Simulation

Simulation: Transactions of the Society for Modeling and Simulation International
89(6) 693–702

DOI: 10.1177/0037549713479037
sim.sagepub.com
SAGE SCS

Rainwater harvesting system using a non-parametric stochastic rainfall generator

Lawrence V Fulton[1,3], Nathaniel D Bastian[2,3], Francis A Mendez[1,3] and R Muzaffer Musal[1,3]

Abstract

With water becoming an even scarcer resource, rainwater harvesting (RWH) systems are becoming increasingly more commonplace as mechanisms to capture and store rainwater for both agricultural and domestic use. Three important engineering considerations associated with the construction of RWH systems are the capture surface area, the tank volume required for specific demand levels, and the number of expected occupants. The purpose of this work is to evaluate the engineering design of a RWH system in a semi-arid Texas region using a non-parametric stochastic rainfall generator based on 64 years of data and to provide engineering charts and equations for future use. We model the RWH system using simulation techniques in order to estimate requirements for building a system capable of providing a family with 100% of its water requirements with demand never exceeding available supply (100% demand satisfaction).

Keywords

rainwater harvesting, stochastic modeling, systems simulation, sustainability development

1. Introduction

1.1 Background

Water is a scarce resource, and in many areas, it is becoming even more so.[1] A solution that is becoming more commonplace to address this scarcity is rainwater harvesting (RWH). RWH includes the capture and storage of rainwater for agricultural and possibly domestic use.[2] The practice dates back at least as far as 4000 years ago in Jordan, where surface runoff was harvested.[3] The Qing and Han dynasties used RWH over 2000 years ago,[4] and the use of RWH by the ancient Greeks,[5] Romans,[6] Indians,[7] and other cultures is well documented. RWH is both feasible and increasingly under consideration by many countries.[8–12]

The quality and quantity of water from RWH have both been investigated,[13–17] and operational design characteristics proposed.[12] RWH design for Six Sigma methods have been discussed.[18] Typical systems modeling considerations include water demand, catchment area capacity, forecasted rainfall, cistern capacity, and system efficiency. Sizing of a RWH system is based on these considerations, and the uncertainty associated specifically with supply has resulted in many classes of tools, such as catchment versus storage curves,[19] sequential peak analysis,[20] and stochastic precipitation generators.[21] Of particular importance is the work of Basinger et al.,[10] who provide a reasonable, non-parametric, stochastic rainfall methodology given historical rainfall data. The advantage of the non-parametric approach is that parametric distributional assumptions need not apply, as they often do not hold.[22] Basinger et al.'s model provides for seasonality in a range around the day of the year but does not investigate trends over the 25-year time horizon required and assumes that one time lag is sufficient for modeling.

1.2 Study purpose

The purpose of this paper is to evaluate the design of a RWH system in a semi-arid Texas region using a

[1]Department of Computer Information Systems & Quantitative Methods, McCoy College of Business Administration, Texas State University, USA
[2]Department of Mathematics and Statistics, The Undergraduate School, University of Maryland University College, USA
[3]System Modeling and Renewable Technology (SMART) Laboratory, Department of Industrial Engineering, Texas State University, USA

Corresponding author:
Nathaniel D Bastian, MSc, Department of Mathematics and Statistics, The Undergraduate School, University of Maryland University College, Adelphi, MD 20783, USA. Current address: 94–109 Wali Pl Apt F201, Waipahu, HI 96797, USA
Email: nathaniel.bastian@fulbrightmail.org

stochastic rainfall generator based on historical data. The stochastic rainfall generator is an implementation of the non-parametric framework designed by Basinger et al.,[10] while the simulation (implemented in commercial simulation software) loosely follows the flowchart in Mun and Han.[12] Design of experiments (DOE) was conducted to manipulate cistern size (based on procurement availability constraints), capture area (roof size), and expected number of occupants. These three engineering considerations were important considering that the analysis supported the construction of a RWH system providing 100% of the primary author's domestic water. In addition, the results of the DOE provided engineering charts and equations for use in future RWH projects.

In terms of design features, the key distribution of interest was the cistern's contents (in gallons), with an emphasis on the distribution of the minimum (an order statistic). The owners did not want to purchase water, and RWH providers need information regarding minimum requirements for sustaining various family sizes. The measure of interest is termed ''demand satisfaction,'' meaning that daily supply at least meets daily demand, and the goal is 100%.

The quality of the water was to be improved by use of first flush, removing the first 100–200 gallons from the system, and subsequently processing the rainwater through a series of gross filters (sock and basket types), a 50-micron filter, and a 5-micron carbon filter. After filtration, the water was exposed to ultraviolet light for purification. The filter and ultraviolet array are shown in Figure 1, while the design mock-up for the RWH system is shown in Figure 2.

The importance of this study is multi-fold. This study represents the first of its type to use both a rainfall generator and a simulation in a prospective rather than a retrospective capacity, and it is the first of its type to be implemented for Central Texas, a semi-arid region. The study provides decision support (through DOE) for determining minimum RWH engineering design characteristics (roof size and cistern size) necessary to ensure that demand never exceeds supply given the number of occupants in any area provided daily, historical rainfall stream. The study provides charts and equations for use by civil engineers to estimate minimum requirements for future construction. It extends the data capture of the non-parametric rainfall generator proffered by Basinger et al.[10] from 25 to 64 years. This research also demonstrates a real application of simulation for Central Texas, a region that has limited and ever scarcer water resources, so much so that rationing is now commonplace.

