Proceedings of the 2012 Winter Simulation Conference
C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher, eds

CONSTRUCTION ANALYSIS OF RAINWATER HARVESTING SYSTEMS

Lawrence V. Fulton

Texas State University
601 University Drive
San Marcos, TX, USA

Rasim Muzaffer Musal

Texas State University
601 University Drive
San Marcos, TX, USA

Francis A. Méndez Mediavilla

Texas State University
601 University Drive
San Marcos, TX, USA

ABSTRACT

We present the results of a simulation to assess the optimal design characteristics of rainwater harvesting systems to be used in a semi-arid region of the United States. The simulation leverages a stochastic, non-parametric rainfall generator based on 64-years of daily historical data. The assumption of non-stationarity of rainfall is also thoroughly investigated for this paper. Of specific interest to this simulation was the estimate of roof capture space and cistern capacity required for a 100% reliable system capable of supporting family sizes of two or three for a 30-year time horizon. Considerations included rainfall supply, system capture efficiency, household occupancy, as well as individual demand variation. The optimal design characteristics in terms of roof surface area and cistern volume necessary for 100% reliability are presented using two response surface plots and separate multiple regression modeling based on expected occupancy.

1 INTRODUCTION

Rainwater harvesting systems (RWH) date back to at least 4,000 years ago (Abdullah and Al Shareef 2009), yet the ability to estimate rainfall based on historical data continues to be the topic of much discussion (Lall and Sharma 1996; Guo and Betz 2007; Basinger, Montalto, and Lall 2010). The difficulty in estimating rainfall is magnified when attempting to derive an optimal design of RWH, a system which is largely dependent upon the ability to model both supply, demand and storage effectively. In this study, we evaluate the construction parameters of an RWH system to be built in a semi-arid region of Texas. We modify the nonparametric, stochastic rainfall generator of Basinger et al. (2010), increasing the daily data collection time frame from 25 years to 64 years while retaining the assumption that daily rainfall probabilities and distributions are contingent upon knowledge of a 30-day centered moving average around the previous day's information. We gather daily rainfall data from of United States National Oceanic and Atmospheric Administration (NOAA) freely available datasets (NOAA, 2012a) . As an additional extension over previously submitted work, we also evaluate the possibility of non-stationarity, the idea that water variability is non-constant over time, which recent studies indicate to be a serious concern (Milley et al. 2008). In this study, we were concerned only with the stationarity or non-stationarity of supply (rainfall) rather than the larger analysis of regional water availability. Of primary interest is the required roof surface area and cistern volume to design a system capable of supporting all of a small family's water needs with 100% reliability (defined as the complement of the percent of time the tank is empty) in a semi-arid region of Texas and based on analyses of supply, demand, efficiency, occupancy, and other distributions. The results of the analyses are summarized in a response charts and a fitted multiple regression equations, another useful extension, which is also potentially useful for water planners in this region of Texas.

The significance of this study is multifold. First, the study extends previous non-parametric rainfall generators longitudinally. Second, the study evaluates non-stationarity of rainfall as part of a potential rainfall estimator. Third, it provides a mechanism for determining the optimal roof surface area and cistern size for the construction of an RWH that is 100% reliable. Fourth, the conclusions from the study informed the real-world construction of an author's RWH.

This study also expands and improves a paper currently under journal consideration in several major ways. First, we extend the analysis of rainfall generators to explore for non-stationarity of rainfall in the geographic region, a unique contribution in itself. Second, we provide separate surface response plots for given family sizes (rather than an analysis for a single family unit), which are in themselves more useful to RWH system planners. (While we only report family sizes of two and three, additional runs will be reported.) Third, we provide curve fitting models via regression that allow planners to have estimates of what building requirements are likely to be needed. Fourth, we evaluate the portability of models across family sizes (demand ranges). None of this work has been published or presented previously by the authors.

This paper proceeds as follows. First, the methods and materials are presented in Section 2. Section 2.1 (Study Design) elucidates the balance equations of the simulation model, while Section 2.2 (Study Setting) provides a portrait of the geography, underscoring the importance of the study. Section 2.3 (Rainfall Generator) discusses the rainfall generator model, while Section 2.4 (Simulation Flowchart) provides a discussion of the flowchart used for the continuous simulation. Section 3 provides the results and discussion of all models (including the response surface chart and fitted multiple regression model), while Section 4 provides conclusions.

