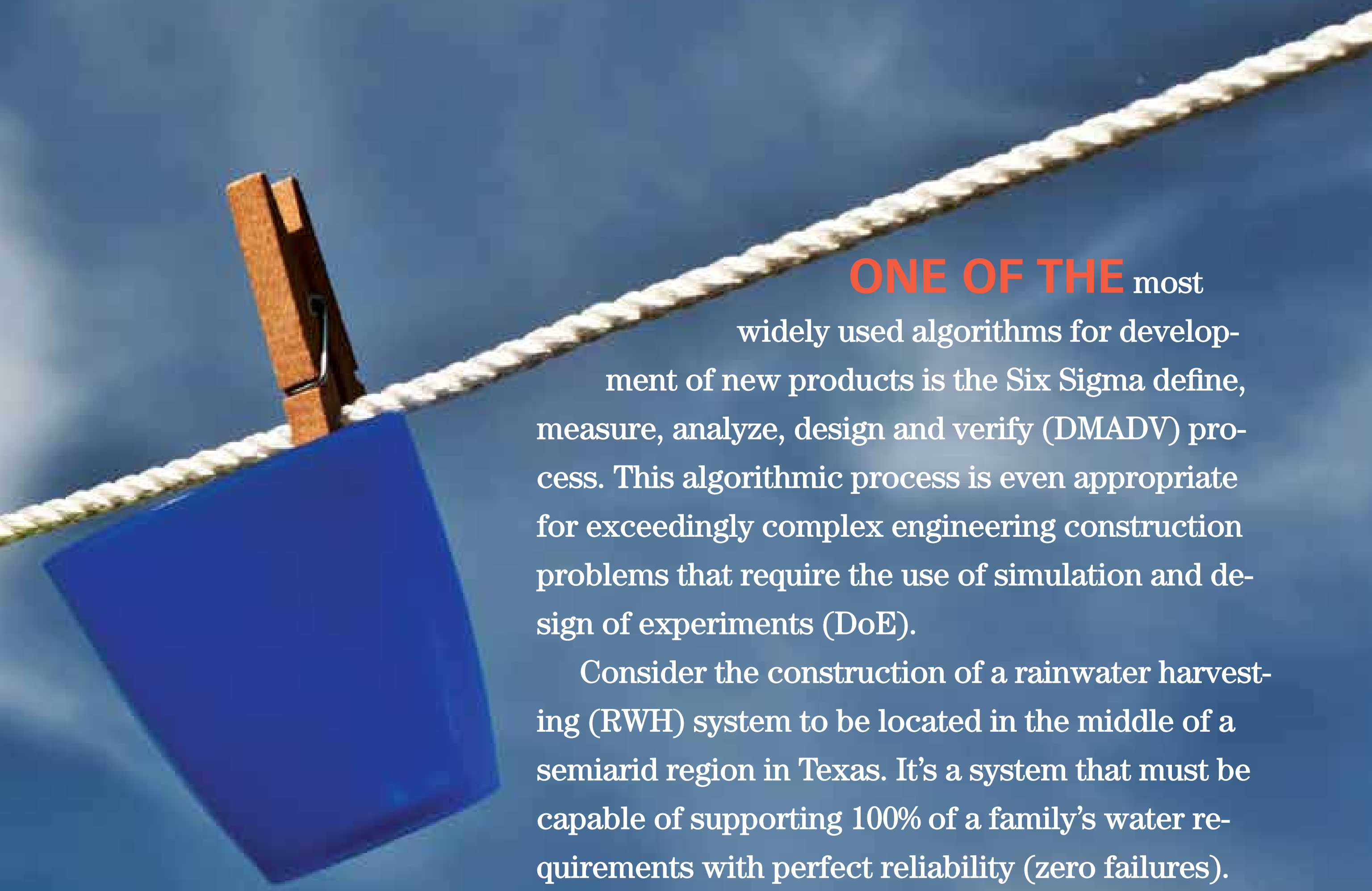
Rain Gauge

Six Sigma is used to design a **rainwater harvesting system** for a sustainable home

by Lawrence V. Fulton

In 50 Words Or Less

- The define, measure, analyze, design and verify (DMADV) process is a useful approach to solving complex engineering construction problems that require simulation and design of experiments.
- In this case study, DMADV and simulation were applied to design and construct a new rainwater harvesting system in a sustainable home in Texas.