2. Materials and methods

2.1 Study design

The design of the simulation for this RWH system is based on balance equations, where captured supply of rainwater from the roof (C), demand for purified water (D), and overflow (O) are the primary considerations for the volume (V) in the cistern. On any given day indexed with $t \in T$, the volume in the cistern is expressed as shown in Equation (1). Note that t is a daily index rather than monthly or yearly. Equation (2) illustrates a minimum/maximum approach for modeling tank capacity (Cap), which ignores the amount of overflow:

$$V_t = Max\{0, V_{t-1} + C_t - D_t - O_t\} \tag{1}$$

$$V_t = Min\{Max\{0, V_{t-1} + C_t - D_t\}, Cap\} \tag{2}$$

Units of measure in this study are provided in feet and gallons due to the construction and supply considerations of building a RWH system in the United States. Volume is

Figure 1. The filter and ultraviolet (UV) system for the rainwater harvesting system is shown. From right to left: 50-micron filter, 5-micron carbon filter, UV light for bio-contamination purification.

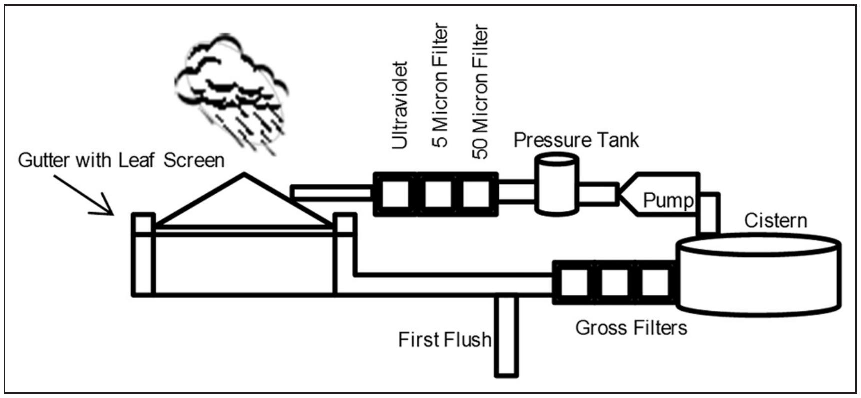

Figure 2. Typical design of a rainwater harvesting system.

measured in gallons, and capture area is measured in square feet. Historical rainfall data (in inches) derive from the National Oceanic and Atmospheric Administration (NOAA).[23] Daily rainfall data for the selected observation station are freely available from September 1946 to July 2011, but to have complete years of data (important for the stochastic rainfall generator), only data from July 1947 to July 2011 were included (23,407 observations over 64 years). The observation station used in this study is the San Antonio, Texas, International Airport. This station is in close proximity to the location of the RWH site.

2.2 Study setting

San Antonio is in the heart of Texas and relies on water from the Edwards Aquifer as its primary source. The 30-year ''normal'' rainfall is 32.27 inches with an all-time minimum of 10.17 inches (1917) and maximum of 52.28 inches (1973) according to NOAA statistics.[24] The median annual rainfall is 28.53 inches. From 2005 to 2011, San Antonio experienced three instances where rainfall was below 20 inches and one instance just above it. Due to fast population growth, drought, and other factors, rationing occurs rather routinely. The Edwards Aquifer Authority provides groundwater stewardship for the Aquifer,[25] and with an increasing population, water will become even more scarce. Scarcity has been such a concern that the construction of desalinization plants has been discussed.[26] Feasible alternative sources of water need to be investigated and documented, and design characteristics associated with these sources need to be specified.

2.3 Stochastic rainfall generation

Building the rainfall generator required first estimating whether a simulation day would have any measurable rainfall (defined as a ''wet'' day) based on a conditional probability distribution generated for that specific day and derived from historical rain data conditioned on whether the previous day experienced rainfall. (The exact method for generating this distribution is discussed shortly.) If and only if a simulation day was estimated to experience rain (a ''wet'' day), then the expected amount of rain was then sampled from a second probability distribution based on historical data.

To generate a rainfall stream, the procedure of Basinger et al.[10] was modified as follows. Firstly, the marginal probabilities of ''wet'' and ''dry'' for each individual day were calculated based on 64 years of data and a window of 15 days preceding through 14 days following the target day. For example, the marginal probability for a ''wet'' day on 16 January would be the sum of all the wet days from 1 January to 30 January for the 30-year window divided by the sum of all days in this same time period. This mechanism inherently accounts for seasonality of rainfall (albeit not trend). Secondly, conditional probabilities for each pair-wise combination of days (day t and day $t-1$) were calculated (although the Basinger et al. article depicts them as joint probabilities). Equations (3)–(8) illustrate the stochastic rainfall generator probability distributions, with the index t indicating that each day {1, 2, 3,…, 365} has its own distribution:

$$P(W_t) = \frac{W_t}{W_t + D_t} \tag{3}$$

$$P(D_t) = 1 - W_t \tag{4}$$

$$P(W_t|W_{t-1}) = \frac{W_t W_{t-1}}{W_t W_{t-1} + D_t W_{t-1}} \tag{5}$$

$$P(W_t|D_{t-1}) = \frac{W_t D_{t-1}}{D_t D_{t-1} + W_t D_{t-1}} \tag{6}$$

$$P(D_t|W_{t-1}) = \frac{D_t W_{t-1}}{D_t W_{t-1} + W_t W_{t-1}} \tag{7}$$

$$P(D_t|D_{t-1}) = \frac{D_t D_{t-1}}{D_t D_{t-1} + W_t D_{t-1}} \tag{8}$$

After determining the outcome for a specific day (marginal if day 1, conditional otherwise), then the expected value for the rain is calculated based on the distribution of rain over the same, centered, rolling 30-day period for each day. The amount of rain (S_t) is then calculated based on a Bernoulli trial for rain with the probability of success equal to either Equation (3), (5), or (6) (depending on which applies) multiplied by the distribution of rain for that day (see Equation (9)):

$$S_t = XY, X \sim B(P(W_t | \ldots), Y \sim Empirical\,Distribution \quad (9)$$

To implement this generator in our simulation, we built conditional probability tables for each day {1 January, 2 January…31 December} given the possible conditions of the previous day {wet, dry}. Each daily distribution was based upon the rain experienced within 30 days from the day of interest (centered). In other words, the conditional probability of rain for 1 January was calculated using data from mid-December to mid-January for all 64 years of observations. We then built distributions of the amount of rainfall for each of the 366 days. In doing so, we were able to determine if a particular day experienced rain (also based on a centered, 30-day probability calculation). If so, we sampled from the distribution of the amount of rain usually experienced on this day (using a rolling, centered moving average of 30 days).

2.4 Rainwater harvesting system simulation

To determine requirements for building a RWH system capable of providing a family with 100% of its water requirements with 100% ''demand satisfaction'' based on 64 years of historical data, a simulation was undertaken using stochastically generated rainfall. The simulation was designed to incorporate fully flexible parameters for all distributions, so as to be useful for others. Figure 3 illustrates the flowchart used for the development of the simulation.

The simulation begins by initializing variables, setting indices for iterations and days to one, and assigning DOE factors. Based on architectural plans, the owner estimated that the 3200-square foot house could be engineered to have between 3000 and 5000 square feet of usable roof space, depending on design, so five levels were selected for comparison: *Roof* = {3000, 3500, 4000, 4500, 5000}. Cistern capacity was available in increments of 5000 gallons, *Cap* = {15,000, 20,000, 25,000, 30,000, 35,000, 40,000}, for a total of six levels. The number of occupants was estimated to be two; however, for planning considerations, this parameter was evaluated as follows: *Num* = {1,2,3}. By manipulating this parameter, more robust engineering planning charts and equations become available for future construction. The simulation assumed that the tank was at half-capacity for start-up (specifically since water would be needed to test plumbing during construction), but this parameter was flexible.

For demand, previous research suggests that houses designed with low-flow appliances (such as the one here) experience about 40.8 gallons usage per person per day, while standard houses use 69.3 gallons per person per day.[27] Given responsible design (e.g., dual-flush toilets, low-flow showers, etc.), a conservative distribution of individual domestic water usage (Z) lies between the two values. Hence, a uniform distribution provides a reasonable distribution of individual demand: $Z_t \sim U(40.8, 69.3)$. Given daily individual demand and occupancy, the calculation for daily consumption (D) is as follows: $D_t = Z_t \times Num$.

After estimating daily demand, rainfall (S) was generated stochastically as discussed previously. For the stochastic demand process, probabilities for day one were generated by sampling from the marginal distribution for that day. For each subsequent day, the single-step Markov chain probabilities were used (Equations (3)–(8)) to calculate the probabilities for the Bernoulli trial, while the empirical distribution of rain was used to estimate the amount of rainfall (if needed).

If rainfall occurred on any given day, the amount captured was subject to an efficiency adjustment. Typical systems advertise between 75% and 90% efficiency (E) when factoring in all inefficiencies, such as flushing, overflow, leakage, etc.,[28] and so a uniform distribution around these ranges is reasonable for estimation: $E_t \sim U(.75, .90)$. After estimating efficiency, the capture rate was a simple process of multiplication: $C_t = E_t \times S_t$. Estimates for overflow, shortages, and volume in the tank were trivial calculations based on balance equations previously discussed and depicted in the flowchart. In addition, failures (demand exceeding supply on any given day on any iteration) were tracked, as the system was intended to preclude demand exceeding available supply at any given time. A single failure was unacceptable.