2 MATERIALS AND METHODS

2.1 Study Design

Figure 1 is an original diagram of a commonly accepted rainwater harvesting system, and this diagram provides the basis for the study design. Rainwater flows from the roof and is largely (but not completely) captured by guttering systems usually equipped with some sort of debris filter system (such as screens). Some of the initial rainwater is flushed out of the system in order to eliminate contaminants (such as bird feces) that accumulate on the roof. Water then passes through one or more gross filters on the way to storage in a cistern. A pump transfers water from the system to a pressure tank (possibly through additional filtration first). On demand, water is transferred through multiple filters of decreasing size (e.g., 50 micron, 5 micron, and perhaps, 3 micron). One or more of these filters are likely to be charcoal based. For potable water systems, the water is typically exposed to ultraviolet light prior to entering the house. Excess water is ejected out of the system, sometimes to non-potable tanks for irrigation of plants. This design is typical of RWH systems (*Texas Manual on Rainwater Harvesting Manual* 2005).

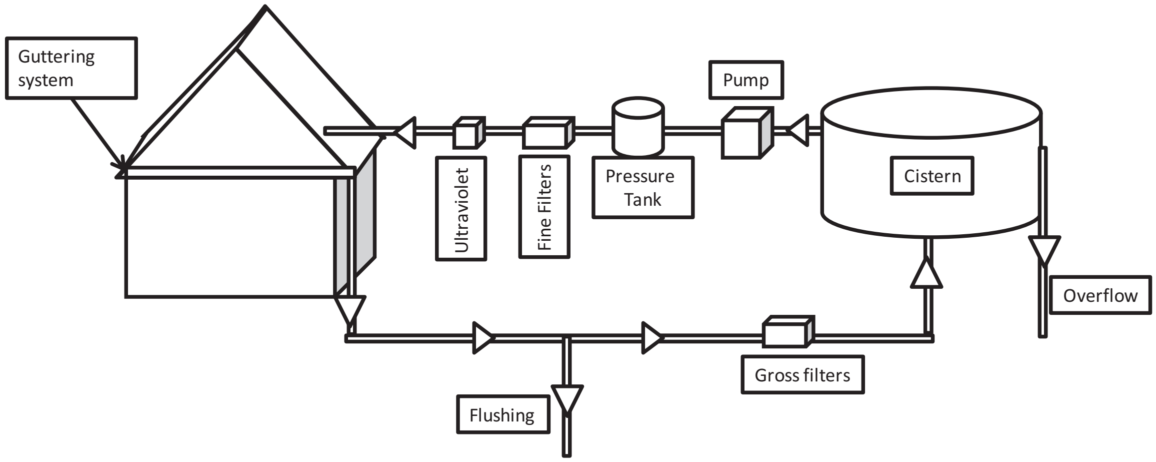

Figure 1: A typical RWH system is shown.

Although generating quality water is a major concern of RWH systems, it is also well-researched (Evans et al., 2006 and Magyar et al., 2007). This study focuses on estimating the proper design necessary to meet demand quantity. One of the most common issues in constructing a RWH system is that of determining the optimal capture surface size and storage space especially given uncertainty regarding supply (rainwater frequency). Insufficient capture surface will generate insufficient water for the required demand. Insufficient storage capacity will result in inability to provide water during periods of no rain. Because of the complexity associated with the analysis of these two elements and the requirement to solve for balance equations associated with supply, demand, and overflow, this study leveraged continuous simulation. The balance equation for the simulation is Equation (1).

$$V_t = \sum_{i=0}^{i=t} Max\{0, V_{t-1} + C_t - D_t - O_t\} \quad (1)$$

In this equation, V is the volume in the tank, C is the supply of rainwater from the roof, D is the demand for purified rainwater, and O is the overflow. These variables are indexed by time measured in days, t. Therefore, the volume at time t is the maximum of zero (volume is obviously positive, semi-definite) or the sum of the previous day's cistern water volume plus the current day's rainfall minus the current day's demand minus any overflow (a function of tank capacity).

2.2 Study Setting and Data

This study's setting is San Antonio, Texas, as one of the authors was to build a home using 100% rainwater recapture. Data for the study were freely available from the NOAA (2012a), and the closest reporting station (the San Antonio International Airport) was selected as most representative.

The authors sought to ensure that the system was 100% reliable; and to do so required careful analysis, as San Antonio is a semi-arid region that routinely experiences droughts and has a rainfall median of only 28.53 inches based on the past 64 years of data (NOAA, 2012b). The concern for having reasonable rainfall estimates was clear. While historical data provided excellent base runs, the stochastic nature of rainfall made the need for rainfall generators that better reflected the randomness of the system necessary. An investigation of rainfall supply stationarity was an addition necessary to evaluate whether additional tank capacity or capture surface might be needed in the future; and perhaps, forecast when it might be needed.

2.3 Rainfall Generator

Two potential rainfall generators were therefore evaluated. The first generator was based on Basinger et al. (2010). The probability of rainfall on day *t* was a function of day *t-1* only; however, a yearly trend was not evaluated. More specifically, a 30-day centered analysis of the probability of rain on day *t-1* was used to determine if rainfall would occur on day *t*. If it did occur, then the frequency of rain was evaluated. Each day of the year had its own separate Bernoulli probability of rain (one-step Markov Chain); however, the model assumed stationarity of rainfall.

