Points about Averages

1. ‘Instant’ base data are not ‘continuous’
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need base data interval, window length, step length and time-label position all specified

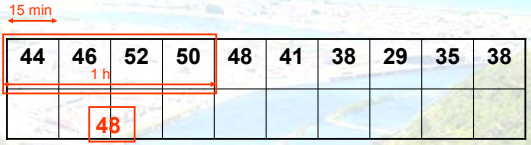


Add up the 15-minute ‘instant’ values and divide by 4.



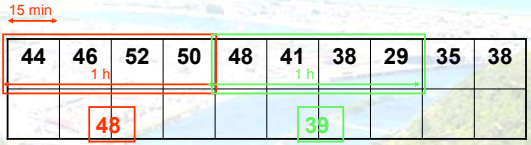
Move 1 step, add (48-44)/4 = 1 added to 48 to get 49.

And this 1 h window, mid-labelled moving average with a 15-minute step length



is quite different from

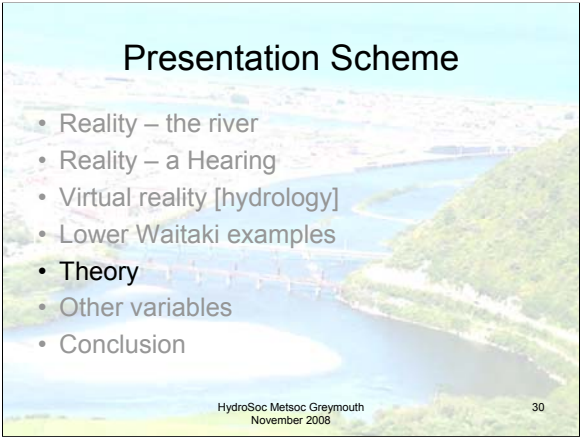
And this 1 h window, mid-labelled moving average with a 15-minute step length



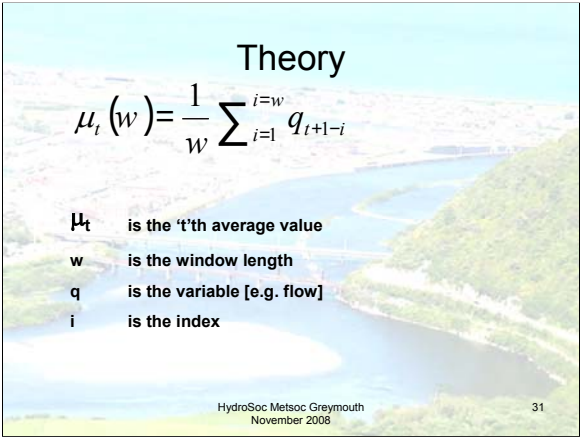
is quite different from

a 1-hour sequential average

of the same data

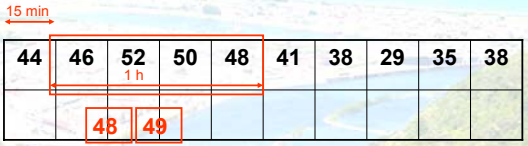


This is the 90 seconds of maths!



i.e. add them up and divide by how many.

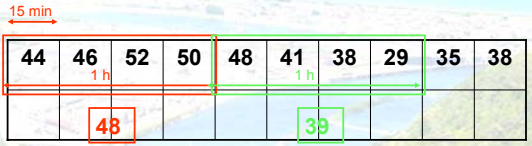
In this example, step length 1:



Window length = 4

$\mu_1(4) = 48$; $\mu_2(4) = 49$; etc.

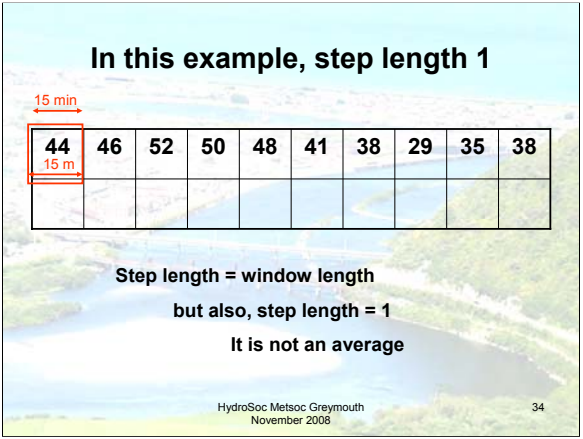
In this example, step length 4:



Window length = step length

It is not a moving average

but a sequential average



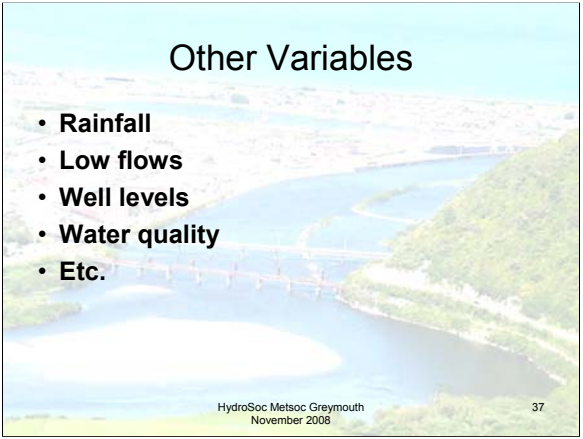
It is not an average because there is only one number in the window.

Simple Averages

- Are not always simple
- Are not always averages

5 Important Points about Averages

1. ‘Instant’ base data are not ‘continuous’
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need base data interval, window length, step length and time-label position all specified



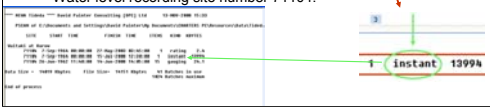
Or bank balances, or share prices, or …

Waitaki Catchment Water Allocation Regional Plan 2005

Table 3: Environmental flow and level regimes for water bodies in the Waitaki catchment

Water bodies	Environmental Flow regimes
xvii. Lower Waitaki River	d. All flows in the Lower Waitaki River determined for the purpose of this item xvii are to be based on measurements at the Kurow recorder[21] and based on 1-hour rolling averages

[21] Water level recording site number 71104.

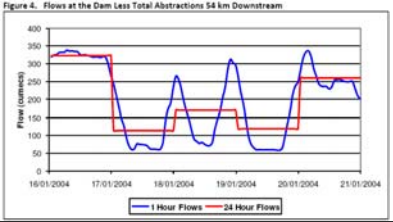


This was what gave rise to my ‘lemma’ to evidence at the Lower Waitaki Hearing. 1-hour rolling averages, of 15-minute data, step length 1, centre-labelled [probably].

Dave Stewart evidence to Lower Waitaki Hearing August 2008

32. The solution may be that the minimum flow becomes a mean flow over a longer time period such as 24, 48 or 72 hours. While the Plan states the minimum flow is 150 m3/s and it is based on a 1-hour rolling average, it does not state that the minimum flow is the 1-hour rolling average. There is no reason why the minimum flow cannot be the 24, 48, or 72-hour mean flow for the whole Lower Waitaki River based on the 1-hour rolling average flow.

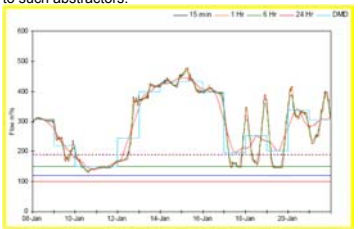
Figure 4. Flows at the Dam Less Total Abstractions 54 km Downstream



It seems he means the minimum flow could be 24-hour sequential, illustrated in the figure.

Rob Potts evidence to Lower Waitaki Hearing September 2008

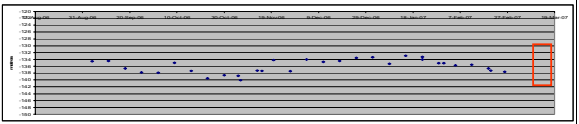
5 The concept suggested is based on two rolling averages: A 24-hour rolling average to provide trends and to put abstractors on notice that restrictions are possible; and a 6-hour rolling average that triggers the level of restriction. This would operate independently from any input from Meridian but is intended to give downstream abstractors some forewarning that restrictions may be imposed – a matter which may be desirable to such abstractors.



DMD in the figure is a sequential average. 15 min is the base 'instant' data. The others are 1-step rolling averages based on 15-min data with the window lengths shown. Rob originally used end-labelled averages, but revised those to centre-labelled after my evidence.

