

Supplementary information

Assisting global rainwater harvesting practitioners: a decision support tool for tank sizing method selection under uncertainty

D. Vargas¹, I. Dominguez¹, S. Ward^{2*}, E.R. Oviedo-Ocaña¹

1) Escuela de Ingeniería Civil

Facultad de Ingenierías Físico-mecánicas

Universidad Industrial de Santander

Carrera 27 Calle 9, Bucaramanga, Colombia

2) Centre for Water, Communities and Resilience

Department for Geography and Environmental Management

Faculty of Environment and Technology

University of the West of England, Bristol

Frenchay Campus, Coldharbour Lane, Bristol, BS16 1QY

* Corresponding author: sarah10.ward@uwe.ac.uk

Explanation of sizing methods for the rainwater storage tank

1. Simplified methods

In this category, four methods suggested by institutions were applied: a) Portuguese Association for Quality and Efficiency in Building Services (ANQIP), b) German Institute for Standardization (DIN), c) Environmental Agency (EA), and d) International Water and Sanitation Centre (IRC). These methods are defined by relationships between easily measured variables. Table 1 includes a description of the methods.

Values adopted during storage volume simulation, using simplified methods were taken as suggested by the respective guides (ANQIP, 2015; DIN 2002; EA, 2010, IRC, 2007) and consistent with the definitions for the study sites.

Table 1. Simplified methods used for storage tank sizing

Method	Description	Equation	Variables
ANQIP	It is used for sizing tanks where consumption is nearly uniform and residence times range between 20 and 30 days, except when the desired end-use is irrigation (90 days).	$V = \text{Min} \{V_1 \text{ o } V_2\}$ $V_1 = (0.0015 * N) * P * A$ $V_1 = (Z_1) * P * A$ $V_2 = (0.003 * N) * U * C_{AE}$ $V_2 = (Z_2) * U * C_{AE}$	V: storage tank volume (L) V ₁ : annual rainwater yield (L) V ₂ : annual required volume (L) N: maximum number of retention days (i.e. 20 to 30 days) A: catchment area (m ²) U: inhabitants per household P: mean multi-annual rainfall (mm) C _{AE} : annual percapita demand (L) η: hydraulic coefficient of filter efficiency C: runoff coefficient k: factor that depends on the climatic zone and the design objective (was set at 8 days) Z _i : empirical constant characteristic of each method
DIN	It is used for sizing tanks from small-scale systems.	$V = \text{Min} \{V_1 \text{ o } V_2\}$ $V_1 = (C * \eta * 0.06) * P * A$ $V_1 = (Z_3) * P * A$	
EA ¹	Like DIN (2002), with an exception on the calculation percentage.	$V_2 = (0.06) * U * C_{AE}$ $V_2 = (Z_4) * U * C_{AE}$	
IRC	It is designed to store 80% of the rainwater captured and assumed that the user adopts an adaptive demand strategy (i.e. uses one third more water per day when the tank is close to overflow and uses one third less per day when the tank is almost empty).	$V = \left(C * k * \eta * \frac{1}{365} \right) * P * A$ $V = (Z_7) * P * A$	

Note: ¹ in the case of EA, only the coefficient changes (i.e. 0.05 by 0.06) and the empirical constant Z_i are presented.

2. Continuous simulation of mass balance methods

In this category, five methods were tested: Yield After Storage (YAS), Yield Before Storage (YBS), theta (θ), Ghisi (via the computer model Neptune), and Rippl. In these methods, changes in water storage in a previously fixed tank volume were selected through mass balance equations (Liaw & Tsai, 2004). Except for the Rippl method, for all continuous simulations, system performance was assessed through efficiency for different tank sizes. All the models addressed correspond to continuous simulations performed on a daily scale, differing in the mass balance equation that describes the behaviour of the amount of water stored, which for the initial interval was assumed to be zero (i.e. simulations assumed tank starts empty).

a) YAS, YBS and θ :

Three water balance equations known as YAS, YBS and a general form for the parameter θ were applied (Fewkes, 1999b). The YAS equation consists of removing the demand after the rain has been added to the tank and the spill has occurred, while the YBS equation assumes that the demand is removed before the spill occurs. The last simulation algorithm was a generalized equation for the variable θ , which can take values between zero (i.e. YAS) and one (i.e. YBS). This variable was defined with a value of 0.5 to represent an intermediate method between YAS and YBS. Table 2 presents the mass balance equations for storage tank sizing.