As a separate analysis, we decomposed the 64 years of rainfall by year and by month. This decomposition provided a secondary demand stream for analysis; however, no trend emerged, although monthly seasonality was confirmed. While annual rainfall variability appeared to increase somewhat, this change was not statistically meaningful. Since the first model adapted from Basinger (2010) already accounted for monthly effects (seasonality) by using a 30-day centered moving average, this component of the decomposition model was already evaluated. Therefore, the modified Basinger (2010) model was accepted as reasonable. Figure 2 shows the median rainfall as well as the 5^{th} and 95^{th} percentiles for all years containing full monthly data. Figure 3 illustrates the monthly distribution of rainfall.

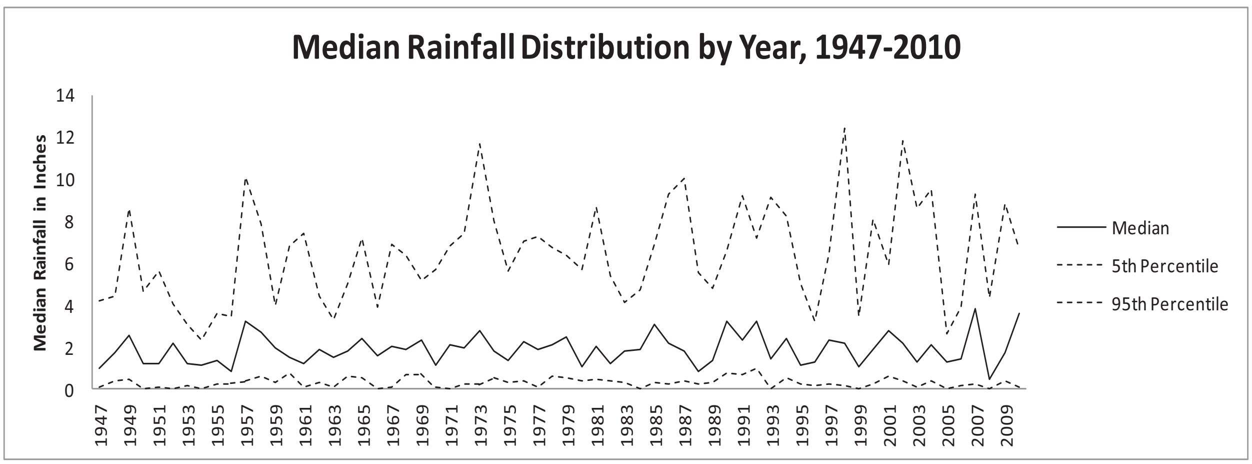

Figure 2: Rainfall distribution by year is shown.

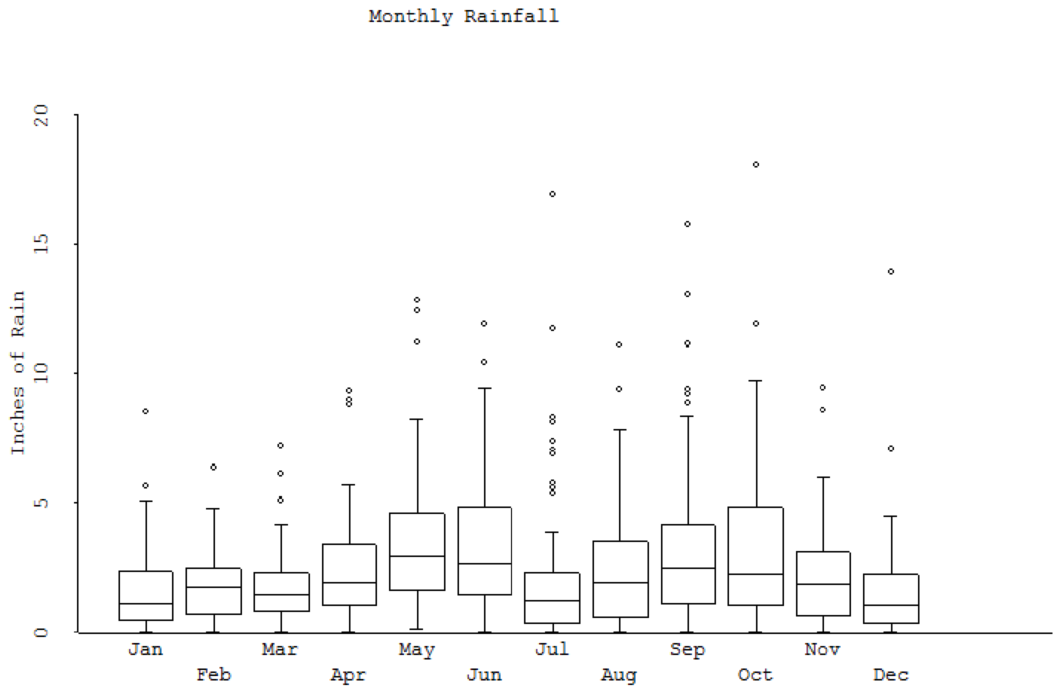

Figure 3: Rainfall distribution by month is shown.