The time horizon for the simulation (T) was set to 30 years or 10,957 days, which corresponds to the estimated years of occupancy of the owner and a typical 30-year mortgage. An initial 30 replication of a 30-year simulation run for a house with a 5000-square foot roof and a 40,000-gallon cistern capacity initialized at 20,000 gallons of water and two occupants resulted in near-immediate balance of the cistern, an average of 38,127 gallons in the cistern, and a standard deviation of about 146.61 gallons in the cistern. To establish a 95% confidence interval (CI) about the sampling distribution of means for V with a 50-gallon margin of error (100 gallons total), an estimated 36 runs per experiment were required, as shown in Equation (10):

$$t_{\alpha,n} = \frac{(\bar{x} - \mu)\sqrt{n}}{s} \rightarrow t_{.05,36} \cong \frac{(50)\sqrt{36}}{146.61} \quad (10)$$

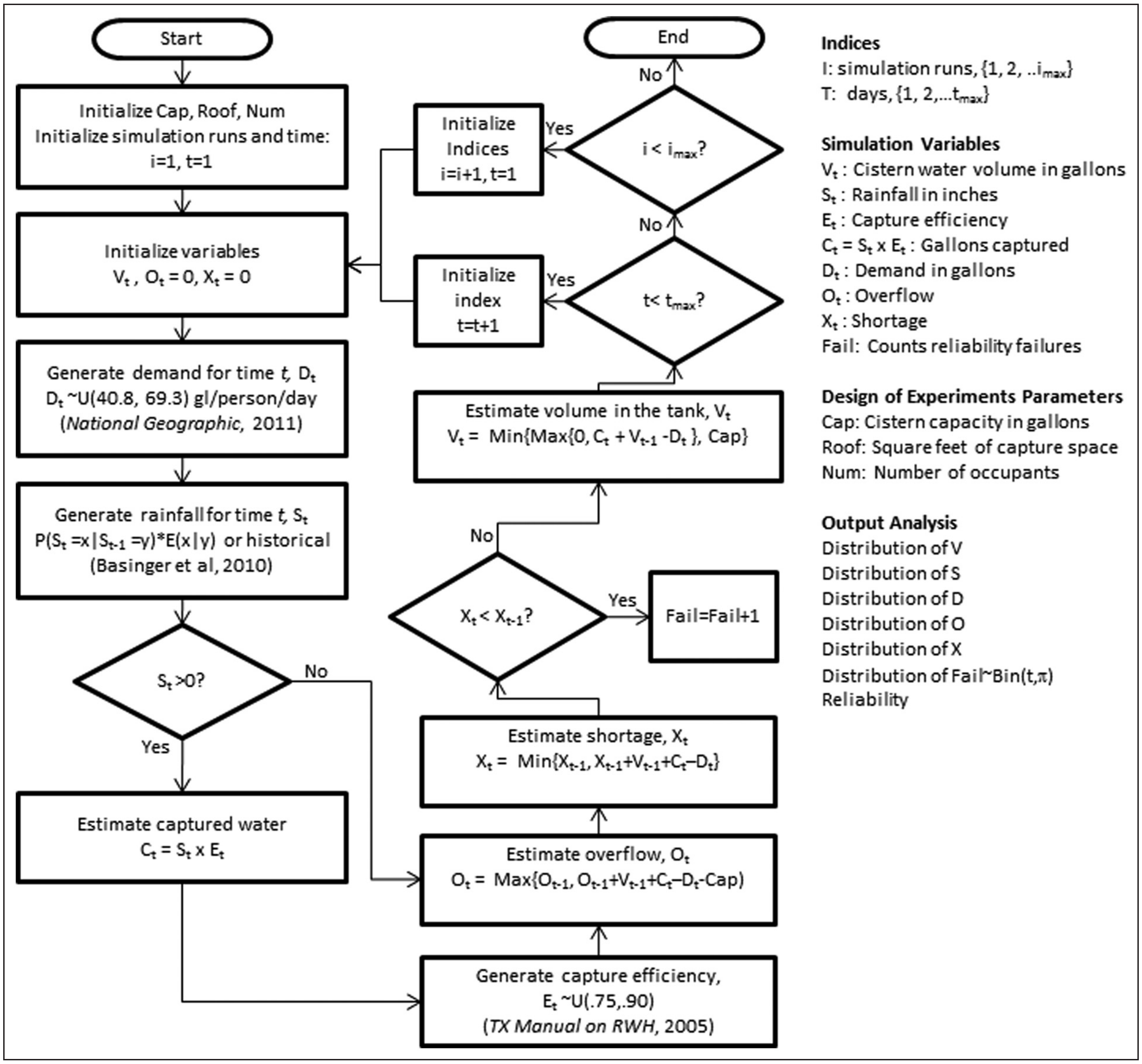

Figure 3. Rainwater harvesting system simulation flowchart.

2.5 Model implementation

The simulation was initially programmed in Microsoft Excel but was subsequently programmed using ProModel® simulation software, which uses a programming language to generate both discrete and continuous simulations.[29] In this case, the researchers invoked the ''Tank'' submodule, which provided a collection of procedures capable of tracking tank transactions (continuous rather than discrete event). One significant advantage of ProModel® is the ability to leverage and record random number streams, removing unwanted variation across experimental design conditions. Being able to see the effect of design parameters rather than just random variation is important when comparing design options. While generating random number streams in Excel is possible, it is not nearly as simple. The graphical model of the ProModel® simulation is shown in Figure 4.

