12-2011

Modeling passive solar distillation production in Las Vegas, Nevada

Noe I. Santos
University of Nevada, Las Vegas

MODELING PASSIVE SOLAR DISTILLATION PRODUCTION IN LAS VEGAS, NEVADA

by

Noe I. Santos

Bachelor of Science in Engineering
University of Nevada, Las Vegas
2009

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science in Engineering

Department of Civil and Environmental Engineering
Howard R. Hughes College of Engineering
The Graduate College

University of Nevada, Las Vegas
December 2011

THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Noe I. Santos

entitled

Modeling Passive Solar Distillation Production in Las Vegas, Nevada

be accepted in partial fulfillment of the requirements for the degree of

Master of Science in Engineering
Department of Civil and Environmental Engineering

Aly Said, Committee Chair

David James, Committee Member

Sajjad Ahmad, Committee Member

Robert Boehm, Graduate College Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies
and Dean of the Graduate College

December 2011

ABSTRACT

Modeling Passive Solar Distillation Production in Las Vegas, Nevada

by

Noe I. Santos

Dr. Aly M. Said, Examination Committee Chair
Assistant Professor, Civil and Environmental Engineering
University of Nevada, Las Vegas

A study has been performed to examine the effects of daily weather on the performance of commercial solar distillation basins (solar stills). The objectives of this study were to evaluate the long term performance of solar stills, to instrument two solar stills and record sub-hourly thermal properties, to evaluate existing heat transfer modeling methods for hourly production, and to create new models to predict daily production using experimental distillate production and local weather data by utilizing artificial neural networks, genetic algorithms, and multivariate regression. A system dynamics model was also created to determine the required basin area and storage volume to produce enough water to meet year round potable water demand.

Solar still production was measured between January 2011 and September 2011. The average daily yield of solar still #1-A (SS1-A) and solar still #1-B (SS1-B) ranged from 2.11 ± 0.35 L/m^2 and 2.00 ± 0.46 L/m^2 (winter season) to 5.53 ± 1.01 L/m^2 and 5.64 ± 1.06 L/m^2 (summer season), respectively.

The artificial neural network model performed with a mean absolute error as low as 9.4% with up to 92.4% of production predictions within 0-20% of the actual daily production. The genetic algorithm model performed with a mean absolute error as low as 11% with up to 91% of production predictions within 0-20% of the actual daily

production. The multivariate regression model performed with a mean absolute error as low as 9.7% with up to 94.1% of production predictions within 0-20% of the actual daily production.

Analysis of the sub-hourly performance data indicated that large distilland volumes resulted in a greater proportion of production occurring during the night compared to smaller distilland volumes. Hourly temperature data was used to calculate heat transfer coefficients which could predict hourly distillate production with a mean absolute error between 26% and 53%.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my family and friends for their never ending support throughout my entire collegiate education. I am very grateful for all of the support and guidance that my advisors, Dr. David E. James and Dr. Aly M. Said, have provided. Dr. James and Dr. Said devoted tremendous time and effort to make this study a success. I also appreciate the help of Dr. Sajjad Ahmad and Dr. Robert Boehm for their assistance with system dynamics and heat transfer methods, respectively. I greatly appreciate the UNLV Center for Energy Research for providing the long term solar data that was used for this study. I would also like to thank the late Harold Hay for his passion for solar distillation and for financially supporting the original goals of this study. This study would not have been possible without his support.

I would also like to thank the United States Geological Survey and the National Institutes for Water Resources for their financial support. I also appreciate the assistance from Chris Carrier and Oscar Quiroz for their help in collecting daily solar still production data when I was not available. And last but certainly not least, I would like to thank the Western Alliance to Expand Student Opportunities (WAESO) for their funding to support undergraduate research assistants throughout this study. With support from WAESO, Khloe Campos, Eric Mata, Jessica Menchaca, Sean Robinson, and Pablo Vazquez were able to assist me with the data processing and allowed me to focus on the various model developments for this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

With anticipated future increases in energy costs, water purification processes such as multistage flash, multiple effect, vapor compression, reverse osmosis, electrolysis, phase change, and solvent extraction will see their price per unit of water increase drastically over time. Furthermore, rising energy prices will also increase the costs required for pumping desalinated water and transporting it to the desired location. One low cost, point-of-use alternative to energy-intensive approaches for purification of brackish, saline, or polluted waters is passive solar distillation (Fath, 1998). Solar distillation is a simple and clean technology that can be used to distill brackish/polluted water into drinkable water and can also be used to reduce the fossil fuel dependence that exists with current large scale desalination methods. Being able to predict solar still performance from long-term daily varying solar irradiance, air temperature, wind speed, wind direction, and cloud cover data could allow for the appropriate sizing of solar distillation facilities. This could allow for the determination of the correct level of investment needed to produce the correct amount of potable water to supply individuals or a community.

1.1 Overview

Solar stills can be considered to perform like a miniature watershed; producing varying amounts of potable water in response to fluctuating meteorological conditions. Adequate prediction of solar still output using conventionally obtainable meteorological data would allow for a cost effective and reliable design of a solar still system as a water supply. This would allow the system to generally produce sufficient yields to meet a community's potable water demand.

Solar stills could be widely implemented in Nevada because of its combination of ample sunlight and, in many locations, supplies of surface or groundwater that require treatment before becoming potable (State of Nevada Department of Conservation, 2009). These types of conditions also exist in many arid locations worldwide where groundwater supplies are contaminated with arsenic or fluoride. Worldwide, many people suffer chronic health impairments due to the inability to treat contaminated water effectively and economically. In the southwestern U.S., contaminated surface runoff or ground water could be purified for use as a potable supply, crop irrigation, or for landscape irrigation. Furthermore, in an urban location, solar stills could also be implemented into Leadership in Energy and Environmental Design (LEED) projects for onsite treatment of some types of wastewater and runoff. Doing so could qualify a project for credits towards United States Green Building Council (USGBC) accreditation.

According to the World Health Organization (2005), the ability to invest in drinking water and sanitary conditions worldwide would have a variety of different economic benefits in several forms. The benefit of having access to potable water could save $7 Billion (U.S.) a year for health agencies and $340 Million (U.S.) for individuals (World Health Organization, 2005). Furthermore, the world population would see 320 million productive days gained each year for individuals in the 15-59 years of age group, an additional 272 million school attendance days a year, and an additional 1.5 billion healthy days for people under 5 years of age. The value for such an improvement in living conditions and productivity is estimated to be worth $9.9 Billion (U.S) per year (World Health Organization, 2005). According to the World Health Organization (2005), their investment goal of $11.5 Billion (U.S.) per year would see a total payback of $84 Billion

(U.S.) per year.

Renewable energy systems are capable of producing energy from sources that are freely available and are also characteristically environmentally friendly (Kalogirou, 2005). Although renewable energy powered desalination systems cannot compete with conventional systems in terms of the cost per unit of water produced (at current fossil fuel prices), they are applicable in certain areas and are likely to become accepted as a feasible solution in the near future (Kalogirou, 2005).

1.1.1 Passive Solar Distillation

A review on solar distillation technologies and costs concluded that, because of fairly low energy fluxes from sunshine, space requirements for solar stills are high compared to other technologies (Kalogirou, 2005). At the current state of solar still technology, for daily water yields ranging from 1 to 7 L/m^2, a medium sized community requiring 200 m^3/day (Kalogirou, 2005) would require 3 to 20 hectares of still area. Due to the high capital costs involved with solar distillation, primarily due to land and equipment costs, accurate prediction of daily production is vital to the success of a new system. Accurate predictions would allow for the optimization of capital expenses and would minimize the risk involved with such an investment.

Research into solar distillation goes as far back as the fourth century B.C. when Aristotle described a method to evaporate and condense polluted water for potable use (Tiwari, Singh, & Tripathi, 2003). However, the earliest documented work on solar distillation came from Arab alchemists in the 16^{th} century (Tiwari et al., 2003) and later with the use of wide earthen pots exposed to the sun (Tiwari et al., 2003).

Conventional means of desalination such as steam distillation and reverse osmosis

both require significant quantities of energy to separate sea salt and water. Due to the high recurring energy costs to perform desalination, few of the water short areas of the world, besides some countries in the Middle East that have enough money to perform desalination due to oil income, can afford conventional desalination approaches (Kalogirou, 2005).

Foster, Eby, and Amos (2005) have been applying passive solar distillation in the colonias, unorganized/incorporated communities with limited infrastructure development, along the U.S.-Mexico border for ten years. Foster et al.'s (2005) main effort was to develop solar distillation technology and apply it as an effective method to purify water and meet potable water demands. Foster et al. (2005) have also performed water quality tests and have analyzed the economics and payback periods for installing a new solar still system. Over the course of their work, Foster et al. (2005) have acquired one of the most extensive solar still daily production datasets available.

1.1.2 Single Effect Solar Still

Figure 1.1 illustrates a schematic example of a single basin solar still that was operated by Venkatesh (2007) between February 2006 and July 2007. Single basin solar stills are a common solar device capable of converting brackish/polluted water into potable water by utilizing solar energy (Kabeel & El-Agouz, 2011). Since single basin solar stills rely on passive methods to produce distilled water, a variety of different factors impact performance and efficiency. These factors include location, orientation, solar radiation intensity, ambient temperature, depth of water in the basin, glass cover material, insulation materials and their respective thickness, inclination angle of the glass cover, and the heat capacity of the still (Kabeel & El-Agouz, 2011).

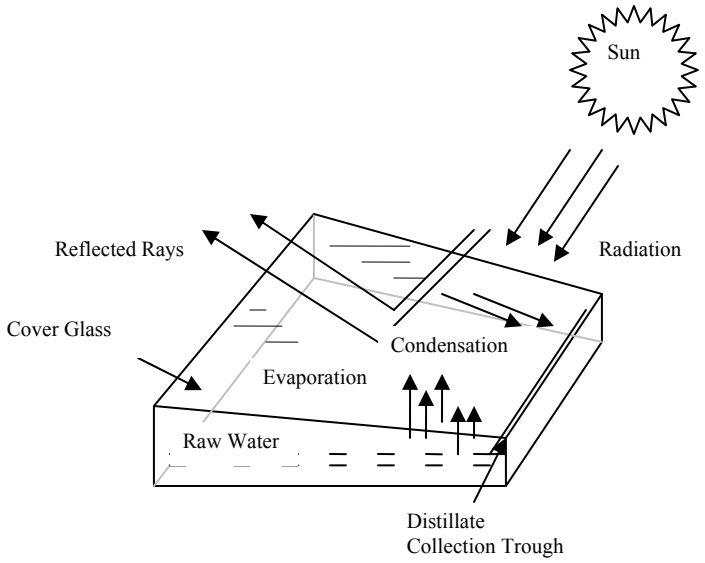


Figure 1.1 Example schematic of a single basin solar still (Venkatesh, 2007)

Out of these factors, the solar radiation intensity is the main supplier of energy into the solar still system. Because of this, the angle of the glass cover plays an important role in terms of transmitting and reflecting the sun's rays (Kabeel & El-Agouz, 2011). It is usually recommended that the glass cover inclination angle be equal to the latitude of the location to allow the transmittance of the Sun's rays normal to the glass cover throughout most of the year (Kabeel & El-Agouz, 2011). Besides the angle of the cover, the cover material plays a role with regards to light transmittance, service life, and ease of use. Glass is the preferred material due to its high transmittance and long service life; however, due to the fragile nature of glass, polyethylene (plastic) covers may be used where transporting glass covers proves to be too hazardous (Kabeel & El-Agouz, 2011).

The temperature differences that occur between the solar still and the environment are a key driving force with regards to hourly and daily production (A.K. Tiwari & G.N. Tiwari, 2006). As the temperature difference between the distilland (water in the basin undergoing distillation) and the glass cover increases, the circulation of air and vapor

inside the solar still increases as well. As a result of this process, the evaporative and convective heat transfer between the distilland and the surface cover increases (Kabeel & El-Agouz, 2011).

The depth of water in the basin also has an effect on the productivity of the solar still and has been found to be inversely proportional to the daily productivity (Kabeel & El-Agouz, 2011). Kabeel and El-Agouz (2011) also state that the deeper the distilland in the basin, the less susceptible it will be to a drop in temperature due to sudden solar intensity variation as a result of changes in cloud coverage for short periods of time. In situations where there is a decrease in solar intensity, the energy stored in the distilland is released allowing for production to remain continuous. This same phenomenon allows for solar stills with greater distilland depths to continue to produce water after sunset and throughout the night (Kabeel & El-Agouz, 2011).

1.2 Solar Still Performance Modeling

Kalogirou (2005) cited previous research studies that predicted solar still performance including computer simulation (Cooper, 1969), thermic circuit and sankey diagrams (Frick, 1970), periodic and transient analysis (Sodha, Navak, Tiwari, & Kumar, 1989; Tiwari, & Rao, 1984), iteration methods (Toure & Meukam, 1997), and numerical methods (Sartori, 1987; Log, Eibling, & Blowemer, 1961).

Despite the different numerical techniques, all of the above cited methods rely on mechanistic, internal heat transfer models which were first applied and published by Dunkle in 1961 and subsequently revisited by other researchers such as Tiwari and Tiwari (2006). However, use of these models usually requires simplifying assumptions regarding the relative magnitude of several components of heat transfer.

The heat transfer model for solar distillation, as developed by Jakob (1949), applied by Dunkle (1961) and later converted to S.I. units by Tiwari and Tiwari (2006), relies on many variables such as density, specific heat, thermal conductivity, kinematic viscosity, coefficient of thermal expansion, latent heat of vaporization, partial saturated vapor pressure of water, and heat transfer coefficient relationships derived from convective heat transfer research in the 1920s and 1930s (Jakob, 1949; McAdams, 1954; Fishenden, 1957). Experimental correlations are derived from intensive data logging of the solar still's thermal characteristics including the outer glass temperature, inner glass temperature, vapor temperature, distilland temperature, internal solar still humidity, distillate output, and also environmental data such as ambient air temperature, ambient air velocity, and total and diffused radiation (Dunkle, 1961; Tiwari & Tiwari, 2006). Due to the large amount of high temporal resolution data needed to validate a heat transfer distillation model, the ability to accurately forecast distillate production is limited by the ability to measure the variables needed to determine experimental correlations for the heat transfer model. While the heat transfer model has been used successfully in the past, the amount of time, data storage, and the complexity of the calculations may put this approach out of reach in many parts of the developing world.

1.2.1 Dunkle and Tiwari's Heat and Mass Transfer Methods

Dunkle (1961) applied Jakob's (1949) derived empirical relationships, as cited in McAdams (1954), for internal heat transfer processes that occur between parallel plates to solar distillation. The basic equation for convective heat transfer is $Nu = C \cdot [Gr \cdot Pr]^n$, where Jakob (1949) has found "n" to be 1/3 for turbulent flows and "C" to be 0.075 for dry air convection between parallel horizontal plates with the lower plate being

warm and the upper plate cool. Dunkle (1961) simplified this relationship for a limited range of values where the Grashof number is between 3.2 x 10^5 and 1 x 10^7. Equation 1.1 details Jakob's (1949) original relationship between the Nusselt, Grashof, and Prandtl numbers.

$$Nu = \frac{h_c' \cdot x}{k_f} = 0.075 \cdot \left[\frac{(x^3 \cdot g \cdot \beta \cdot \Delta T')}{V_f^2} \cdot Pr_f\right]^{\frac{1}{3}}$$

Equation 1.1 Jakob's (1949) relationship between Nusselt, Grashof, and Prandtl numbers

Dunkle revised Equation 1.1 by assuming an operating temperature of 30°C to evaluate the various temperature based properties of dry air. The simplified form resulted in the expression of the convective heat transfer coefficient as a function of $\Delta T'$. Equation 1.2 shows how Dunkle (1961) evaluated his $\Delta T'$ term.

$$\Delta T' = (T_w - T_g) + \frac{T_w \cdot (P_w - P_g)}{39 - P_w}$$

Equation 1.2 Dunkle's (1961) $\Delta T'$ term evaluation

h_c'	Convective heat transfer coefficient (BTU/Hr-ft^2-°F)	$\Delta T'$	Adjusted temperature difference (R)
x	Distance between surfaces (ft)	Pr_f	Prandtl Number
k_f	Thermal conductivity (BTU/Ft-Hr-°F)	T_w	Distilland temperature (R)
V_f	Kinematic Viscosity (ft^2/s)	T_g	Inner glass temperature (R)
g	Gravitational acceleration (ft/s^2)	P_w	Partial pressure of water vapor (psi)
β	Expansion Coefficient (1/R)	P_g	Partial pressure of water vapor at inner glass temperature (psi)

Dunkle (1961) simplified Equation 1.1 by evaluating the properties of dry air at 30°C Equation 1.3 shows the simplification performed by Dunkle (1961) by evaluating the thermal conductivity, kinematic viscosity, expansion coefficient, and the Prandtl number at 30°C.

$$0.075 \cdot k_f \cdot \left[\frac{(g \cdot \beta)}{V_f^2} \cdot Pr_f\right]^{\frac{1}{3}} = 0.128$$

Equation 1.3 Dunkle's (1961) simplified relationship for the heat transfer coefficient

The coefficient in Equation 1.1 changed from 0.075 to 0.128 as a result of evaluating the various temperature based properties of air at 30°C, as shown in Equation 1.3. As a result of simplifying Equation 1.1 by evaluating certain air properties, Dunkle (1961) created a correlation for convective heat transfer (h_c') solely in terms of the $\Delta T'$ term. Dunkle's (1961) fully simplified convective heat transfer function is shown as Equation 1.4.

$$h_c' = 0.128 \cdot \left[(T_w - T_g) + \frac{T_w \cdot (P_w - P_g)}{39 - P_w}\right]^{1/3}$$

Equation 1.4 Dunkle's (1961) simplified relationship for the heat transfer coefficient

According to Tiwari and Tiwari (2006), Dunkle's (1961) simplified relationship has basic limitations and is only valid for cavities that have parallel evaporative and condensing surfaces, coefficients must be independent of the cavity volume, internal temperature conditions must be low or within 45°-50°C, and the temperature difference between evaporative and condensing surfaces must be within 17°C (Tiwari & Tiwari, 2006).

Dunkle's (1961) simplified relationship was converted for use in metric units by Tiwari and Tiwari (2006) and is presented in Equation 1.5. Equations 1.6 and 1.7 were reported by Tiwari and Tiwari (2006) to estimate evaporative heat transfer and hourly production from the convective heat transfer coefficient relationship developed by Dunkle (1961). In this thesis, the author independently calculated the source of the coefficients in Equations 1.4 and 1.5 and found them to be correct when using values for dry air at 30°C.

$$h_{cw} = 0.884 \cdot \left[(T - T_{ci}) + \frac{(T + 273) \cdot (P - P_{ci})}{268.9\ x\ 10^3 - P}\right]^{\frac{1}{3}}$$

Equation 1.5 Tiwari and Tiwari's (2006) version of Dunkle's (1961) simplified relationship for the heat transfer coefficient in S.I. Units

$$h_{ew} = 0.01623 \cdot h_{cw} \cdot \frac{P - \varphi \cdot P_{ci}}{T - T_{ci}}$$

Equation 1.6 Evaporative heat transfer coefficient relationship reported by Tiwari and Tiwari (2006)

$$m_{ew} = \frac{h_{ew} \cdot (T_w - T_{ci}) \cdot A_w \cdot t}{\Delta h_v}$$

Equation 1.7 Mass transfer (kg) relationship reported by Tiwari and Tiwari (2006)

P_{ci}	Partial saturated vapor pressure at condensing surface ($N/m^{2)}$	P	Partial saturated vapor pressure at evaporation surface (N/m^2)
T_{ci}	Inner condensing cover temperature (°C)	T	Water Temperature (°C)
φ	Relative humidity	Δh_v	Enthalpy of evaporation of water (J/kg)
t	Time (seconds)	A_w	Evaporative surface area (m^2)
Gr	Grashof number	Pr	Prandtl number
L_v	Dimension of condensing cover (m)		

Tiwari and Tiwari's (2006) h_{cw} is the metric form of Dunkle's (1961) h'_c. Tiwari and Tiwari's (2006) method for determining the convective heat transfer coefficient is performed by substituting Equation 1.6 into Equation 1.7 and simplifying. Further simplification is accomplished by substituting the relationship between h_{cw} and the Grashof and Prandtl numbers as shown in Equation 1.8. The final simplified equation is shown as Equation 1.9. Equation 1.9 was used by Tiwari and Tiwari (2006) to determine values of "C" and "n" for Tiwari and Tiwari's (2006) particular solar still design.

$$h_{cw} = \frac{\lambda}{L_v} \cdot C \cdot (Gr \cdot Pr)^n$$

Equation 1.8 Convective heat transfer coefficient's relationship to Grashof and Prandtl number

$$\frac{m_{ew}}{R} = C \cdot (Gr \cdot Pr)^n; where\ R = \frac{0.01623}{\Delta h_v} \cdot \frac{\lambda}{L_v} \cdot A_w \cdot t \cdot (P_w - \varphi \cdot P_{ci})$$

Equation 1.9 Distillate output's relationship to Grashof and Prandtl numbers

When analyzing experimental data with Equation 1.9, the constants "C" and "n" can be found by computing Gr and Pr and taking the logarithm of both sides of Equation 1.9. Data for $\mathrm{Ln}\left(\frac{\mathrm{m_{ew}}}{\mathrm{R}}\right)$ is then plotted against $\mathrm{Ln}(\mathrm{Gr} \cdot \mathrm{Pr})$ to fit it to the standard linear equation form y = mx + b. "C" and "n" can then be found by using least squares regression. The corresponding equation would result with the following.

$$y = ln\left(\frac{m_{ew}}{R}\right), b = ln(C), x = ln(Gr \cdot Pr), and\ m = n$$

Once the regression analysis is completed, "*n*" is equal to the coefficient for the slope of the linear function and "*C*" is equal to the exponential value of the y-intercept ($\exp^b$).