ONE OF THE most widely used algorithms for development of new products is the Six Sigma define, measure, analyze, design and verify (DMADV) process. This algorithmic process is even appropriate for exceedingly complex engineering construction problems that require the use of simulation and design of experiments (DoE).

Consider the construction of a rainwater harvesting (RWH) system to be located in the middle of a semiarid region in Texas. It’s a system that must be capable of supporting 100% of a family’s water requirements with perfect reliability (zero failures). DMADV is an appropriate algorithm for assessing the design of this new construction, and simulation is an appropriate tool given the complexity of the RWH system.

The minimum water in the tank is of great importance. Reliability of the system—based on availability of water to satisfy demand—is obviously of utmost concern, as running out of water would defeat the purposes of the owners: independence and sustainability.

Capturing water

Typical RWH systems intended to provide potable water for homes have several well-known and well-documented components.[1] Rainfall from the roof surface (often galvanized steel or another appropriately engineered surface) funnels to gutters, which are normally equipped with debris protection screens.

Because some of the first water harvested from a dry roof surface might contain unwanted biological residue (for example, bird excrement and associated contamination), an initial quantity of rainwater is flushed out of the system in a procedure typically termed "first flush."

The water then proceeds through several gross filters and rests in a cistern for future use. A pump forces the water from the cistern to a pressure tank, and then the water is forced through two fine, charcoal filters (for example, 50 micron and five micron). Finally, the water is exposed to ultraviolet light for biological purification prior to entering the house's plumbing.

The method for obtaining quality water compliant with local standards has been refined over many years. Figure 1 illustrates a basic RWH system.

Knowing how to construct a viable system is not equivalent to understanding basic design considerations. Two of the most basic considerations for an RWH designed to support 100% of a family's needs with perfect reliability are the size of the cistern and the surface area of the roof.

Given daily estimates of rainwater supply and individual demand, cistern capacity and roof area are of primary concern, as collection and storage must be sufficient to cover times of drought. In semiarid regions of Texas, a drought is not a rare event. Use of desalinization plants has even been discussed recently based on current and anticipated water shortages.[2]

Applying DMADV

In this case study, the use of DMADV is applied to RWH construction associated with the National Association of Homebuilder's highest-scoring (Emerald class) sustainable home, one which my wife and I had built by master builder Darrel McMaster.[3-5]

The home was engineered to be point source, meaning it acquires all of its power and water from the environment. The water engineering component was one of the complexities, but the DMADV algorithm provided a reasonable mechanism for assessing design considerations.

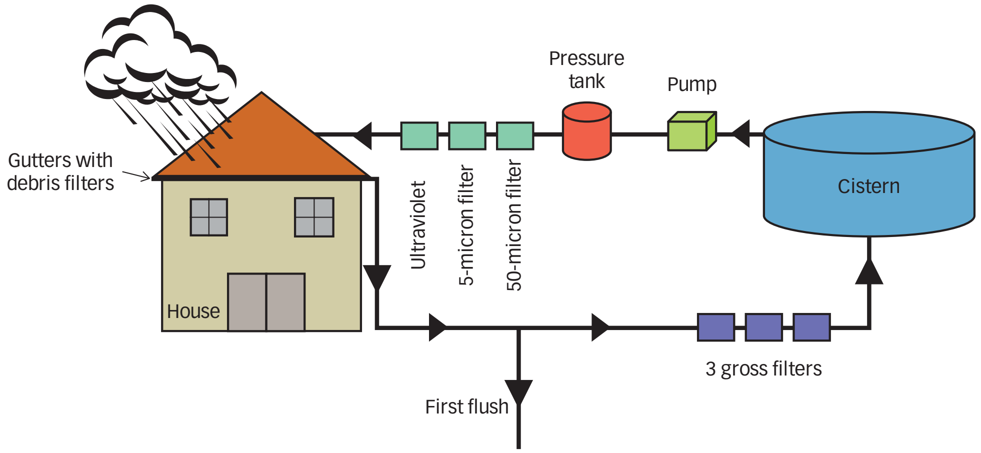
Before detailing the components of DMADV, you need a perspective of the math behind rainwater capture. A cubic foot of rain (12 inches of rainfall over one square foot) equates to 7.48 gallons. Assume you have a roof with surface area of 3,000 square feet. Additionally, assume average rainfall during the year is similar to that in this case study: 32 inches per year. Finally, assume the RWH system is 85% efficient. The annual capture would be:

$$\frac{7.48\text{ gal}}{\text{ft}^3} \times \frac{1\text{ ft}}{12\text{ in.}} \times \frac{3{,}000\text{ ft}^2}{1} \times \frac{32\text{ in.}}{\text{yr}} \times 0.85 = 50{,}864\ \frac{\text{gal}}{\text{yr}}$$

For a two-person family, assume an average consumption of 60 gallons per person per day for a total of 120 gallons per day and about 43,800 per year (ignoring leap years). You can see that deterministic math appears to support the use of an RWH system because the mean yearly supply of 50,864 gallons exceeds the yearly mean demand of 43,800 gallons.

You cannot, however, see how often the tank might be empty based on daily variations. What is the likelihood you might run out on a given day? These individual variations clearly illustrate why simulation might be useful. They also illustrate the complexity associated with designing RWH systems. You cannot use a simple equation because rainfall is a chaotic process.

Typical rainwater harvesting system / FIGURE 1



Simulation and DoE can be used to provide response surfaces of reliability versus cistern volume and roof square footage to support the new construction. Through response surfaces, decision-makers have a tool with which to estimate future design characteristics given similar rainfall, occupancy and demand variation. Let's take a closer look at each component of DMADV.

Define

Careful problem definition is always one of the most vital design considerations. Solving the wrong problem is useless. In this case, the research question is straightforward: What are the optimal design parameters for cistern volume (in gallons) and the roof surface area (in square feet) to develop a 100% reliable RWH system capable of supporting a two-person family given daily historical rainfall data from the National Oceanic and Atmospheric Administration[6] and estimates of daily water consumption from the National Geographic Association?[7]

The definition scopes the problem in a way that two engineering parameters emerge: cistern capacity and roof surface area. Essentially, the analysis should help decision-makers understand what cistern size should be purchased and how much extra roof surface area should be built (using additional patios, overhangs and garage space) to address water needs of the home. The solution to this problem should also provide information for future construction based on the parameters of interest.

Measure

The carefully scripted problem definition helps determine the parameters and variables of interest. In this particular case, the construction variables of cistern volume and roof surface area are of interest. Manipulating these variables might help you understand overall system reliability.

One of the primary concerns evident from the research question is the 30-year system reliability, which is operationally defined as the probability that daily supply plus volume in the cistern would be greater than or equal to daily demand on any given day for 30 years (hopefully, the minimum of the owners' anticipated life spans).

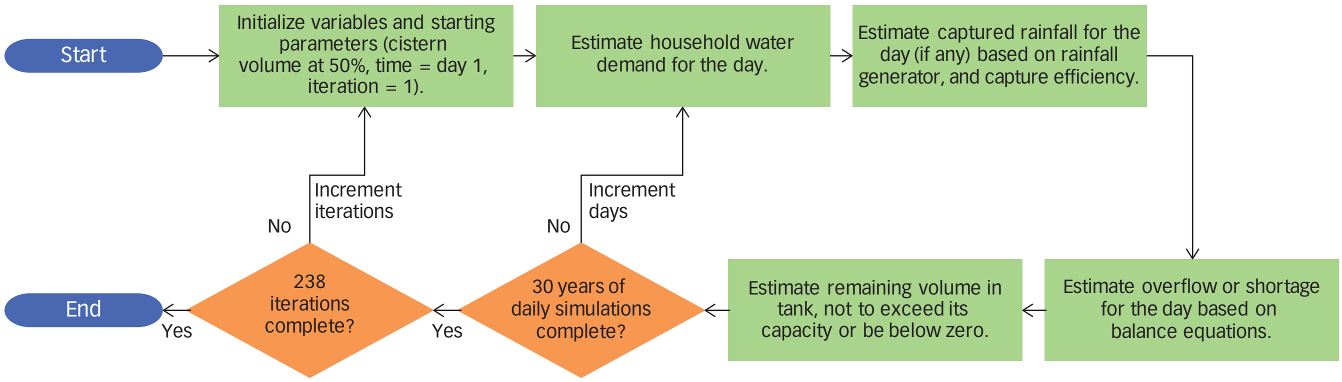
Reliability derives from the balance equations associated with supply and demand. In this case, the distribution of the minimum water in the tank is more important than the distribution of the median or mean.

