Kim Iles …

Not a Scaler

Forest Biometrician
Sampling Consultant / Teacher

My Points today

1) The form of the numbers stated is misleading

2) Sampling Error is also computed wrong in scaling and timber cruising

A "Sampling Error"
is multiple
("t") Standard **Errors**

SE is about **variability**
of an **average** around a **line**

The historical "**probable error**"
is the 50% level or **t = 0.7**

It was the "coin-flip" answer
(just like the average is)
±4%, for example

The 95% or t=2 level is :

The amount of error you are pretty sure you **DO NOT** have.
±12%, for example

Silly – right ???

The **probable error** is the most “honest” level to use in my opinion.

It is about 1/3 the 95% level.

12% vs. **4%** - **think about it**.

At least use the **t=1 level** **6%**

The Sampling Error is supposed

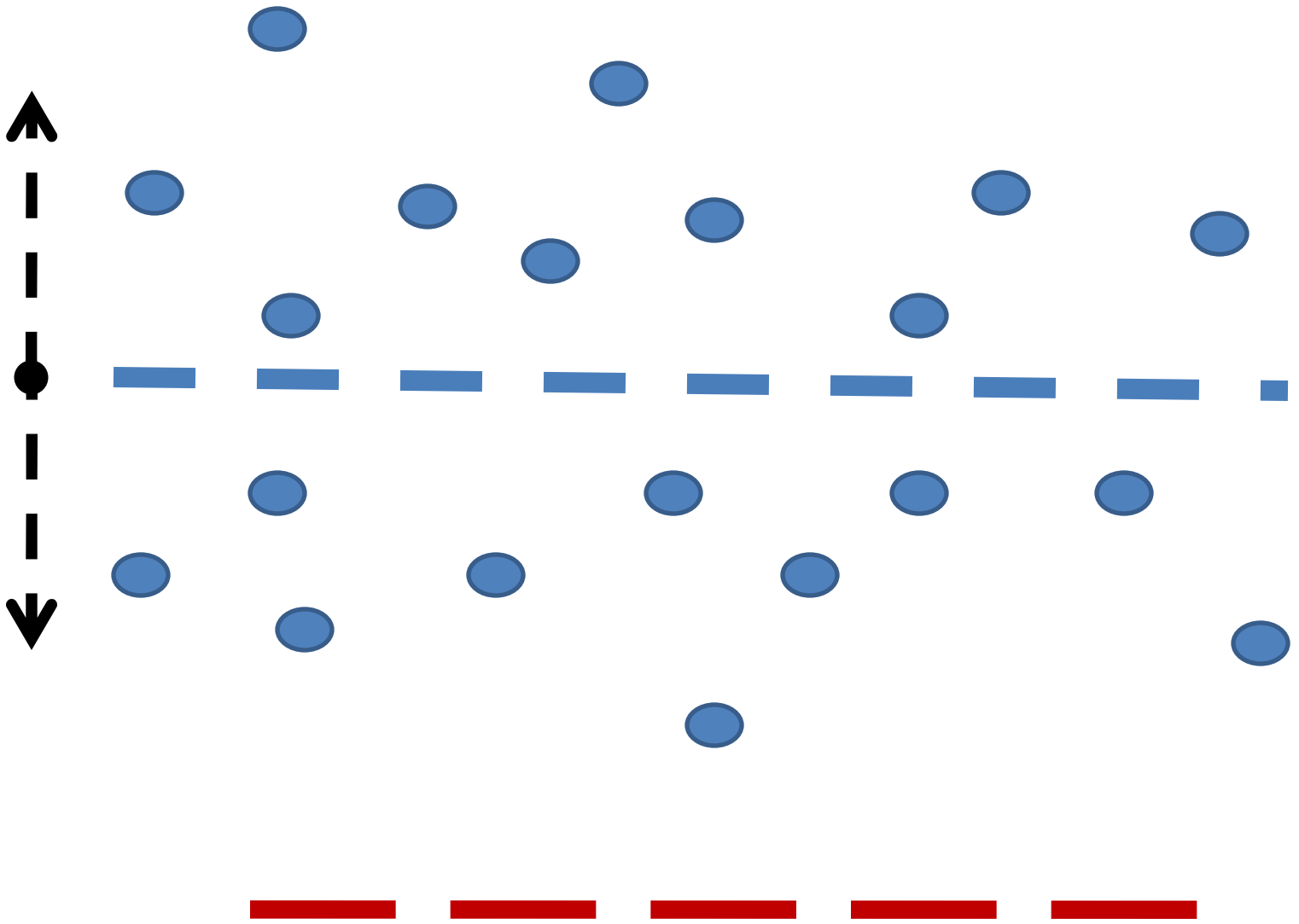
to indicate how close you are
to a total or average.

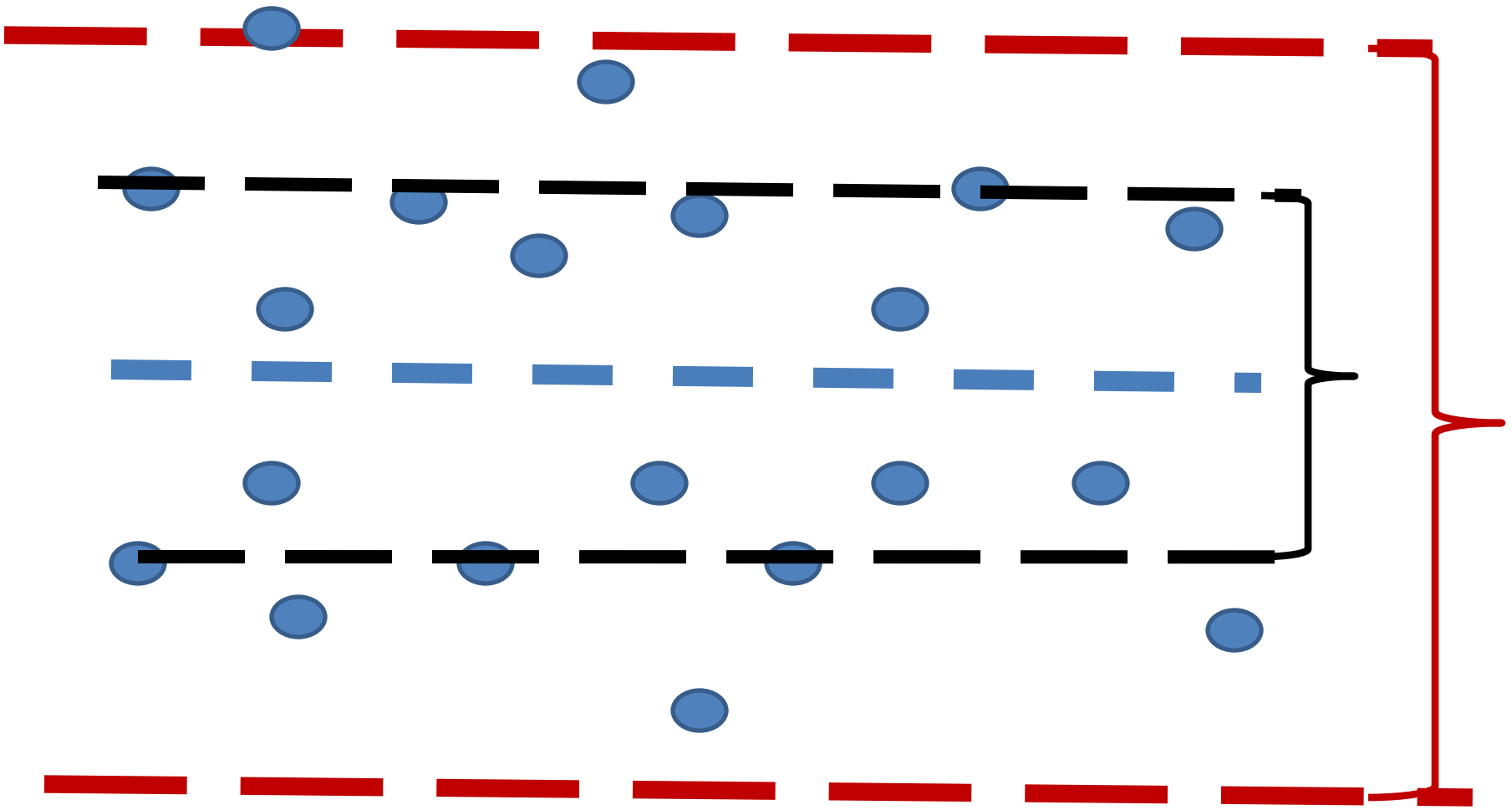
How far is “typical” ??

