Mean vs Median

Analysing Stormwater Drainage Results with ARR 2016

Bob Stack
Director, Watercom Pty Ltd, Wolli, NSW, Australia
E-mail: bobstack@watercom.com.au

Benjamin Kus
Director, Kustom Engineering Pty Ltd, Warilla, NSW, Australia
E-mail: ben@kustomengineering.com.au

INTRODUCTION

The latest edition of Australian Rainfall and Runoff (ARR) 2016 has introduced new methods for performing stormwater analysis with the introduction of a new Initial Loss & Continuing Loss model, updated rainfall intensities and temporal patterns, and using ensembles of ten rainfall patterns per duration. The authors of this paper have been involved in the incorporation of ARR 2016 methods in the DRAINS program, and in providing information on the new methods to program users.

Since the official release of the ARR 2016 guidelines, there has been mixed advice provided by ARR for selecting critical results with regards to the mean or median storm in each ensemble. Unofficial discussions at the recent 2018 Stormwater Australia National conference indicate a preference leaning towards the mean being adopted for all results.

Here, the authors present information relating to the mean versus the median with reference to both stormwater and non-stormwater relates resources, and how certain scenarios common to stormwater drainage are likely to favour the median over the mean.

EXTRACT FROM THE ARR 2016 GUIDELINES

Section 5.2.2 in Chapter 5 of Book 2 (ARR 2016) states that '*it is well recognised that temporal patterns exhibit significant variability between rainfall events of similar magnitude, and the adopted pattern can have a significant effect on the estimated peak flow.*' Further the guideline states that *'Most of the historical research on temporal patterns has assumed that the central tendency of the pattern is more important than the variability, with the aim of producing a typical, representative or median pattern.*'

SYNOPSIS

The ARR 2016 Guidelines identify the importance of adopting an appropriate temporal pattern, and raise the topic of central tendency being more important than the variability. The definition of Central Tendency from Wikipedia is:

'In statistics, a central tendency (or measure of central tendency) is a central or typical value for a probability distribution [1]. It may also be called a center or location of the distribution. Colloquially, measures of central tendency are often called averages. The term central tendency dates from the late 1920s [2]. The most common measures of central tendency are the arithmetic mean, the median and the mode.'

[1] Weisberg H.F (1992) *Central Tendency and Variability*, Sage University Paper Series on Quantitative Applications in the Social Sciences, ISBN 0-8039-4007-6 p.2

[2] Upton, G.; Cook, I. (2008) *Oxford Dictionary of Statistics*, OUP ISBN 978-0-19-954145-4 (entry for "central tendency")

The attached Appendix 1 contains further extracts from several articles/papers discussing when to use the mean and when to use the median. There seems to be general agreement that if a distribution is skewed the median is the better choice. If it is not skewed then the choice is not so clear. It is also noted that the mean is influenced more by outliers.

The last extract in Appendix 1 makes the useful point that if the mean and median are considerably different, this indicates that the data are skewed (i.e. they are far from being normally distributed) and the median generally gives a more appropriate idea of the data distribution. In addition to this, the extract states that if the mean and median are not too different use the mean for discussion of the data, because almost everybody is familiar with it.

The Author's do not believe this final statement is a valid argument and that perhaps we could say it doesn't matter if the mean is similar to the median. This makes the choice simple – use the median which is good for both skewed and non-skewed distributions, and deals better with outliers.

URBAN DRAINAGE PHILOSOPHY

An urban drainage system is designed to reduce surface flooding to an acceptable level. The philosophy behind the Queensland Urban Drainage Manual design procedure is to carry as much water as we safely can on the surface and divert the excess into underground pipework. The design criteria then deals with safe values for surface flows (depth and velocity x depth). So, in assessing whether an urban drainage system is satisfactory, we are primarily interested in surface flows.

Peak flows coming off sub-catchments may not be skewed (although Appendix 3 suggests they may be), however, it appears that splitting the flows into surface and sub-surface flows has the effect of skewing the surface flows. Figure 1 shows an example of (surface) flow (in an overflow route) taken from a DRAINS model. The green and brown horizontal lines represent the mean and median values for 9 different ensembles (each ensemble has 10 storms with the same burst duration, and burst durations range from 10 minutes to 2 hours).