It is a well known fact, in the field of distillation, that the daily still production increases as the temperature of the distilland increases and condensing cover temperature decreases (Tiwari & Tiwari, 2006). Furthermore, the temperature difference between the distilland and condensing cover (ΔT) also plays a role with the convective mass transfer.

A study performed by Tiwari and Tiwari (2006) indicated that the methods for calculating h_{cw} and h_{ew} with Dunkle (1961) and Tiwari and Tiwari's (2006) methods are close in agreement as long as the distilland temperature was less than 50°C. Tiwari and Tiwari's (2006) method involved solving for new "C" and "n" values for their own solar still since their still's glass cover had a slope of 30°; which is significantly different from the assumption of parallel plates used by Dunkle (1961).

A maximum deviation of 75% and 75.6% was found for h_{cw} and h_{ew}, respectively, between the methods of Dunkle (1961) and Tiwari and Tiwari (2006) during the early morning when distilland temperatures were at their lowest. The mornings values

notwithstanding, the values for h_{cw} and h_{ew} were in close agreement. Tiwari and Tiwari (2006) calculated h_{cw} to vary between 1.4 to 1.8 W/m^2-°C while Dunkle's method calculated h_{cw} to vary between 0.4 to 2 W/m^2-°C while operating a solar still at a depth of 4 cm over the course of one day. While Dunkle's (1961) and Tiwari and Tiwari's (2006) method calculated different results for h_{cw}, both methods calculated h_{ew} to vary between 1 to 22 W/m^2-°C over the course of one day and were close in agreement.

1.2.2 Improved Heat and Mass Transfer Correlations

Because of the various approaches to modeling heat and mass transfer, the results from different studies are considered dissimilar and may even contradict each other (Hongfei, Zhang, Jing, & Yuyuan, 2002). According to Hongfei et al., one of the most noticeable issues with the original Dunkle (1961) model is that the approximating assumptions that were made to establish his correlation prevents the model from being applicable in other operating scenarios. Some of these approximations include assuming 50°C as the average air-vapor temperature, assuming a 17°C temperature differential between the distilland and surface cover, and not including the characteristic dimension (x_1) in the Nusselt correlation (Hongfei et al., 2002). Another issue that prevents the adoption of any one set of values for heat and mass transfer models (HMTs) is the assumption that the relationship between convective and evaporative heat transfer coefficients will change as the temperature conditions change (Hongfei et al., 2002).

One improvement for the HMT method was performed by Chen, Ge, Sun, & Bar (1984) using the following empirical relationships.

$$h_c = 0.2 \cdot Ra^{0.26} \cdot \frac{k_f}{x_1} \qquad (3.5x10^3 < R_a < 10^6)$$

Equation 1.10 Convective heat transfer coefficient (Chen et al., 1984)

$$h_m = \frac{h_c}{p_f \cdot C_{paf} \cdot Le^{1-n}}$$

Equation 1.11 Convective mass transfer coefficient (Chen et al., 1984)

$$m_e = h_m \cdot (p_w - p_g)$$

Equation 1.12 Evaporation rate per unit area of evaporation surface (Chen et al., 1984)

k_f	Thermal conductivity of air (W/m-°C)
x_1	Characteristic space between evaporation and condensation surfaces (m)
Ra	Rayleigh number = (Gr · Pr)
h_m	Convection mass transfer coefficient (m/s)
h_c	Free convective heat transfer coefficient (W/m^2-°C)
p_f	Density of humid air (kg/m^3)
C_{paf}	Heat capacity of humid air (J/kg-°C)
Le	Lewis number
p_w	Saturation vapor pressure of water at evaporation surface (N/m^2)
p_g	Saturation vapor pressure of water at condensation surface (N/m^2)
n	1/4 ($1x10^4$ < Gr < $2.51x10^5$), 1/3 ($2.51x10^5$ < Gr < $1x10^7$)

Hongfei et al. (2002) verified the model presented by Chen et al. (1984) through a series of experiments using an indoor multi-stage stacked tray distiller with an electrical heater as a heat source. Hongfei et al. (2002) found a mean difference of 8.5% between the experimental and calculated yields by using the model presented by Chen et al. (1984).

Hongfei et al.'s 2002 study was performed indoors for two months where the ambient conditions did not fluctuate as they would in an actual field setting. The study performed by Hongfei et al. (2002) concluded that h_c was proportional to $Ra^{0.26}$ and the incorporation of the characteristic size (x_1) overcomes the shortcomings of Dunkle's (1961) study.

1.2.3 Computational Fluid Dynamics

Setoodeh, Rahimi, & Ameri (2011) created a three dimensional, two phase model for

evaporation and condensation processes in a solar still using computational fluid dynamics (CFD) at quasi steady state conditions. The main foundation of their study was based on numerical modeling of unsteady forms of continuity, momentum, energy, volume conservation, and mass transfer equations. Because of the complexity of the model set up and the input requirements, the computation time per simulation would last 4 to 12 hours using parallel processing with eight 2.76 GHz and two 3.00 GHz processors.

Setoodeh et al. (2011) found that the CFD model averaged errors of 7.79% and 14.48% for the production rate and water temperature simulation, respectively. The work performed by Setoodeh et al. (2011) calculated temperature and production rates that were in close agreement to experimental data.

1.2.4 Nocturnal Production of a Solar Still

Malik & Tran (1972) performed a study in which hot feed water was introduced into a solar still during night time operation. By doing so, Malik & Tran (1972) found the effects of several parameters on the night time production. These parameters include the initial distilland temperature, the drop in distilland temperature, and the distilland depth.

The Malik & Tran (1972) study relied on mathematical methods to create the necessary heat transfer relationships to model night time production. This particular study also used the modified relationships developed by Dunkle (1961) to estimate the various heat transfer coefficients. The Malik & Tran (1972) study calculated convective heat loss coefficients, radiation heat loss from the cover to the sky, and conduction heat loss to the ground. Malik & Tran (1972) developed a simplified mathematical model which linearizes the heat transfer relationships and calculates a closed form solution for the

nocturnal output as a function of initial distilland temperature and the drop in distilland temperature.

After two months of experimental testing, Malik & Tran (1972) concluded that night time production of a solar still is affected by the ambient air temperature, wind velocity, water depth, initial distilland temperature, width of the solar still, and time. Malik & Tran (1972) found that lower values of relative humidity, sky temperature, ambient temperature, and high wind velocity will cool the distilland quicker. Moreover, the higher the aforementioned variables are, the slower the distilland will be cooled. However, the amount of distillate produced by cooling the distilland by a certain amount remains the same regardless of the environmental conditions. Malik & Tran (1972) concluded that the environmental factors merely change the rate at which distillate is produced instead of affecting the quantity.

Mathioulakis, Voropoulos, & Belessiotis (1999) also developed a model to predict nocturnal production by utilizing the temperature of the distilland as the main variable. The energy balance of the system as developed by Mathioulakis et al. (1999) follows the premises that the reference period is imposed by the variation of solar energy and coincides with the daily energy cycle, the sum of energy inflows equals the sum of energy outflows minus the environmental losses, the environmental losses can be expressed as a function of the temperature differences between the solar still and the external system, and the produced water represents the outflows from the system. Mathioulakis et al. (1999) developed the following relationships.

$$M_{w,d} = f_{1,d} \cdot H_d + f_{2,d} \cdot \left(T_{w,d} - T_{a,d}\right) + f_{3,d}$$

Equation 1.11 Day water production (Mathioulakis et al., 1999)

$$M_{w,n} = f_{2,n} \cdot \left(T_{w,n} - T_{a,n}\right) + f_{3,n}$$

Equation 1.12 Nocturnal water production (Mathioulakis et al., 1999)

d	Day subscript	T_w	Distilland temperature
n	Night subscript	T_a	Ambient air temperature
f	Linear regression constant	H	Daily Insolation

Mathioulakis et al. (1999) created a linear least squares regression model to solve for the constants (f). However, as is normal for other solar still modeling methods, the implementation of this method requires hourly meteorological data at the minimum. Furthermore, the ability to execute this particular modeling method requires the ability to predict the distilland temperature. To resolve this issue, Mathioulakis et al. (1999) also created a relationship that expresses the energy level of a still on any day (k) as a result of the cumulative contribution of energy from the previous days (n). The results from this approach to predicting distillate temperature show that average daily water temperature can be satisfactorily expressed in relation to the main meteorological data from the day, night, and the previous day (Mathioulakis et al., 1999).

Mathioulakis et al. (1999) stated that the calculated error in yield involved with the method was typically small and did not exceed 1.5 kg. Furthermore, Mathioulakis et al. (1999) also studied the effect of the data set size on the total error. It was found that after 80 days of data, the error involved through this particular modeling method was around 0.01%. Daily errors were not systematic and therefore positive and negative deviations were compensated over a period of time (Mathioulakis et al., 1999).

1.3 Solar Still Performance Enhancements

Most of the current research in solar distillation has focused on modifying the solar still design to introduce components that would allow water to either evaporate or condense faster. Some of the modifications that have been studied include using internal

and external condensers (Fath & Elsherbiny, 1993; Fath & Elsherbiny, 2004; Rubio, Porta, Fernandez, 2000), using black walls with cotton cloth (Fath, 1998), the use of black dye and charcoal in the distilland (Tiwari, Gupta, & Lawrence, 1989; Morcos, 1993; Dutt, Kumar, Anand, & Tiwari, 1989), multi-wick solar stills (Sodha, Kumar, Tiwari, & Tyagi, 1981; Tanaka, 2011; Murugavel & Srithar, 2011), double basin solar stills (Dutt et al., 1989; Sodha, Nayak, Tiwari, & Kumar, 1980; Murugavel & Srithar, 2011) and condensing cover cooling (Abu-Hijleh, 1996; Dhiman & Tiwari, 1990).

The main focus of enhancement studies is to increase the thermal efficiency and the daily production through either active or passive methods (Fath, 1998). Some of the methods used to enhance performance include lowering the depth of water in the basin (Murugavel, Sivakumar, Ahamed, Chockalingam & Srithar, 2010; Tiwari & Tiwari, 2006, Murugavel & Srithar, 2011), reducing heat losses through the walls of the still (Mohamad, Soliman, Abdel-Salam, & Hussein, 1995), and using materials to darken the water mass (Fath, 1998; Morcos, 1993; Nafey, Abdelkader, Abdelmotalip, & Mabrouk, 2002). The design of the surface cover can also be modified either by utilizing a thinner cover (Kabeel & El-Agouz, 2011), optimizing the glass cover angle (Tanaka, 2010; Akash, Mohsen, & Nayef, 2000) or by using a double slope solar still (Garg & Mann, 1976; Al-Karogilou & Alnaser, 2004; Tiwari & Rao, 1984; Rubio et al., 2000; Murugavel & Srithar, 2011). Through various studies, double slope solar stills were found to maximize the absorbed solar radiation; however, single sloped double stills experience less convection/radiation losses and perform better in cold conditions (Fath, 1998).

1.3.1 Flat Plate Collectors and Energy Storage

As previously discussed, the greater the temperature difference between the distilland

and the surface cover, the greater the hourly production due to increased convective and evaporative heat transfer. Because of this phenomenon, past research studies have experimented with various ways to increase the temperature of the distilland by using either flat plate collectors or solar reflectors (Kumar & Sinha, 1995; Tanaka, 2011; Morcos, 1993; Esfahani, Rahbar, & Lavvaf, 2011).

In a study performed by Morcos (1993), several methods of performance enhancement were tested including the addition of black dye to the distilland and coupling a flat plate collector to a still with black dye operating on the thermosyphon circulation mode. The results of this study showed the addition of a flat plate collector with thermosyphon circulation performed the best along with the addition of blackened jute cloth floating in the distilland and a small addition of black dye.

The use of energy storage materials was investigated by Murugavel et al. (2010) in a study which compared solar still performance with the addition of quartzite rock, red brick pieces, concrete pieces, washed stones, and iron scraps. Out of all these materials, it was found that 3/4" quartzite rock was the most effective at increasing daily yields. Murugavel et al. (2010) reported a 4.8% increase in daily production by using the 3/4" quartzite rock over not using any energy storage materials. Sakthivel, Shanmugasundaram, & Alwarsamy (2010) utilized jute cloth to increase the evaporation surface area. The use of jute cloth resulted in a 12% increase in daily yield compared to a conventional still without jute cloth (Sakthivel et al., 2010). Sakthivel et al. (2010) points out that jute cloth is a low cost material which is easily installed.

El-Sebaii, Yaghmour, Al-Hazmi, Faidah, Al-Marzouki, & Al-Ghamadi (2009) also performed a study on energy storage materials to enhance still production by utilizing

locally available sand. The results of this study showed that efficiency increased from 27% to 37.8% and daily yield increased from 2.85 kg/m^2 to 4.00 kg/m^2 (40% increase).

1.3.2 Optimizing Surface Cover Inclination

The ability of a solar still cover to maximize light transmission and minimize light reflection allows the still to produce an optimum amount of distillate. Akash et al. (2000) performed a study that varied the glass angle of a double slope solar still in Amman, Jordan (31.95° N, 35.93° E). The study by Akash et al. (2000) tested daily solar still production at various cover angles (15°, 25°, 35°, 45°, 55°) during the month of May. Akash et al. (2000) discovered that the peak distilland temperature occurred between 13:00 and 14:00 and the maximum distillate production occured with a cover angle of 35°. This result is in agreement with the practice of having the glass cover inclined at an angle similar to the latitude of the region.

A similar study by Tanaka (2010) found that the optimum cover angle was 10° during the summer season and 50° during other seasons in Kurume, Fukuoka, Japan (33.32° N, 130.52° W). Khalifa (2011) performed a review on multiple glass cover scenarios done by other authors. According to the review done by Khalifa (2011), it was found that out of 20 different studies, 10 of the studies concluded that a glass angle of approximately 30° optimizes solar still production in countries between 30° - 40° N latitude. The remaining studies claimed that smaller cover inclination angles, around 10° - 20° would optimize production depending on the season of operation.

1.3.3 Distilland Depth Effects

Khalifa & Hamood (2009) performed a summary review detailing 24 past studies which analyzed the effect of distilland depth on daily solar still yield. Khalifa &

Hamood's 2009 study found that all 24 studies showed a significant decrease in daily yield as the depth of the distilland increased. Their studies experimented with distilland depths from 0.5 cm to 30 cm, covered locations between 32° S and 31° N of the equator, and occurred through various seasons.

Khalifa & Hamood (2009) also introduced a correlation to predict daily yield based on the distilland depth. They derived the correlation based on data available from the reviewed studies. Khalifa & Hamood's (2009) correlation is shown below.

$$P = 3.259 \cdot Depth^{-0.19}$$

Equation 1.13 Daily yield (L/m^2) as a function of distilland depth (cm) (Khalifa & Hamood, 2009)

From the nature of the equation, it can be seen that estimated daily production is inversely proportional to the distilland depth. The developed correlation was developed for double/single slope solar stills with cover angles between 10° and 35° and distilland depths between 0.5 cm and 30 cm. Developed from a variety of published studies, this particular function had with a coefficient of determination (R^2) of 0.129 (12.9% of variance is accounted for by the distilland depth).

1.4 Objective and Scope of Work

The purpose of this study was to measure the long term performance of two identical solar stills and to evaluate several different modeling methods to predict daily solar still yield using local weather data. From the literature review, it can be seen that the current state of the art in solar still modeling methods is mainly founded on heat and mass transfer (HMT) methods. While these methods have seen decades of research, application and retooling, they are still difficult to apply to predict long term solar still production which could allow for the adoption of solar stills as a true water resource worldwide.

Authors such as Mathioulakis et al. (1999) and Khalifa & Hamood (2009) have researched and published correlations that attempt to move away from the necessity of heat and mass transfer methods. As long as heat and mass transfer methods are used, there will be a necessity to instrument solar stills with temperature sensors and data logging equipment to perform heat transfer modeling. In order to make solar stills a more attractive method for water purification, there needs to be an effective method to accurately predict solar still production by using widely accessible meteorological data and simple/reliable modeling methods. The objectives of the proposed research include:

1. Operate two identical solar stills and record daily distillate production and sub-hourly distillate production and thermal properties associated with each respective still.
2. Acquire solar still production data from a 2006-2007 study by Venkatesh (2007) and create two artificial intelligence models by using Artificial Neural Networks (ANNs) and Genetic Algorithms (GAs)
3. Validate ANN and GA models developed with Venkatesh (2007) with new daily production data collected between January 2011 and September 2011
4. Compare the quality of predictions for ANNs and GAs to predictions from conventional multivariable regression
5. Verify the heat transfer model developed by Jakob (1949) as applied by Dunkle (1961) and converted to S.I units by Tiwari and Tiwari (2006)
6. Evaluate a stored heat model developed from Mathioulakis et al. (1999).
7. Develop a System Dynamics model to determine storage volume needs in response to predicted fluctuations in still yield based on local weather, user

demand, basin area, and reservoir size.

1.5 Research Significance

The work performed in this study will allow for the development of several solar still production models that will allow for the accurate determination of daily distillate yield by using easily accessible weather data. Use of these models could reduce the risks involved with investing in a solar distillation system and could, therefore, make solar distillation a more reliable water resource and potentially increase its worldwide adoption.

1.6 Organization of Thesis

This thesis is organized into six chapters. The first chapter contains an introduction and literature review on passive solar stills, heat and mass transfer models, and performance enhancements to increase distillate yield.

The second chapter includes the details regarding the materials and methods involved for the current 2011 study.

The third chapter contains an overview of the collected data for the current 2011 study as well as an analysis of the hourly and daily data for distillate production and the thermal characteristics of the two solar stills.

The fourth chapter focuses on the artificial neural network (ANN) modeling method and verifying the model built from the 2006-2007 data with the newly collected 2011 study data.

The fifth chapter focuses on the genetic algorithm (GA) and multivariable regression (MVR) modeling methods and verification of the models built with the 2006-2007 data with the newly collected 2011 study data. The fifth chapter will also cover a system

dynamics (SD) model and the sensitivity of a solar still system to various variables.

Finally, the sixth chapter presents conclusions, recommends the best models based on accuracy and ease of use, and makes recommendations for future additional work.

CHAPTER 2

MATERIALS & METHODS

2.1 Overview

This study examines the performance of solar distillation basins from data collected during two separate testing periods. The first data collection period ran between February 2006 and July 2007 using two different commercial solar distillation basins known as Solar Still #1 (Sunwater) and Solar Still #2 (SolAqua) (Venkatesh, 2007). During this data collection period, daily collected distilled water production was measured and the effect of varying distilland volume on daily production was examined. One solar still was instrumented with a rain gauge data logger to measure and record the real time production of distilled water. Two stills were also instrumented with HOBO® pendant temperature data loggers to measure distilland temperature. In this study, distillate refers to the water produced as a result of evaporation (end product) and distilland refers to the water in the basin undergoing distillation.

The second data collection period ran between January 2011 and September 2011 and used two identical models of Solar Still #1 (SS1-A and SS1-B) that were used during the previous study. Solar Still #2 (SS2) was not used in 2011 due to issues with excessive vapor leaks through its seals. During this data collection period, daily collected distilled water production was measured and the effect of varying distilland volume on daily production was also examined.

SS1-A and SS1-B were both instrumented with rain gauge data loggers to measure real time distillate production and YCT brand temperature data loggers to measure distilland, vapor, inner glass, and outer glass temperatures at five minute intervals.

HOBO® pendant temperature and light intensity data loggers were also used to measure and record the light intensity and ambient temperature conditions on site.

Tables 2.1 and 2.2 summarize the monthly average weather conditions that were experienced during the 2006-2007 and 2011 studies, respectively. Historic data were obtained from the McCarran International Airport National Weather Service (NWS) station via the Weather Underground (2011) online database. Hourly cloud cover data were converted to daily values by averaging the hourly cloud cover data.

Table 2.1 Weather conditions at McCarran International Airport, 2011 study period

Month	Average Temperature (°C)	Average Wind Speed (m/s)	Average Wind Direction (Degrees from North)	Average Cloud Cover (Fraction of Sky Area)
Feb-06	12.2	2.68	132	0.357
Mar-06	12.8	3.58	144	0.471
Apr-06	18.9	4.02	146	0.291
May-06	27.2	3.13	146	0.308
Jun-06	32.2	3.13	154	0.253
Jul-06	34.4	3.13	149	0.413
Aug-06	32.8	3.58	147	0.189
Sep-06	27.2	3.13	135	0.243
Oct-06	20.0	2.68	143	0.301
Nov-06	15.0	2.68	128	0.365
Dec-06	8.9	2.68	124	0.315
Jan-07	7.8	3.13	132	0.208
Feb-07	12.8	3.58	132	0.352
Mar-07	18.3	3.13	141	0.311
Apr-07	21.7	4.02	138	0.391
May-07	27.2	3.58	142	0.231
Jun-07	32.2	3.58	153	0.211
Jul-07	35.6	3.13	139	0.379

During the 2006-2007 study, July was the month with the highest average

temperature and January was the month with the lowest average temperature. The highest daily average wind speed was observed during the month of April and the winds predominantly originated from the southeast of the Las Vegas valley.