What was interesting to note is the absence of a reasonable trend effect in any model using monthly and yearly data . Figure 3 depicts the monthly distribution of rainfall (inches) and illustrates relatively larger variability of rainfall during the summers and falls. Figure 4 shows the rain frequency by day and 30-day centered moving averages. A noticeable drop in the rain probability occurs during the summer months.

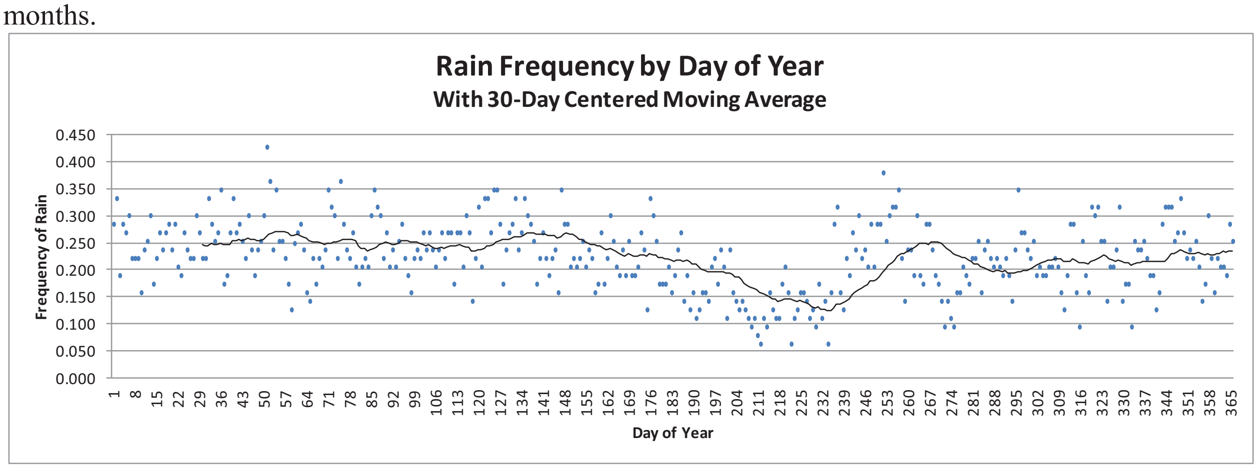

Figure 4: Rainfall frequency by day and 30-day centered moving averages (solid line) are shown.

When rainfall does occur, the distributions by day are depicted in Figure 5. A high degree of variability in the distributions exist between the months of March and October.

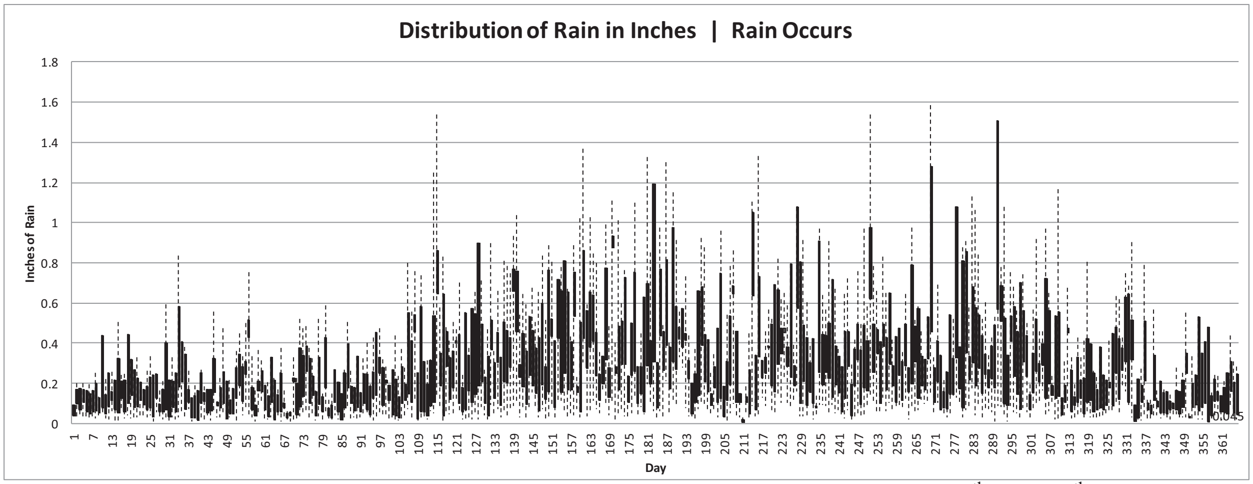

Figure 5: Daily distributions of rain by day are shown. Dashed lines represent 25th and 75th percentiles.