This example is a perfect case in which focusing strictly on measures of central tendency is ill-advised. Sometimes, you might care about the minimum or maximum—or another percentile. These distributions are called order statistics because their importance derives from their position in an ordered array.

Analyze

This RWH problem is complex and requires quite a few basic and non-basic quality tools. Specifically, the simulation leverages historical rainfall data that might be used either deterministically or stochastically (probabilistically) in some sort of rainfall generator. In this

Simplified flowchart for the simulation / FIGURE 2



case study, stochastic methods are discussed because they routinely produced more conservative estimates than deterministic methods and because rainfall is unpredictable.

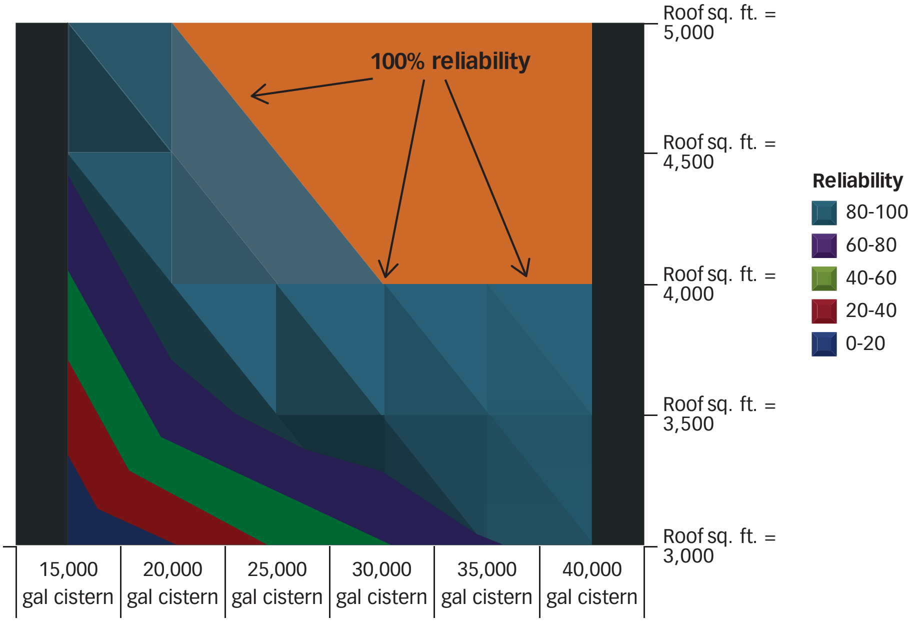
The simulation also applies stochastic demand based on estimates of individual use as provided by the National Geographic Association.[8] Based on construction guidelines—and potential future work—the simulation evaluates a set of feasible cistern sizes (set at increments of 5,000 from 15,000 to 40,000 gallons) and a set of feasible roof surface areas (set at increments of 500 from 3,000 to 5,000 square feet). These parameters may be adjusted for specific design considerations, too.

In cases requiring operational decisions in which you have well-defined and repetitive processes, interdependent and variable processes, high complexity and a reasonable cost-benefit relationship, simulation is often the best method.[9] Such is the case in this study.