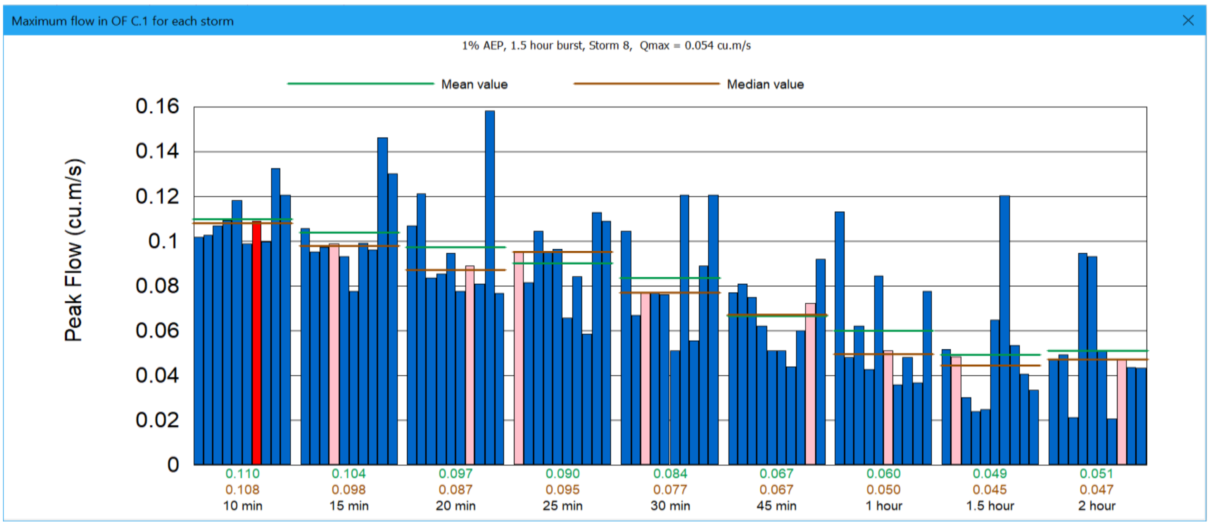

Figure 1 – Overflow route results from a DRAINS model

DISCUSSION ON URBAN DRAINAGE SCENARIOS

Appendix 2 provides an example of how splitting the flows between surface and subsurface flows at a pit can skew the surface flow. In this situation all the statistical literature indicates the median is a better indicator of central tendency than the mean.

An analysis of peak flows from sub-catchments in Appendix 3 suggests that the results may be skewed. If so, the literature clearly indicates that the median is the better choice.

It may be that the flow from a sub-catchment (all surface flow) is not skewed. However, with a sample size of 10 (i.e. 10 storms per ensemble) it is possible that we can get an outlier that will distort the sample mean. This makes it a poor indicator of the population mean. The median is immune to this.

SUMMARY

In summary, the median is a more robust indicator of central tendency and this is why we prefer it for urban drainage. With an even number of storms (10) the median peak value is defined as the mean of the peak values for the Rank 5 and Rank 6 storms. DRAINS uses the storm above the median (i.e. Rank 6) so that in fact it is slightly more conservative than simply using the median.

Appendix 1 – Extracts from Various Papers and Guidelines

Extract from ARR2016 Section 3.3.2

The ensemble event method results in an increase in computational requirements. Rather than adopting typical fixed values of inputs in the hope of achieving probability neutrality, modelled inputs are selected from an ensemble of inputs and the simulation results are based on the central tendency of the outputs (i.e. the average or the median, as judged appropriate for the degree of non-linearity involved).

Extract from ARR Revision Project 3 – Part 3 – Preliminary Testing of Temporal Pattern Ensembles – Stage 3 Report P3/S3/013 December 2015, Section 7.6

The study adopted the mean peak flow estimate from each ensemble for each duration. As a sensitivity test the median was used and compared to the FFA. The effect on the quantile estimates was minimal other than on catchments with low flows particularly those with high losses.