2.4 Simulation Flowchart

The rainfall simulation itself was relatively straightforward. Initially implemented in Excel using VBA script, it was ported over to ProModel® as part of a simulation class and designed using the "tank" model subroutines that allow quantitative, continuous distribution modeling. Figure 6 depicts the simulation flowchart, which is similar to that of Mun & Han (2012), although derived independently prior to construction of the RWH system.

In this simulation, variables are initialized, and design of experiments factors associated with roof square footage and cistern volume in gallons are set to {3000, 3500, 4000, 4500, 5000} and {15000, 20000, 25000, 30000, 35000, 40000}, respectively. Additionally, the expected number of daily occupants are evaluated on the set {2, 3}. The starting cistern volume for all simulation runs was set to 50% of tank capacity, as priming the cistern is necessary for plumbing checks regardless of rainfall availability.

Individual water demand was based on National Geographic (2011) analysis (uniform between 40.8 and 69.3 gallons per person per day). This consumption rate coupled with the expected number of individuals assumed to occupy the home on any given day (2 or 3).

The capture efficiency of rainfall was estimated to be uniformly distributed between .75 and .90 based on the Texas Manual on Rainwater Harvesting (2005). Because of the estimated occupancy of the owner, the simulation horizon was set to 30 years. Given the necessity to bracket the volume in the tank tightly, a 100-gallon margin of error and a 95% confidence interval were selected, resulting in a 238-run requirement given an initial standard deviation of 782.42 gallons for a 30-year scenario run.

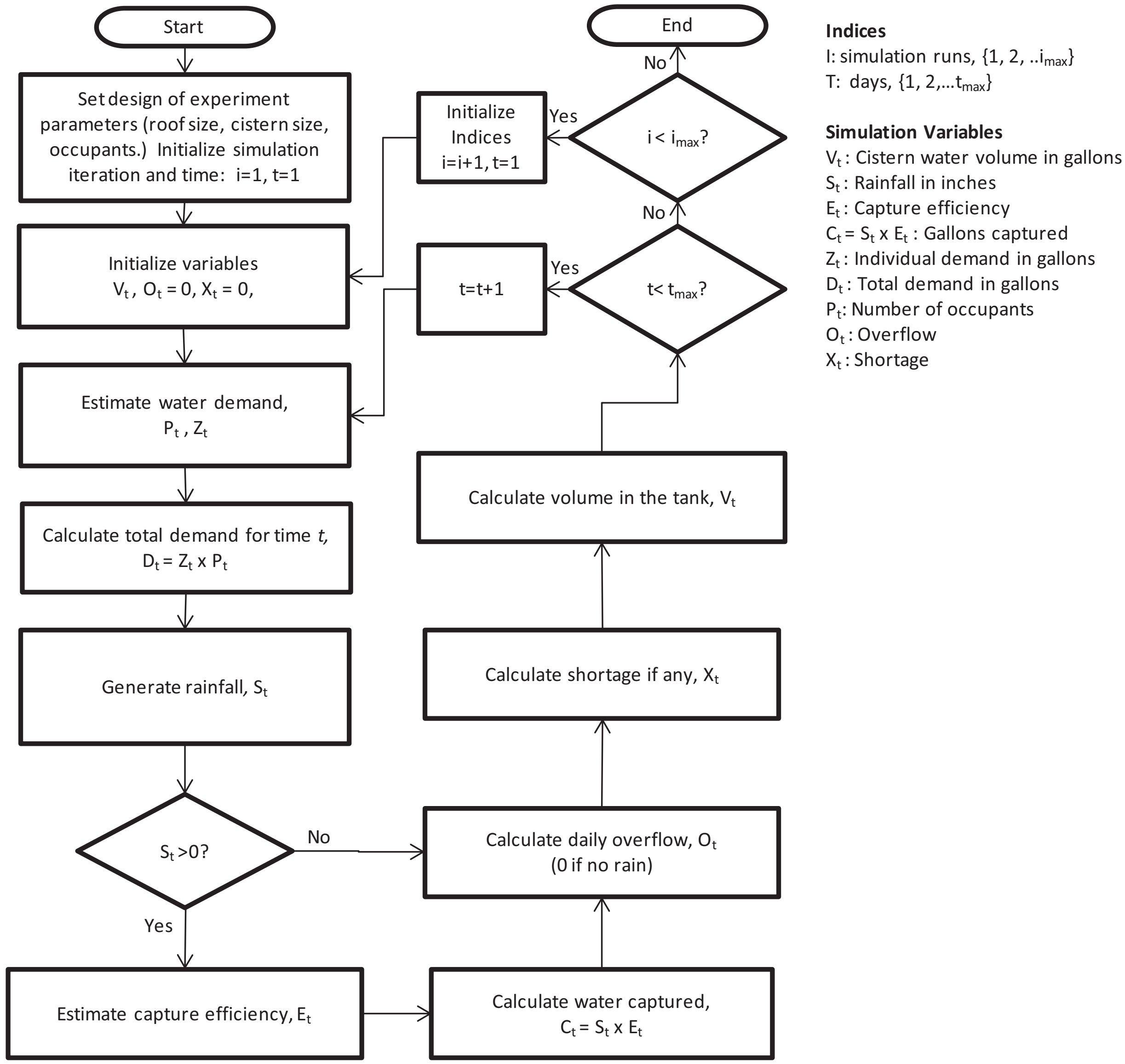

Figure 6: The flowchart for the simulation is shown.

