Technical Note: T15-1

SIMPLIFIED RAINFALL STATISTICS FOR ON-SITE WASTEWATER MANAGEMENT : Which statistic applies?

Dr Robert Patterson
FIEAust, CPSS(3), CAg

16th December 2020

Simplified Rainfall Statistics for On-site Wastewater Management

Dr Robert A. Patterson Lanfax Laboratories, Armidale

1. Introduction

The role of an on-site wastewater system designer, when using a water balance model, is to understand the sensitivity of the wastewater generation and disposal regime to the rainfall and evaporation statistics for the locality, constrained by the practical and physical limits of the land application area. While a suitable model may provide a numerical outcome, the on-site designer may need to modify that outcome for local knowledge and environmental or regulatory constraints. The model is simply a tool in which the water balance algorithms are automated and the sensitivity of the various inputs (wastewater and rainfall) and outputs (drainage, runoff, evapotranspiration) is better understood. It is critical that these variables match the proposed management system as accurately as possible, that is the numbers match reality.

The aim of this paper is to compare various rainfall statistics for towns in several climate zones in New South Wales to show the differences between annual and monthly data summaries. One objective is to show whether the median, mean or some other ranking is the most appropriate monthly value to choose for a conservative water balance outcome.

2. Existing Water Balance Rainfall Inputs

This technical note addresses only monthly statistics, acknowledging that daily models may use many years of historical data for a locality and account for a wider range of variables. Compared to daily models, monthly models are known to be conservative, more suited to small area domestic wastewater systems of sub-soil dispersal or sub-surface irrigation. However, some professional water balance models are more suited to large scale agriculture rather than a small rural residential allotment.

A ***water balance model*** is simply a computer aided set of calculation of a range of inputs and losses to the hydrologic cycle. Water comes into the system as wastewater and rainfall, leaving as runoff, deep drainage, capillary flow in all directions, evaporation from exposed soil surfaces, transpiration of the vegetation in contact with the water in the soil, and changes to the soil water storage. These variables are altered by the daily activities in the household (wastewater generation) and slight changes to the output mechanisms as influenced by local climate and soil characteristics unique to the location.

Monthly water balance models are in wide use in the design of on-site wastewater management systems. Both the *Environment and Health Protection Guidelines* (DLG *et al*., 1998: pages 155 & 159)) for domestic on-site systems and the NSW *Environmental Guidelines: Use of Effluent by Irrigation* (DEC, 2004, page 57) guidelines, for low strength effluent re-use in agricultural setting, suggest the use of 50th percentile (median) rainfall figures in monthly modelling. The source of the research that supports this statistic is not reported in either document. The NSW Environment Protection Authority (NSW EPA) follow the DEC (2004) guidelines. Local Councils are required to consider the DLG guidelines in regulating on-site wastewater management policies (*Local Government (General) Regulations 2007*, clause 29(3)).

That the 50th percentile rainfall is considered appropriate for the planning of on-site wastewater disposal or re-use fails to understand that such a statistic encourages acceptability of a failure once in every two years; an ill-considered outcome. When the water balance selected the 50th percentile monthly value as the input rainfall variable, again, ignoring the chance that the chosen monthly values may add to be different to the 50th percentile annual value. The risk of higher frequency failure has increased to the detriment of the land application area.

AS/NZS 1547:2012 *On-site domestic wastewater management* (Standards Australia & Standards New Zealand, 2012) fails to mention the use of any rainfall statistic, obviously ignoring the significant differences that arise between low and high return intervals as mentioned in the Standard (Appendix Q, page 180), leaving such selection entirely to the designer. Such omissions are likely to encourage disagreements between designers and regulators.

3. Available Rainfall Data

While the Bureau of Meteorology (www.bom.gov.au) provides access to climate data from recording stations across Australia, the website summaries only provide a narrow range of statistics. By accessing the monthly records for a particular station, the designer is able to perform percentile ranking to calculate any rank of interest. For this project, the monthly rainfall values were downloaded for the 16 stations shown in Table 1.

This document was originally written in 2015 and the data for the stations have not been undated. The explanation of the significance of choosing a more appropriate rainfall statistic is the purpose that would not be further enhanced by updating to the last full year of records (2019). There may be slight variation in the values due to the very dry period 2017-2019, but the additional re-calculation is excess to the purpose of this article.

It is assumed that readers will explore the statistics for their own areas based upon the examples here.

Table 1. Sixteen recording stations selected to show variation across NSW from coastal to inland.