Extracts from Australian Bureau of Statistics on Central Tendency

Advantage of the median: The median is less affected by outliers and skewed data than the mean, and is usually the preferred measure of central tendency when the distribution is not symmetrical.

Limitations of the mean: As the mean includes every value in the distribution the mean is influenced by outliers and skewed distributions.

Extracts from Laerd Statistics

(https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php)

When not to use the mean:

The mean has one main disadvantage: it is particularly susceptible to the influence of outliers. These are values that are unusual compared to the rest of the data set by being especially small or large in numerical value. For example, consider the wages of staff at a factory below:

Staff	1	2	3	4	5	6	7	8	9	10
Salary	15k	18k	16k	14k	15k	15k	12k	17k	90k	95k

The mean salary for these ten staff is $30.7k. However, inspecting the raw data suggests that this mean value might not be the best way to accurately reflect the typical salary of a worker, as most workers have salaries in the $12k to 18k range. The mean is being skewed by the two large salaries. Therefore, in this situation, we would like to have a better measure of central tendency. As we will find out later, taking the median would be a better measure of central tendency in this situation.

Another time when we usually prefer the median over the mean (or mode) is when our data is skewed (i.e., the frequency distribution for our data is skewed). If we consider the normal distribution - as this is the most frequently assessed in statistics - when the data is perfectly normal, the mean, median and mode are identical. Moreover, they all represent the most typical value in the data set. However, as the data becomes skewed the mean loses its ability to provide the best central location for the data because the skewed data is dragging it away from the typical value. However, the median best retains this position and is not as strongly influenced by the skewed values. This is explained in more detail in the skewed distribution section later in this guide.

Skewed Distributions and the Mean and Median

Mean vs Median

We often test whether our data is normally distributed because this is a common assumption underlying many statistical tests. An example of a normally distributed set of data is presented in Figure 2 below:

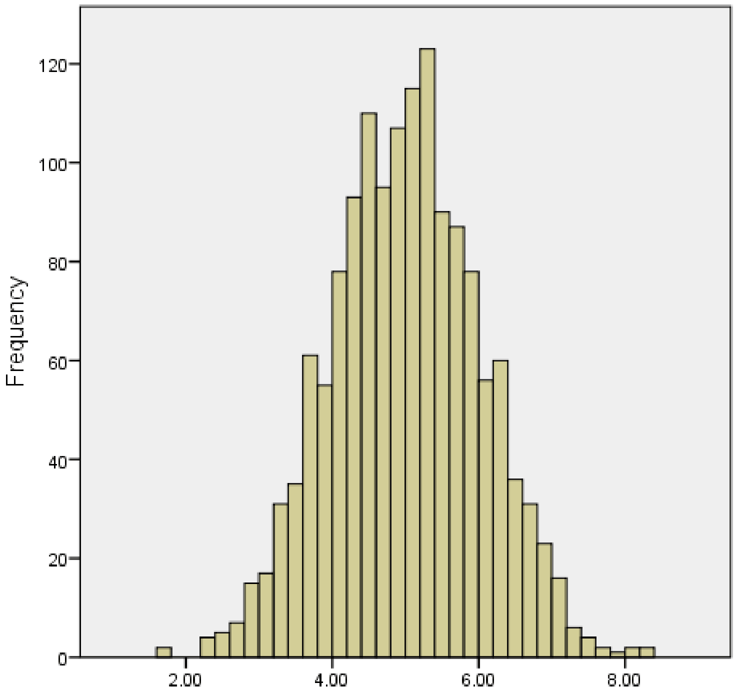

Figure 2 – Normally distributed set of data

When you have a normally distributed sample you can legitimately use both the mean or the median as your measure of central tendency. In fact, in any symmetrical distribution the mean, median and mode are equal. However, in this situation, the mean is widely preferred as the best measure of central tendency because it is the measure that includes all the values in the data set for its calculation, and any change in any of the scores will affect the value of the mean. This is not the case with the median or mode.