3 RESULTS AND DISCUSSION

Verification of the system design was conducted with the owners and by use of previous literature (Mun & Han, 2012). Validation of the simulation revealed that no statistically significant or practically relevant differences existed between the a priori and posterior distributions. Convergent validity was achieved by programming the simulation using two different tools and by having four student teams conduct simulations, all of which reached similar conclusions.

An effective way to display the results of this simulation are via response surface charts. Figure 7 represents the contour plot for cistern volume and roof surface area against reliability with the expected number of occupants set to 2. This figure depicts the optimal configurations for roof are and cistern size for the distributions discussed earlier. Specifically, any combination of roof sizes greater than or equal to 4,000 square feet coupled with cistern sizes of greater than 30,000 gallons would result in a 100% reliable system. In fact, 5,000 square feet of roof space and 40,000 gallons of cistern capacity resulted in an absolute minimum of more than 10,000 gallons in the tank after 238 iterations of 30-year runs.

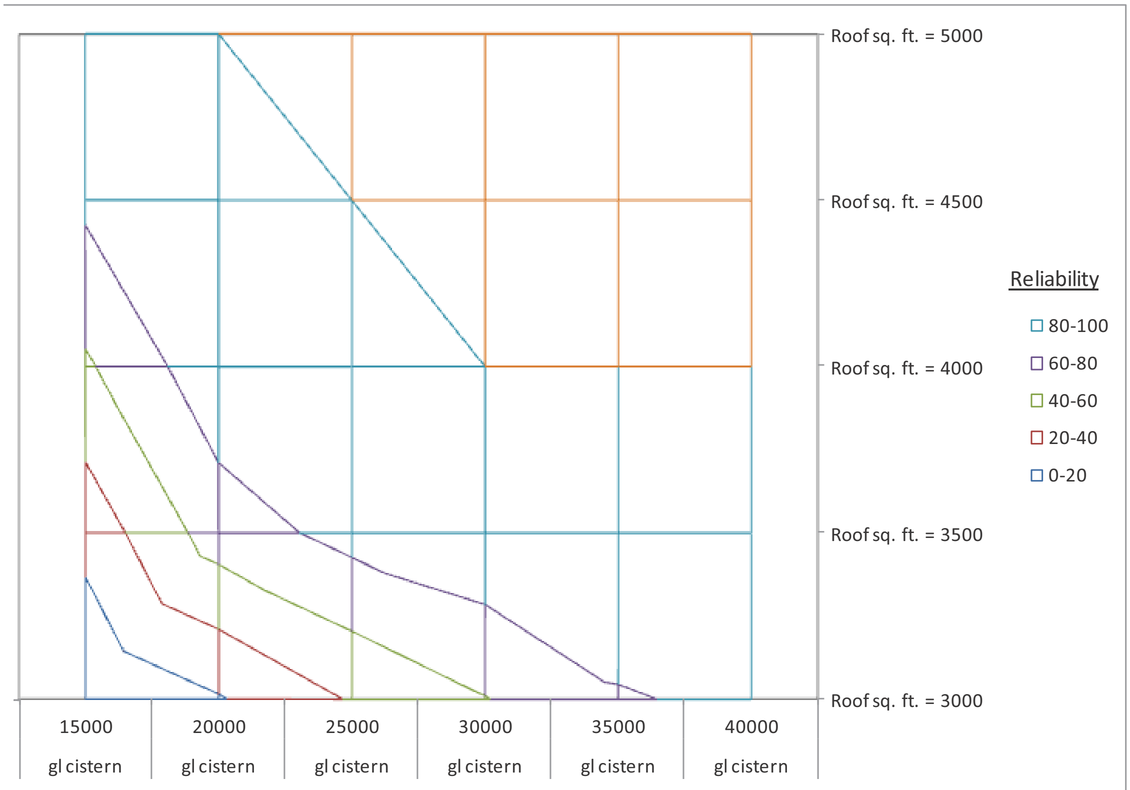

Figure 7: Contour plot of rainfall versus cistern volume and roof surface area for two occupants is shown.