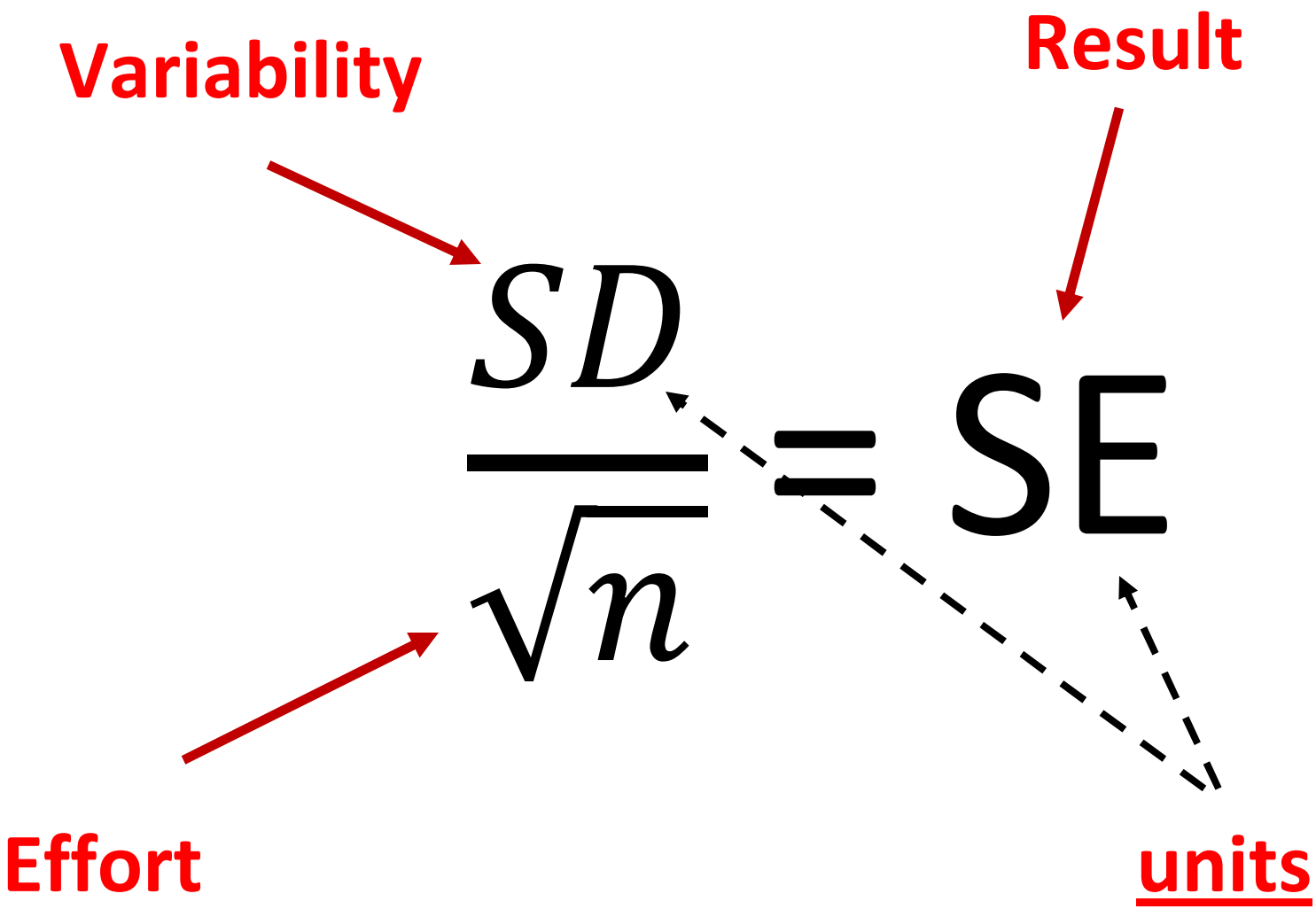
Where would YOU draw the
line to indicate “the” error
in this series of check cruises ???