However, when our data is skewed, for example, as with the right-skewed data set shown in Figure 3 below we find that the mean is being dragged in the direct of the skew. In these situations, the median is generally considered to be the best representative of the central location of the data. The more skewed the distribution, the greater the difference between the median and mean, and the greater emphasis should be placed on using the median as opposed to the mean. A classic example of the above right-skewed distribution is income (salary), where higher-earners provide a false representation of the typical income if expressed as a mean and not a median.

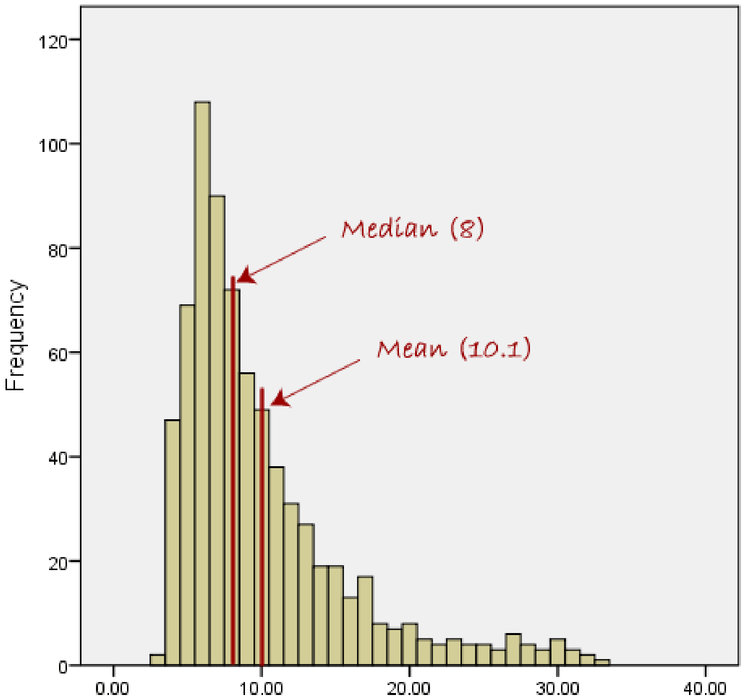

Figure 3 – Skewed set of data

If dealing with a normal distribution, and tests of normality show that the data is non-normal, it is customary to use the median instead of the mean. However, this is more a rule of thumb than a strict guideline. Sometimes, researchers wish to report the mean of a skewed distribution if the median and mean are not appreciably different (a subjective assessment), and if it allows easier comparisons to previous research to be made.

Watercom comment: For a non-skewed distribution the arguments for/against the mean and median get very statistical (see following extract)

Extract from R. Serfling (2009) 'Asymptotic Relative Efficiency in Estimation' University of Texas at Dallas

(Page 3 - http://www.utdallas.edu/~serfling/papers/Serfling_ARE_estimation.pdf)

"...for sampling from a Normal distribution, the sample mean performs as efficiently as the sample median using only 64% as many observations," but for the double exponential, "the sample mean requires 200% as many observations to perform equivalently to the sample median."

Extract from Clinfo.Eu: Stuck in the Middle – mean vs median

(http://www.clinfo.eu/mean-median/)

So which one should we use? The best strategy is to calculate both measures.

If they are **not too different**, use the mean for discussion of the data, because almost everybody is familiar with it.

If both measures are considerably different, this indicates that the **data are skewed** (i.e. they are far from being normally distributed) and the median generally gives a more appropriate idea of the data distribution.

Appendix 2 – Flow Split at an on-grade Pit

Consider an on-grade pit with 10 different approach flows:

93, 94, 95, 96, 97, 98, 99, 100, 101, 102 L/s

The mean is 97.5 and the median is also 97.5. Although inlet capacity varies with approach flow, let's simplify the process and assume the pit inlet capacity is 100 and any flow above this value spills down an overflow route. With a typical approach flow of 97.5 we don't expect the pit to spill. No spill during minor flows (10% AEP or 20% AEP) is a common design criterion.

The flows spilling down the overflow route would be:

0, 0, 0, 0, 0, 0, 0, 0, 1, 2

The mean is 0.3 and the median is 0. If we used the mean flow in the overflow route we might say the mean flow is above 0 and therefore the pit is likely to spill. The problem here is that the mean makes no distinction between the values 93, 94 ….. 100. If we use the median flow of 0 we would say the pit is not likely to spill (i.e. the same conclusion as when we looked at approach flows).