The primary system locations in the model included the roof (with infinite capacity), the tank (with capacity determined via manipulation during simulation runs), the overflow unit (with infinite capacity), and the house (with infinite demand capacity). For all simulation runs, modelers assumed that the tank was primed to 50% capacity, a reasonable simplification, as plumbing systems require evaluation during the construction phase. For example, a 40,000 gallon tank would be filled at V_o with 20,000

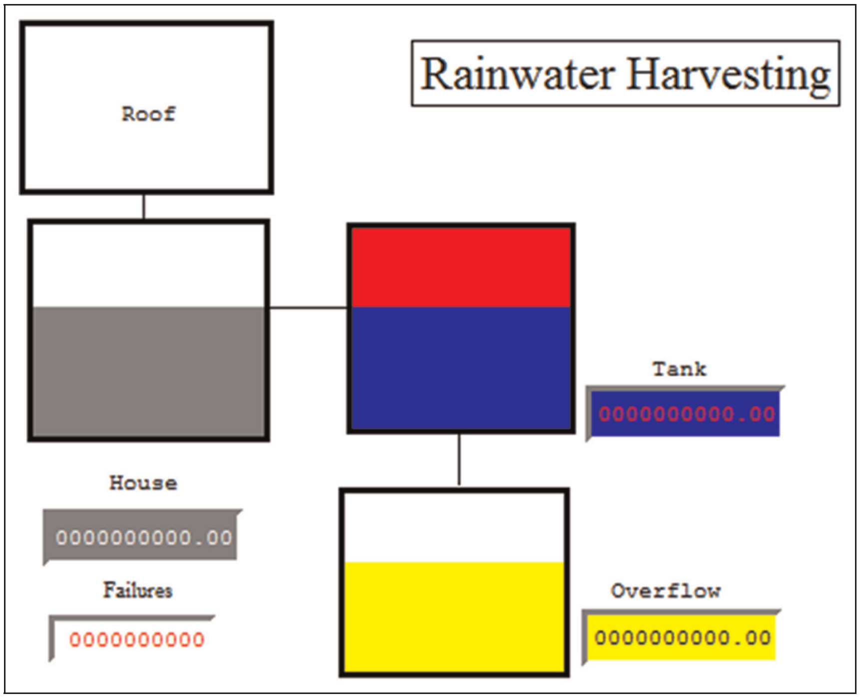

Figure 4. The graphical model for the ProModel® simulation.

gallons. The primary entity, rain, arrived on the roof daily; however, the quantity of that rain in inches was determined by the stochastic rainfall generator and was most often zero (S_t). Processing logic for the arrival of rain followed the flowchart (Figure 3).

2.6 Verification and validation

Model verification and validation (V&V) required obtaining agreement regarding the flow model and hypothesis testing of various a priori distributions with posterior results. We used both qualitative and quantitative analysis to conduct V&V.

To investigate validity, we analyzed and conducted tests of prior and posterior distributions. Firstly, we looked at annual rainfall distribution. To analyze the validity of the annual rain forecasts, we generated an estimate of annual rainfall using the generator and compared it with the actual annual rainfall for the entire 64 years of data. The generator produced a yearly mean of 30.35 inches of rain, while the region of interest actually experienced 30.36 inches over the 64-year period, an irrelevant difference.

To validate the daily forecasts produced by the rainfall generator, we calculated daily expected rainfall using the generator based on data from 1947 to 1976 only. We compared these expectations to actual rainfall data from 1977 to 2011. This ''split-half'' sampling serves to investigate the validity of using the rainfall generator for forecasting. Despite the large number of paired responses (366 including leap years), statistically significant differences between the forecast and the actual expectations were not present based on a Wilcoxon's rank sum test with continuity correction (W = 68,684, p = .551). This analysis was conducted using R Statistical Software.[30] In fact, the median absolute deviation was a nominal .033 inches, and the mean squared error was nearly zero (.004 inches). The median forecast error was nearly zero as well ($-$.005 inches). The generator performed as expected. In Section 3, we discuss how the generator performed as part of the baseline simulation.

We also analyzed the distributions for efficiency and for demand to be sure that posterior distributions matched the a priori distributions. For all runs, these distributions were distributed uniformly with appropriate means and ranges. The next section provides details on these distributions as well.

To be valid for model comparison, we needed to ensure that the random streams were identical across experimental conditions. The use of random number streams in ProModel® removed unwanted model-to-model variation, providing a mechanism for comparing the effects of the DOE factors rather than just randomness.

As part of convergent validity analysis, the ProModel® simulation was compared to a separate simulation built in Microsoft Excel. The results from the second simulation matched the ProModel® simulation closely (convergent validity). Further, four separate student teams designed simulations with similar findings. The distribution of the water in the tank, the overflow, and the shortfall were compared against the capacity to ensure that the simulation was operating as expected. Results of the simulation were compared against builder results of similar RWH system construction, and the results were congruent.

As part of the verification process, the simulation was built graphically and movement throughout the model was tracked. Visible variable counters allowed us to assess model operation, verifying that the model was behaving appropriately.

3. Results and discussion

The initial 36 simulation runs were performed on a baseline scenario using two occupants along with 5000 square feet of roof space and a 40,000-gallon cistern. Several steps were taken to validate the baseline model. The results provided an estimate for V of 37,984.59 gallons and an associated 95% CI of (37,974.89 gl, 37,994.3 gl). Mean demand (D) for the 36 runs was 110.11 gallons each month with an associated 95% CI of (110.06 gl, 110.16 gl). Given 55.05 gallons as the ''average'' use from the U(40.8 gl, 69.3 gl) demand distribution and given two occupants, one would have expected an average daily use of about 110.10 gallons. Our posterior distribution matches our expectation.