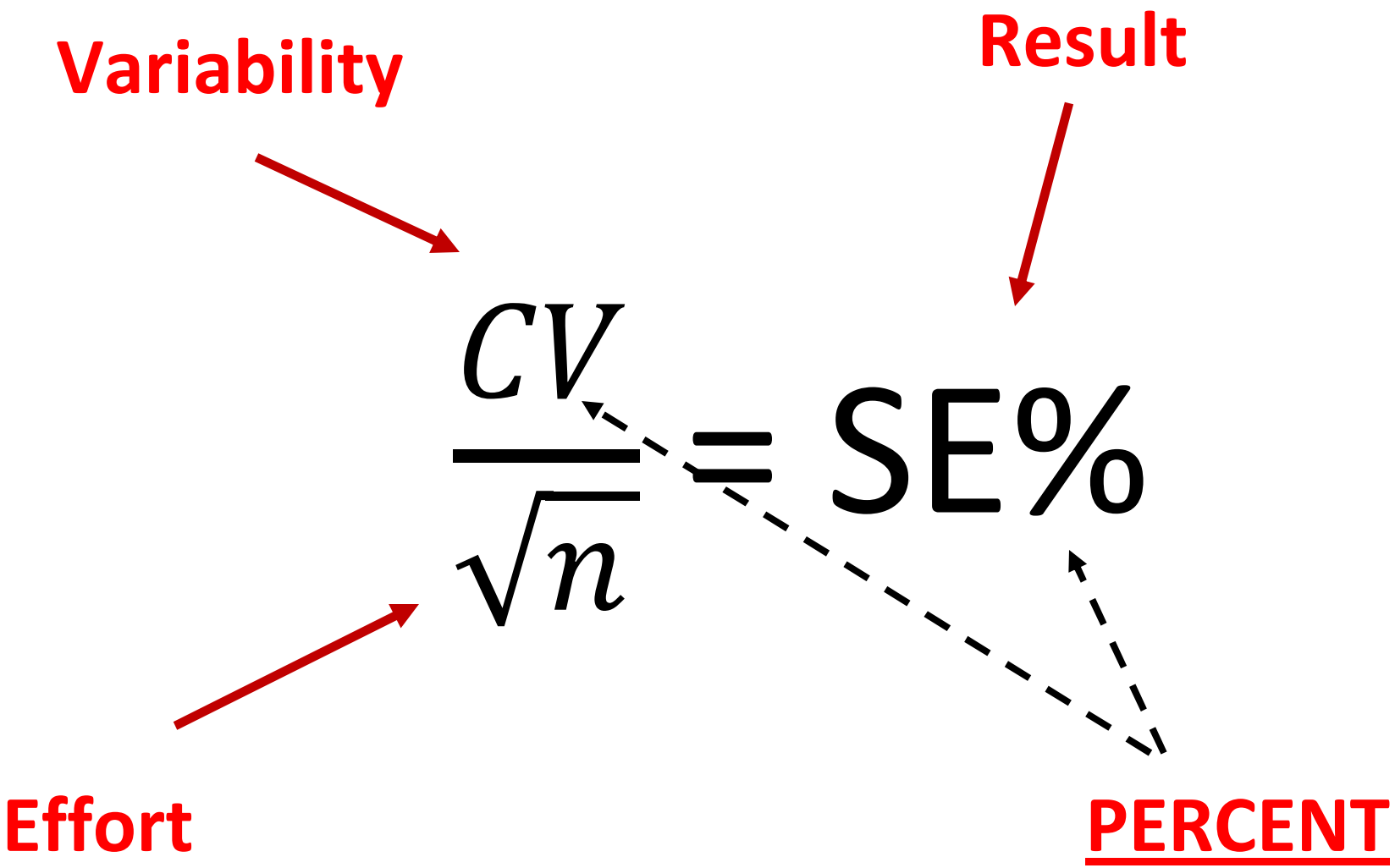
% error

Check scales

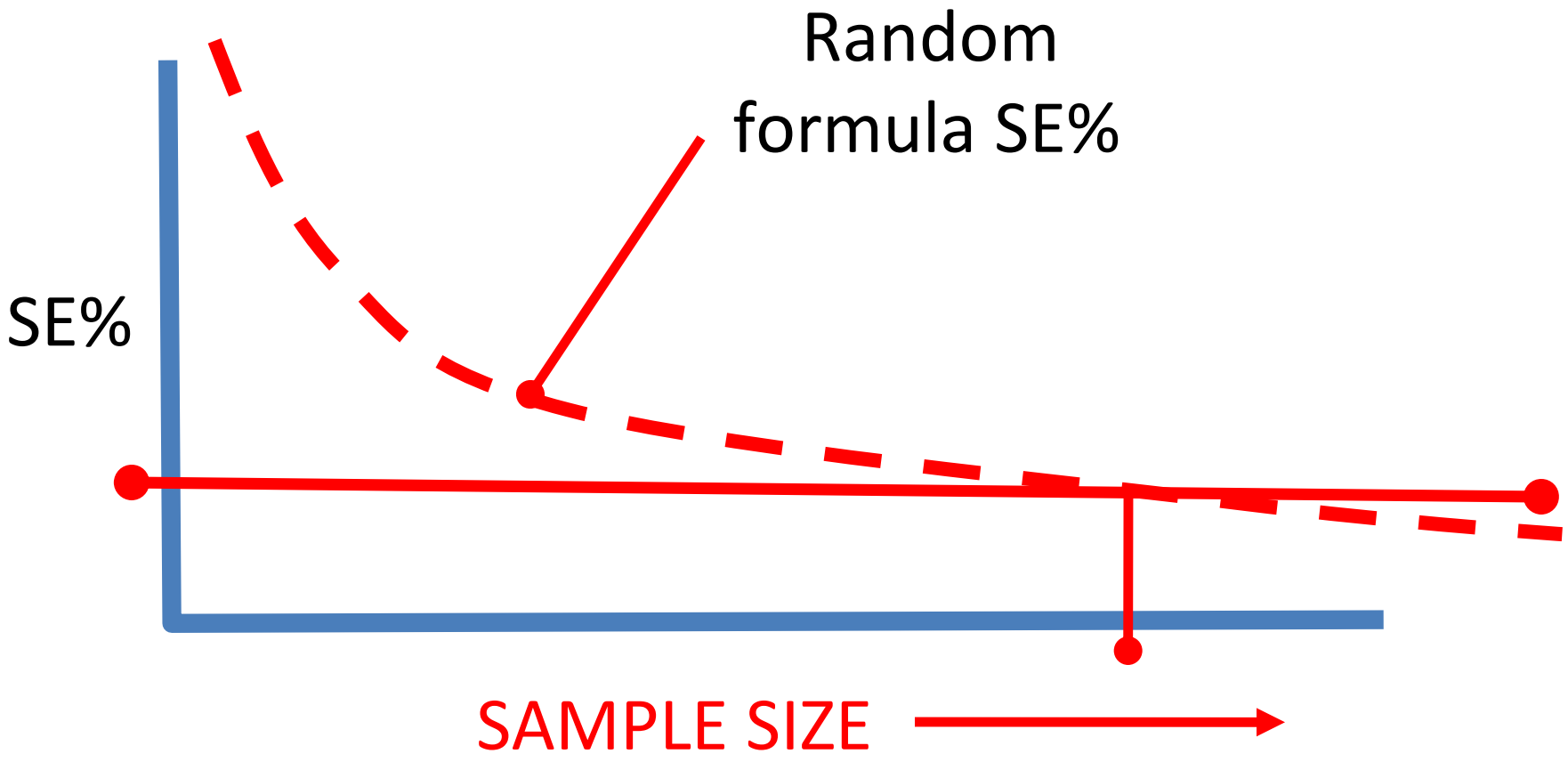





Variability
Result
$\frac{SD}{\sqrt{n}} = SE$
Effort
units


Variability
Result
$\frac{CV}{\sqrt{n}} = SE\%$
Effort
PERCENT

The sampling error computed with the random equation



The **SD** around the line is divided by the square root of sample size to get one **SE**