Table 2.2 Weather conditions at McCarran International Airport, 2011 study period

Month	Average Temperature (°C)	Average Wind Speed (m/s)	Average Wind Direction (Degrees from North)	Average Cloud Cover (Fraction of Sky Area)
Jan-11	12.0	3.26	114	0.292
Feb-11	10.4	4.15	121	0.337
Mar-11	16.1	4.44	123	0.433
Apr-11	19.3	5.32	116	0.333
May-11	22.3	5.34	123	0.317
Jun-11	30.0	5.20	131	0.171
Jul-11	33.2	4.43	134	0.320
Aug-11	34.6	4.72	143	0.240
Sep-11	29.0	3.09	115	0.306

During the 2011 study, August was the month with the highest average temperature and February was the month with the lowest average temperature. The highest daily average wind speed was observed during the month of May and the winds predominantly originated from the southeast of the Las Vegas valley.

2.2 Site Conditions

The test site for this study is located on the roof of the Howard R. Hughes College of Engineering building at the University of Nevada, Las Vegas (UNLV) (36.11°N, 115.142°W, 617 meters (2,024 ft) above sea level). The solar stills were located on the south side of the roof which offered a clear view of the southern sky. This location was chosen after the research team built a platform out of concrete pavers to prevent damage

to the roofing material. There is an 4 m (13 ft) tall utility structure directly to the north of the solar still test site which shields the solar stills from northerly winds; however, there is no other interference caused by this utility structure. Figure 2.1 and 2.2 illustrate the panoramic southern and northern view as seen from SS1-B and the roof parapet, respectively. Any curvature shown in Figures 2.1 and 2.2 is a result of the automated stitching used by the digital camera to create a panoramic view.

Figure 2.1 Southern view as seen from SS1-B

Figure 2.2 Northern view as seen from the parapet

Figure 2.1 shows the parapet located 7.62 m (25 ft) directly south of SS1-B. The parapet has an average height of just 87 cm (2.85 ft). Both SS1-A and SS1-B were placed on tables which were leveled to within 0.1° and had a height of 76 cm (2.5 ft). Because of the low parapet height and the distance from the parapet to SS1-B, there was no shading impact on the solar stills.

2.3 Materials

This study uses data obtained between February 2006 and July 2007 from solar stills manufactured by Sunwater (formerly located in Woodruff, AZ) and SolAqua ("SolAqua," 2011). The Sunwater solar still (SS1), pictured in Figure 2.3, has a rectangular basin area of 0.976 m^2 and has a body composed of aluminum. Sunwater stills are based on a design developed and published by Horace McCracken (1985). The still is insulated with 1" thick polyisocyanurate foam board and is coated with U.S. Food and Drug Administration (FDA) approved silicone sealant in layers with unbounded glass fiber cloth (Venkatesh, 2007). Closed cell foam window gaskets were used to seal SS1's cover glass to the body of the still. The FDA approved silicone sealant was used by the manufacturer to render the produced distillate potable. SS1 has two large diameter inlets, one for distilland delivery and one for drainage, and two smaller outlets for distillate collection. The manufacturer designed and produced SS1 with a fixed glass cover slope of 2°.

Figure 2.3 SS1 in use during the 2011 study

The SolAqua solar still (SS2), shown in Figure 2.4, has a basin area of 0.767 m^2 and has a body composed of fiberglass with foamed in place insulation. The seals are in the form of a hard rubber u-channel molding which wraps around the perimeter of the still clamping the glass against a d-section soft rubber seal bonded to the fiberglass box (Venkatesh, 2007). SS2 has two large diameter inlets for filling and drainage and a single small diameter outlet for distillate collection. The manufacturer designed and produced SS2 with a fixed glass cover slope of 9°.

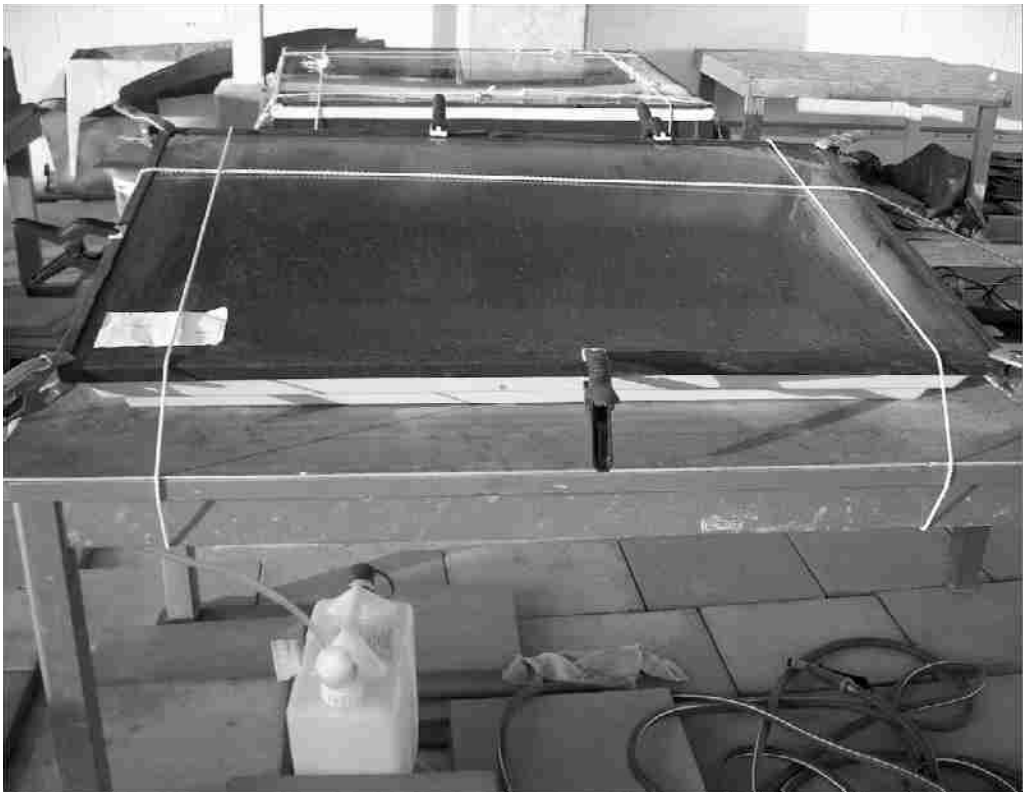


Figure 2.4 SS2 in use during the 2006-2007 study (Venkatesh, 2007)

2.4 Methods

2.4.1 General Procedure

The methodology for the 2011 study was developed from the methodology used for

the 2006-2007 study performed by Venkatesh and was adapted to fit the goals of the 2011 study. The 2011 study used two identical SS1s during the study period that were placed on the tables that were used during the 2006-2007 study. The still closest to the utility structure to the north will be identified as SS1-A and the still to the south of SS1-A will be identified as SS1-B. The distilland volume for each solar still was varied throughout the study period in order to evaluate the effects of distilland volume on daily production. Table 2.3 shows the scheduled distilland volume used throughout the study.

Table 2.3 Distilland volume schedule

	Distilland Volume (L)	
Operation Period	**SS1-A**	**SS1-B**
1/18/11-1/23/11	20	20
1/24/11-2/11/11	26	13
2/12/11-4/3/11	13	26
4/4/11-4/15/11	39	13
4/16/11-5/4/11	13	39
5/5/11-5/27/11	26	20
5/28/11-6/18/11	20	26
6/19/11-7/31/11	30	30
8/1/11-8/26/11	40	25
8/27/11-9/30/11	25	40

Unlike the 2006 study, the 2011 study used only standard clear glass throughout the experiment since the use of low emissivity glass was found to negatively affect the performance of the still (Venkatesh, 2007). Each still was filled with Las Vegas Valley Water District tap water supplied by a spigot on the roof. The average total dissolved solids concentration of the tap water during the 2011 study was approximately 598 parts per million (Las Vegas Valley Water District, 2011).

Each solar still was instrumented with a YCT brand temperature data logger, equipped with fiberglass insulated Type-K thermocouples, to record the distilland, vapor, inner glass, and outer glass temperature at 5 minute intervals throughout the entire period of the study. Figure 2.5 shows the YCT brand temperature data logger in its protective expanded polystyrene cooler. Where the Type-K thermocouple was attached to the glass surface, 3M brand aluminum foil tape was used to create a 5 cm x 2.5 cm strip to protect the sensor from the intense sunlight to prevent the sensor from gaining heat due to light intensity. Figure 2.6 illustrates the thermocouple and protective 3M brand aluminum foil on the inner glass cover surface.

The YCT data logger for each still was kept inside an expanded polystyrene cooler to protect the data logger from the exterior weather conditions. The cooler protected the equipment from intense heat and light during the summer operation and freezing temperatures during the winter operation. The polystyrene cooler can be seen in Figure 2.5. Additional containers of frozen water were placed inside each cooler every morning to maintain interior conditions between 21°C to 27°C (70°F to 80°F) during the summer operation. Moreover, heated water containers were placed inside each cooler at sundown to maintain interior conditions between 15°C to 21°C (60°F to 70°F) during the winter operation. The internal temperature of the polystyrene coolers was recorded with an Onset HOBO® pendant temperature data logger.

Onset HOBO® light and temperature pendant data loggers were also used to measure the ambient air temperature and the local light intensity at five minute intervals. The light data logger allowed for the accurate determination of the time for sunrise and sunset as it is experienced at the test location. The light/temperature sensors were mounted on 54 mm

(0.25 in) of expanded polystyrene to protect the logger from the hot surface of the table. The pendant temperature data logger can be seen in Figure 2.7. In order to record the ambient temperature at the test site, one of the HOBO sensors was placed underneath SS1-B's table to shade the sensor from light; allowing the sensor to read/record the true ambient temperature.

SS1-A and SS1-B each had their distillate outlet drain connected to a tipping bucket rain gauge data logger. The data logger allowed for the recording of real time distillate production which was then verified with the daily production measurements by using a 1 liter graduated cylinder.

Produced distillate from SS1-A and SS1-B was measured by using a Rainwise Rainew® and Onset HOBO® brand rain gauge data logger, respectively. A power drill with a circular saw bit head was used to create the proper size opening on two lids, for the 18.9 L (5 gal) containers, so that the measured distillate from each rain gauge could then be collected inside each container. Each container was tied to the leg of a table and stone pavers were used to brace each container during wind storms to prevent any tipping and loss of collected distillate. Figure 2.8 illustrates the rain gauge data logger set up. Figures 2.9 and 2.10 illustrate the plan and profile view, respectively, of SS1-A/B to demonstrate the location of the YCT brand data logger and the various thermocouples used to measure the distilland, vapor, and inner/outer glass temperatures.

Figure 2.5 YCT brand temperature data logger with ice water containers

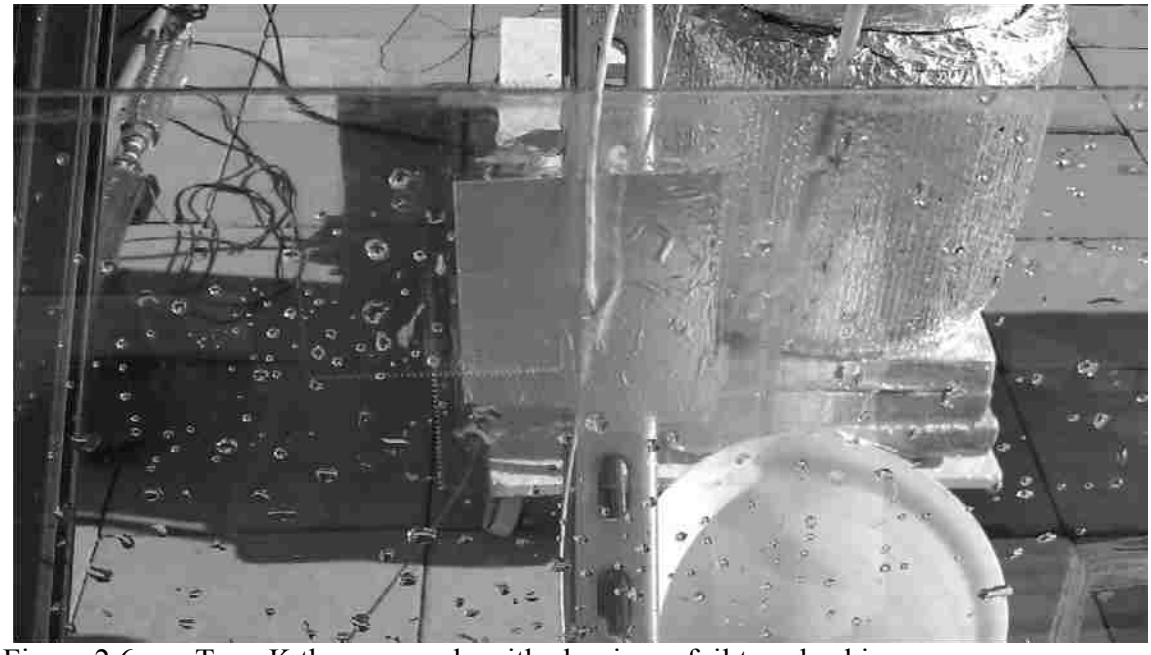
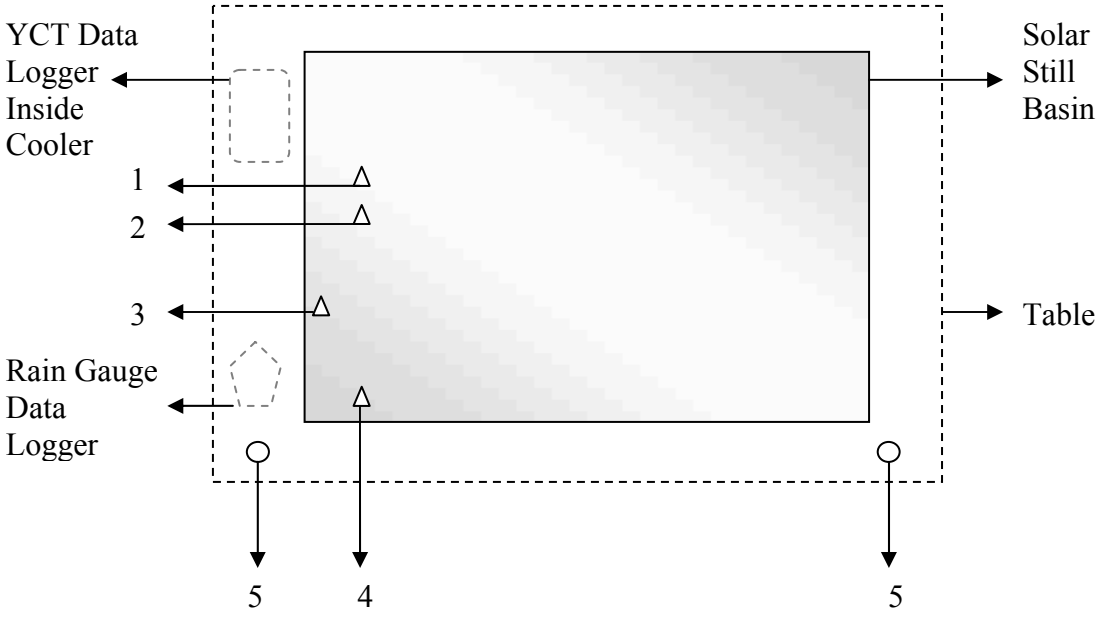


Figure 2.6 Type K thermocouple with aluminum foil tape backing

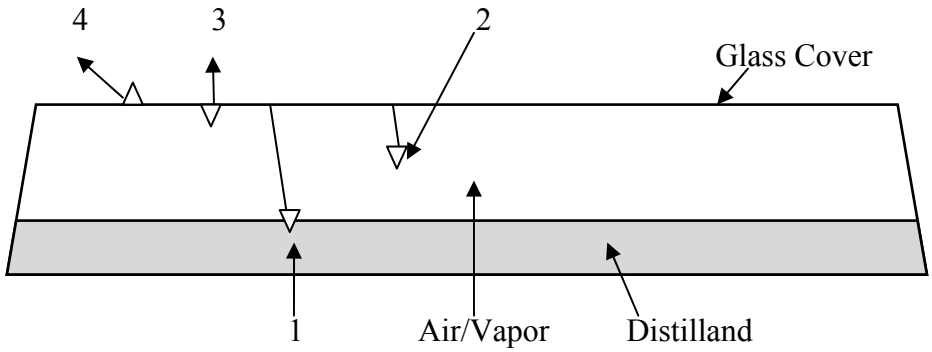
Figure 2.7 Onset HOBO® pendant temperature/light data logger

Figure 2.8 Rainwise (left) and Onset (right) rain gauge data loggers attached to SS1-A and SS1-B, respectively



1. Thermocouple measuring distilland temperature
2. Thermocouple measuring vapor temperature
3. Thermocouple measuring inner glass cover temperature
4. Thermocouple measuring outer glass cover temperature
5. Solar intensity data logger

Figure 2.9 Plan view of SS1-A/B data acquisition set up



1. Thermocouple measuring distilland temperature
2. Thermocouple measuring vapor temperature
3. Thermocouple measuring inner glass cover temperature
4. Thermocouple measuring outer glass cover temperature

Figure 2.10 Profile view of SS1-A/B data acquisition set up

2.4.2 Data Preparation for Modeling

In order to execute the different solar still modeling methods, the daily production data from the 2006-2007 study had to be paired with the corresponding weather conditions that were hypothesized to affect distillate production. Daily average ambient air temperature and daily average wind speed were retrieved from the McCarran International Airport National Weather Service (NWS) station. A summary of the weather conditions can be seen in Table 2.1 and 2.2.

The average daily wind direction was calculated from hourly values that were obtained from the McCarran NWS station and were then modified, for modeling purposes, to range between 0 to 180 degrees from north. This modification was necessary in order to prevent a large numerical change when the prevailing winter and spring wind directions fluctuated between northwesterly and northeasterly.

The average daily cloud cover was calculated from hourly values that were obtained from the McCarran NWS station (Weather Underground, 2011). Each hourly reading provided a text description for the fraction of the sky covered by clouds. The cloud cover readings were described as “Few” (1/8–2/8 cloud coverage), “Scattered” (3/8–4/8 cloud coverage), “Broken” (5/8-7/8 cloud coverage) and “Overcast” (8/8 cloud coverage).

Daily global horizontal insolation data was obtained from the National Renewable Energy Laboratory (NREL) Measurement and Instrumentation Data Center (MIDC) site for UNLV (NREL, 2011). Hourly insolation data rates were multiplied by 3,600 seconds to convert from W-Hr/m^2 to J/m^2 to reflect the amount of energy per square meter. The daily total insolation was then calculated by summing individual values for total hourly insolation.

Table 2.4 summarizes the maximum and minimum production that was recorded by SS1 and SS2 during the 2006-2007 study as well as the production recorded by SS1-A and SS1-B during the 2011 study.

Table 2.4 Performance summary for 2006-2007 and 2011 study

Solar Still Type	Days of Available Data	Maximum Daily Production (L/m^2)	Minimum Daily Production (L/m^2)
SS1 (2006)	453	8.8	0.21
SS2 (2006)	312	5.8	0.31
SS1-A (2011)	243	6.8	0.47
SS1-B (2011)	242	6.8	0.57

Since the original 2006 study used standard glass covers from two manufacturers, due to the need to replace a broken glass cover, a Student's t-test had to be performed to test if there was any significant performance difference between either glass cover for SS1 at a significance level of $\alpha = 0.05$. After comparing the production of SS1 under the different glass covers, it was found that there was no statistically significant production difference ($p \geq 0.05$, two tailed, and assuming unequal variances).

Since SS1 was operated with low emissivity glass from July 2006 to October 2006, production data during this period had to be substituted from a separate still identical to SS1 known as SS1-C. This was done solely to complete the visual depiction of SS1's seasonally varying yield between summer and fall. In Figure 2.11, the daily production and insolation data for SS1 appears as gray squares and open diamonds, respectively. The daily production and insolation data from SS1-C appears as gray circles and open triangles, respectively, for the data markers between July 2006 and October 2006. Data

from SS1-C was solely used to complete a depiction of the seasonally varying pattern in Figure 2.11. Production data from SS1-C were not combined with any data from SS1 for modeling purposes.

Figures 2.11 and 2.12 display the daily production data for SS1, SS1-C and SS2 during the original 2006-2007 study. Distillate production is graphed along with insolation to show the dependence of distillate production on insolation.