Figure 8 shows the effect of adding a third occupant to the demand structure. Unfortunately, the results are not nearly as encouraging as Figure 7 when the number of occupants increase from two to three. In fact, the best reliability achieved in 238 iterations of 30 years is 90.76% (with 5,000 square feet and a 40,000 gallon cistern.) Additional cistern capacity would result in more rapid changes in reliability. In fact, about 5,500 square feet of roof space with a 40,000 gallon cistern is required to achieve a 100% reliable system. (Additional work will investigate capture efficiency rates as well to address this issue.)

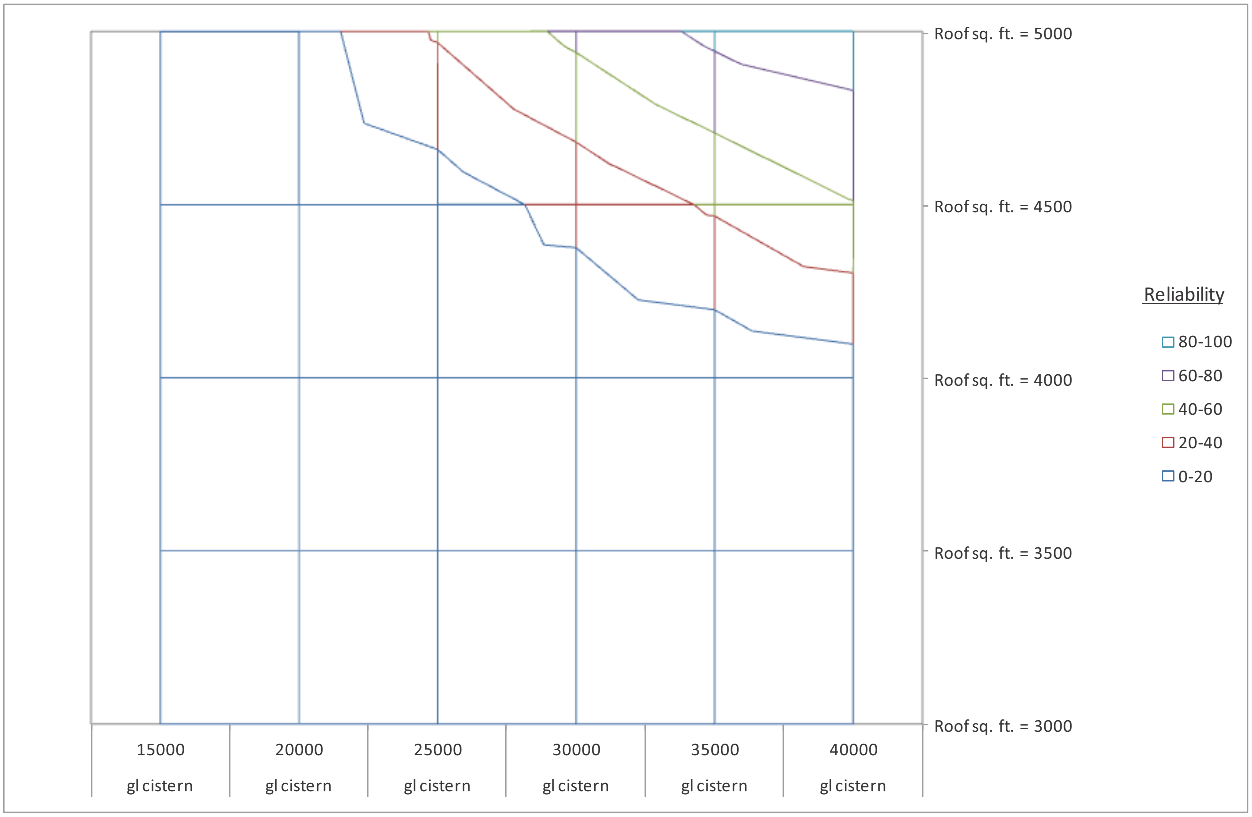

Figure 8: Contour plot of rainfall versus cistern volume and roof surface are for three occupants

Fitting a multiple regression equation to describe the response surface chart for two occupants (Figure 7) was straightforward. Based on the runs, squared terms and an interaction were included in a reasonable model (R^2 = .950, Se = 6.836, F(5, 24) = 90.372, p<.001). Equation 2 is the model (all coefficients statistically significant at the p<.001 level) that describes reliability, *X*, measured on [0, 100].

$$X = -619.95 + .014C + .229R - 9.32 * 10^{-8}C^2 - 1.95 * 10^{-5}R^2 - 1.78 * 10^{-6}C * R \qquad (2)$$

The median absolute deviation for this regression model was 4.334, which would equate to 4.334% reliability. The maximum error was 12.191, and the minimum error was .296 (a near perfect fit).