In this simple case we might conclude that we should look at approach flows to assess whether a pit is likely to spill. With more than one overflow route supplying water to the pit we would need to add the hydrographs and then take the peak flow in the combined hydrograph. This is inconvenient, but not impossible.

However, pit inlet capacity is not the only reason a pit might spill. It may be that the pit outlet pipe does not have capacity to carry the flow and water fills the pit causing it to spill. So perhaps we also need to look at HGL inside the pit to determine if it is likely to spill. But then we have another problem. When the flow spills the HGL is somewhat capped (it would be higher if the pit surface level was higher). So the HGL is not a good indicator to use.

In summary, the median spilling flow provides a convenient and robust measure of whether the pit is likely to spill.

Appendix 3 – Flow from Sub-Catchments

Is the flow from sub-catchments in a DRAINS model using ARR2016 rainfall patterns skewed? If so we would expect to see a trend for the mean to exceed the median (positive skew) or for the median to exceed the mean (negative skew). We checked results of a small DRAINS model for a Sydney suburb. It contains 5 sub-catchments with 9 storm ensembles (burst durations ranging from 10 min to 2 hours). This gives 45 ensembles. Here are the results for one sub-catchment:

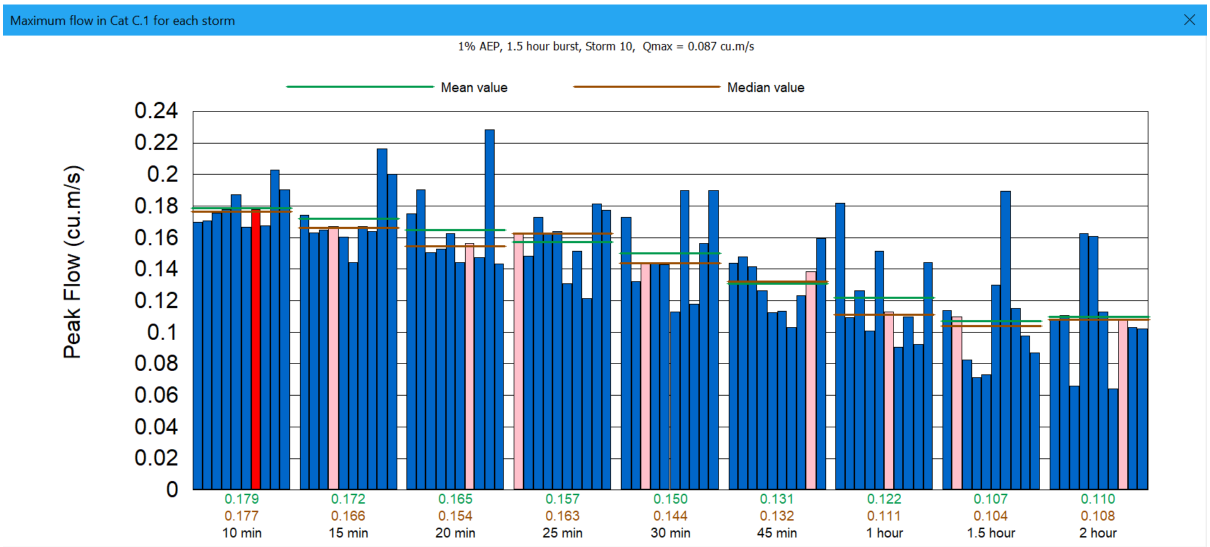

Figure 4 – Catchment flow results from a DRAINS model

Results for the five sub-catchments are summarised in the table below.

Catchment	Mean exceeds Median	Median exceeds Mean	No Difference
A.1	7	1	1
A.2	6	3	0
B.1	5	2	2
C.1	7	2	0
A.3	6	2	1
Total	31	10	4

There seems to be a trend for the mean to exceed the median. Does this indicate the data is skewed? We are not statisticians, but it certainly looks likely. If so, the literature clearly indicates the median would be the better indicator of typical results (central tendency).