Lynton Dairy Ltd Consent CRC022071.1

9 **Should the static water level in monitoring well L36/1689, using a ten day moving average, reach the following trigger levels, then the overall take shall reduce by the values shown:**



NIWA Instrument Systems Update 3 September 2006
"The system records the groundwater level, transmitting these data by cellular modem to Lynton Dairy. This enables the company to manage its water usage and comply with the consent by progressively reducing its intake, should the 10-day moving average water-level fall below specified consent thresholds. Environmental Consultancy Services (Timaru) carries out routine site inspections of the bore and publishes the data* on the web for Lynton Dairy."
[* Logged at hourly intervals]

This is an example, from the Environment Canterbury website of how NOT to use moving averages. Note how many data points are inside the window at any one time. A 10-day moving average of weekly data doesn't make much sense! The explanation [after I pointed this out] was that the original data are hourly. An ECan staffer grumpily told me after my Shanty Town presentation that the condition was written by a Hearing Commissioner!

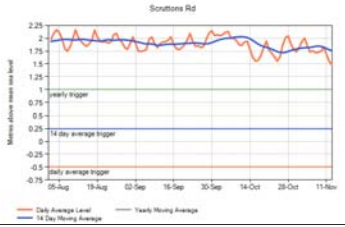
ECan Policy WQN10 Woolston/Heathcote groundwater management

(2) If the water level at the Scrutton's Road monitoring bore is simultaneously below all the groundwater trigger levels (i) and (ii) and (iii) below, all abstractions from the first confined aquifer shall cease until the water level recovers to a level where at least one trigger level is not exceeded:

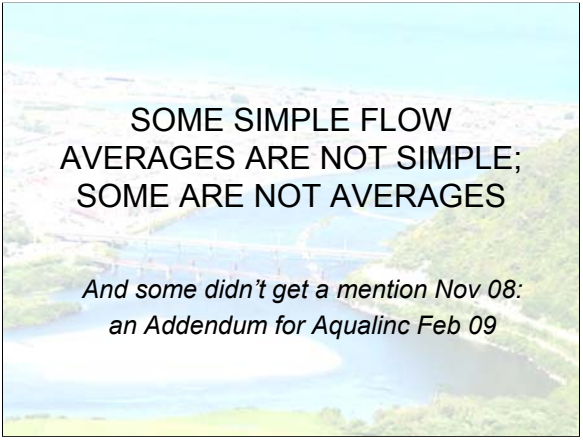
(a) one metre above mean sea level datum when taken as an average over a moving 365-day period;

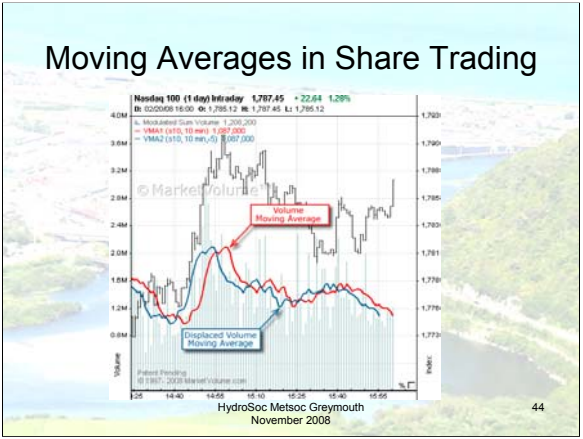
(b) 0.25 metres above mean sea level datum when taken as a 14-consecutive day moving mean; or

(c) 0.5 metres below mean sea level datum when taken as an average over a 24-hour period.

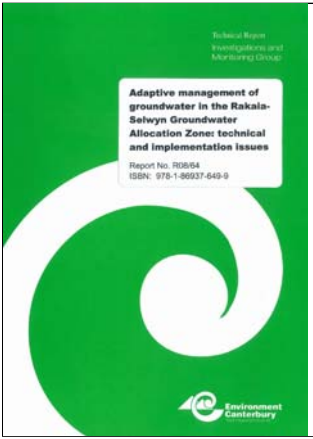


Dave Scott pointed me to this one as a GOOD example of moving average use. Correct. Unfortunately, the 365-day moving average was ‘missing’ from the web page; they found they had an instrument fault. The page still does not specify base data period, step length or label position.