Station/No.	Opened	Altitude	Station/No.	Opened	Altitude
Ainslie (ACT) 70000	1935	585	Ulmarra 58059	1891	5
Armidale 56002/56037	1857/1997	980	Hillston AP 75032	1881	122
Bega 69002	1879	50	Manildra 65022	1888	530
Byron 58007	1892	3	Narrabri 54120	1870	213
Camden 68007	1882	61	Oberon 63063	1888	1053
Pokolbin (Cessnock) 61056	1905	140	Port Macquarie 60026/60139	1942/1995	20/4
Coffs Harbour 59010/59040	1900/1965	21/5	Robertson 68054	1890	758
Gosford 61023/61319	1877/1971	4/35	Wagga Wagga 73127	1898	219

Data sourced from Bureau of Meteorology's website in 2015.

Altitude is reported in metres above average sea level (m ASL). Where possible, record periods in excess of 100 years were sought, with 2013 the last full year of records.

The data for Armidale, Coffs Harbour, Gosford and Port Macquarie are each the composite of two stations, following the closure of the earlier station and the opening of a new location on the second date shown in Table 1.

4. Processing the Data

The data for the stations set out in Table 1 were downloaded as monthly totals for all years of available data. Numerous blank entries were labelled "null" to indicate missing data. While there are many acceptable methods for synthesizing these missing values, particularly where daily data are available, none was used here. The 'null' entry simply meant that the month was ignored in computing the various statistics within the inbuilt capacity of Excel™.

To provide some understanding of the error that may arise from ignoring 'null' entries, for each of the stations the number of such entries was counted and calculated as a percentage of the total number of monthly values. These percentages ranged from 0.4% to 4.7%, for an average of 2.0%. At these low levels, the likely impact upon the ranking process is considered acceptable for the purpose of constructing a water balance.

For this paper, the monthly rainfall records were downloaded into an Excel™ spreadsheet and the monthly statistics for median, mean, 60th, 70th and 90th percentiles were performed; a simple task. The summation of the monthly data was then compared with the same statistic for the historical annual totals and a comparative rank of summed total to actual total derived. It would be inaccurate to assume that the summation of any monthly statistic, other than mean, would sum to the same statistic for the annual ranking.

For each of the station's monthly data the following statistical sets were compiled:

- Median – 50th percentile for each month, and annual median value.
- Mean – average for each month, and annual mean value.
- 60th percentile – the value met in 60% for each month and the annual total rainfall.
- 70th percentile – the value met in 70% for each month and the annual total rainfall.
- 90th percentile – the value met in 90% for each month and the annual total rainfall.

While these statistical values may seem important for risk assessment of failure for effluent disposal to land, by either irrigation or drainage, the monthly models actually only use the relative statistic for each month without any reference to its association with the actual recurrence interval for the annual value. Since the summation of the monthly values is where the water balance utilises each monthly value, it is critical that the chosen statistic is realistic in terms of actual annual recurrence interval, a gauge of the overall projected performance of the system.

Addressing the problem associated with the chosen monthly ranking, and its summation to an actual annual ranking requires that before one can choose an appropriate set of monthly statistics, the likely reflection of those monthly values needs to be kept in perspective with historical data. Does the sum of median monthly values actually total the median annual rainfall? If not, then the wrong monthly values have been chosen.

Each monthly statistic is summed to an annual total and that total is ranked to the actual annual rainfall records. The comparisons for each of these five statistics are set out in Table 2. Is the summation of the 75% monthly value close to the annual 75th percentile rainfall? On the basis of closeness of the chosen value to the actual annual statistic will qualify the suitability of that statistic to the task at hand; otherwise choose a different percentile rank.

5. Statistical Summaries

Following the set of statistical summaries suggested in Section 4, Table 2 shows the values derived for each of the 16 rainfall recording station, and makes a comparison between the summed monthly values and the actual annual values for the same statistic. The percent value of the summed is related to the percentile rank of the actual annual figures of the same rank. The percentage values shown with the summed 90th percentile values are comparisons with the wettest year on record.

Table 2. Statistics of the summation of monthly values compared to percentile values