The marginal probability of rain in the simulation was .23, which is congruent with the empirical probability of .22. Mean rainfall on the roof was approximately 250.97 gl per day with a 95% CI of (247.46 gl, 254.48 gl). Actual rainfall on the 5000-square foot roof given empirical data from 1947 to 2011[22] resulted in an estimate of (.08296 inches/12 inches per foot) $\times$ 5000 square feet $\times$ 7.48 gl per cubic = 258.55 gl per day. The stochastic rainfall estimator slightly underestimates the daily capture (making it a conservative tool).

The average amount captured per day (C_t) was 207.08 gl, reflective of the average capture efficiency (82.5%). The uniform efficiency distribution operated correctly. The base scenario resulted in a system where demand never exceeded supply. Given careful analysis, the model appeared to produce numbers consistent with expectations and inputs.

For the baseline simulation run, the distribution of the minimum volume in the tank was monitored, as this distribution was of significant concern to the owner. The data from the time series for all 36 runs were collected into a spreadsheet, graphed, and a distribution was fit in order to best classify the shape. The observed minimum for all 36 runs was 12,777.81 gl. The histogram of the minimum volume distribution is shown in Figure 5.

After analyzing the baseline simulation to validate that the output appropriately represented the input, the simulation was run for all DOE conditions: *Roof, Cap, Num.* In order to describe the design characteristics that result in a system where modeled demand never exceeds estimated supply, a regression approach was used to fit a planning equation where appropriate. A discussion of the baseline, two-occupant model follows with subsequent discussion of the three- and one-occupant models.

For the baseline, two-occupant model, no combination involving 3000 square feet of roof capture space was able to satisfy demand 100% of the time. Only a cistern size of 40,000 gallons resulted in 100% demand satisfaction for 3500 square feet of capture space. At 4000 square feet, cistern sizes of 25,000 gallons or more resulted in complete demand satisfaction. At 4500 square feet, any cistern size of 20,000 gallons or more was sufficient to satisfy demand. Finally, at 5000 square feet, any cistern size was acceptable. Figure 6 provides a surface response chart for demand satisfaction as a function of roof capacity and cistern capacity.

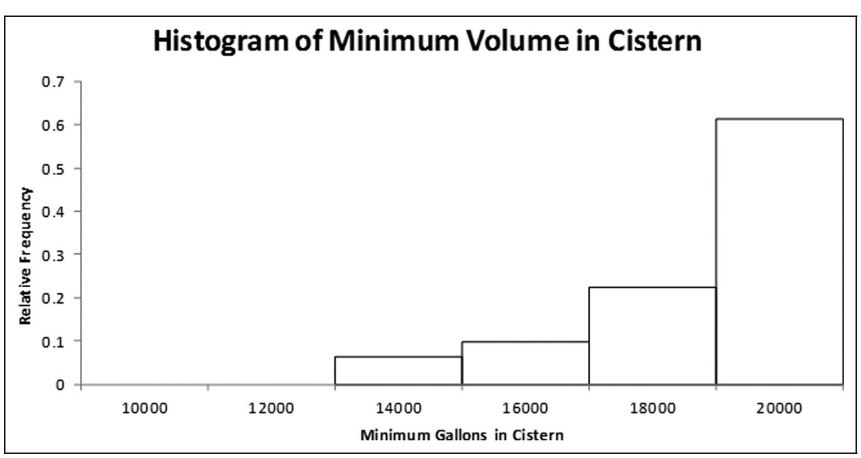

Figure 5. The distribution of the minimum for the baseline run.

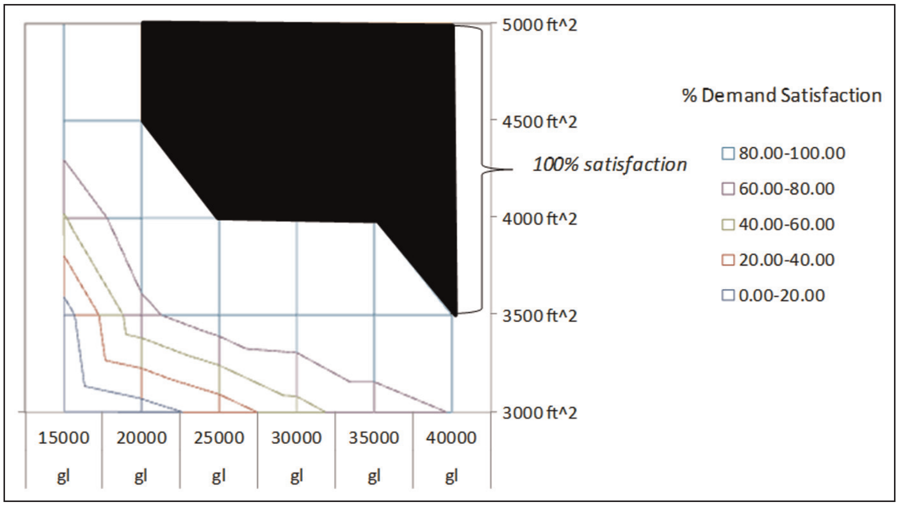

Figure 6. Demand satisfaction for the two-occupant model.