SOME SIMPLE FLOW AVERAGES ARE NOT SIMPLE; SOME ARE NOT AVERAGES
And some didn't get a mention Nov 08:
an Addendum for Aqualinc Feb 09



Displaced volume moving average i.e the label position. These averages also appear in things like oil prices per barrel.



Exponentially weighted moving average or EWMA: A smoothing statistic for monitoring a process that averages time series data in a way that gives decreasing weight to data more distant from the time under consideration. The value of the smoothing coefficient alpha (α) changes according to the need to highlight pre-existing over recent data. In effect, it is a moving mean with a variable 'memory' of previous data.

The EWMA method was superior to simple three-year moving averages in reducing scatter.

This appeared after my Shanty Town presentation. Authors: Howard Williams, Dave Scott and Vince Bidwell. So I provided a brief explanation of how the EWMA relates to a moving average.

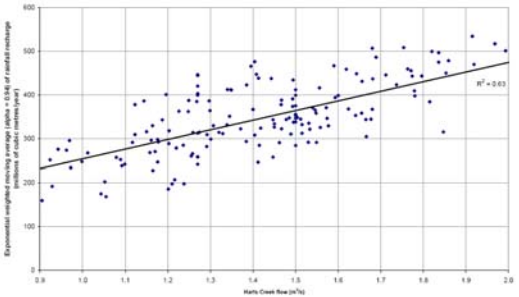


Figure 4-1: Plot of EWMA of rainfall recharge versus Harts Creek flow

The EWMA provided better fit than various moving averages in their report.

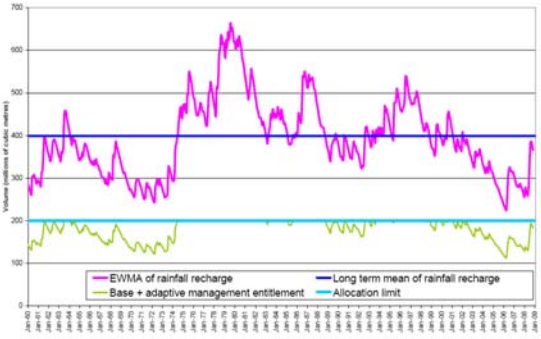
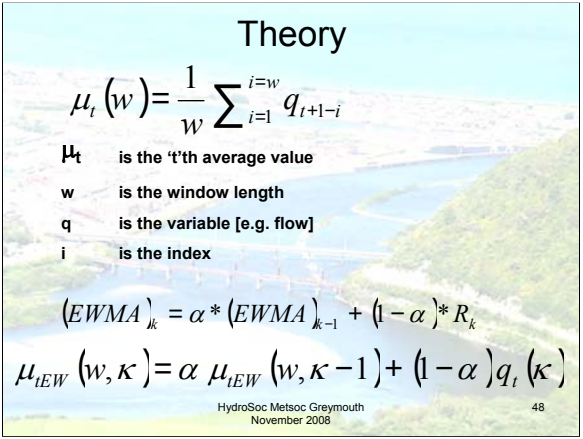


Figure 5-2: **Time series plot (magenta line) of EWMA of annual rainfall recharge volume in cubic metres (alpha = 0.97) and corresponding time series of adaptive allocation (green line)**

The alpha coefficient decides how long the memory window is, providing more or less emphasis on old values compared to recent values.



It is a recursive relationship comparable to my earlier example dropping off 44 and adding 48 [instead of adding the 4 values in the new window and dividing by 4, again]. Of course you have to 'start' somehow.

More Theory

$$\mu_{tEW}(w,\kappa) = \alpha\ \mu_{tEW}(w,\kappa-1) + (1-\alpha)q_t(\kappa)$$

The exponentially weighted moving average is also called 'Brown's Simple Exponential Smoothing'. It is equivalent to an 'ARIMA (0,1,1) model without constant'.

'With constant' allows for a constant linear trend in the data.

Then there are

Brown's Linear Exponential Smoothing

Brown's Quadratic Exponential Smoothing

Thus you can add more and more parameters to get better and better fit until you can fit anything to anything!

Thank you for listening!

1. 'Instant' base data are not 'continuous'
2. Averages reduce variability and deviations
3. Sequential averages do this differently from, and more than, moving averages
4. How moving averages are time-labelled affects comparisons with other information
5. For clarity, moving averages need base data interval, window length, step length and time-label position all specified