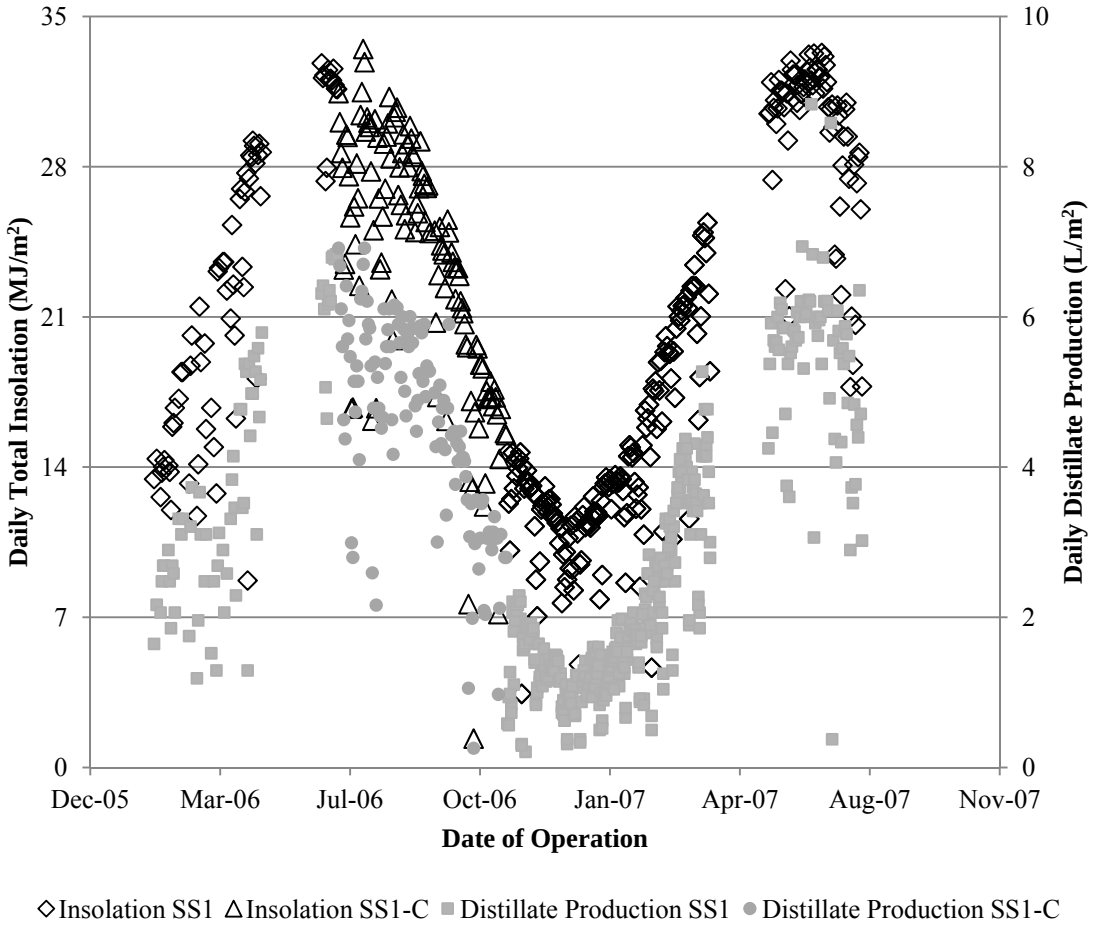


Figure 2.11 SS1 and SS1-C daily production and historical insolation, 2006-2007

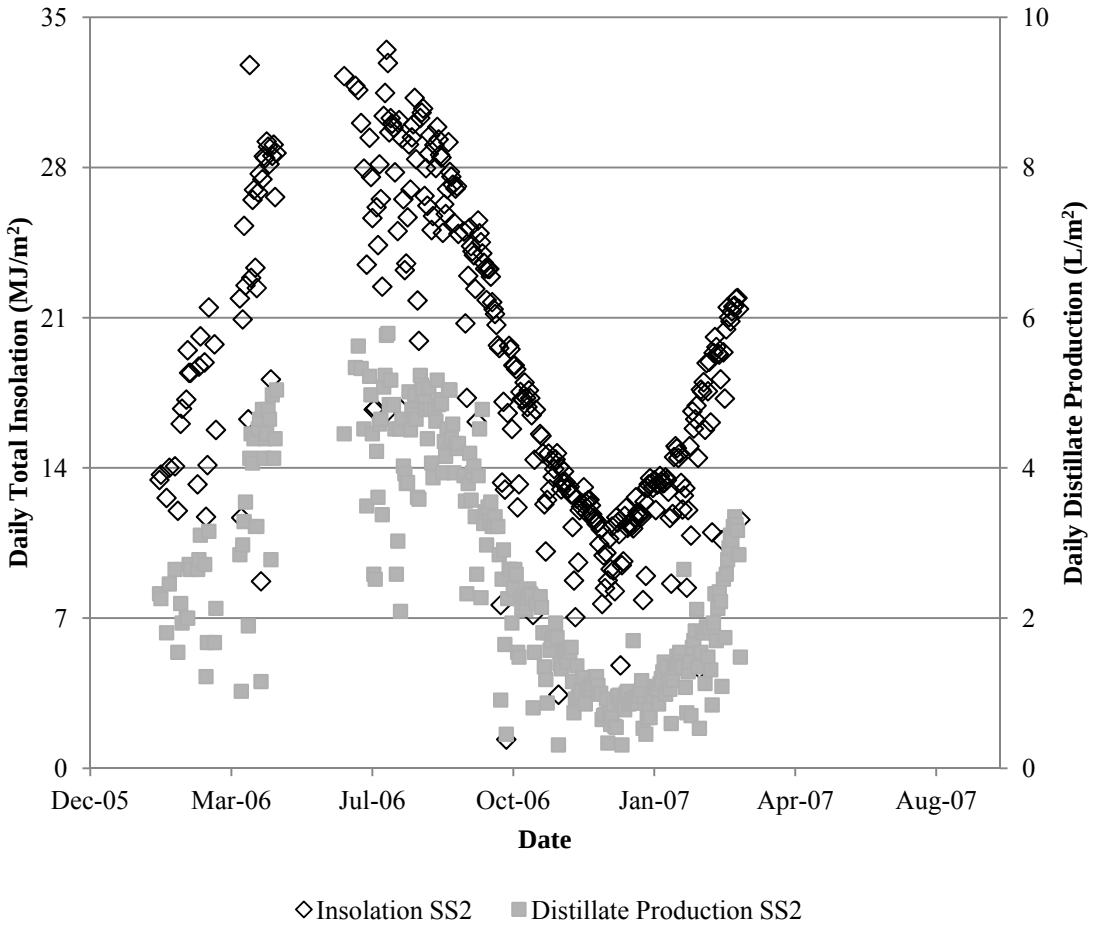


Figure 2.12 SS2 daily production and historical insolation, 2006-2007

CHAPTER 3

DAILY AND HOURLY SOLAR STILL PERFORMANCE

3.1 Overview

This study examined daily solar still production under varying distilland volume scenarios throughout the winter, spring, and summer seasons. Several temperature readings such as the distilland, vapor, and inner/outer cover glass temperatures were recorded at five minute intervals and converted to hourly averages. Real time distillate production was recorded by rain gauges and compiled to correspond to the hourly temperature intervals. Daily cumulative distillate production volumes were measured by graduated cylinder. The combination of the hourly and daily data provides a detailed look into the performance of solar stills as a water resource system whose production responds to changes in distilland volume and to fluctuating weather conditions.

3.2 Daily Production

Figures 3.1 and 3.2 illustrate the daily production as graphed against daily total insolation for SS1-A and SS1-B, respectively. SS1-A and SS1-B's daily production correlations to daily total insolation can be estimated by analyzing the Pearson correlation coefficient (R) which evaluates the relationship between an independent and dependent variable. The Pearson correlation coefficients between daily production and insolation for SS1-A and SS1-B are 0.973 and 0.949, respectively. The lower R value for SS1-B can be seen as a result of the fluctuations in production near the high end of total insolation.

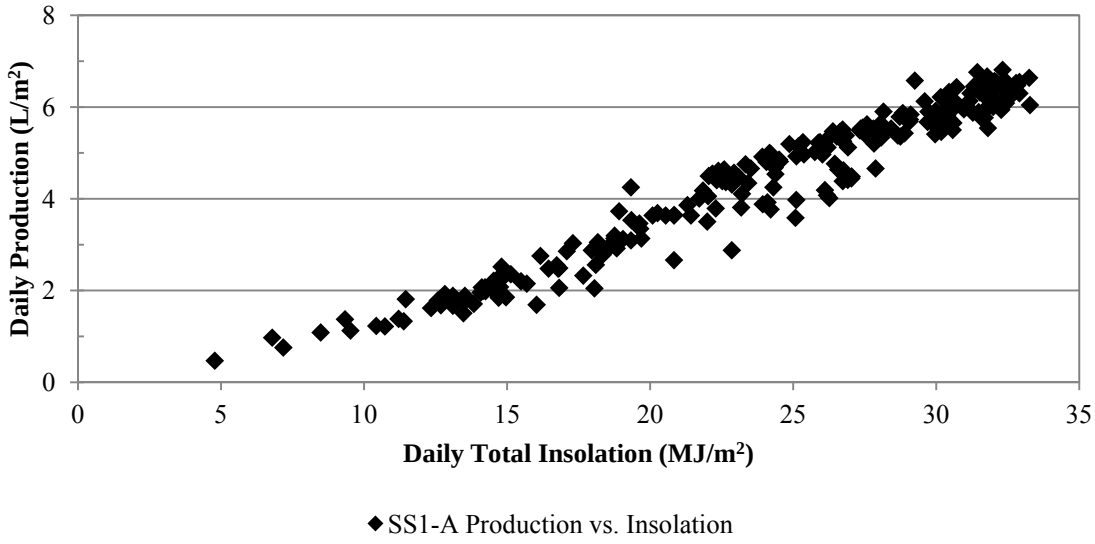


Figure 3.1 SS1-A's daily production vs. daily total insolation between 1/2011-9/2011

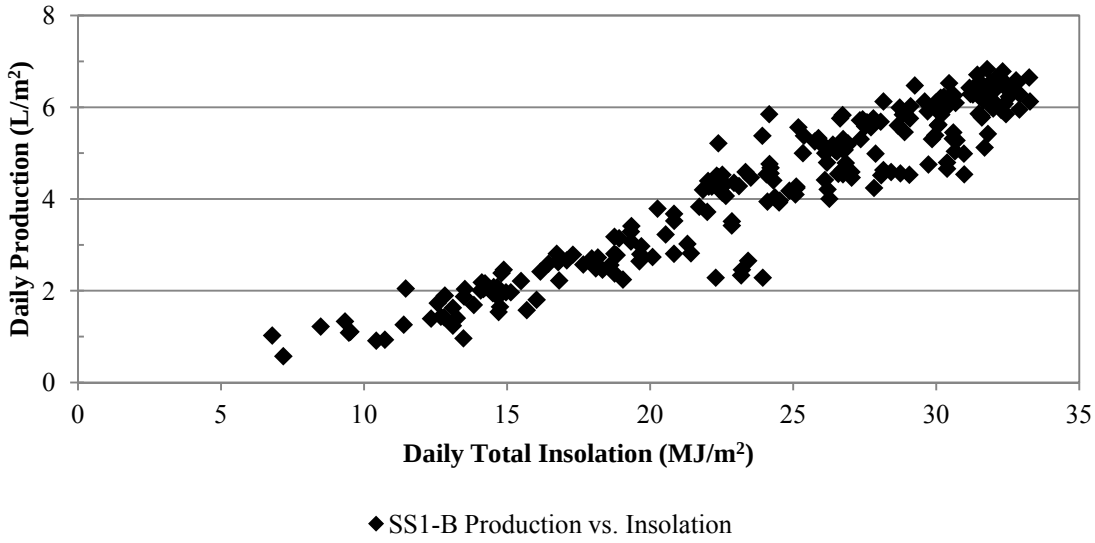


Figure 3.2 SS1-B's daily production vs. daily total insolation between 1/2011-9/2011

The data sets for SS1-A and SS1-B contained 243 and 242 data points, respectively, for daily distillate production. Table 3.1 details the monthly average production for the 2011 study.

Table 3.1 Average daily production by month for 2011 study

	SS1-A		SS1-B	
Month	**Average Daily Production (L/m^2)**	**Coefficient of Variation**	**Average Daily Production (L/m^2)**	**Coefficient of Variation**
Jan-11	1.87 ± 0.25	13.4%	1.76 ± 0.42	24.1%
Feb-11	2.28 ± 0.50	22.0%	2.19 ± 0.51	23.4%
Mar-11	3.28 ± 1.19	36.1%	2.66 ± 0.87	32.6%
Apr-11	4.52 ± 1.17	26.0%	4.20 ± 0.93	22.1%
May-11	5.23 ± 0.99	18.9%	5.24 ± 0.91	17.3%
Jun-11	6.24 ± 0.45	7.2%	6.22 ± 0.45	7.2%
Jul-11	5.28 ± 1.27	24.1%	5.36 ± 1.27	23.7%
Aug-11	5.26 ± 0.96	18.3%	5.39 ± 1.00	18.6%
Sep-11	4.04 ± 1.34	33.1%	3.84 ± 1.24	32.3%

Maximum average production for the 2011 study occurred during the month of June for both SS1-A and SS1-B. Production during the month of June had low variation indicating fewer fluctuations in weather conditions for the month of June. The maximum daily distillate production for SS1-A was 6.81 L/m^2 and occurred on 7/21/11 while the maximum distillate production for SS1-B was 6.83 L/m^2 and occurred on 7/1/11.

Daily solar still production varies throughout the year as a result of the day length, total insolation, and other weather conditions such as ambient temperature, wind speed, and cloud cover. Figure 3.3 illustrates the daily production data for the 2011 study for SS1-A and SS1-B. Figure 3.3 also indicates a strong relationship between daily distillate production and the total insolation as seen in Figures 3.1 and 3.2.

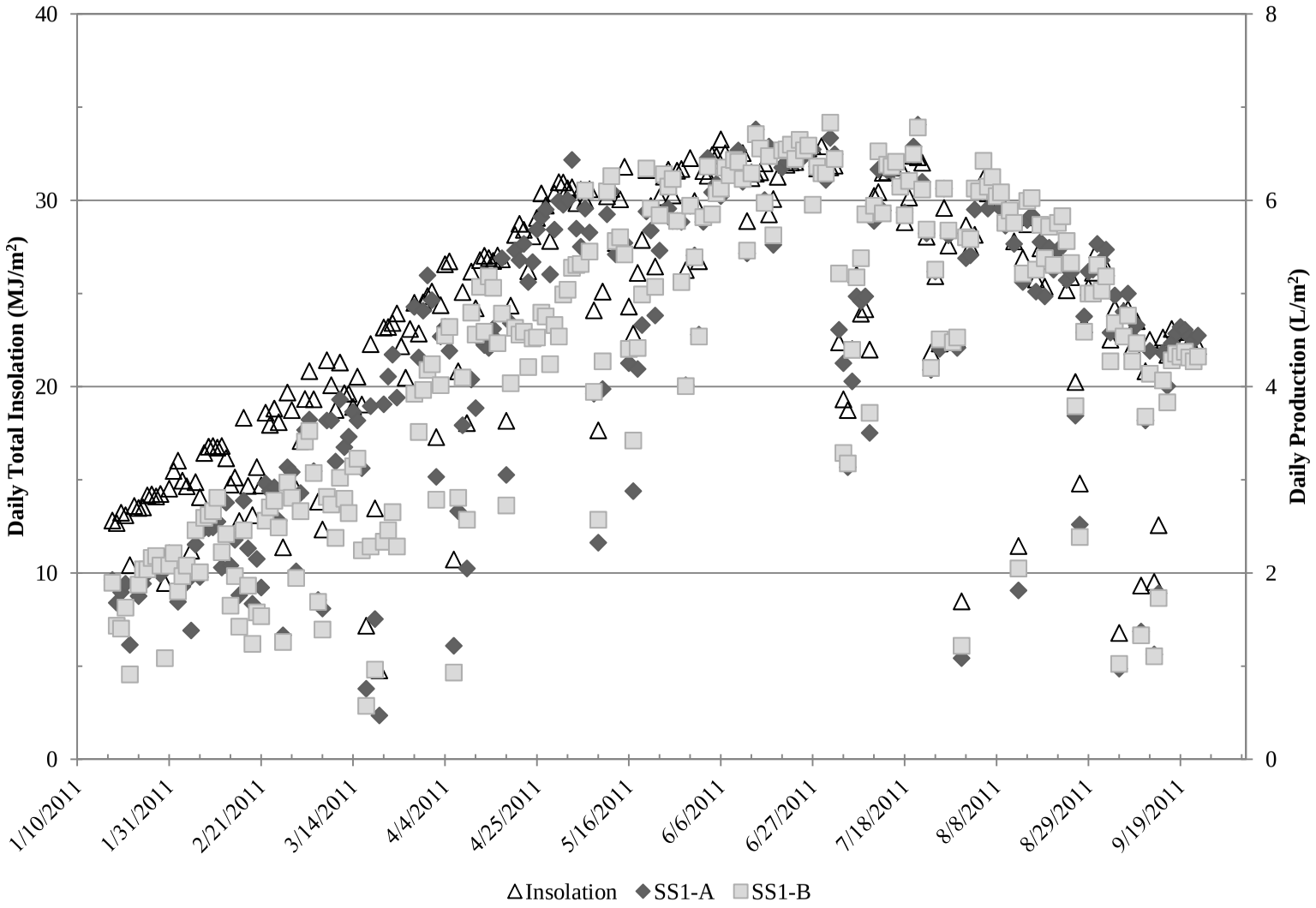


Figure 3.3 Daily production and insolation data for SS1-A and SS1-B, 2011

Minimum average production for the 2011 study occurred during the month of January for both SS1-A and SS1-B. January was the only other month to feature small standard deviations (less than 15%) from average daily production values. Minimum distillate production for SS1-A and SS1-B occurred on 3/17/11 with 0.76 L/m^2 and 0.57 L/m^2, respectively. The minimum production occurred in March due to a cold snap with extremely cloudy skies. The day featuring minimum production, March 17, 2011, had an average temperature that was only 1°C higher than the average temperature in January 2011 and experienced only 7.1 MJ/m^2 of insolation (January experienced a minimum insolation of 9.5 MJ/m^2 on 1/30/11). Not considering the extreme weather on 3/17/11, the minimum production for SS1-A and SS1-B occurred on 1/22/11 with daily productions of 1.23 L/m^2 and 0.91 L/m^2, respectively.

Figures 3.4 and 3.5 illustrate the 5th and 95th percentile values for SS1-A and SS1-B's daily production throughout the 2011 study, respectively.

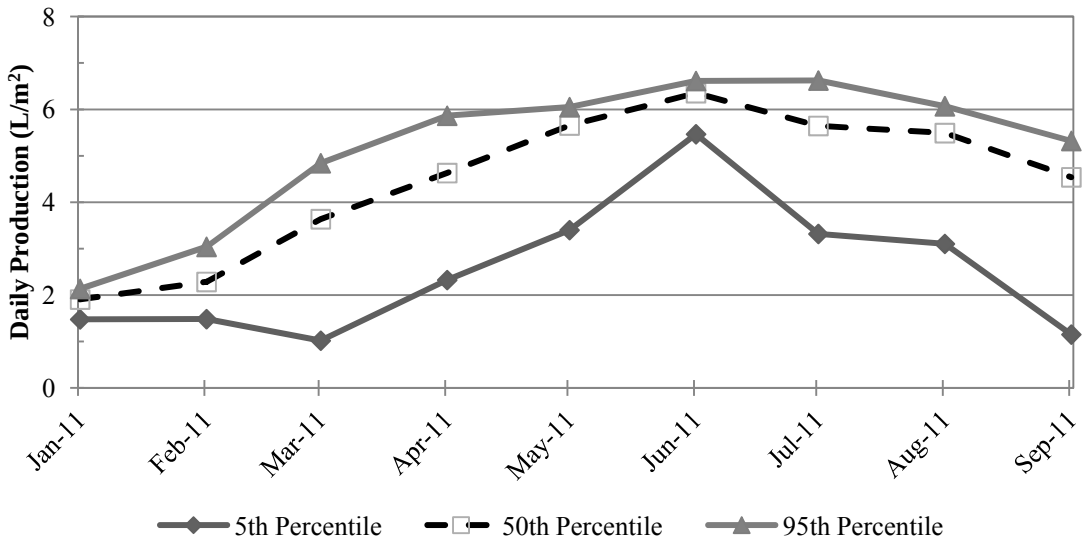


Figure 3.4 SS1-A's 5th, 50th, and 95th percentile daily production by month

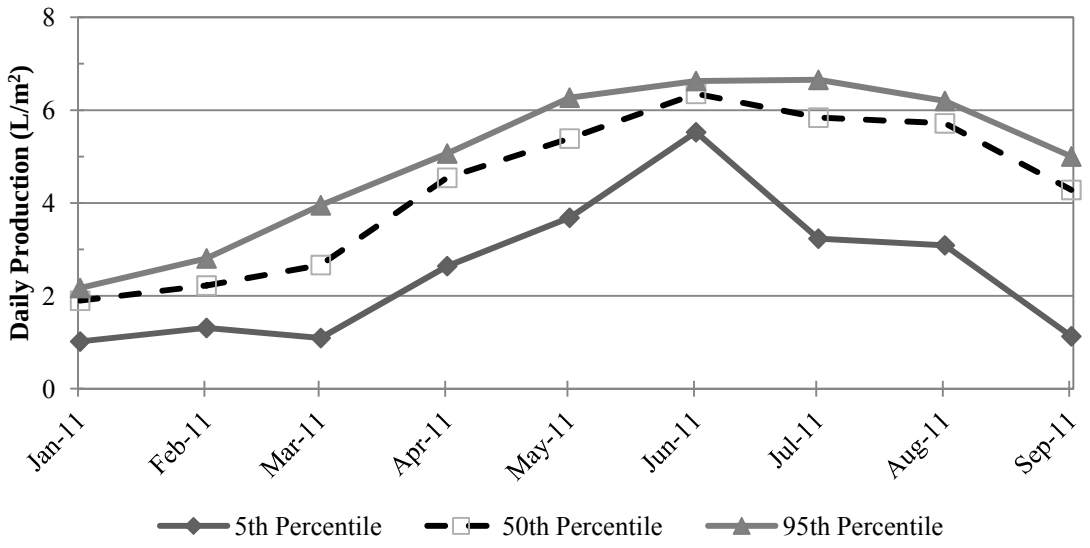


Figure 3.5 SS1-B's 5th, 50th, and 95th percentile daily production by month

The resulting "envelope area," between the 95th and 5th percentile values, indicates the range for 90% of the daily production data for each month. The "envelope area" is generally wide except for the months of January and June. As previously discussed, January and June production values had the smallest deviation from the average production (less than 0.50 L) due to more consistent weather conditions. The large boundary area for other months can be attributed to changing weather conditions such as variable day to day cloud cover, wind speed, and ambient temperature.