Table 2. Methods of continuous simulation of mass balance used for storage tank sizing – YAS, YBS and θ methods

Method	Equation	Variables
YAS	$Q_t = P_t * A * C$ $Y_t = \text{Min} \left\{ \begin{array}{l} D_v \\ V_{t-1} \end{array} \right\}$ $V_t = \text{Min} \left\{ \begin{array}{l} V_{t-1} + Q_t - Y_v \\ S - Y_v \end{array} \right\}$	t: time interval (day) P _t : rainfall in the time interval t (mm) A: catchment area (m ²) C: runoff coefficient Q _t : water collected in the time interval t (L) V _t : storage volume at the end of the time interval t (L) Y _t : rainwater delivered by the tank during the time interval t (L) D _t : rainwater demand for the time interval t (L)
YBS	$Q_t = P_t * A * C$ $Y_t = \text{Min} \left\{ \begin{array}{l} D_v \\ V_{t-1} + Q_v \end{array} \right\}$ $V_t = \text{Min} \left\{ \begin{array}{l} V_{t-1} + Q_t - Y_v \\ S \end{array} \right\}$	
θ	$Q_t = P_t * A * C$	

	$Y_t = \text{Min} \left\{ \begin{array}{l} D_{t'} \\ V_{t-1} + \theta * Q_{t'} \end{array} \right\}$ $V_t = \text{Min} \left\{ \begin{array}{l} (V_{t-1} + Q_t - \theta * Y_t) - (1 - \theta) * Y_{t'} \\ S - (1 - \theta) * Y_{t'} \end{array} \right\}$	S: tank volume (L) V _{t-1} : storage volume at the end of the previous time interval t-1 (L) θ: parameter whose value ranges from zero to one E: efficiency in potable water savings (%)
Efficiency	$E(\%) = 100 * \frac{\sum Y_t}{\sum D_t}$	

b) Neptune

Ghisi et al. (2007) proposed the computer program Neptune to estimate the potential for potable water savings in four states of the Southeast region of Brazil. Table 3 describes the water balance equations to assess tank level behaviour over time (Ghisi & Marcel, 2014).

Table 3. Methods of continuous simulation of mass balance used for storage tank sizing – Neptune method.

Method	Equation	Variables
Neptune	$V_{ac}^t = P_t * A * C$ $V_{inf\ in}^t = \text{Min} \left\{ \begin{array}{l} V_{inf} \\ V_{inf\ fim}^{t-1} + V_{ac}^t \end{array} \right\}$ $V_c^t = \text{Min} \left\{ \begin{array}{l} D_{pluv}^t \\ V_{inf\ in}^t \end{array} \right\}$ $V_{inf\ fim}^t = \text{Min} \left\{ \begin{array}{l} V_{inf\ in}^t - V_c^t \\ V_{inf} - V_c^t \end{array} \right\}$	t: time interval (day) V _{ac} : volume of water that runs through the catchment area at a given time t (L) P _t : rainfall in the time interval t (mm) A: catchment area (m ²) C: runoff coefficient V _{inf in} ^t : volume available in the tank at the beginning of the time interval t (L) V _{inf} : storage tank volume (L) V _{inf fim} ^{t-1} : volume available at the end of the previous time interval t-1 (L) V _c ^t : volume of rainwater consumed in the time interval t (L) D _{pluv} ^t : rainwater demand for the time interval t (L) V _{inf fim} ^t : volume available at the end of the time interval t (L) E: efficiency in potable water savings (%)
Efficiency	$E(\%) = 100 * \frac{\sum V_c^t}{\sum D_{tot}^t}$	

c) Rippl method

The Rippl algorithm calculates the volume of storage necessary to regulate the tank outputs allowing a constant supply to the system. The method calculates the water surplus required to be stored during the rainy season to compensate for the shortage in dry season (Tomaz, 2005; Quadros, 2010). Tank volume was the maximum cumulative (positive) difference

between rainwater needed and collected (Matos et al., 2013). Table 4 summarizes the analytical approach used by Sanches et al. (2015) for this method.

Table 4. Methods of continuous simulation of mass balance used for storage tank sizing - Rippl method.