Fitting a multiple regression equation for demand satisfaction scaled [0,100] versus cistern size, roof size, squared terms, and interactions resulted in a reasonable planning equation, $F(5,24) = 45.22$, $p < .001$, $R^2 = .90$. Specifically, Equation (11) shows all terms in the model, each one significant at $p < .001$. This equation may be used for planning two occupancy residences in the same region:

$$X = -711 + 14.23C + 26.97R - .09C^2 - .24R^2 - .19C \times R \tag{11}$$

For the three-occupant model, the results are somewhat more problematic. Specifically, no combination of roof and cistern size was sufficient to satisfy demand 100% of the time. For the largest capacity model (roof size of 5000 square feet and cistern size of 40,000 gallons), 3 out of 36 runs resulted in demand exceeding supply, although the average minimum was 8887 gallons. Because no combination of planning factors achieved 100% ''demand satisfaction,'' we omitted the regression analysis.

For the one-occupant model, the design considerations become simpler. No matter the design options considered, the RWH system experienced 100% demand satisfaction. In fact, the minimum in the tank for a single occupant over the 36 iterations and 30-year runs with a 3000 square foot capture area and a 15,000 gallon tank was 4617 gallons. Indeed, smaller roof capture space would be sufficient here, and additional studies should expand to smaller capture surfaces and smaller cistern sizes.

The design implications are clear. RWH systems are highly sensitive to expected occupancy. A single additional occupant in this scenario results in increased risk of demand exceeding supply, although the risk was small (3 out of 36) for the largest capacity considerations. Fundamentally, the design of a RWH system in this region must carefully consider occupancy constraints.

4. Concluding remarks

RWH – the capture and storage of rainwater for agricultural and possibly domestic use – is becoming a more commonplace method for addressing the issue of water scarcity. One of the major considerations associated with the construction of RWH systems, however, is the capture surface area and the tank volume required for specific demand levels, given various occupancy estimates. In an effort to model this system, we implemented a non-parametric stochastic rainwater generator and simulation to determine requirements for building a RWH capable of providing a family with 100% of its water requirements with demand never exceeding supply.

Although the simulation results suggested that a 30,000-gallon cistern with 4000 square feet of roof capture area would suffice, one of the authors of this study decided to over-engineer the cistern and roof capacities because of the relatively small difference in cost, the decision to incorporate a covered deck (additional roof space), and the availability of acreage for the home site. The owners incorporated a 5000-square foot roof structure with a 40,000-gallon cistern. After priming the cistern with approximately 10,000 gallons during the construction phase, no additional water has since been required. The owners plan on using a secondary catchment tank for irrigation of vegetable gardens because of the amount of overflow experienced. Given current demand, the cistern hovers at capacity; however, the owners are tracking volume daily to confirm the initial planning factors. Figure 7 shows a photograph of the operational cistern.

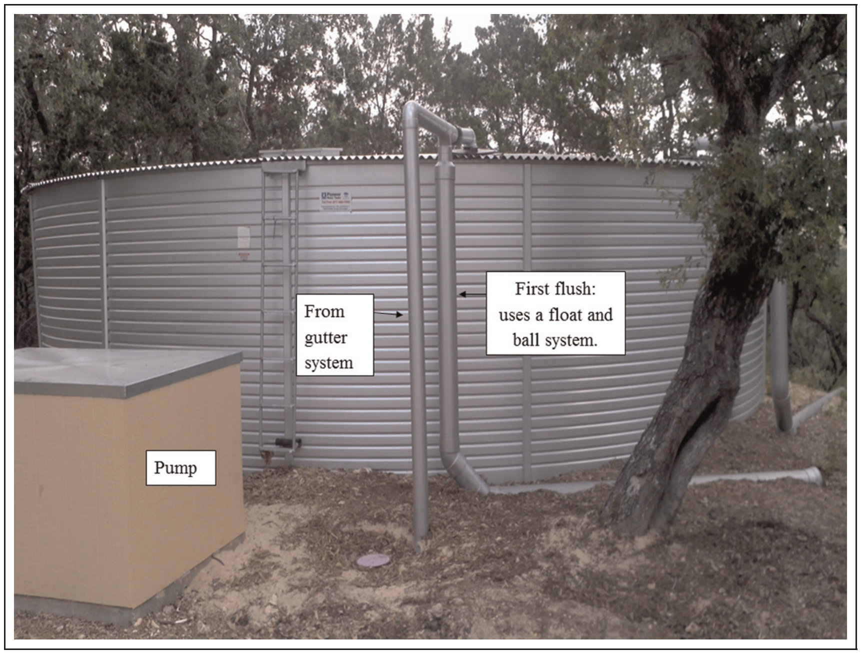

Figure 7. The operational rainwater harvesting system.

The limitations of this study are clear. Firstly, this study applies only to semi-arid regions. Secondly, it is appropriate only for the DOE factors included. Extrapolation beyond these parameters is not appropriate. Future work will expand the analysis to different roof and cistern capacities, as well as occupancies. Additional demand sensitivity analysis will be forthcoming.