3.3 Yields and Efficiency

The efficiency of a solar still details how effective the still is at absorbing solar energy and evaporating and collecting water. Equation 3.1 illustrates the method used to calculate solar still efficiency.

$$\eta = \frac{M \cdot \Delta H_v}{I_{Total}}$$

Equation 3.1 Overall solar still efficiency

M Daily Distillate Production (kg/m^2)
ΔH_v Enthalpy of Vaporization for Water (MJ/kg)
I_{Total} Total Global Insolation (MJ/m^2)

The efficiencies of SS1-A and SS1-B were calculated between January 2011 and September 2011 using Equation 3.1. Figure 3.6 illustrates the efficiencies of SS1-A and SS1-B for every distilland volume operating scenario.

The different distilland volume scenarios previously listed in Table 2.3 are shown as labels for each column. With one exception, Figure 3.6 indicates a higher efficiency for each still when operated with a smaller distilland volume. The only exception to this pattern was the period between 1/18/11-1/23/11. Since this was the first week where the solar stills were operated, an error in the setup of the window seals and a short averaging period could have had roles in causing SS1-B to operating less efficiently.

Figure 3.6 also indicates that as the ambient temperature and insolation increased from the winter season to the summer season, the difference in efficiency between different distilland volume scenarios decreased substantially. The maximum efficiency difference between SS1-A and SS1-B, during winter and spring, was 6.9% (2/12-4/3 and 4/16-5/4) while the maximum efficiency difference during summer was only 3% (8/27-9/30). Since the summer season in Las Vegas is reliably warm and sunny, SS1-A and SS1-B experience fewer fluctuations in summer ambient weather conditions that may negatively affect solar still performance. The winter/spring seasons offer more variable conditions in cloud cover and wind speed while providing lower levels of insolation.

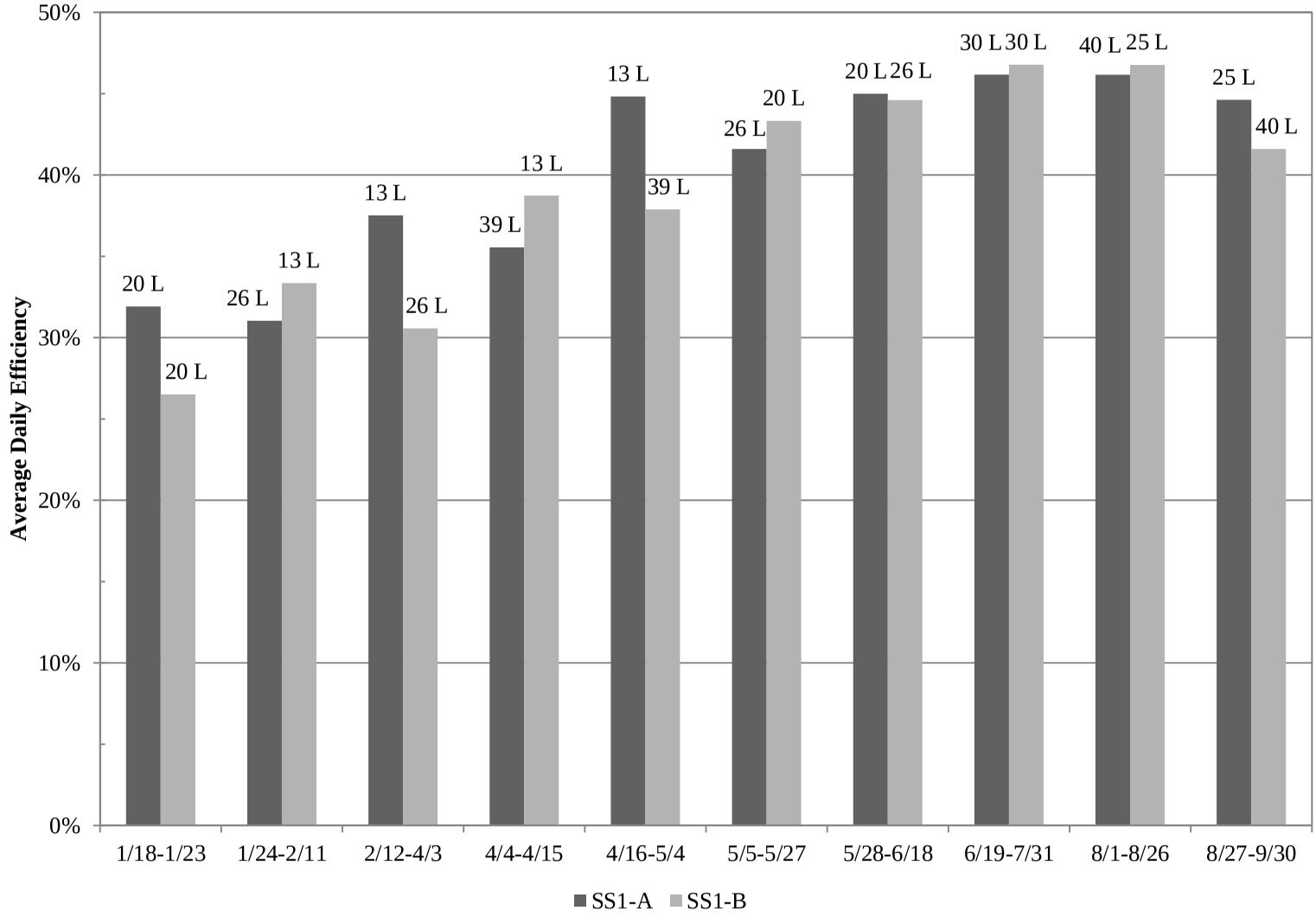


Figure 3.6 Average daily efficiencies for each operating distilland volume scenario, 2011

3.4 Hourly Performance

3.4.1 Seasonal Hourly Insolation

The nature of passive solar distillation results in a strong dependence on the amount of hours in the day and the quantity of solar energy received throughout the day. Figure 3.7 illustrates the hourly patterns in total insolation during the spring/fall equinox and the summer/winter solstice. The greatest amount of incident energy is available during the late spring, summer, and early fall seasons.

Data from the NREL MIDC site for UNLV recorded 15 hours of sunlight on the summer solstice with a peak hourly total of 3.70 MJ/m^2 and a daily total of 31.9 MJ/m^2, 10 hours of sunlight on the winter solstice with a peak hourly total of 1.91 MJ/m^2 and a daily total of 11.2 MJ/m^2, 13 hours of sunlight on the spring equinox with a peak hourly total of 3.26 MJ/m^2 and a daily total of 23.4 MJ/m^2, and 13 hours of sunlight on the autumn equinox with a peak hourly total of 3.01 MJ/m^2 and a daily total of 22 MJ/m^2.

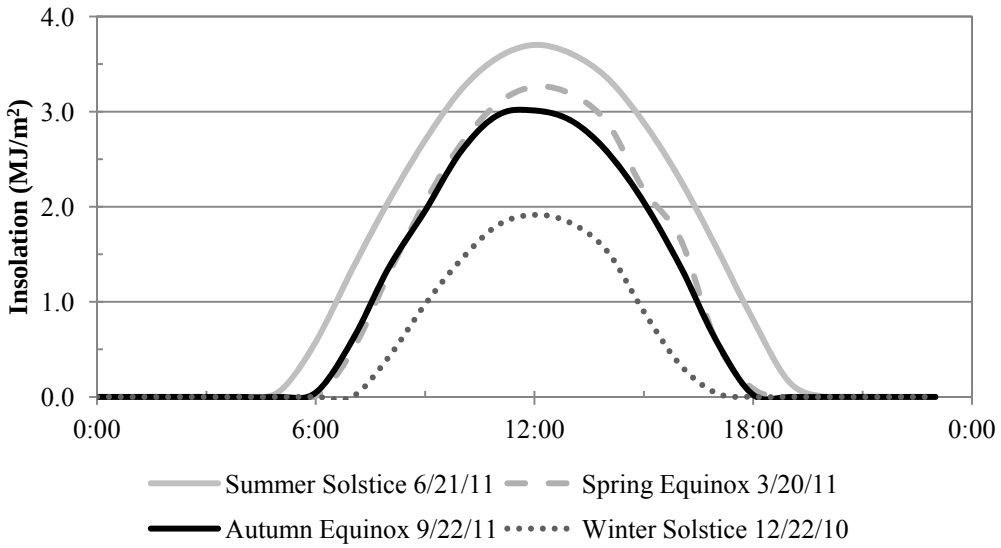


Figure 3.7 Measured insolation patterns for different seasons

3.4.2 Hourly Distillate Production

Figure 3.8 illustrates an example of hourly distillate production and insolation for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11.

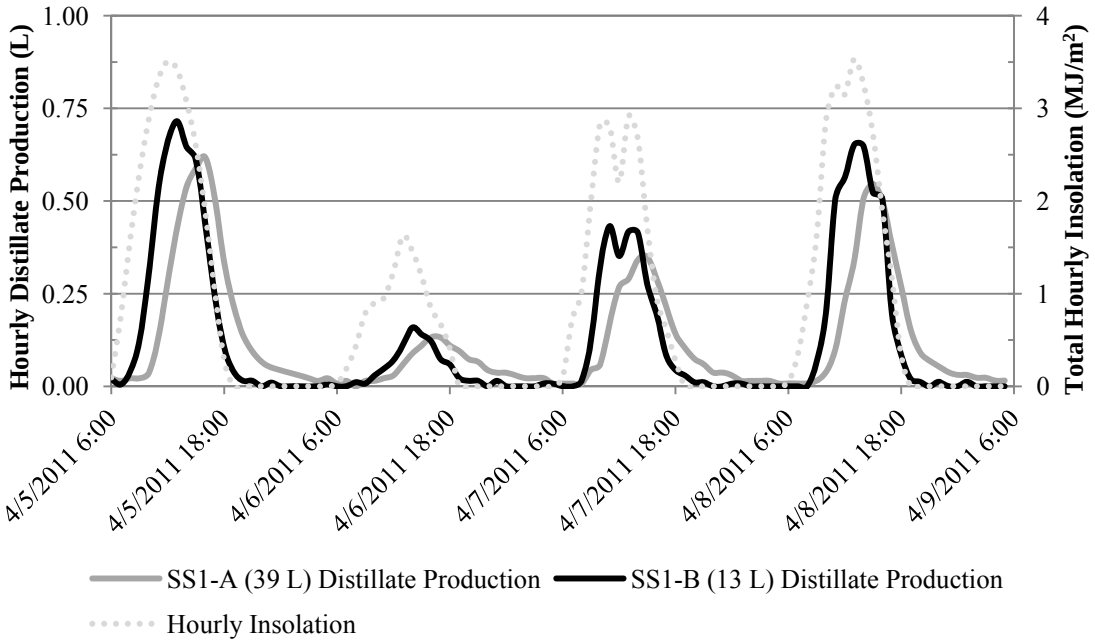


Figure 3.8 Hourly distillate production and insolation for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11

Figure 3.8 indicates a one to two hour lag between peak insolation and peak distillate production. The difference in distilland volume resulted in a lower overall production for SS1-A compared to SS1-B. Even though Figure 3.8 is a short duration example, the one to two hour lag between peak insolation and peak distillate production was seen throughout the entire study and was also observed by Venkatesh (2007).

The amount of lag between production and insolation for SS1-A and SS1-B is also dependent on the distilland volume. Figure 3.8 shows the peak production for SS1-A (39 L) occurring later than the peak production than SS1-B (13 L) which appears to more

closely follow insolation over time.

Figure 3.8 also indicates that the solar still with a larger distilland volume tends to continue producing distillate later into the evening without any interference from sudden decreases in insolation. Between 0900 and 1200 on 4/7/11, there was a sudden decrease in insolation. The sudden decrease in insolation is mirrored in the hourly production decrease of SS1-B (13 L); however, SS1-A (39 L) does not demonstrate any decrease in production due to the decrease in insolation. Once again, even though Figure 3.8 is an example, the same behavior between the sensitivity of production to insolation for smaller distilland volumes was found to occur throughout the study. Figure 3.9 illustrates the hourly distillate production and insolation for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11.

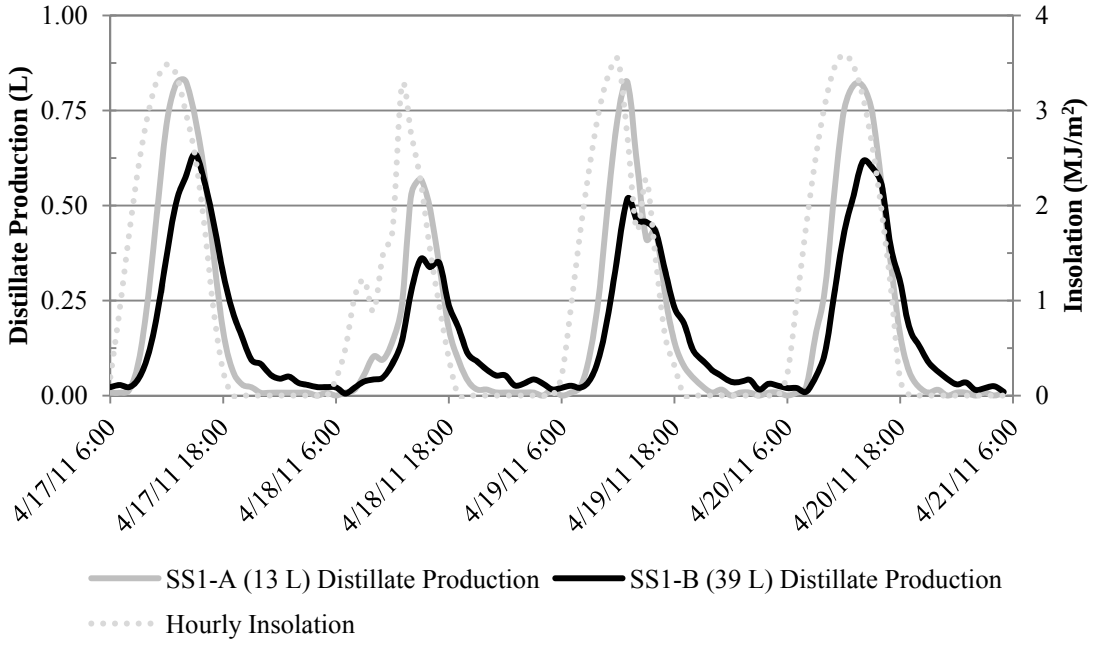


Figure 3.9 Hourly distillate production and insolation for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11

After swapping the operating distilland volume compared to the period 4/5/11 and 4/9/11, Figure 3.9 still exhibits some of the same properties shown in Figure 3.8. The reversal of the operating distilland volume between SS1-A and SS1-B results in the continuation of the observed patterns with regard to production lagging insolation and lower daily production for the still operating with a larger distilland volume.

3.4.3 Average Hourly Temperature Readings

Figures 3.10 and 3.11 are example illustrations of the average temperature data for SS1-A (39 L) and SS1-B (13 L) between 4/4/11 and 4/15/11, respectively.

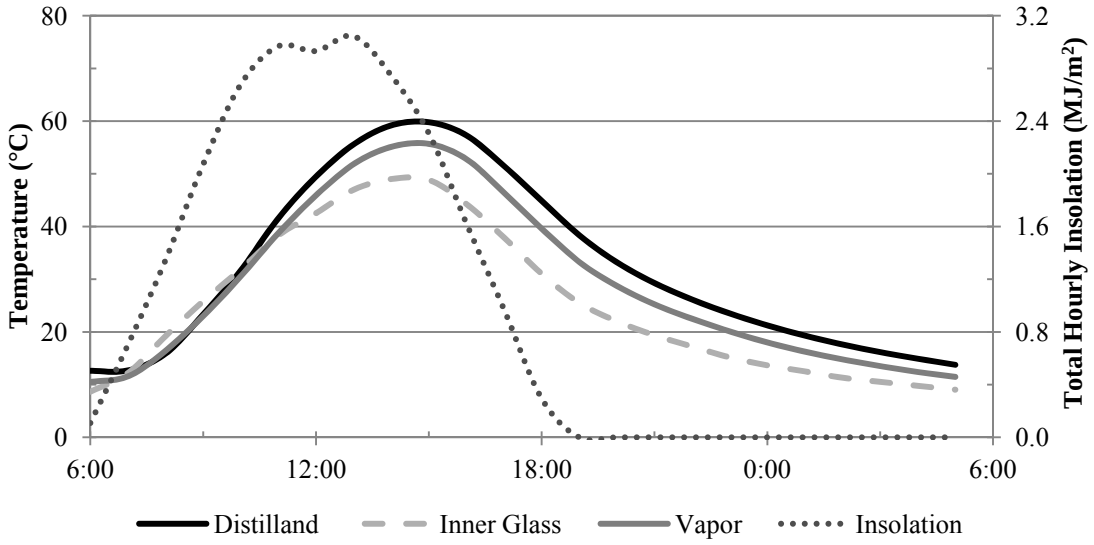


Figure 3.10 SS1-A (39 L) average hourly temperature readings between 4/4/11 and 4/15/11

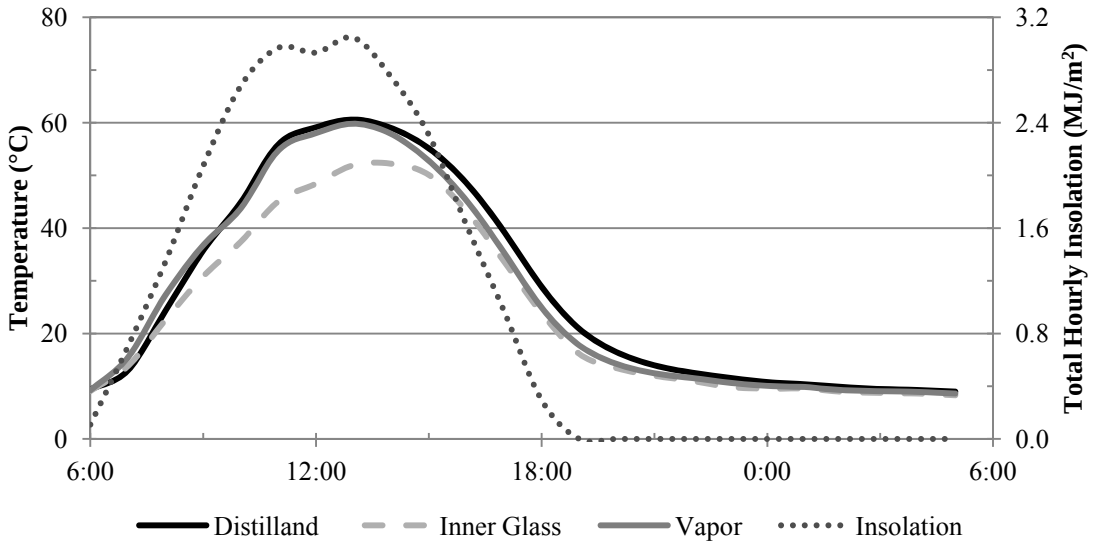


Figure 3.11 SS1-B (13 L) average hourly temperature readings between 4/4/11 and 4/15/11

Figures 3.10 and 3.11 illustrate the average temperature data during early spring while SS1-A and SS1-B were operated under large and small distilland volume scenarios, respectively. Comparison of the large and small distilland volume scenarios, for the identical time period, reveals SS1-A (39 L) had a peak distilland temperature of 59.8°C at 15:00 while SS1-B (13 L) had a peak distilland temperature of 60.6°C at 13:00. The inner glass and vapor peak temperatures occurred at the same time as the distilland for SS1-A (39 L) and SS1-B (13 L). SS1-A had a peak vapor temperature of 55.7°C and a peak inner glass temperature of 48.8°C. SS1-B had a peak vapor temperature of 59.8°C and a peak inner glass temperature of 51.9°C. Figures 3.10 and 3.11 indicate that stills operated with a deeper distilland have cooler daytime operating temperatures for the distilland, vapor, and inner glass cover.

SS1-B (13 L) proved to be slightly warmer than SS1-A (39 L) during the day; however, SS1-A (39 L) managed to stay warmer than SS1-B (13 L) during night time

operation. By midnight, the distilland temperature for SS1-A (39 L) is 21.2°C while the distilland temperature for SS1-B (13 L) is 10.8°C; a 10.4°C difference. The vapor and inner glass temperature for SS1-A (39 L) managed to stay 7.9°C and 4.1°C warmer, respectively, compared to SS1-B (13 L) at midnight.