Station/I.D.	**Median**		**Mean**		**60th percentile**		**70th percentile**		**90th percentile**	
	summed	actual	summed	actual	summed	actual	Summed	actual	summed	actual
Ainslie (ACT) 70000	545/30%	665	641/45%	644	659/49%	691	874/80%	741	1294/121%	856
Armidale 56002/56037	691/30%	772	788/56%	791	826/63%	814	951/84%	861	1494/100%	1008
Bega 69002	527/11%	859	858/57%	859	703/36%	886	951/70%	959	2086/114%	1330
Byron 58007	1538/24%	1851	1872/52%	1874	1847/50%	1966	2318/85%	2130	3738/129%	2408
Camden 68007	537/22%	702	747/55%	743	701/50%	786	884/72%	867	1682/103%	1076
Pokolbin 61056	579/18%	751	749/49%	757	705/43%	816	889/80%	846	1579/110%	1009
Coffs Harbour 59010+	1292/21%	1576	1654/57%	1648	1606/53%	1675	2019/83%	1821	3416/101%	2253
Gosford 61023+	1013/21%	1269	1304/55%	1308	1268/50%	1335	1527/78%	1433	2908/124%	1794
Ulmarra 58059	768/17%	982	1030/56%	1032	968/48%	1074	1236/79%	1163	2167/113%	1437
Hillston AP 75032	267/19%	359	370/54%	369	356/50%	384	465/77%	422	845/103%	529
Manildra 65022	552/28%	664	665/50%	668	679/55%	710	816/81%	748	1335/102%	891
Narrabri 54120	490/23%	636	644/52%	646	638/50%	708	800/79%	751	1451/111%	875
Oberon 63063	721/30%	823	841/57%	841	883/60%	888	1089/86%	959	1607/110%	1118
Port Macquarie 60026+	976/19%	1211	1314/60%	1317	1265/54%	1319	1603/75%	1484	2812/96%	1851
Robertson 68054	1180/18%	1625	1660/51%	1674	1483/42%	1746	1854/72%	1840	3846/122%	2417
Wagga Wagga 73127	440/32%	498	524/57%	524	533/58%	543	642/80%	594	1063/110%	752

+ two stations combined – Refer to Table 1. Rainfall in millimetres Data sourced from Bureau of Meteorology's website

6. Interpretation of Data

Table 2 presents the statistics that may be used in constructing a water balance based upon monthly rainfall for particular localities. The stations selected are simply representative of various regions in NSW. It is expected that each designer/regulator will have data for the localities in which they work and may have compiled other statistical summaries to assist in understanding the random nature of rainfall, but also historical averages of those events.

While the NSW Guidelines (DLG *et al.*, 1998) and the DEC (2004), work on median values, it can be seen that for all the selected localities, the sum of the median monthly values varies between 11% and 32% of the actual annual median value derived from historical data, meaning that in all cases these values can be expected to be exceeded 89%-68% of the years. Few planners would consider a pass mark of less than 30% any encouragement of success.

The selection of the median values represents a significant risk, yet deemed by the regulators to be acceptable for designing domestic on-site effluent management systems. Did the authors of these two texts understand the significance of their requirements?

The average values for the summed monthly values are very close to the historical average, simply because the arithmetical calculation leads to the same point – allowing for rounding errors in the original data.

Let us assume that an acceptable failure rate is that the designed irrigation area, or trench system may fail once every four years, rather than three failures every four years by choosing the median rainfall. From Table 2 we seek out the rainfall statistic that will result in a probability of 75%, that is, a 75 % success rate.

So what of a higher statistical success, say 90%? In Table 2 we need to scan the two columns in 90th percentile rank. When we observe the values for Ainslie (ACT), the summation of the monthly 90th percentile rainfall is 1294 mm, 124% higher than the statistical annual 90th percentile. The choice of this statistic would simply confirm the designer does not understand basis statistics. Overdesigning any system leads to other problems, not the least the additional cost of installing and maintaining the system.

7. Conclusion

While the regulators appears to be content with the choice of median monthly rainfall values being used for domestic wastewater systems, a simple examination of a range of statistics for locations in NSW shows otherwise. The median values result the likely under-design of the land application area, in the order of 11% to 30% of the actual annual median rainfall. Such inaccuracy in modelling will result in a possible failure three years in every four years; totally unacceptable, yet encouraged by government guidelines.

The obvious solution is to target a higher performance by choosing a monthly statistic that produces an annual value close to about 70% probability of occurrence. While the inputs to the model are historical values that may never reflect future rainfall monthly totals, at least better than a 50% chance is achievable.

A more details appraisal of rainfall statistics is provided at www.lanfaxlabs.com.au/rainfall_statistics.htm

8. References

Bureau of Meteorology Climate Data Online access from http://www.bom.gov.au/climate/data/index.shtml

DEC (2004) *Environmental Guidelines: Use of Effluent by Irrigation* . Dept. Environment and Conservation, Sydney

DLG *et al.,* (1998) *NSW Environment and Health Protection Guidelines: On-site sewage management for single households* NSW Dept Local Govt, NSW Environ. Protection Authority, NSW Health, NSW Land & Water Conservation and Dept Urban Affairs & Planning.

Lanfax Laboratories, accessed from www.lanfaxlabs.com.au/rainfall_statistics.htm

Patterson, R.A. (1994) *On-site treatment and disposal of septic tank effluent.* Ph.D. thesis. University of New England, Armidale.

Standards Australia & Standards New Zealand (2012) AS/NZS 1547:2012 *On-site Domestic Wastewater Management*. Standards Australia, Lidcombe and Standards New Zealand, Wellington