Fundamentally, engineering a RWH system to provide 100% of domestic use of water even in semi-arid regions may be assisted by the use of simulation techniques. The importance of this study is clear. It demonstrates how a stochastic, non-parametric rainfall generator might be incorporated into a simulation model. The model shows clearly how sensitive RWH systems are to occupancy. Finally, the study provides a clear indication of some of the considerations necessary when planning a RWH system that is capable of sustaining a family with 100% demand satisfaction.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Rosegrant MW, Cai X and Cline SA. *Global water outlook to 2025: averting an impending crisis*. Columbo, Sri Lanka: International Water Management Institute, 2002.
2. Boers TM and Asher JB. A review of rainwater harvesting. *Agric Water Manage*1982; 5: 145–148.
3. Abdulla FA and Al-Shareef AW. Roof rainwater harvesting systems for household water supply in Jordan. *Desalinization* 2009; 243: 195–207.
4. Li F, Cook S, Geballe GG, et al. Rainwater harvesting agriculture: an integrated system for water management on rainfed land in China's semiarid areas. *J Hum Environ* 2000; 29: 477–483.
5. Phoca I and Valavanis P. *Rediscovering Ancient Greece: architecture and city planning*. Athens: Kedros Books, 1999.
6. Monteleone MC, Yeung H and Smith R. A review of ancient Roman water supply exploring techniques of pressure reduction. *Water Sci Technol* 2007; 7: 113–120.
7. Radhakrishna BP. Rainwater harvesting, a time-honored practice: need for revival. *Curr Sci* 2003; 85: 1259–1261.
8. Hatt BE, Delectic A and Fletcher TD. Integrated treatment and recycling of stormwater: a review of Australian practice. *J Environ Manage* 2006; 79: 102–113.
9. Han MY, Mun JS and Kim TI. Climate change adaptation strategy by multi-purpose, proactive rainwater management and case studies in Korea. *J Korean Soc Water Wastewater* 2009; 23: 223–230.
10. Basinger M, Montalto F and Lall U. A rainwater harvesting system reliability model based on nonparametric stochastic rainfall generator. *J Hydrol* 2010; 392: 105–118.
11. Rygaard M, Binninga PJ and Albrechtsena JH. Increasing urban water self-sufficiency: new era, new challenges. *J Environ Manage* 2011; 92: 185–194.
12. Mun JS and Han MY. Design and operational parameters of a rooftop rainwater harvesting system: definition, sensitivity and verification. *J Environ Manage* 2012; 93: 147–153.
13. Evans CA, Coombes PJ and Dunstan RH. Wind, rain and bacteria: the effect of weather on the microbial composition of roof-harvested rainwater. *Water Res* 2006; 40: 37–44.

14. Magyar MI, Mitchell VG, Ladson AR, et al. An investigation of rainwater quality and sediment dynamics. *Water Sci Technol* 2007; 56: 21–28.
15. Lye DJ. Rooftop runoff as a source of contamination: a review. *Sci Total Environ* 2009; 407: 5429–5344.
16. Han MY and Mun JS. Particle behavior consideration to maximize the settling capacity of rainwater storage tanks. *Water Sci Technol* 2008; 56: 73–79.
17. Amin MT and Han MY. Roof-harvested rainwater for potable purposes: application of solar collector disinfection (SOCO-DIS). *Water Res* 2009; 43: 5225–5235.
18. Fulton L. Rain gauge: Six Sigma is used to design a rainwater harvesting system for a sustainable home. *Quality Progress* 2013; 46: 30–36.
19. Gould J and Nissen-Petersen E. *Rainwater catchment systems for domestic supply: design, construction and implementation*. London: IT Publications, 1999.
20. Guo Y and Baetz BW. Sizing of rainwater storage units for green building applications. *J Hydrol Eng* 2007; 12: 197–205.
21. Sharma A and Lall U. A nonparametric approach for daily rainfall simulation. *Math Comput Simulat* 1999; 48: 361–371.
22. Lall U and Sharma A. A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resour Res* 1996; 32: 679–693.
23. National Oceanic and Atmospheric Administration. http://www.ncdc.noaa.gov/cdo-web/. Accessed April 21, 2012.
24. National Oceanic and Atmospheric Administration, http://www.srh.noaa.gov/images/ewx/sat/satmonrain.pdf. Accessed April 19, 2012.
25. Edwards Aquifer Authority, http://www.edwardsaquifer.org/index.php. Accessed April 20, 2012.
26. McDonald C. Forum discusses water challenges facing SA, http://www.mysanantonio.com/news/environment/article/Forum-discusses-water-challenges-facing-S-A-691164.php (6 October 2010). Accessed April 21, 2012.
27. National Geographic. Water footprint calculator methodology and tips, http://environment.nationalgeographic.com/environment/freshwater/water-calculator-methodology/ (2011). Accessed April 19, 2012.
28. *Texas manual on rainwater harvesting*, 3rd ed. Texas Water Development Board, 2005.
29. Promodel, http://www.promodel.com (2011). Accessed April 24, 2012.
30. R Development Core Team. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2008. Available at http://www.R-project.org.

Author biographies

Lawrence V Fulton is an Assistant Professor of Quantitative Methods at Texas State University. His research focuses on sustainability, medicine, and the military.

Nathaniel D Bastian is an Adjunct Instructor of Mathematics at the University of Maryland University College. His research focuses on energy and environmental modeling, military medicine, and healthcare systems engineering.

Francis A Mendez is an Associate Professor of Quantitative Methods at Texas State University. His research focuses on computational statistics, nonparametric methods, and machine learning.

R Muzaffer Musal is an Assistant Professor of Quantitative Methods at Texas State University. His research focuses on computational and Bayesian statistical modeling.