A Student's t-test was conducted to determine if the null hypothesis assuming equal mean daily distillate production during 4/4/11-4/15/11 for SS1-A (13 L) and SS1-B (39 L) could be rejected. With a significance of $\alpha = 0.05$, the null hypothesis was accepted with a p-value of 0.460 (two tailed, assuming unequal variances). The Student's t-test indicates that while hourly performance is different for different operating distilland volumes, the total daily production was not significantly different.

Figures 3.12 and 3.13 are example illustrations of the average temperature data for SS1-A (30 L) and SS1-B (30 L) between 6/19/11-7/31/11. The data presented in Figures 3.12 and 3.13 summarize the peak temperature conditions for both stills.

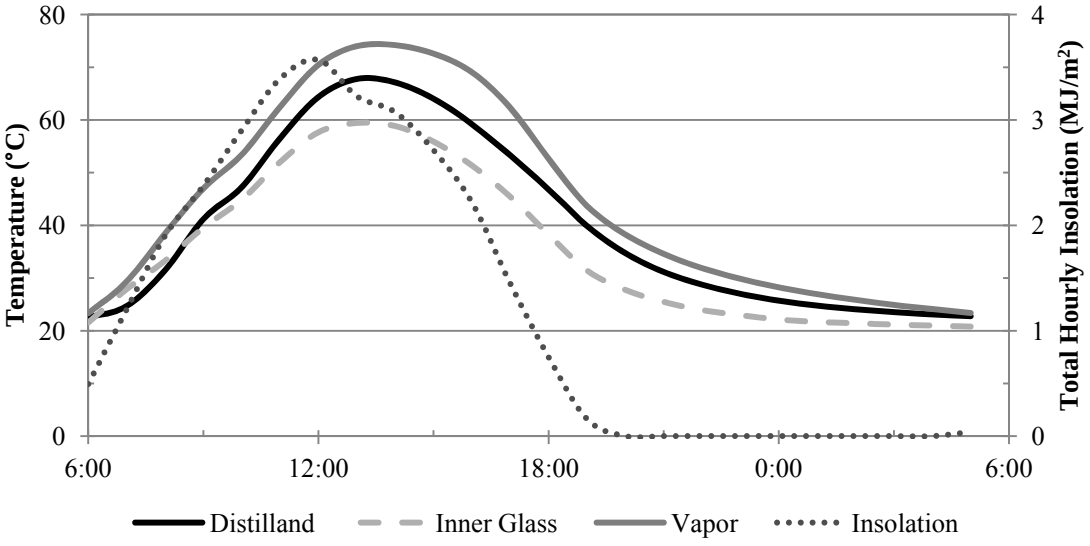


Figure 3.12 SS1-A (30 L) average temperature readings between 6/19/11 and 7/31/11

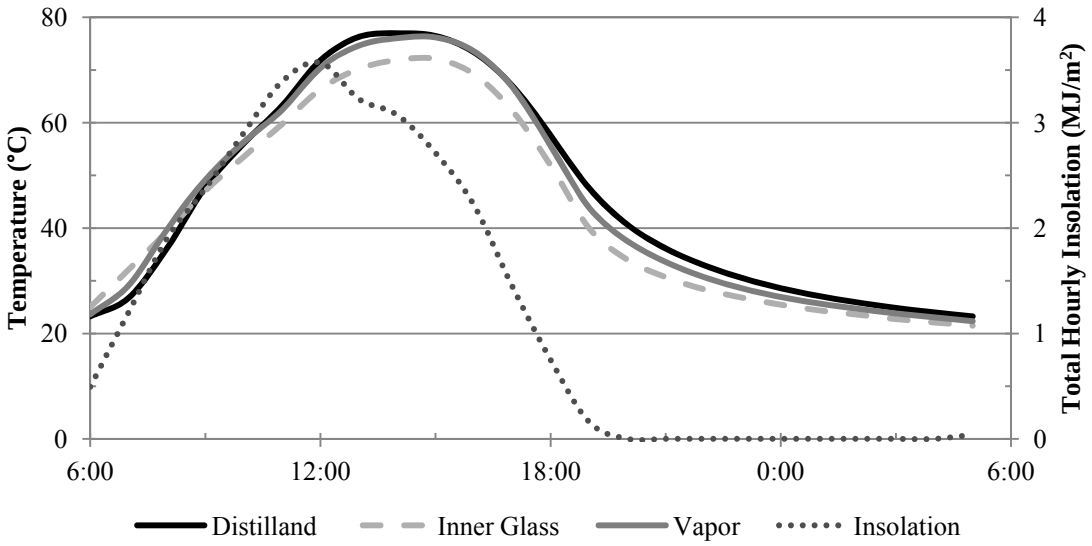


Figure 3.13 SS1-B (30 L) average temperature readings between 6/19/11 and 7/31/11

Figures 3.12 and 3.13 illustrate the average temperature data during the summer season while SS1-A and SS1-B were both operated with a distilland volume of 30 L. This particular period shows SS1-A performing with an average vapor temperature higher than the distilland temperature throughout the peak summer season.

SS1-A (30 L) experienced the average peak distilland temperature at 13:00, the average peak vapor temperature at 14:00, and the average peak inner glass cover temperature at 13:00. SS1-B experienced the average peak distilland temperature at 14:00, the average peak vapor temperature at 15:00, and the average peak inner glass cover temperature at 15:00.

The summer results indicate that peak temperature values occurred at different times for each still and indicate SS1-A (30 L) was performing differently than SS1-B (30 L). A Student's t-test was conducted to determine if the null hypothesis assuming equal mean daily distillate production during 6/19/11-7/31/11 for SS1-A (30 L) and SS1-B (30 L) could be rejected. With a significance of $\alpha = 0.05$, the null hypothesis was accepted with

a p-value of 0.815 (two tailed, assuming unequal variances). The Student's t-test indicates that despite differences in hourly temperatures, there was no statistical difference in the daily distillate production between SS1-A (30 L) and SS1-B (30 L).

SS1-A (30 L) had an average peak distilland, vapor, and inner glass cover temperatures of 67.8°C, 72.6°C, and 59.4°C, respectively. SS1-B (30 L) had an average peak distilland, vapor, and inner glass temperatures of 76.9°C, 76.2°C, and 72.1°C, respectively. SS1-A (30 L) experienced an 8.4°C temperature difference between the average distilland and average inner glass cover temperature while SS1-B (30 L) experienced a temperature differential of only 0.7°C. By 05:00, SS1-A (30 L) and SS1-B (30 L) had an average distilland temperature of 22.8°C and 23.2°C. SS1-A (30 L) and SS1-B (30 L) had an average distilland temperature within 1°C starting at approximately 02:00.

Figures 3.14 and 3.15 illustrate the average temperature data for SS1-A (20 L) and SS1-B (26 L) between 5/28/11-6/18/11. The data shown in Figures 3.14 and 3.15 summarizes the average temperature conditions for SS1-A (20 L) and SS1-B (26 L) during the late spring season.

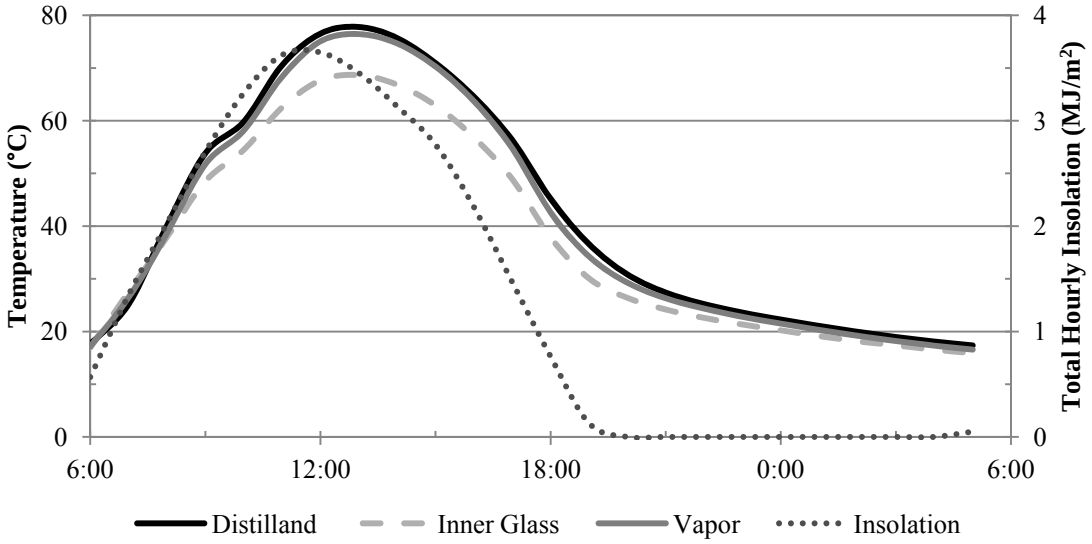


Figure 3.14 SS1-A (20 L) average temperature readings between 5/28/11-6/18/11

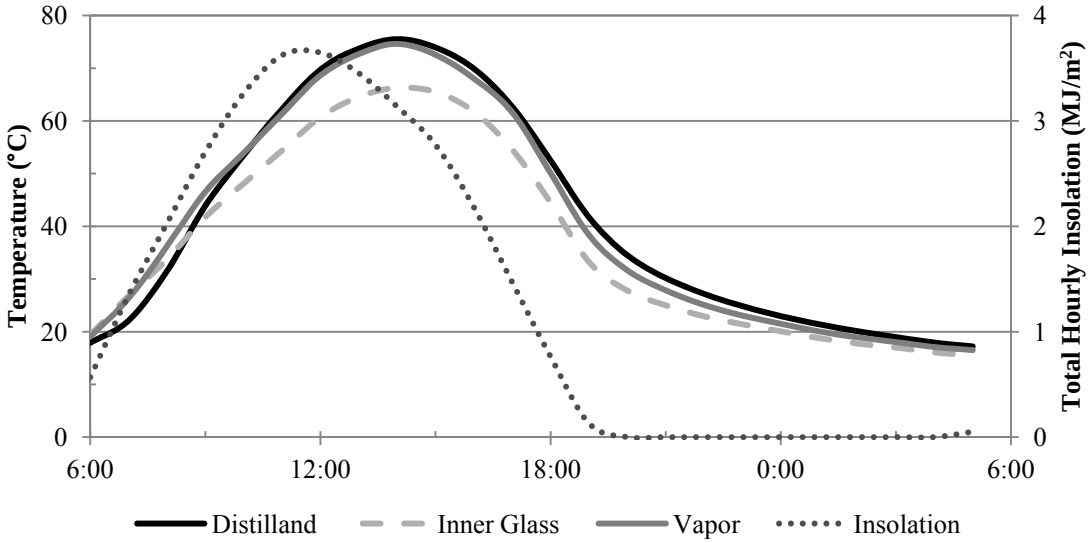


Figure 3.15 SS1-B (26 L) average temperature readings between 5-28/11-6-18/11

From Figure 3.14, it can be noticed that the average distilland and average vapor temperatures are extremely close to each other for SS1-A (20 L); unlike the results shown in Figure 3.12. Figure 3.14 indicates that there might have been a leak with the closed vinyl foam cell seals for SS1-A (30 L) during the 6/19/11-7/31 operating period. Faulty

seals would result in the loss of vapor and heat preventing the maximum distilland and vapor temperature from being reached.

Figure 3.15 illustrates the vapor temperature being higher than the distilland temperature during the early morning for SS1-B (26 L); however, the distilland temperature remains higher than the vapor temperature throughout the afternoon and evening. Figures 3.14 and 3.15 demonstrate how a faulty seal could have contributed to the low distilland temperature for SS1-A (30 L) between 6/19/11-7/31/11.

3.4.4 Hourly Temperature Readings

Figure 3.16 illustrates the hourly ambient and distilland temperature between 4/5/11 and 4/9/11. Figure 3.17 illustrates the distilland temperature for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11. Figure 3.18 illustrates the distilland temperature for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11.

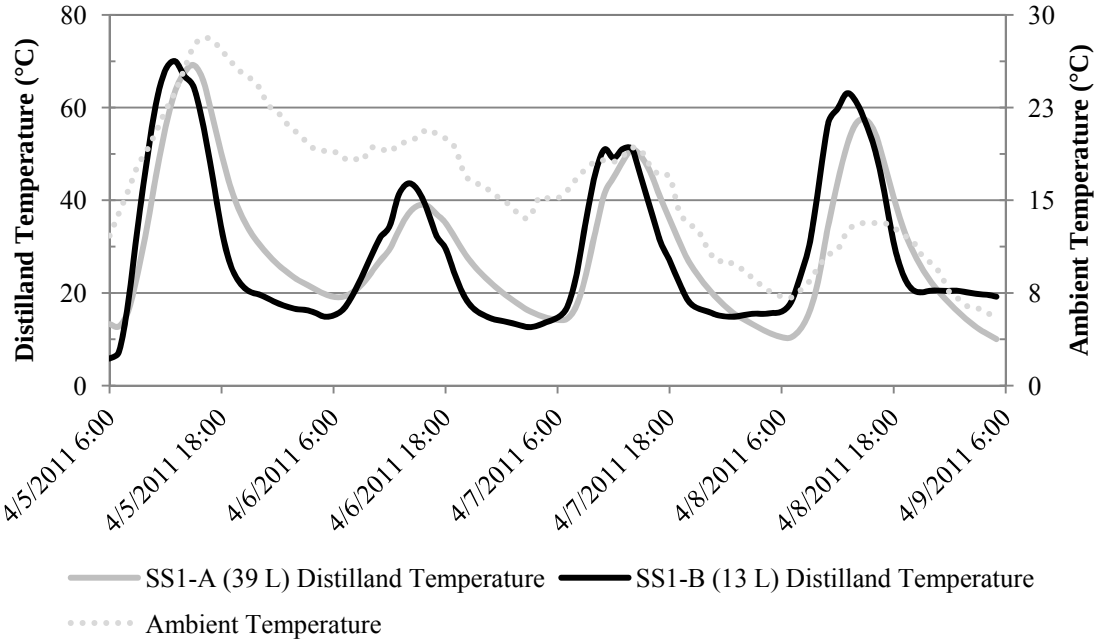


Figure 3.16 Hourly ambient and distilland temperature for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11

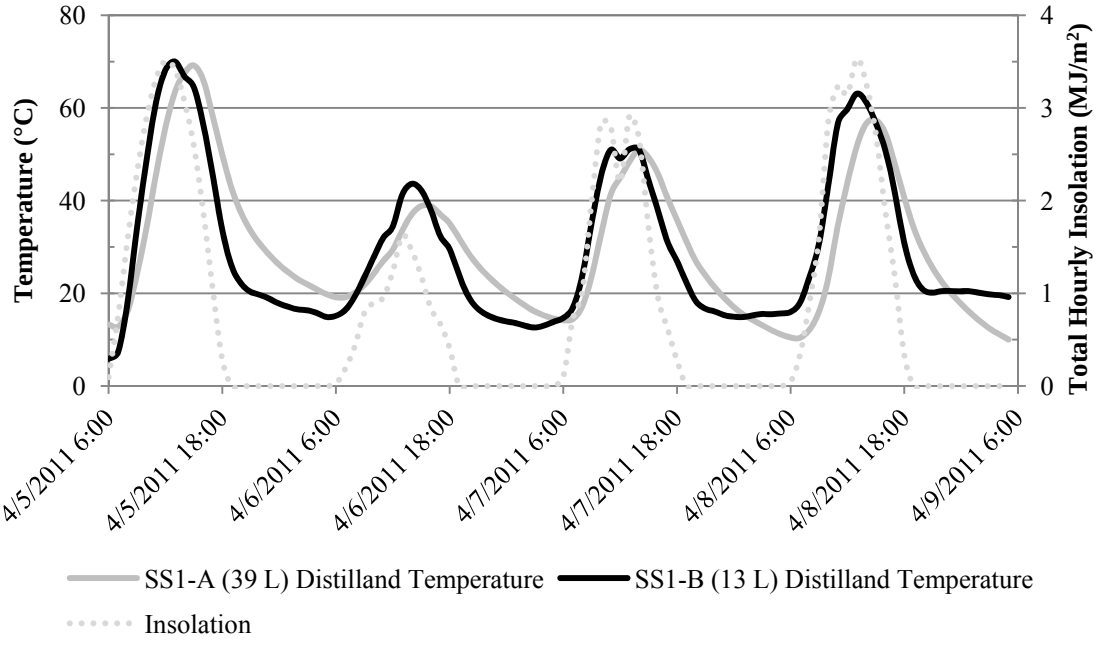


Figure 3.17 Hourly insolation and distilland temperature for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11

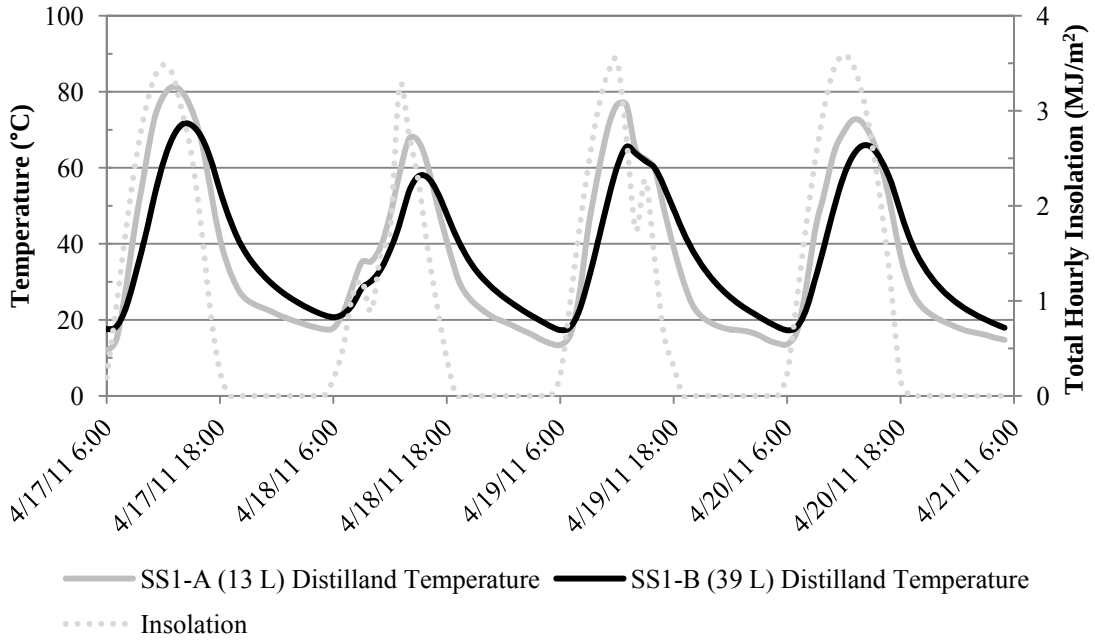


Figure 3.18 Hourly insolation and distilland temperature for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11

Figure 3.16 illustrates the distilland temperature for SS1-A (39 L) and SS1-B (13 L) as well as the ambient temperature between 4/5/11 and 4/9/11. Figure 3.16 illustrates that the ambient temperature has no effect on the distilland temperature. However, Figure 3.17 indicates a stronger relationship between the distilland temperature and the total hourly insolation.

Figure 3.17 illustrates the effect of distilland volume on the timing of the peak distilland temperature. SS1-B (13 L) experienced its maximum distilland temperature sooner than SS1-A (39 L) and more closely matched the change in insolation over time. The distilland temperature for SS1-B (39 L) lagged by one to two hours and did not reach the same peak temperature as SS1-A (13 L).

Figure 3.18 shows a similar behavior seen in Figure 3.17. The reversal of the operating distilland volume between SS1-A and SS1-B results in the continuation of the observed patterns with regard to the peak temperature lag and lower distilland temperature for the still operating with a larger distilland volume.

Figure 3.19 illustrates the difference in temperature between the distilland and the inner glass cover for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11.