The continuous simulation presented here consists of four major variable components:

1. A stochastic, nonparametric rainfall generator that leverages 64 years of historical data, similar to that described in an article published in the *Journal of Hydrology*.[10] This simulation assumes each day of the year had its own probability distribution and associated expected value for rainfall amount given information about only the previous day's distribution. The daily probability distributions are based on 30-day periods, centered on the day in question. Historical data also were used in simulation runs, but the results for historical data were generally just a bit less conservative than the stochastic method.
2. Uniformly distributed water demand based on estimates provided by the National Geographic Association (40.8 to 69.3 gallons per occupant per day).[11]
3. Discrete triangular distribution of occupants with a minimum of one, a mode of two and a maximum of three. (Individual response surface charts might have been modeled given a fixed number of occupants).
4. Uniform rainwater capture efficiency between 75 and 90% based on the *Texas Manual on Rainwater Harvesting*. Not all of the rainfall hitting the roof is captured due to overflow and spillage.[12]

Contour plot of reliability given cistern volume and roof surface area / FIGURE 3



gal = gallon

sq. ft. = square feet

Manipulating the cistern volume and roof capacities allowed for the generation of response surface (contour) plots of reliability as a function of these two construction parameters. The design of the simulation made the parameters particularly easy to change for future analysis.

Developing a reasonable simulation is nontrivial. The first step is often graphical modeling of both the system (Figure 1, p. 32) and its associated process flowcharts (Figure 2, p. 33). The concept of bounded rationality is important in this process—model the important few, not the trivial many. In this case, the filtration systems and pumps were of little importance to the primary design considerations—although immensely important to water quality—and were thus omitted from the simulation.

Programming the simulation model requires some skill and knowledge of statistical techniques, but this model was readily designed in Microsoft Excel and ProModel[13] because of their relatively straightforward components and minimal logic processing.

One of the major requirements for the design of any simulation is verification and validation, an often overlooked requirement. Verification and validation of the model require assessing whether the model captures the conceptual idea (verification) and whether the model reflects the real-world system using reasonable data (validation).

This model followed the rainwater harvesting construct as depicted in previous research[14] (third-party verification), as well as the owner's and builder's analyses of the process (first-party verification). In addition, four separate graduate student teams also modeled the capture process and derived similar system architecture.

In terms of validation, after-the-fact comparisons of water demand reflected the previously assumed uniform distribution, and the output replicated the input. Surprisingly, analysis of rainfall showed that the rainfall estimator slightly underestimated rainfall statistically but not practically. The results from the Excel and ProModel simulations were similar, which provided convergent validity. After some moderate level of effort, the model worked as expected and reflected reality closely.

Using the daily rainfall capture standard deviation from an initial 30-year simulation run (782.42 gallons per day), an estimated 238 iterations would be sufficient to bracket the mean rainfall within the cistern within 100 gallons at the 90% confidence interval level.

This calculation is based on the t distribution with 237 degrees of freedom and derives from solving for the sample size in the t-distribution statistic:

$$n = \left\lceil \frac{(t_{.05,237})^2 s^2}{\varepsilon^2} \right\rceil$$

At this point, it was simply necessary to set the starting parameters and allow the simulations to run.

The simulation had two factors, cistern volume and roof surface area. Cistern volume had six options (or levels, as commonly termed in DoE) beginning with 15,000 gallons with increments of 5,000 and thus a maximum of 40,000 gallons.

The surface area had five options with 500 square feet increments beginning with 3,000 square feet and ending with 5,000 square feet. The total number of 238-iteration, 30-year simulation runs required to conduct this analysis was six levels for cistern volume multiplied by five levels for roof area, totaling 30.

With current computing power, the number of total daily iterations required for all runs (238 iterations x 365 days x 30 years x 30 different simulations = 78,183,000) is nonproblematic. Note that the water volume in the tank was initialized to be 50% of capacity, as some initial

EARLY CONSTRUCTION OF the final tank and recapture system for a rainwater harvesting system.

While this simulation is **relatively simple,** it certainly illustrates the **art of the possible.**

priming is necessary to test plumbing during the construction phase.