Method	Equation	Variables
Rippl	$V_t = C * P_t * A$ $\left\{ \begin{array}{l} S_t = 0, C_t - V_t \leq 0 \\ S_t = C_t - V_t, C_t - V_t > 0 \end{array} \right. \quad t = 1$ $\left\{ \begin{array}{l} S_t = 0, S_{t-1} + C_t - V_t \leq 0 \\ S_t = S_{t-1} + C_t - V_t, S_{t-1} + C_t - V_t > 0 \end{array} \right. \quad t > 1$ $V_r = \text{Max}\{S_t\}, \quad 1 \leq t \leq n$	<p>t: time interval (day)</p> <p>V_t: volume of rainwater collected (L)</p> <p>A: catchment area (m²)</p> <p>P_t: rainfall in the time interval (mm)</p> <p>C: runoff coefficient</p> <p>S_t: stored volume during the time interval t (L)</p> <p>S_{t-1}: stored volume for the previous time interval t-1 (L)</p> <p>C_t: rainwater consumption proposed for the time interval t (L)</p> <p>n: total number of intervals considered</p> <p>V_r: volume of response, associated with the size of the tank that should be installed (L)</p>

3. Cost functions methods

One method that includes information on system costs was applied. The method is based on maximizing the return on investment on an annual scale from savings obtained by stopping using potable water from the conventional system (López-Patiño et al., 2011). A ratio of unamortized investment costs (asset value not taking into account depreciation) was built according to tank size and it was contrasted against a curve associated to financial savings due to stopping using water from the public mains water service. Table 5 describes two cost functions required for applying the method.

The amortized investment cost function was prepared considering the investment costs required for building a system for a defined tank size. To describe the function, we used a regression model that fits the investment-volume behaviour.

The drinking water saving curve was generated using the continuous simulation models YAS, YBS, θ and Neptune, because for each volume they provide the value of the associated efficiency. The efficiencies found by the continuous simulations were multiplied

by the annual demand, according to the proposed uses, to obtain the amount of water supplied by the RWH system per year. Subsequently, using the unit price of drinking water set by the utility, money savings due to stop buying water from the utility were estimated.

Table 5. Method with costs functions used for storage tank sizing

Function	Equation	Variables
Function of amortized costs	$CI(V) = \alpha * C(V) \rightarrow$ $CI(V) = \left(\frac{(1+r)^N * r}{(1+r)^N - 1} \right) * C(V)$	α : amortization ratio $C(V)$: investment costs function without amortization (US\$) r : real interest rate (%) N : amortization period (years)
Function of potable water savings costs	$WS(V) = P_w * D * E(V)$	V : tank volume (L) P_w : unit price of water that includes supply, sanitation and discharge (US\$/L) D : annual demand of the sector (L) $E(V)$: efficiency as a function of volume

The cost-function method required to determine costs of the system elements. For this, market prices of storage tanks from a local company (Homecenter-Sodimac, Bucaramanga) were collected. With this information, an adjustment curve that describes the investment in the tank according to its volume was prepared (See Figure 3 in the main paper). Additional considerations were:

- Drinking water unit price: 1.21 (US\$ / m3) (local bills) (1 US\$ = 2,859.17 COP)
- Real interest rate: 2% (López-Patiño et al., 2011)
- Amortization period (lifetime): 15 years (López-Patiño et al., 2011)

4. Statistical methods

This category comprised two methods: a) Nonparametric stochastic rainfall, and b) Probabilistic model. Both methods used the continuous mass balance equations YAS, YBS, θ and Neptune, since they allow assessing the performance for different tank sizes, but the methods enable a rainfall input provided by statistical methods.

a) Method based on a non-parametric stochastic rainfall generator:

The proposal by Basinger et al. (2010) was applied. This proposal described an algorithm for the generation of a stochastic rainfall based on nonparametric techniques. This method uses probabilities to describe rainfall occurrence. However, these are derived directly from local observations, which allowed generating a portable model that works for any historical precipitation period. The tool "Estimation of Storage and Reliability" (in its acronyms SARET) was used, which consists of a series of steps programmed in a spreadsheet in Visual Basic for Applications (VBA) (Table 6).

Table 6. Steps used by SARET for tank sizing from stochastic rainfall.

Paso	Description
I	Daily rainfall records for several years are required. Leap days should not be included.
II	Data should be sorted in a table in such a way that rows labels correspond to the days of a year, and columns labels for years. Mobile windows must be used. These windows are arrays that contain the information of the 15 previous records and the 14 after the target day through the years (i.e. if there are 25 years available, then there will be 365 mobile windows of 30 days x 25 years size). To consider the mobile windows of the initial and final days of the year, the information of the adjacent years must be used (years of the extremes are considered contiguous).
III	Each daily record must be classified as wet (W) or dry (D), wet when rainfall was greater than zero, and dry when rainfall did not occur.
IV	For each mobile window, the number of events must be counted as follows: W (wet day), D (dry day), WW (wet day after a wet day), WD (dry day after a wet day), DD (dry day after a dry day) and DW (wet day after a dry day).
V	The first-order Markov chain model is completed for each of the 365 days. This is achieved by calculating the following probabilities: $P_i(W) = \frac{\text{Number of days } W}{\text{Number of days } W + \text{Number of days } D}; P_i(D) = 1 - P_i(W);$ $P_i(WW) = \frac{\text{Number of days } WW}{\text{Number of days } WW + \text{Number of days } WD}; P_i(WD) = 1 - P_i(WW);$ $P_i(DD) = \frac{\text{Number of days } DD}{\text{Number of days } DD + \text{Number of days } DW} \text{ and } P_i(DW) = 1 - P_i(DD).$