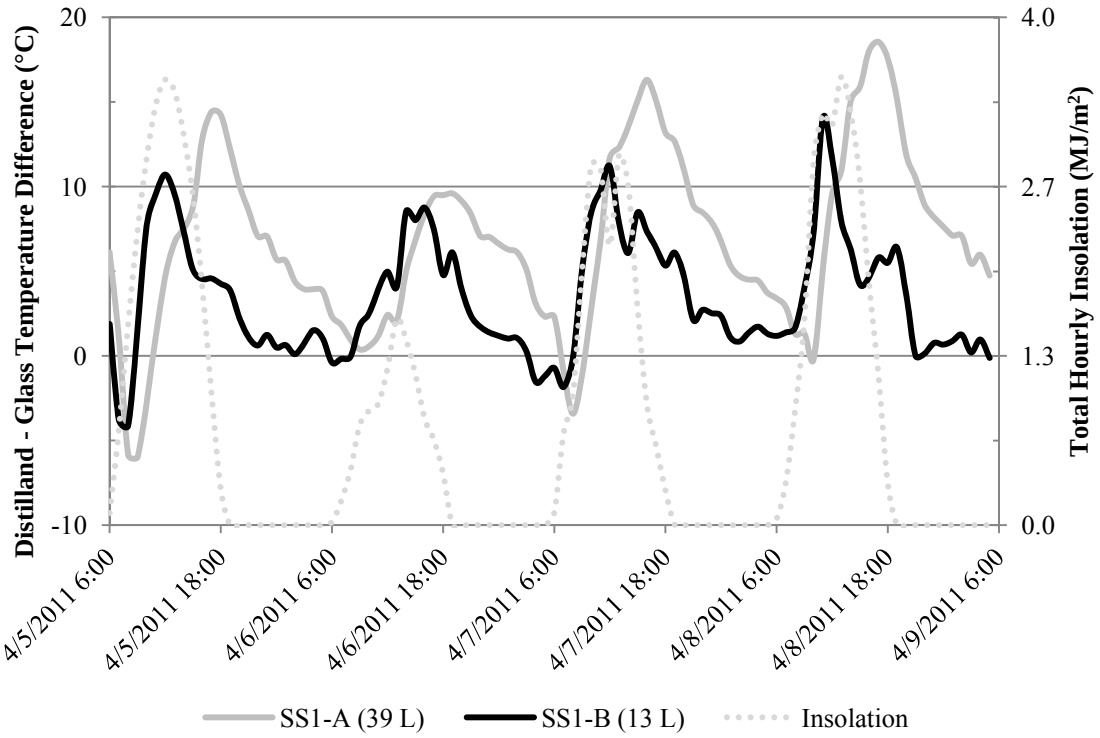


Figure 3.19 Distilland to inner glass cover temperature difference for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11

Figure 3.19 illustrates the difference in temperature between the distilland and the inner glass cover temperature. A positive difference indicates the distilland was warmer than the inner glass cover. Figure 3.19 indicates that SS1-B (13 L) reached its peak temperature difference close to the same time as peak insolation. Furthermore, SS1-A (39 L) reached its peak temperature difference one to two hours after peak insolation occurs. The timing of the peak temperature difference for the two different distilland volumes shows the same behavior as the peak distilland temperature as shown in Figure 3.17. Since SS1-B had a smaller distilland volume than SS1-A, SS1-B reached its peak temperature difference sooner than SS1-A. Figure 3.20 illustrates the difference in temperature between the distilland and the inner glass cover temperature for SS1-A (13L) and SS1-B (39 L) between 4/17/11 and 4/21/11.

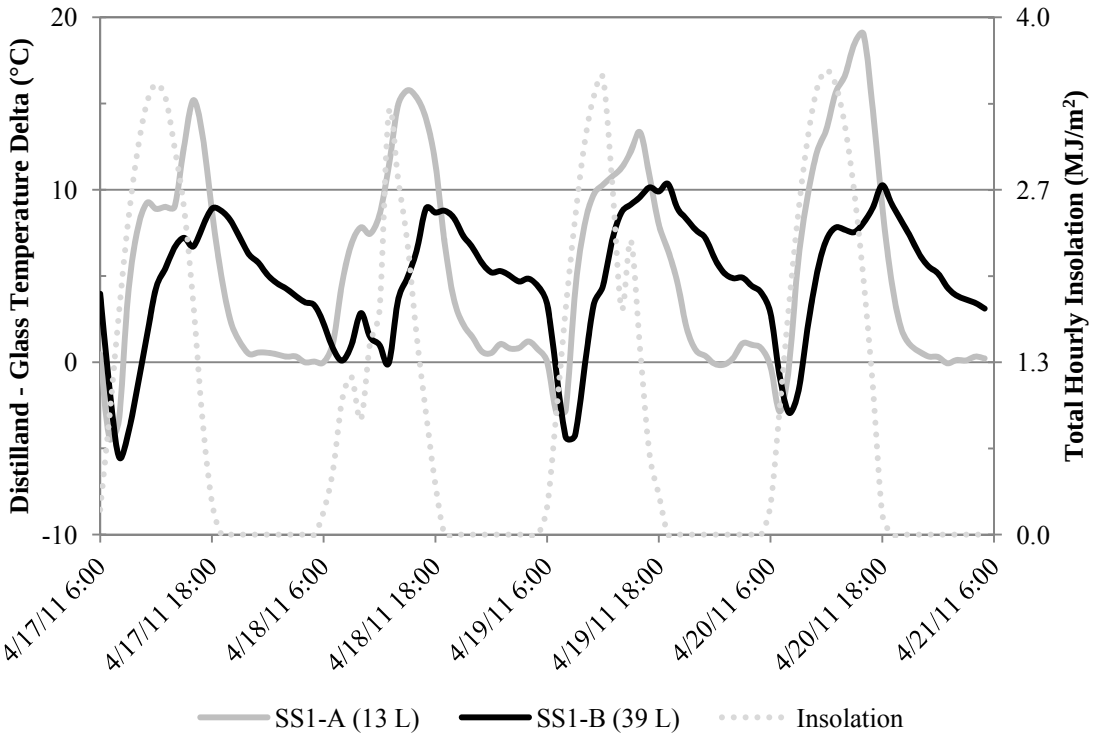


Figure 3.20 Distilland to inner glass cover temperature difference for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11

Figure 3.20 illustrates the magnitude of the difference between the distilland and the inner glass cover is larger for SS1-A (13 L) than SS1-B (39 L). Figure 3.20 indicates that despite the reversal of the operating distilland volume between SS1-A and SS1-B, SS1-A still performs with a higher distilland to inner glass cover temperature difference than SS1-B despite the operating distilland volume. However, the trend regarding the lag in peak temperature difference for the larger distilland volume is still noticeable. Figure 3.21 illustrates the magnitude of the difference between the distilland and the inner glass cover for SS1-A (30 L) and SS1-B (30 L) between 6/20/11 and 6/24/11.

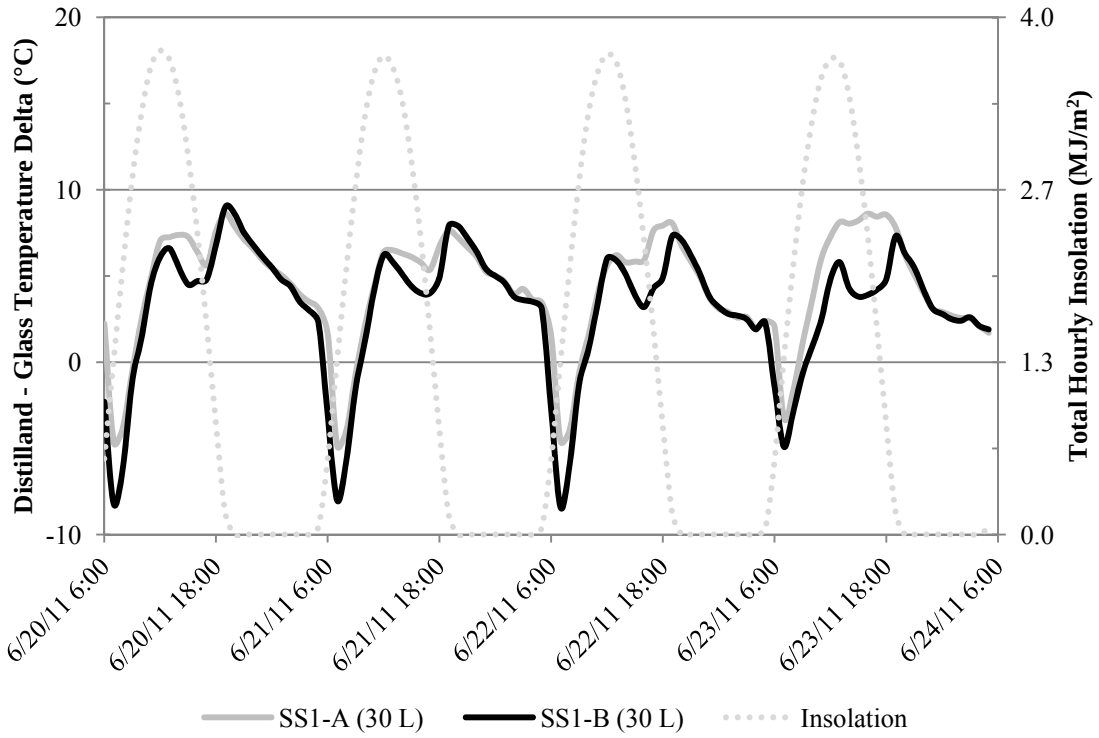


Figure 3.21 Distilland to inner glass cover temperature difference for SS1-A (30 L) and SS1-B (30 L) between 6/20/11 and 6/24//11

Figure 3.21 indicates that SS1-A performs with a higher distilland to inner glass cover temperature difference than SS1-B despite operating at the same distilland volume. However, SS1-B is able to meet the same maximum temperature difference as SS1-A for two out of the four presented days. Figures 3.19, 3.20, and 3.21 indicate that there is a difference in performance between SS1-A and SS1-B which results in SS1-A performing with a greater distilland to inner glass cover temperature.

3.4.5 Average Hourly Production

Figure 3.22 illustrates the temperature differences between SS1-A (13 L) and SS1-B (39 L). A positive temperature difference indicates the temperatures for SS1-B (39 L) are warmer than SS1-A (13 L). The negative temperature differences occur during the morning until 1400 for the distilland, 1300 for the vapor, and 1600 for the inner glass.

The peak negative temperature difference occurs during the peak insolation in the afternoon. After the peak insolation is observed, the magnitude of the negative temperature differences decreases until the peak positive temperature difference occurs in the evening. The peak positive temperature differences occur at 1900 for the distilland/inner glass and 1700 for the vapor.

Since SS1-A and SS1-B were operated simultaneously under different distilland volume scenarios, it was important to analyze the production distribution curves to identify hourly production characteristics. Figure 3.23 illustrates the average cumulative production distribution for SS1-A (13 L) and SS1-B (39 L). Hourly rain gauge data, between 4/16/11 and 5/4/11, was averaged to create Figure 3.23.

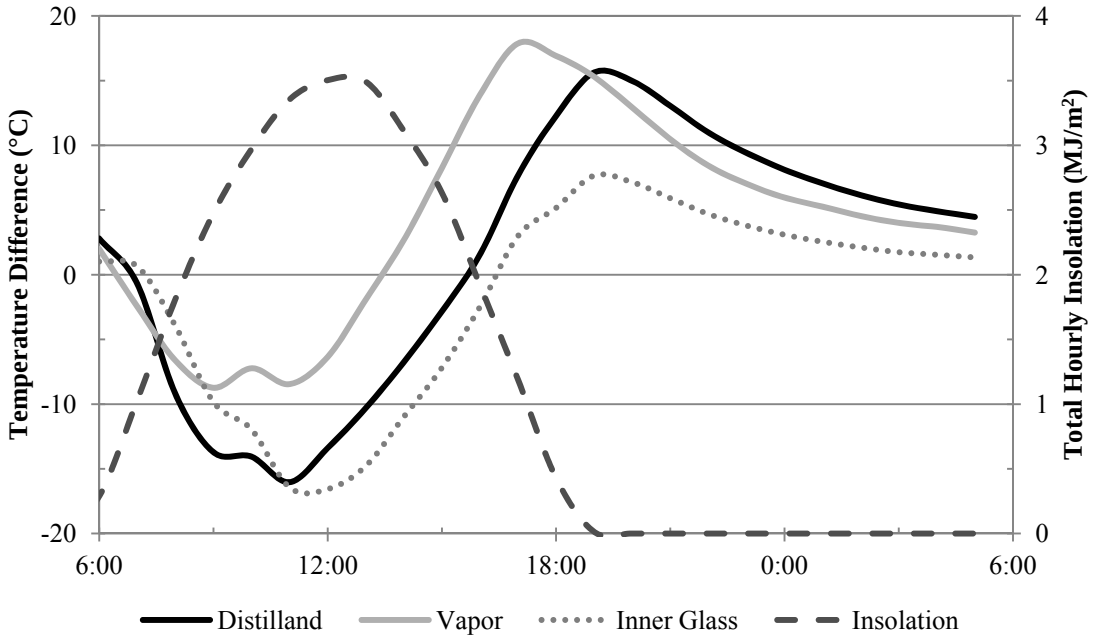


Figure 3.22 Average temperature differences for SS1-B (39 L) minus SS1-A (13 L) between 4/16/11 and 5/4/11

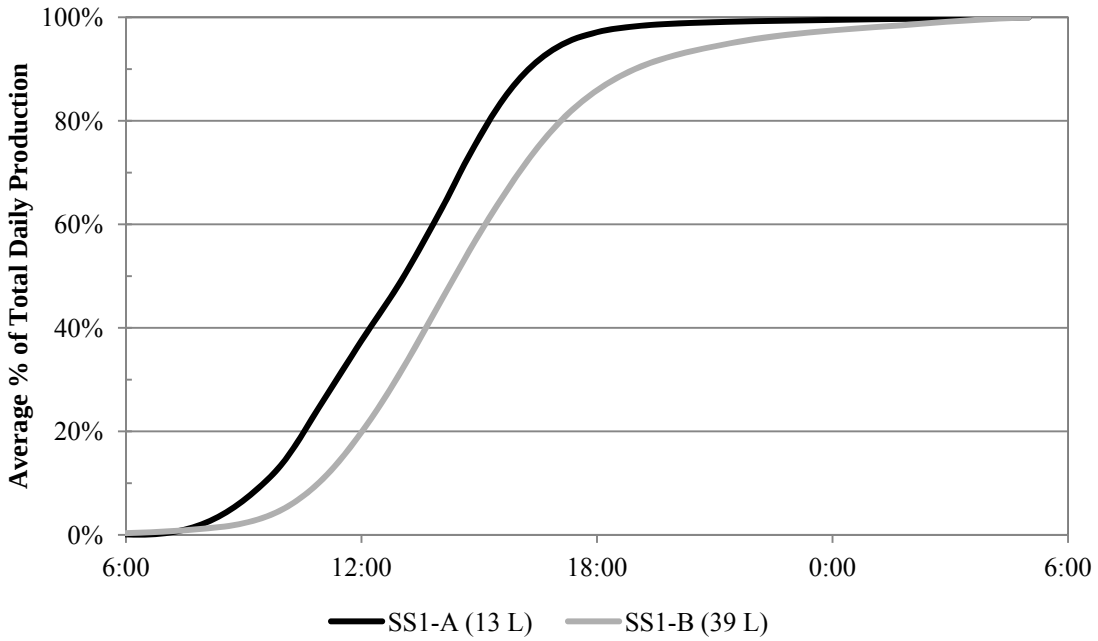


Figure 3.23 Average cumulative production for SS1-A (13 L) and SS1-B (39 L) between 4/16/11 and 5/4/11

Figure 3.23 illustrates the relationship between hourly production and the distilland volume. Figure 3.23 shows that SS1-B (39 L) produced an average of 90.1% of its daily distillate by sunset compared to 98.3% of daily distillate production by sunset for SS1-A (13 L), an 8.2% difference.

Figures 3.24 and 3.25 illustrate the average cumulative distillate produced for SS1-A and SS1-B, respectively, throughout the 2011 study. Figures 3.24 and 3.25 were created by compiling the recorded hourly distillate production data for each respective operating distilland volume scenario. The cumulative hourly data for each scenario were then averaged to create Figures 3.24 and 3.25. Figures 3.24 and 3.25 continue to show the effect of distilland volume on day and night production as shown in Figure 3.23.

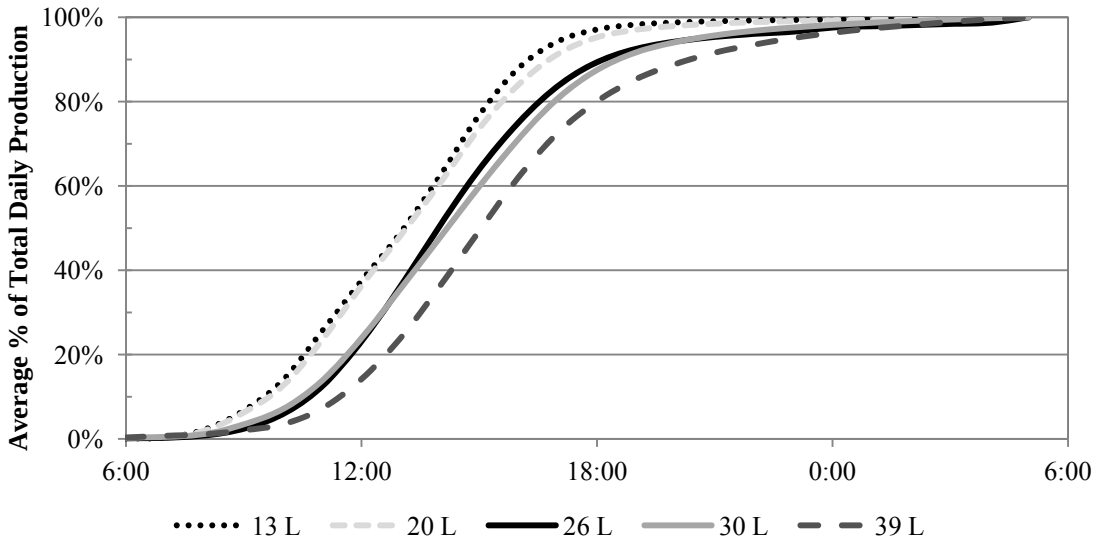


Figure 3.24 Average cumulative distillate production distribution for SS1-A

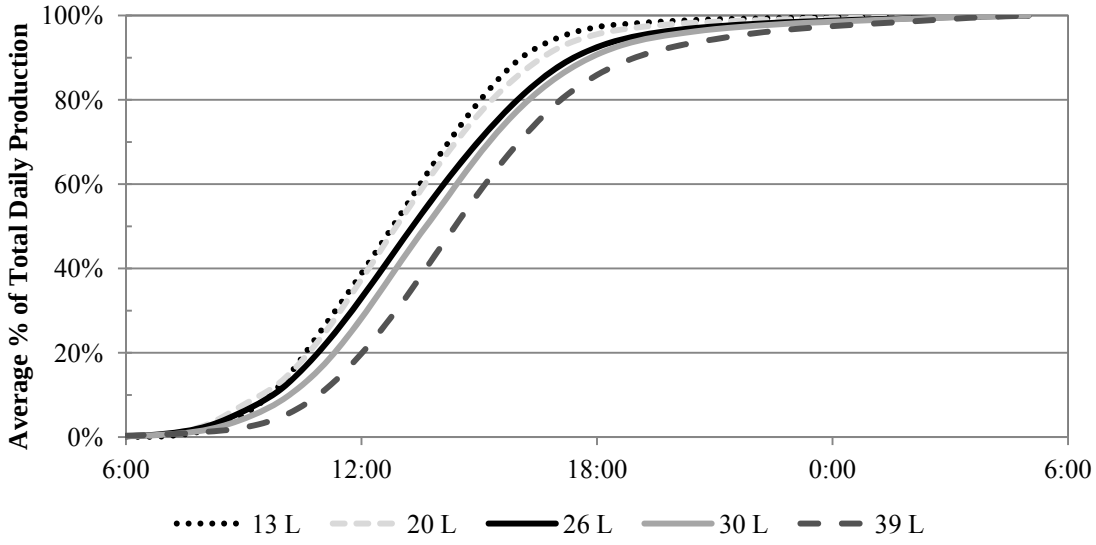


Figure 3.25 Average cumulative distillate production distribution for SS1-B

Scenarios with a distilland volume of 13 L, for SS1-A and SS1-B, produced more of their distillate during daylight hours and less after sunset. Scenarios operating with a starting distilland volume of 39 L, for SS1-A and SS1-B, produced less of their distillate during daylight hours and produced more after sunset when compared to the distilland

volume scenarios operating with 13 L.

Figure 3.24, for SS1-A, illustrates a slightly wider gap in the cumulative distribution when comparing the 13 L and 39 L scenario when compared to SS1-B as seen in Figure 3.25. Table 3.2 details the average percent of total daily distillate produced by sunset for SS1-A and SS1-B.

Table 3.2 Average percent of total daily distillate produced by sunset

	SS1-A		SS1-B	
Date	**Distilland Volume**	**% of Daily Distillate Produced by Sunset**	**Distilland Volume**	**% of Daily Distillate Produced by Sunset**
4/4-4/15	39 L	80.1%	13 L	97.3%
4/16-5/4	13 L	97.1%	39 L	85.9%
5/5-5/27	26 L	89.4%	20 L	95.6%
5/28-6/18	20 L	95.3%	26 L	92.5%
6/19-7/31	30 L	87.5%	30 L	90.7%

Table 3.2 shows a 17% decrease in day production between the 13 L and 39 L distilland scenario and a 7.7% decrease in day production between the 13 L and 26 L distilland scenario for SS1-A. Table 3.2 also shows an 11.4% decrease in day production between the 13 L and 39 L distilland scenario and a 4.8% decrease in day production between the 13 L and 26 L distilland scenario for SS1-B.

Although the previously stated cumulative production decrease occurs during different times of the year for each respective still, a comparison in cumulative production for SS1-A and SS1-B for the same time period reveals the same trend. SS1-A produced 11.2% more distillate by sunset when operated with a distilland volume of 13 L compared to SS1-B when operated with a distilland volume 39 L for the same time period between 4/16/11-5/4/11.