The results of the simulation are best expressed as a contour plot, as shown in Figure 3 (p. 34). As the plot illustrates, reliability is 100% for 20,000 gallons and a 5,000 square foot roof. Cistern values greater than 4,000 gallons with associated surface area of 4,000 or more square feet are also at 100% reliability.

An analysis of the distribution of the minimums for all perfectly reliable configurations confirmed that minimum volume in the tank was smaller for configurations increasingly closer to the 100% reliability line. In other words, overengineering results in a larger minimum statistic. Recall that the most important consideration of this simulation was the distribution of the minimum and that engineering considerations that were 100% reliable but experienced low volume are rightfully of concern.

Design

The design phase of the process began with the decision regarding cistern volume and roof surface area. The roof size on the house was chosen to be 5,000 square feet for a final house size of 3,200 square feet. Given those dimensions, the house required only a 20,000-gallon cistern. But a 40,000-gallon cistern ensured the minimum in the tank (for all 238 iterations) would not fall below 25% of holding capacity, or 10,000 gallons. Due to the conservative nature of the owners, the 40,000-gallon option was chosen. The photo on p. 35 shows early construction of the final tank and recapture system.

Verify

Today, the cistern is completely full, and the overflow is of such quantity that a secondary cistern for non-potable water to support an above-ground garden is under consideration. The system is performing better than expected due to the recent volume of rainfall, and the engineering specifications appear to producing the quantity of water expected.

From this case, the relevance of DMADV and simulation to complex, new product development is evident. While this simulation is relatively simple, it certainly illustrates the art of the possible. Quality and reliability engineering of projects require algorithmic approaches, and DMADV certainly provides the necessary structure. QP

REFERENCES

1. Texas Water Development Board, *Texas Manual on Rainwater Harvesting*, third edition, 2005.
2. Colin McDonald, "Forum Discusses Water Challenges Facing SA," *San Antonio Express-News*, Oct. 6, 2010, www.mysanantonio.com/news/environment/article/Forum-discusses-water-challenges-facing-s-a-691164.php.
3. Jennifer Hiller, "Getting off the Grid," *San Antonio Express-News*, May 25, 2012, www.mysanantonio.com/real_estate/article/getting-off-the-grid-3431059.php.
4. Herman K. Trabish, "Texans Saving Energy and Cash-in Net-zero Homes," *GreentechMedia*, Feb. 24, 2012, www.greentechmedia.com/articles/read/texans-saving-energy-and-cash-in-net-zero-homes/#.
5. Steve Browne, "'Greenest' Home in U.S. located in Bulverde," KSAT 12 NewsTelevision, Feb. 3, 2012, www.ksat.com/news/-Greenest-Home-in-U-S-located-in-Bulverde/-/478452/8589544/-/kugici/-/index.html (case sensitive).
6. National Oceanic and Atmospheric Administration, NNDC Climate Data Online, 2012, www.ncdc.noaa.gov/cdo-web.
7. National Geographic, Water Footprint Calculator Methodology and Tips, 2011, http://environment.nationalgeographic.com/environment/freshwater/water-calculator-methodology.
8. Ibid.
9. Charles Harrell, Royce Bowden and Biman K. Ghosh, *Simulation Using Promodel*, New York, 2011.
10. Matt Basinger, Franco Moltanto and Upmanu Lall, "A Rainwater Harvesting System Reliability Model Based on Nonparametric Stochastic Rainfall Generator," *Journal of Hydrology*, Oct. 2010, pp. 105-118.
11. National Geographic, Water Footprint Calculator Methodology and Tips, see reference 7.
12. Texas Water Development Board, *Texas Manual on Rainwater Harvesting*, see reference 1.
13. ProModel Corp., ProModel Version 11, 2011, www.promodel.com.
14. J.S. Mun and M.Y. Han, "Design and Operational Parameters of a Rooftop Rainwater Harvesting System: Definition, Sensitivity and Verification," *Journal of Environmental Management*, Jan. 2012, pp. 147-153.

LAWRENCE V. FULTON is an assistant professor of quantitative methods at Texas State University in San Marcos. He has a doctorate from the University of Texas at Austin. Fulton is a member of ASQ and an ASQ-certified quality engineer and Six Sigma Black Belt.