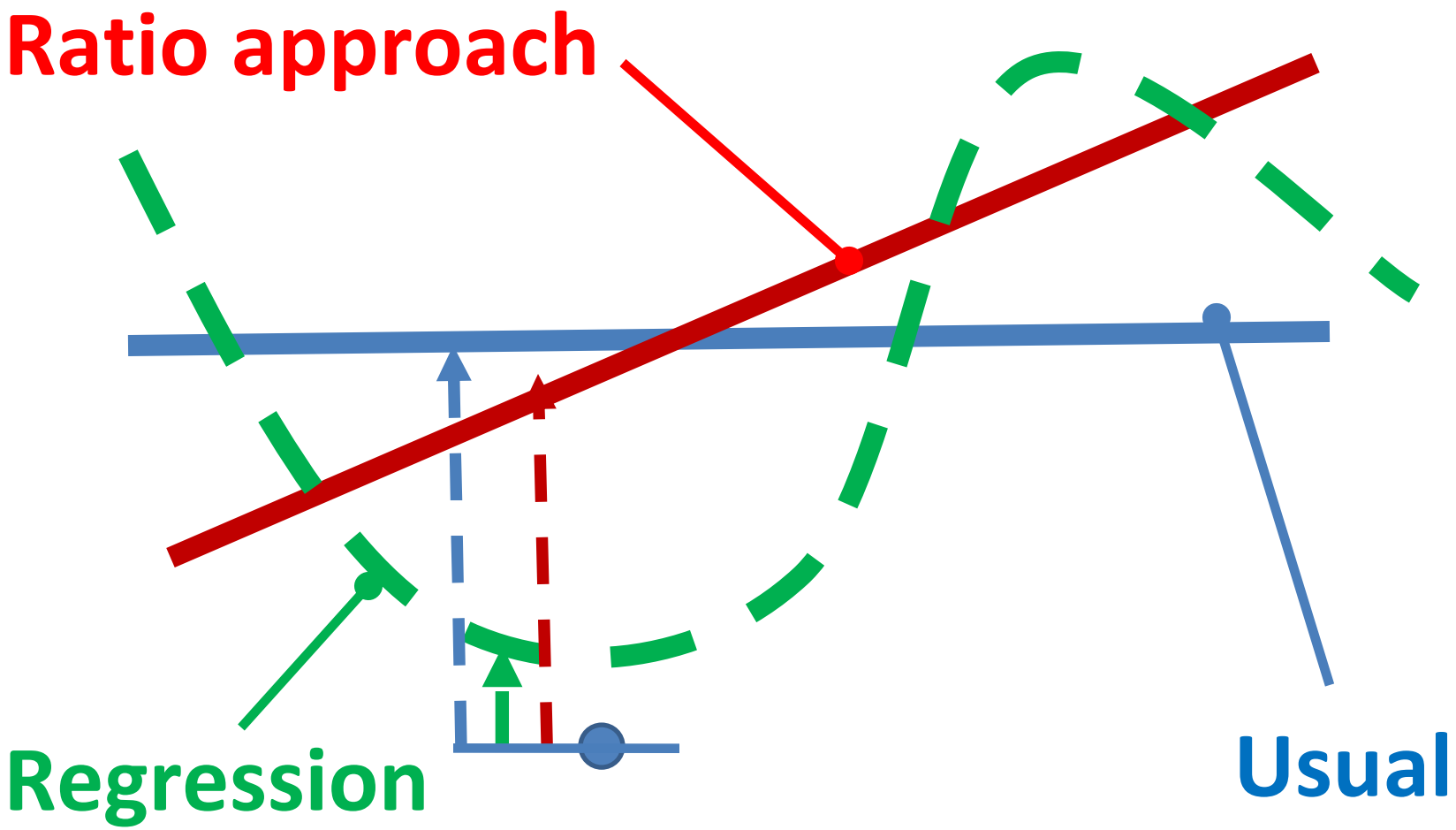
SD is the variability of **items** around a line

If SD or variability is

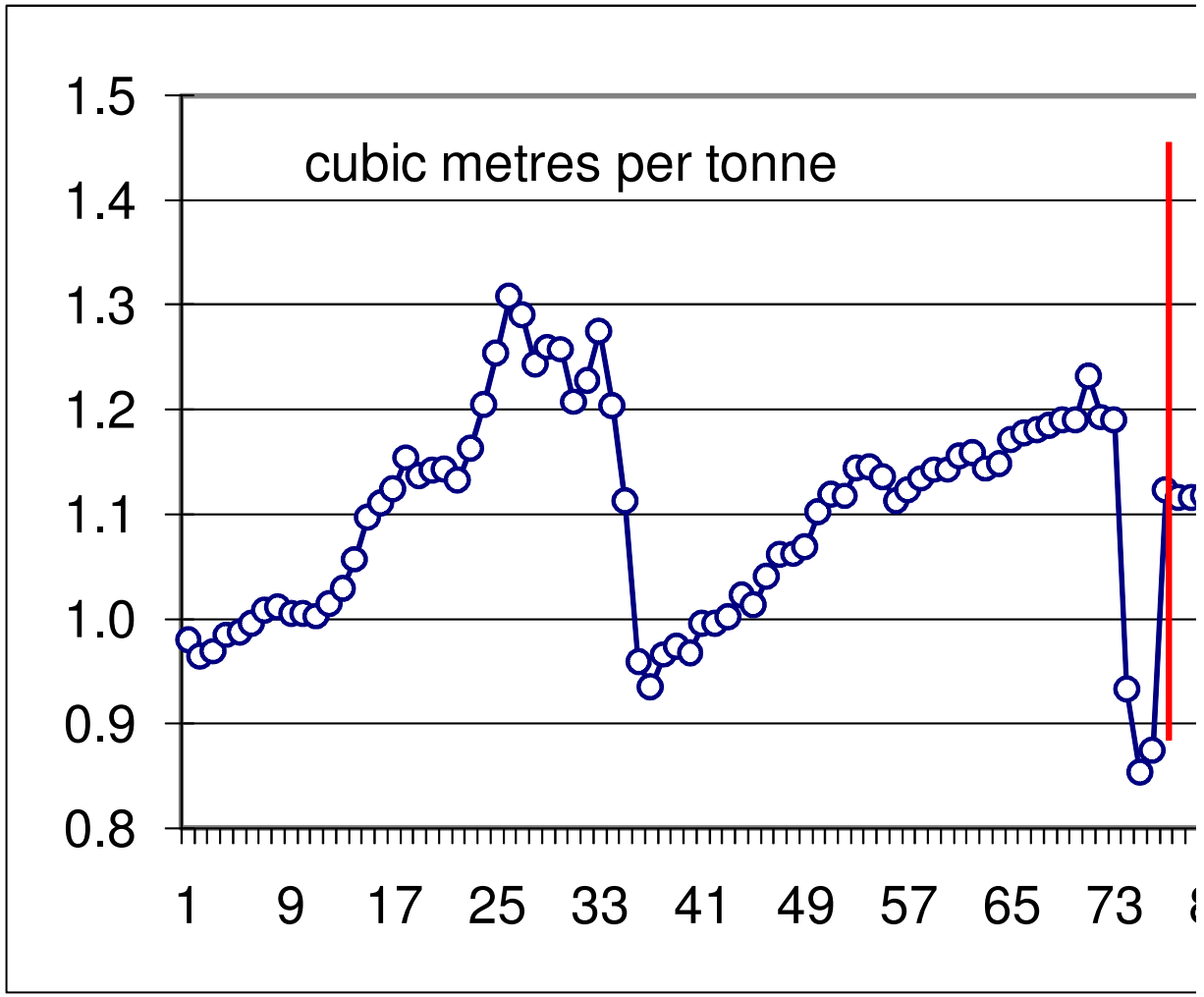
“around a line”