A separate model was analyzed to describe the three occupant contour plot (Figure 8). The model produced was equally interesting with R^2 =.941 ., Se =7.189, F(5, 24) =76.507, p<.001. Again, all coefficients shown in Equation 3 were significant at the .001 level except for the square of the cistern volume, which is not significant (p=.626). Including this term in the model (although it might reasonably be removed), the regression model follows.

$$X = 412.054 - .008C - .190R + 1.044 * 10^{-8}C^2 + 1.950 * 10^{-5}R^2 + 2.098 * 10^{-6}C * R \qquad (3)$$

Again, the fit here was tight, with a median absolute deviation of 5.09%. The maximum absolute deviation was 12.85%, and the minimum absolute deviation was .340%, nearly perfect.

4 CONCLUSIONS

Several conclusions became apparent from this analysis. From an engineering perspective, any combination of roof surface area and cistern volume that rest in the 100% reliability region of Figure 7 is likely to be sufficient for engineering an RWH system, but for houses with expected occupants of three or more, even a 40,000 gallon cistern and 5,000 square feet of roof space is not quite sufficient for 100% reliability. Second, non-stationarity of rainfall is not an important consideration for this particular region of Texas. Although it is likely a consideration for aquifer and stream management given that demand may outpace supply due to population growth, rainfall changes are not readily apparent over the last 64 years in this region. Initially, we had expected otherwise. Third, models that use reasonable distributions provide the ability for researchers to generate response RWH surface contour plots as shown in Figure 7 and 8. These plots are useful in designing future RWH systems in any region for which supply data are available. Fourth, fitting multiple regression models to the plots may provide a simple method for initial engineering estimates. Fifth, the results of simulations like this one may be used for planning purposes, as in the case of this study.

Our work to improve RWH planning and execution is ongoing, and we will continue developing both rainfall generation models as well as simulations to support construction efforts. Future work will expand the number of occupants and manipulate demand considerations as well as efficiency rates.

REFERENCES

Abdulla, F.A., & A. W. Al-Shareef. 2009. "Roof Rainwater Harvesting Systems for Household Water Supply in Jordan." *Desalinization* 243:195-207.

Basinger, M., F. Montalto, and U. Lall. 2010. "A Rainwater Harvesting System Reliability Model based on Nonparametric Stochastic Rainfall Generator." *Journal of Hydrology* 392:105- 118.

Evans, C.A., P.J. Coombes, and R.H. Dunstan, RH. 2006. "Wind, Rain and Bacteria: The Effect of Weather on the Microbial Composition of Roof-Harvested Rainwater." *Water Research* 40: 37-44.

Guo, Y. and B.W. Baetz. 2007. "Sizing of Rainwater Storage Units for Green Building Applications." *Journal of Hydrologic Engineering* 12(2): 197-205.

Lall, U., and A. Sharma. 1996. "A Nearest Neighbor Bootstrap for Resampling Hydrologic Time Series." *Water Resources Research* 32(3): 679-693.

Magyar, M.I., V.G. Mitchell, A.R. Ladson, and C. Diaper. 2007. "An Investigation of Rainwater Quality and Sediment Dynamics. *Water Science & Technology* 56(9) 21-28.

Milley, P.C.D., J. Betancourt, M. Falkenmark, R. Hirsch, Z.W. Kundzewicz, D.P. Lettenmaier, and R.J. Stouffer. 2008. "Stationarity is Dead: Whither Water Management?" *Science Magazine*, 319: 573-574.

National Geographic. 2011. "Water Footprint Calculator Methodology and Tips. Accessed 10 March. http://environment.nationalgeographic.com/environment/freshwater/water-calculator-methodology/

National Oceanic and Atmospheric Administration. 2012a. NNDC Climate Data Online. Accessed 7 March. http://cdo.ncdc.noaa.gov/pls/plclimprod/poemain.cdobystn?dataset=DS3240&StnList=41 7945

National Oceanic and Atmospheric Administration. 2012b. Monthly/Annual/Average Precipitation San Antonio, Texas. Accessed 5 April. http://www.srh.noaa.gov/images/ewx/sat/satmonrain.pdf

Texas Manual on Rainwater Harvesting, 3rd Edition. 2005. Texas Water Development Board. Accessed 30 March. http://www.twdb.state.tx.us/publications/reports/RainwaterHarvestingManual_3rdedition.pdf

AUTHOR BIOGRAPHIES

LAWRENCE V. FULTON holds a doctorate from the University of Texas, Austin. He currently teaches quantitative methods in the McCoy School of Business, Texas State University. His email address is lf25@txstate.edu.

RASIM MUZAFFER MUSAL holds a doctorate from George Washington University. He currently teaches quantitative methods in the McCoy School of Business, Texas State University. His email address is rm84@txstate.edu.

FRANCIS A. MENDEZ MEDIAVILLA holds a doctorate from Rutgers University and is an associate professor of quantitative methods in the McCoy School of Business, Texas State University . His email address is fm16@txstate.edu.