VI	<p>The creation of stochastic precipitation begins by assigning a dry day (D) or a wet day (W) for each of the days (i.e. for 25 years it would be 9125 days). The following criteria must be followed:</p> $iB > P_1(W)? \begin{cases} YES \rightarrow Day_1 = D \\ NO \rightarrow Day_1 = W \end{cases} \leftarrow \text{Format for the first day}$ $iDay_{i-1} = W? \begin{cases} YES \rightarrow iB > P_i(WW)? \begin{cases} YES \rightarrow Day_i = D \\ NO \rightarrow Day_i = W \end{cases} \\ NO \rightarrow iB > P_i(DD)? \begin{cases} YES \rightarrow Day_i = W \\ NO \rightarrow Day_i = D \end{cases} \end{cases} \leftarrow \text{Format for the rest of days}$ <p style="text-align: center;">Where B corresponds to a random number between 0 and 1.</p>
VII	<p>If a day was defined as a dry day, zero precipitation will be assigned, but if a day was defined as a wet day, the k-th highest precipitation value of the data set should be chosen from the mobile window associated with the target day; that is to say, one of the 750 precipitations (i.e. for windows of 25 years) that actually happened is chosen. K-th value corresponds to an approximation to the lower integer of the calculated value as follows</p> $k - th = (Random\ between\ 0\ and\ 1) * (Amount\ of\ W\ in\ the\ mobile\ window) + 1$
VIII	<p>At least 100 stochastic simulations are applied under any mass balance method. From the 100 tank sizes, it is recommended to choose the largest volume conservatively</p>

b) Method based on a probabilistic model:

Probabilistic relationships between tank capacity and supply deficit rates described by Su et al. (2009) were applied. For this, a procedure that consists of two steps was followed: a simulation model of mass balance and a probabilistic modelling.

In the continuous simulation model, water balance equations YAS, YBS, θ and Neptune were used, since they allowed to calculate annual deficits for different volumes. Deficit in the supply is described as a missing percentage of the efficiency to reach a total fulfilment of the demand (i.e. if for a tank volume efficiency is 60%, deficit corresponds to 40%). As a result, two samples sizes were obtained from recorded years, that is 25 and 15 for A and B respectively.

From the choice of a probability distribution that fits the data, the behaviour of the samples associated with each tank size was described. The concept of exceedance probability or return period was applied to study the critical events in which the deficit rates are exceeded. For this, the probability distributions were integrated to generate a graph that describes the deficits based on the tank size for a certain return period (Figure 1). The probabilistic behaviour of the random variables was described by using the normal distribution, as proposed by Su et al. (2009).

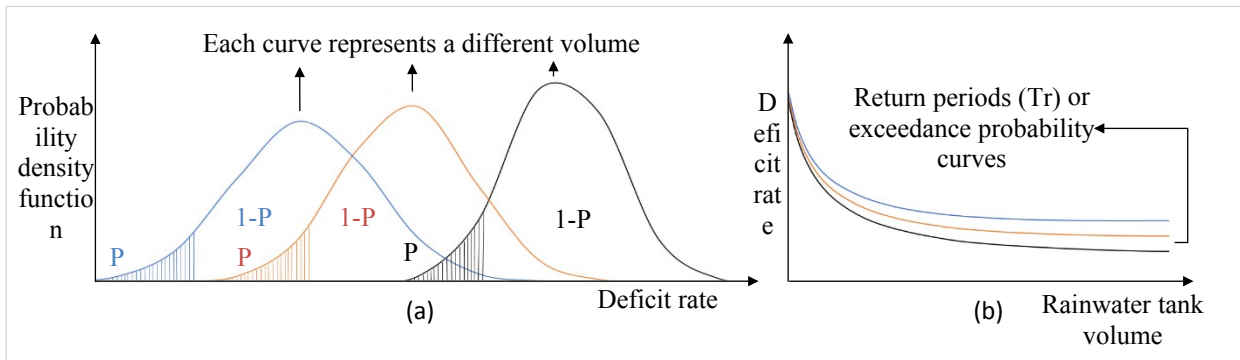


Figure 1. Probabilistic method. (a) Probability density functions. (b) Deficit relations and storage volume according to exceedance probability.