then

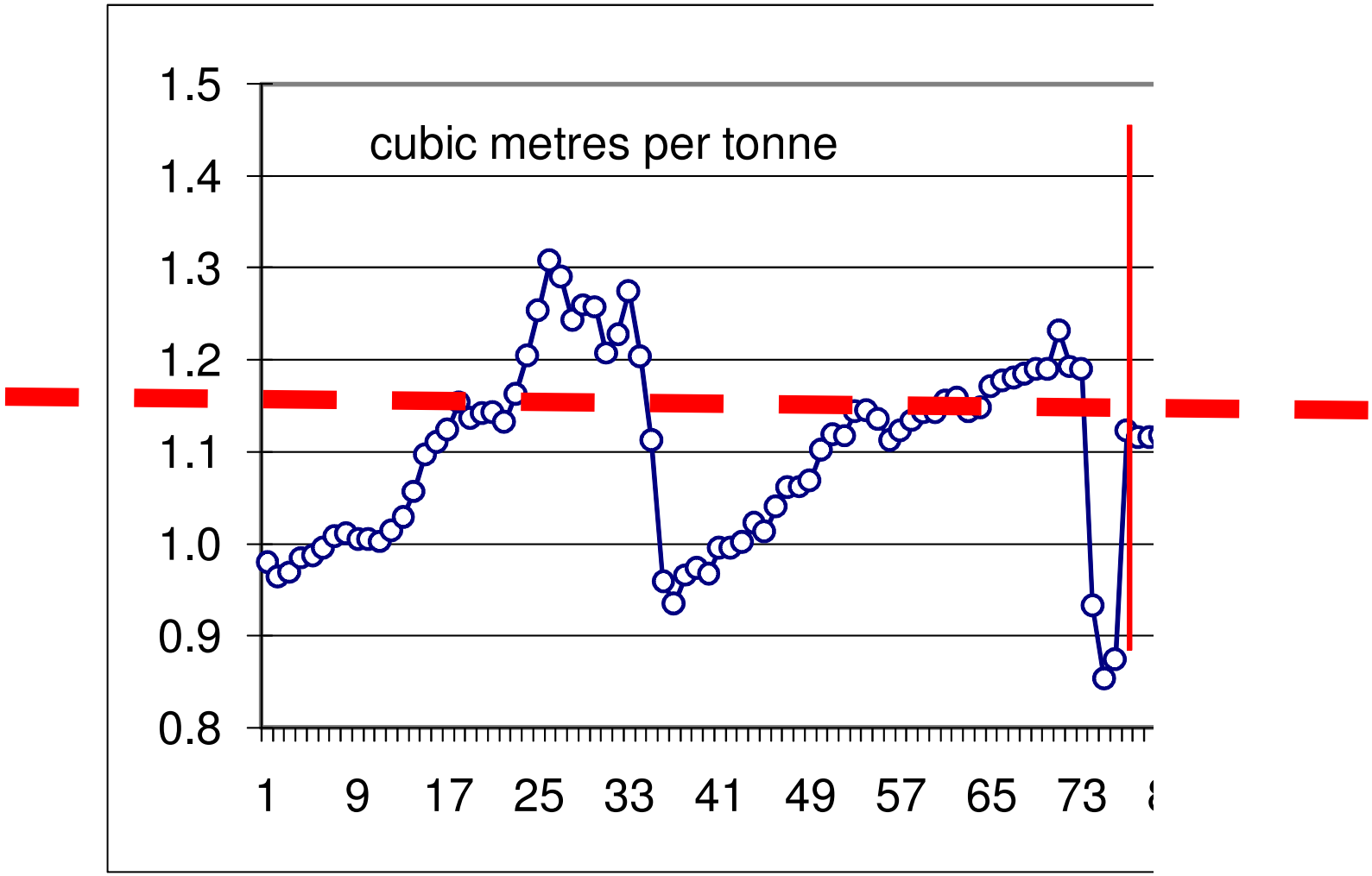
what line ???


Ratio approach
Regression
Usual

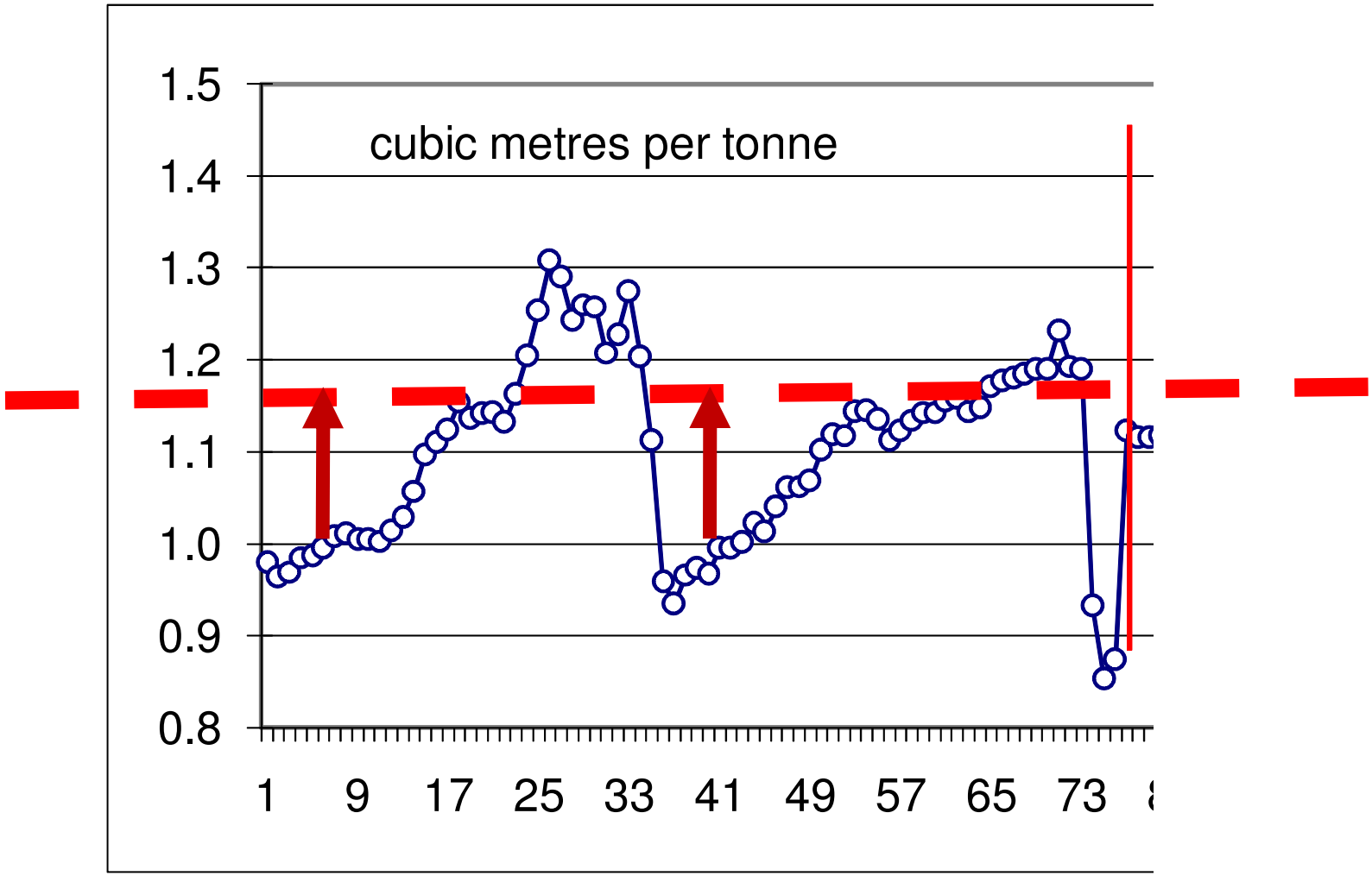
How would you assign values to the new loads of logs ??

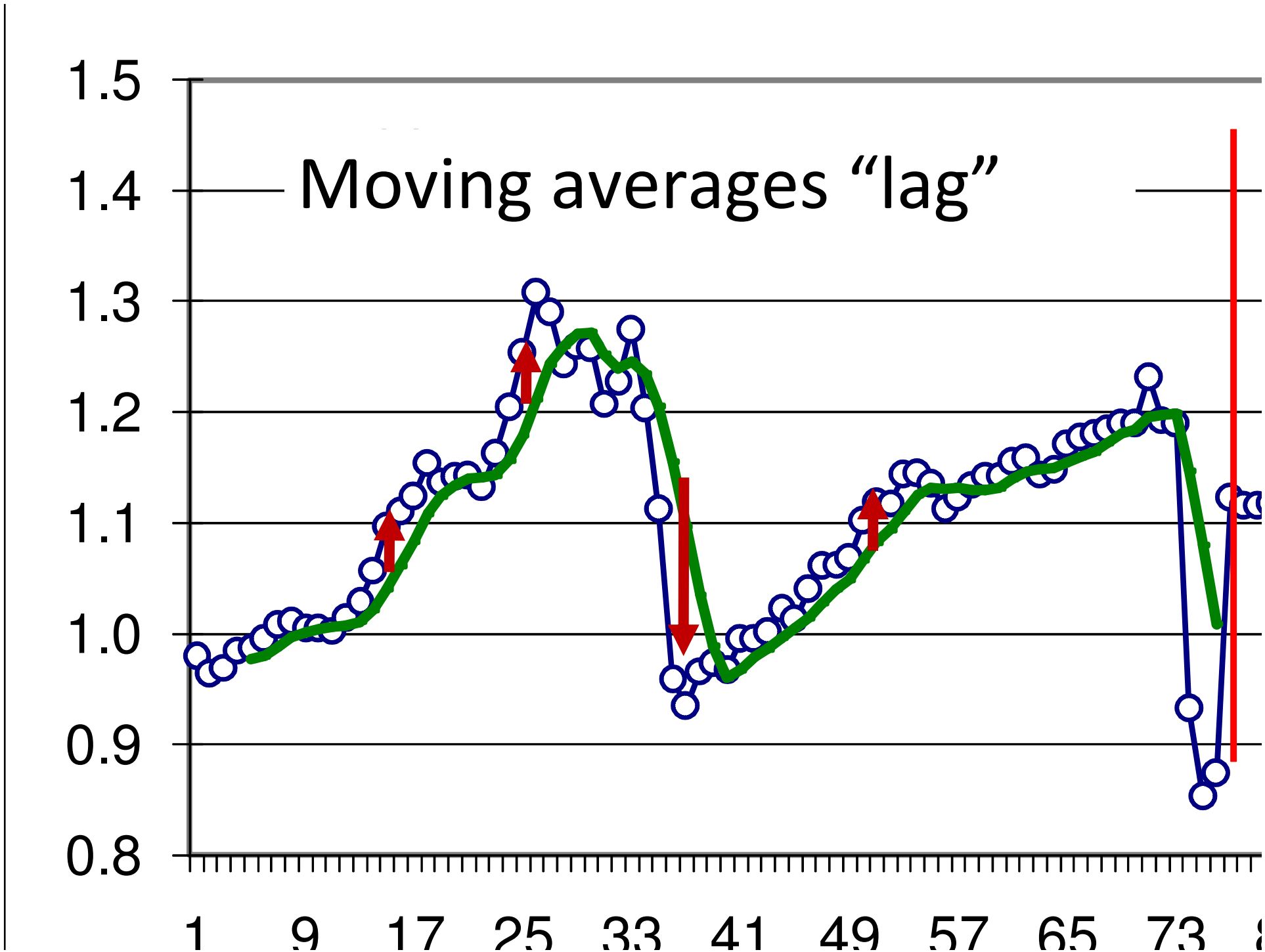


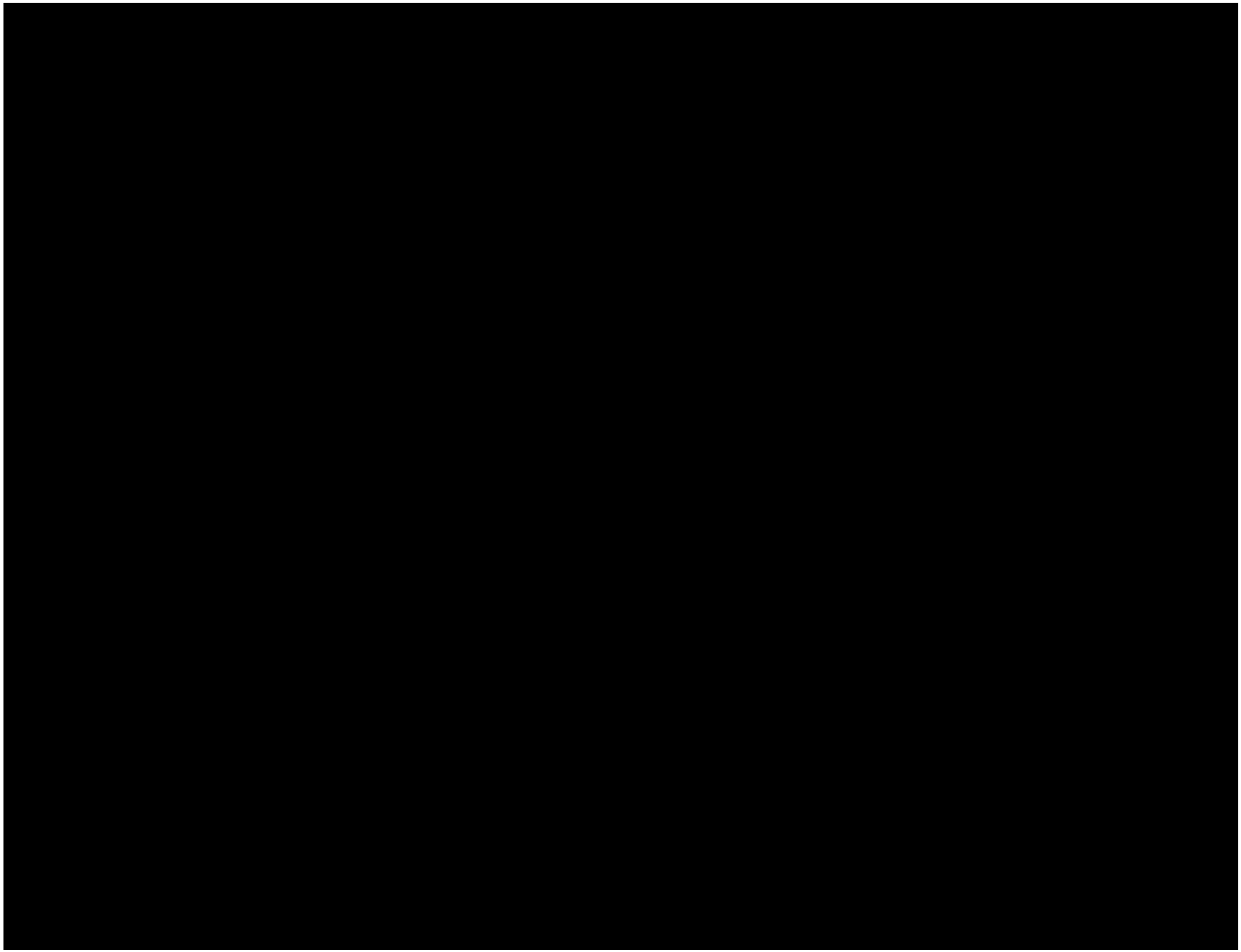
Like This ??


cubic metres per tonne
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
1
9
17
25
33
41
49
57
65
73

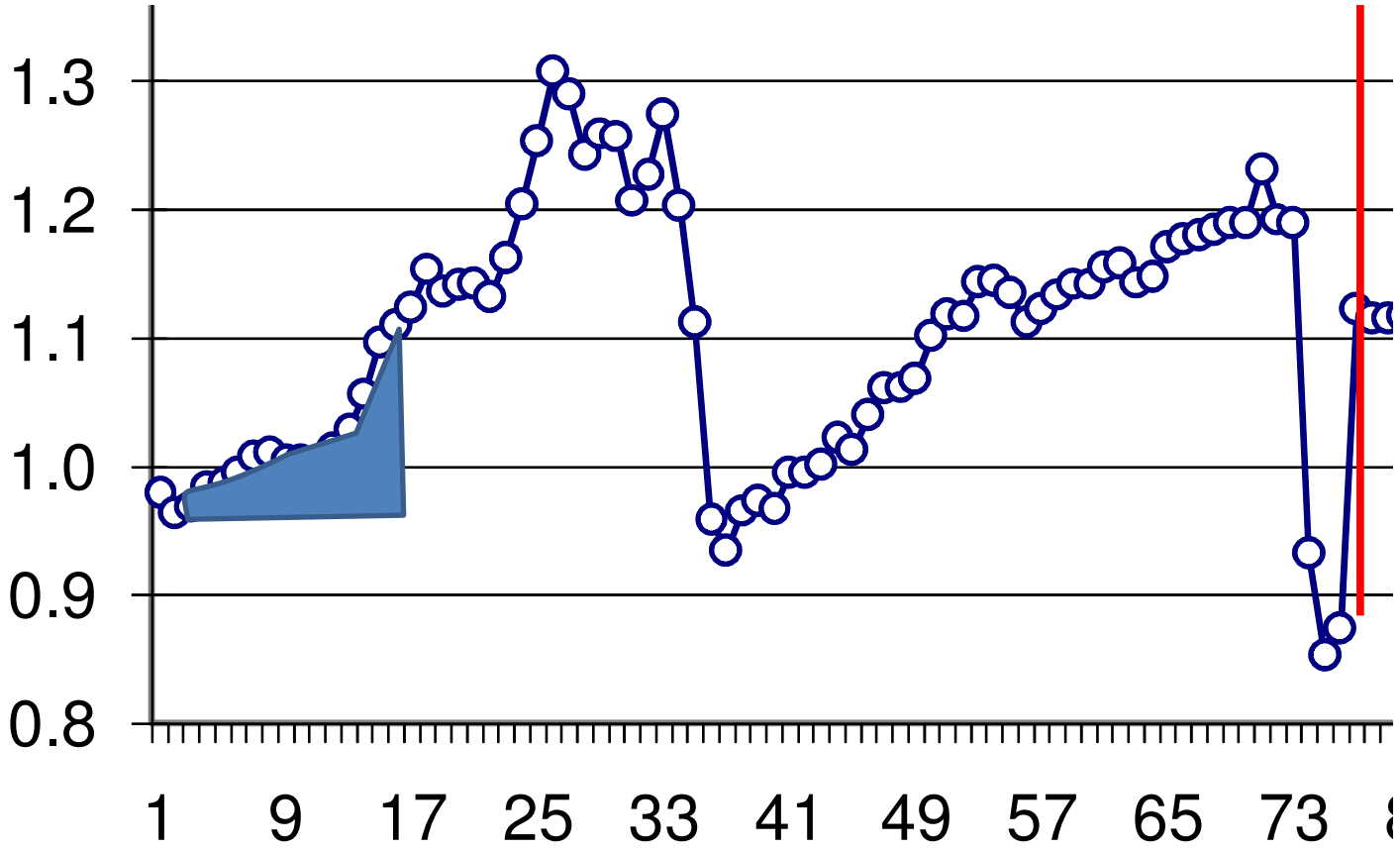
With these errors ??




Moving averages “lag”
1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
1
9
17
25
33
41
49
57
65
73



How about using the last measurement ??
"Prior load expansion"

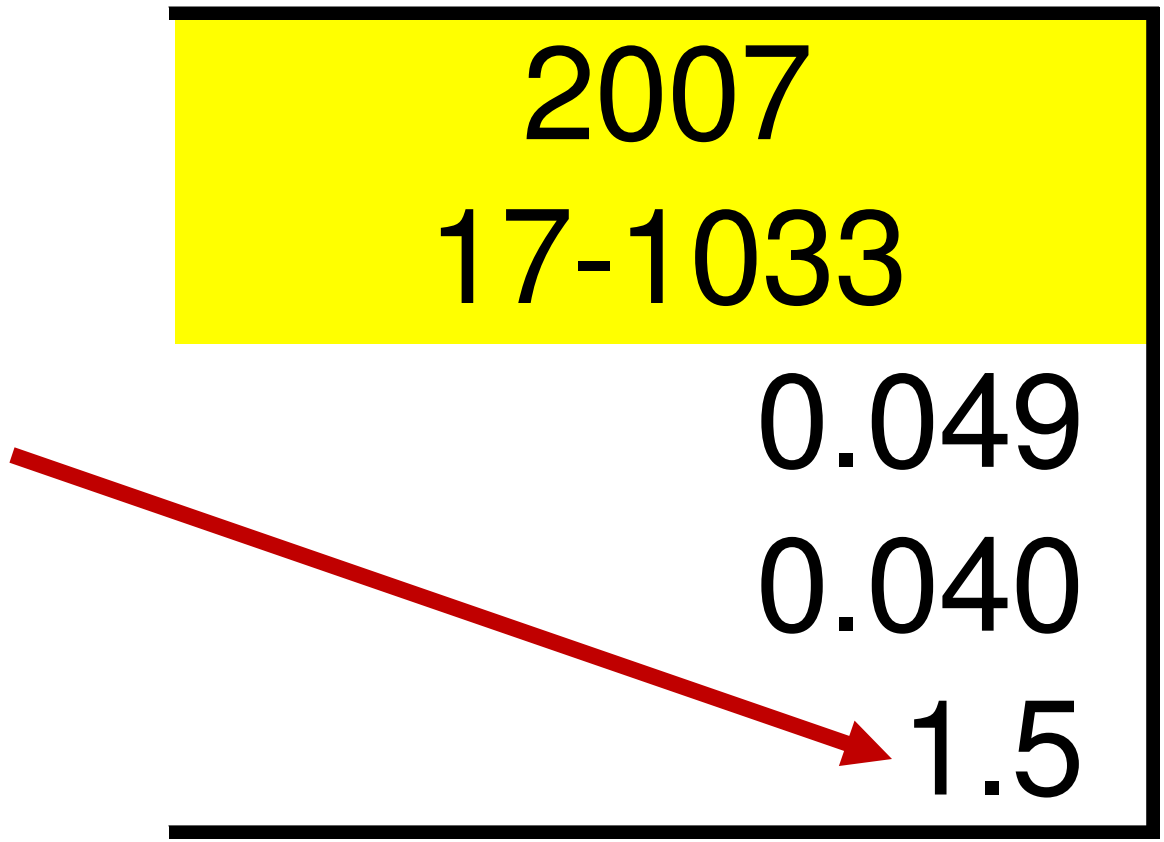


You could easily see how to make smaller errors than with the last sample load, **but SE is NOT the difference from the overall mean**.

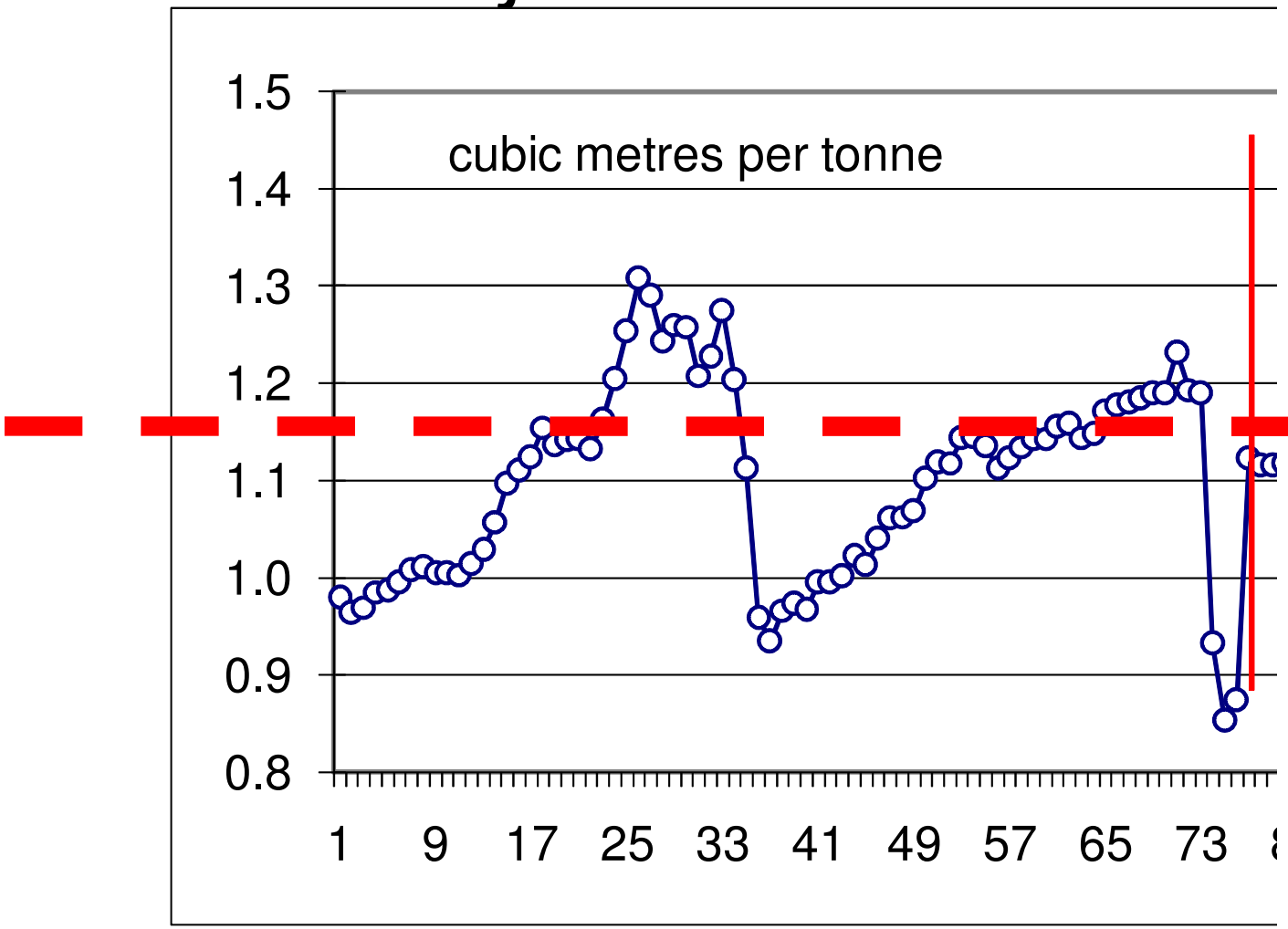
How much difference would it make ?

	2006-7 22-1030
Usual SD	0.1034
Line SD	0.0439
efficiency	5.6

The least efficient (average=2.5)

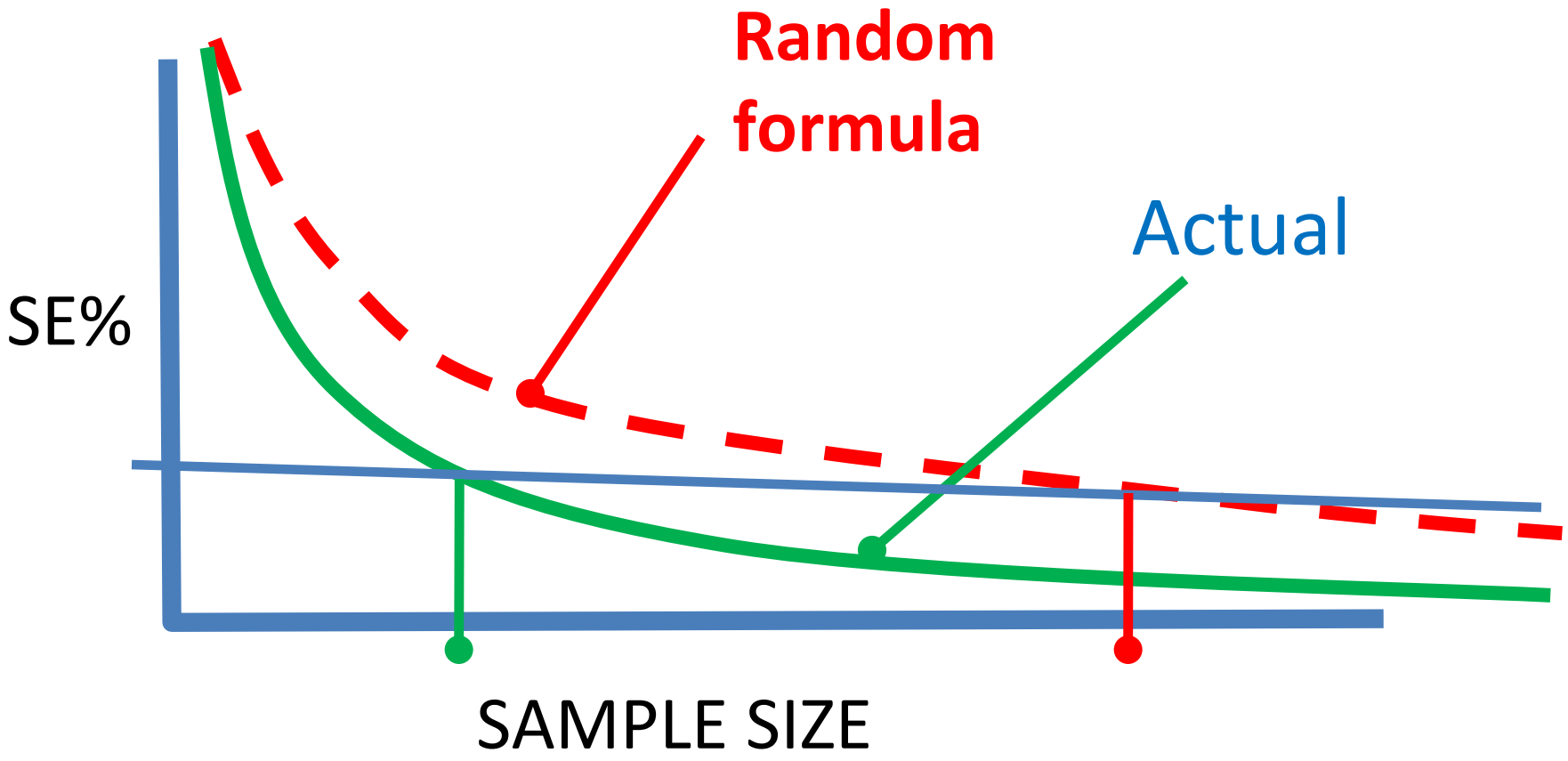


The point is that the usual RANDOM sample equation Is just WRONG



Exactly the same thing is true with forest sampling with grids or other systematic samples.

The actual sampling error vs the random equation



- 1) The actual sampling error is overestimated with random sample formulas
- 2) The “95% confidence limit” is misleading to almost everyone.
- 3) Most weight scaling is too intensive.

Thanks for having me here.