3.4.6 Modeling Day and Night Production

A study performed by Mathioulakis et al. (1999) is one of the few studies on solar stills that focus on modeling hourly, day, and night production. The model developed by Mathioulakis et al. (1999) was previously described in Equations 1.11 and 1.12. Verification of the Mathioulakis et al. (1999) method was carried out using the insolation and temperature data for SS1-A between 4/5/11 and 7/31/11. Equations 3.2 and 3.3 illustrate the original Equations 1.11 and 1.12, respectively, with the calculated coefficients. The coefficients were calculated using the data gathered for the 2011 study and performing a multivariable regression analysis.

$$M_{w,d} = 0.245 \cdot H_d - 0.014 \cdot \left(T_{w,d} - T_{a,d}\right) - 1.768$$

Equation 3.2 Daytime water production correlation (Mathioulakis et al., 1999) with calculated coefficients

$$M_{w,n} = 0.012 \cdot \left(T_{w,n} - T_{a,n}\right) + 0.320$$

Equation 3.3 Nocturnal water production correlation (Mathioulakis et al., 1999) with calculated coefficients

d	Day subscript	T_w	Average distilland temperature
n	Night subscript	T_a	Average ambient air temperature
w	Water subscript	*H*	Daily Insolation

Figures 3.26 and 3.27 illustrate the predicted day vs. actual day production and the predicted night vs. actual night production comparison, respectively. The day production predictions had a mean absolute error of 9.4% and featured an R^2 value of 0.854. The night production predictions had a mean absolute error of 86.1% and featured an R^2 value of 0.070.

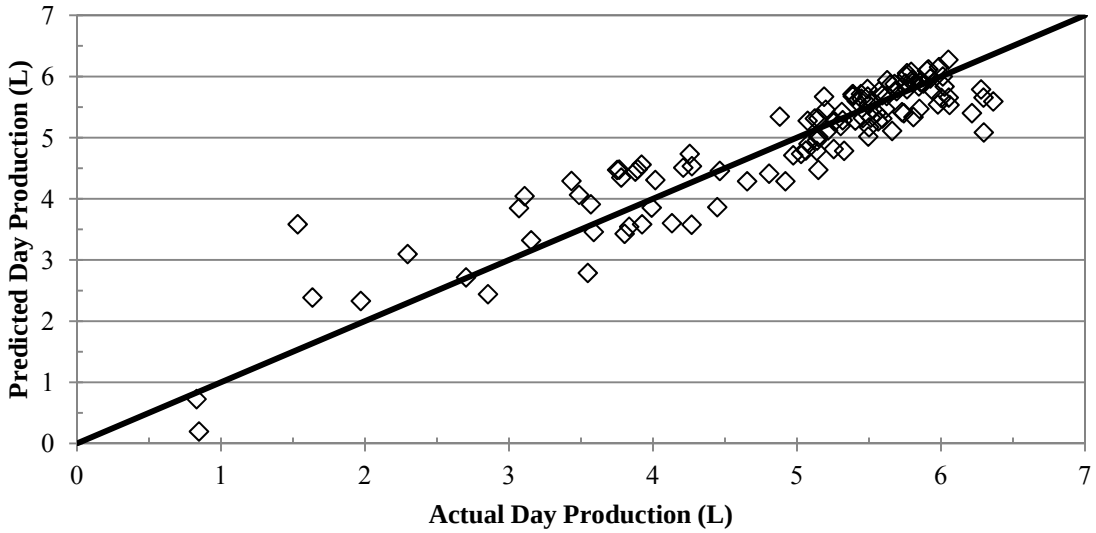


Figure 3.26 Predicted day vs. actual day production using Mathioulakis et al.'s method (1999)

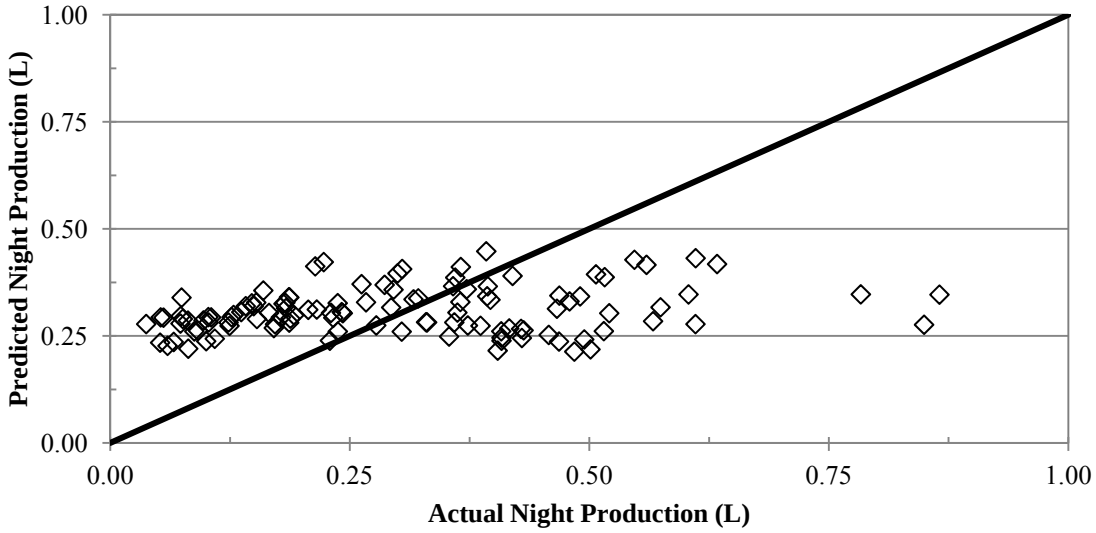


Figure 3.27 Predicted night vs. actual night production using Mathioulakis et al.'s method (1999)

Because of the high observed error for the night predictions, it was hypothesized that the inclusion of the distilland's stored energy with Equation 1.12 would decrease the error observed in the night production predictions. Equation 3.4 shows the modified

Mathioulakis el al. (1999) equation to include distilland stored energy (C). The stored energy was found my multiplying the night time distilland volume (operating distilland volume minus actual day production) by the heat of vaporization for water at the respective average night time distilland temperature. Figure 3.28 illustrates the results of the predicted night vs. actual night production for the modified Mathioulakis et al. (1999) method.

$$M_{w,n} = 0.007 \cdot C - 0.006 \cdot \left(T_{w,n} - T_{a,n}\right) - 0.074$$

Equation 3.4 Modified nocturnal water production correlation (Mathioulakis et al., 1999) with calculated coefficients

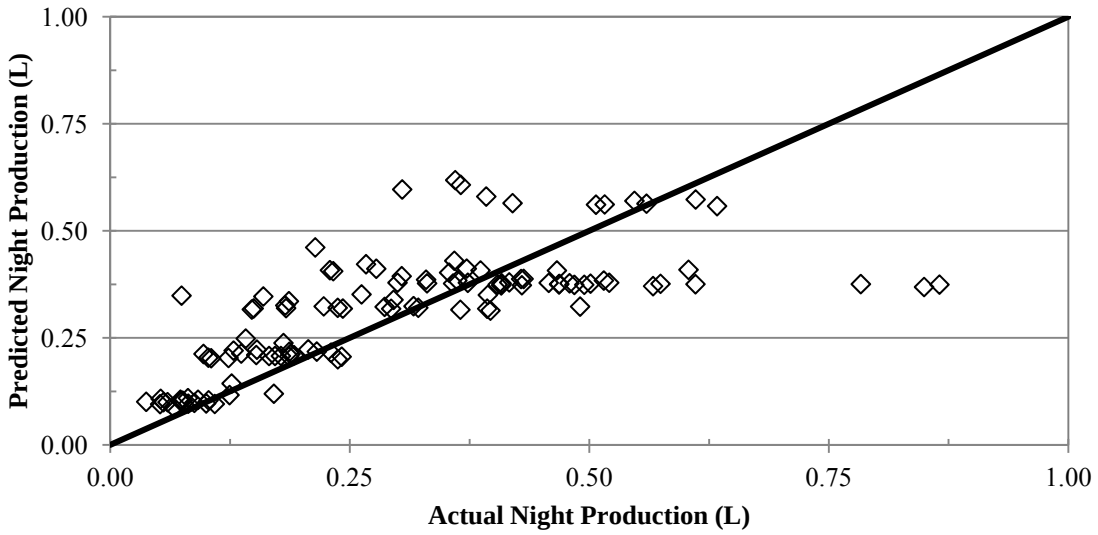


Figure 3.28 Predicted vs. actual night production using modified Mathioulakis et al. (1999) method

By including the distilland's stored energy, the mean absolute error for the night production predictions fell from 86.1% to 34.9% (a 51.2% decrease). The reduced error can be seen by comparing Figure 3.27 and 3.28 and the fact that the predicted data points align more closely with the origin bisector for the modified night time production

correlation than with the original night time production correlation. The R^2 value for the modified night time production correlation was 0.444 while the original night time production correlation had an R^2 value of 0.070.

Figures 3.29 and 3.30 illustrate the predicted daily production using the addition of the predictions for day and night production according to the original and the modified Mathioulakis et al. (1999) method, respectively.

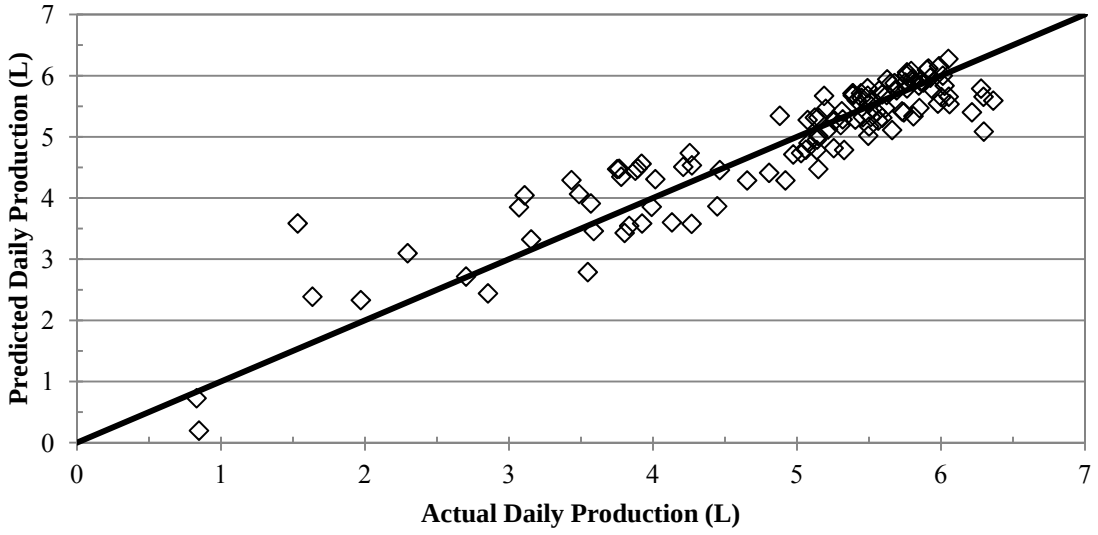


Figure 3.29 Mathioulakis et al. (1999) total predicted daily vs. actual daily production

The total daily production prediction, illustrated in Figure 3.29, performed with a mean absolute error of 7.5% and featured an R^2 value of 0.876. Figure 3.29 also shows how tightly the predictions were distributed along the origin bisector when graphed against the actual daily production.

Figure 3.30 illustrates the daily production predictions, using the combined results from the daytime production and the modified night time production method, performing with a mean absolute error of 8.8% and features an R^2 value of 0.748.

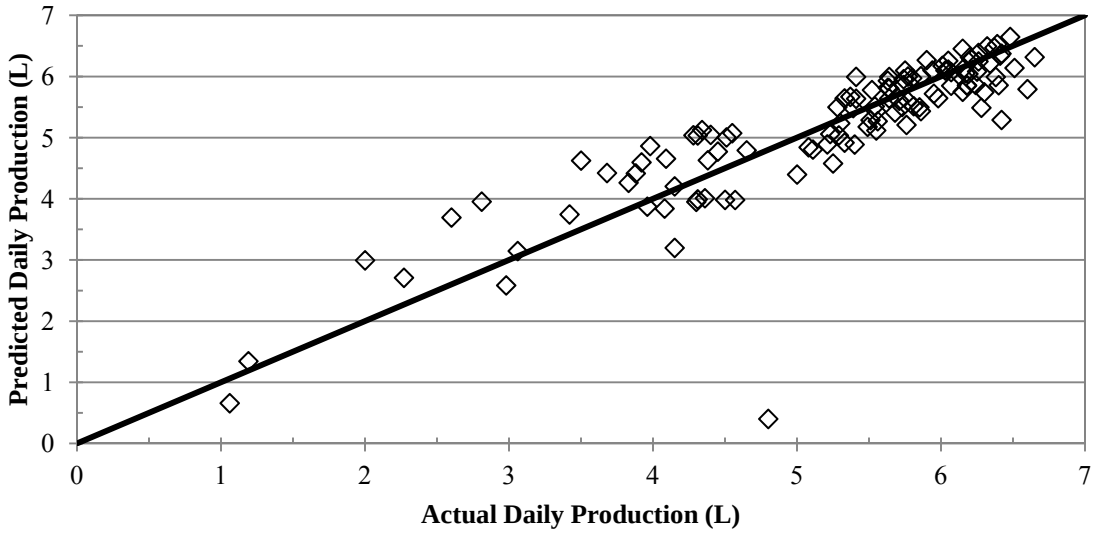


Figure 3.30 Modified Mathioulakis et al. (1999) total predicted daily vs. actual daily production

Despite predictions having low error for the night time production and having total daily predictions distributed tightly around the origin bisector when graphed against actual daily production, the modified Mathioulakis et al. (1999) method is slightly less accurate than the original Mathioulakis et al. (1999) method when comparing the overall daily production predictions.

The daily production prediction method designed by Mathioulakis et al. (1999) performs with low error and a high R^2 value; however, the method still relies on direct measurement of the hourly distilland and ambient air temperature to generate input data.

3.5 Heat Transfer Coefficients

3.5.1 Hourly Heat Transfer Coefficients

The hourly heat transfer coefficients for SS1-A and SS1-B were calculated based on the Jakob (1949) correlation as used by Dunkle (1961) and later converted into S.I. units by Tiwari and Tiwari (2006). The heat transfer coefficients and their calculation methods

were originally presented in Equations 1.5 and 1.6. Measured distilland to glass ΔTs were used to calculate Dunkle's (1961) ΔT′ from Tiwari and Tiwari's (2006) metric conversion of Dunkle's ΔT′ correlation. Once the ΔT′ was calculated, the values of the h_{cw} could be determined. Figures 3.31 and 3.32 illustrate the distilland temperature and h_{cw} for SS1-A (39 L) and SS1-B (13 L), respectively between 4/10/11 and 4/13/11. Figures 3.33 and 3.34 illustrate the distilland temperature and h_{ew}, for SS1-A (39 L) and SS1-B (13 L), respectively, between 4/10/11 and 4/13/11.

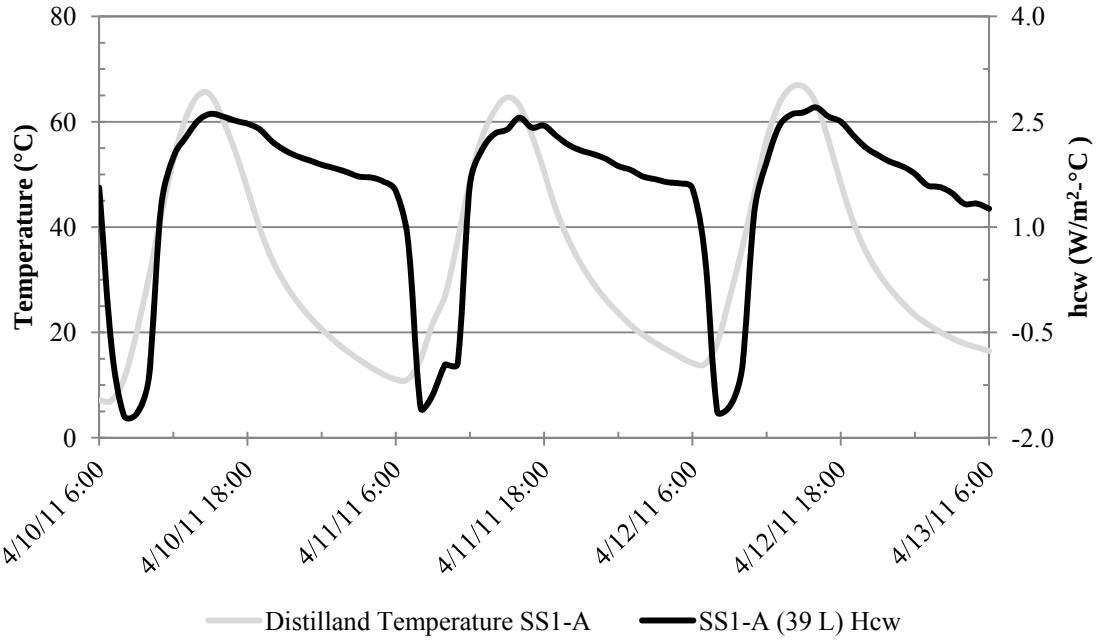


Figure 3.31 Hourly h_{cw} and distilland temperature for SS1-A (39 L) between 4/10/11 and 4/13/11

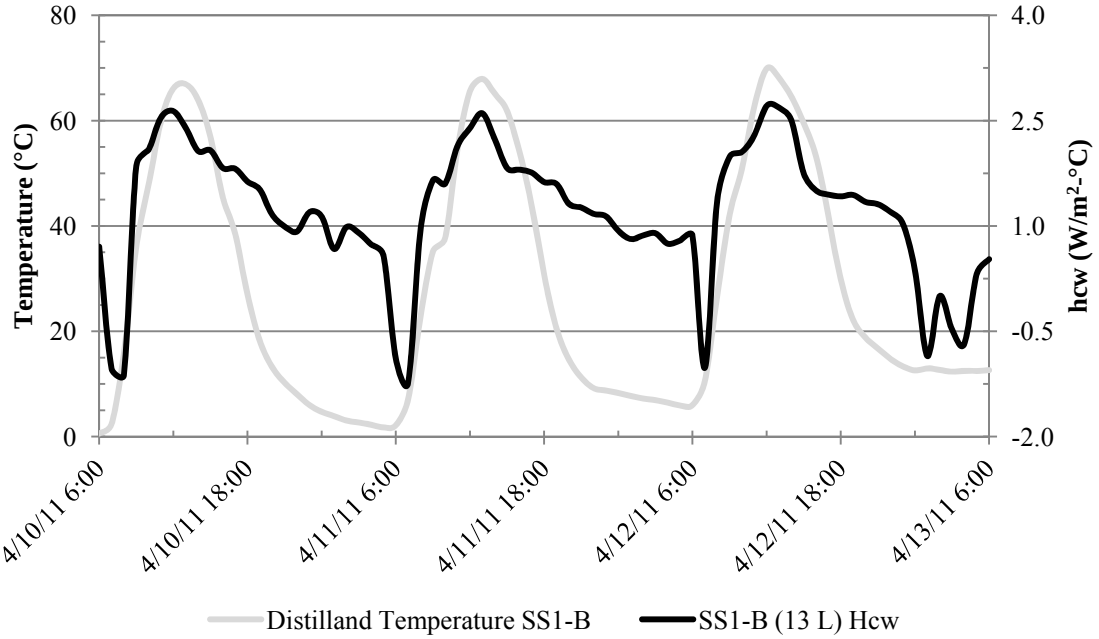


Figure 3.32 Hourly h_{cw} and distilland temperature for SS1-B (13 L) between 4/10/11 and 4/13/11

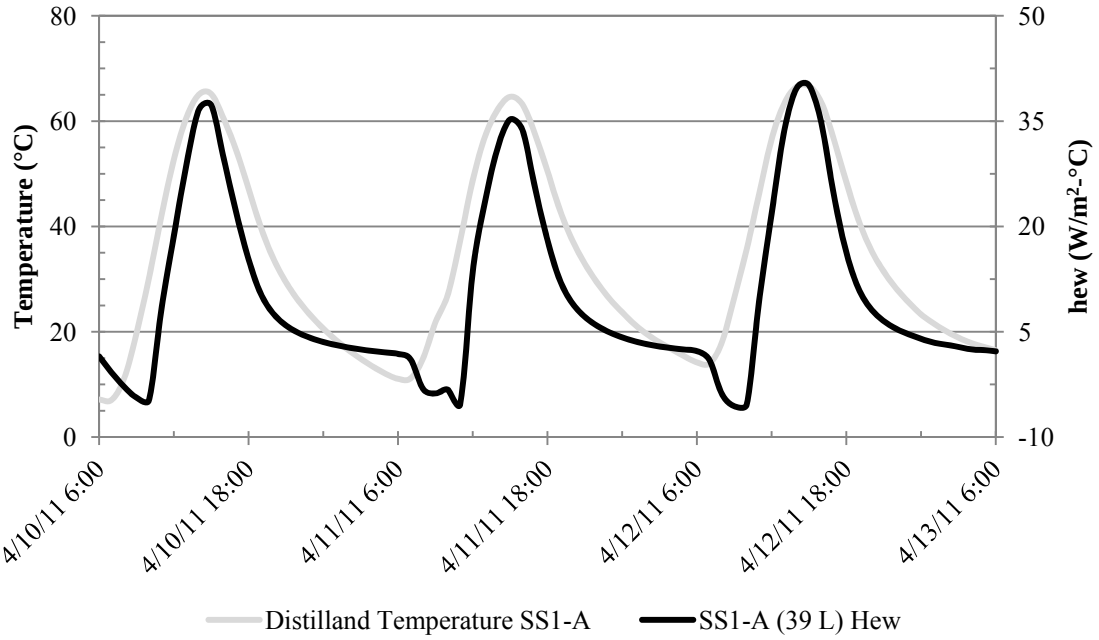


Figure 3.33 Hourly h_{ew} and distilland temperature for SS1-A (39 L) between 4/10/11 and 4/13/11

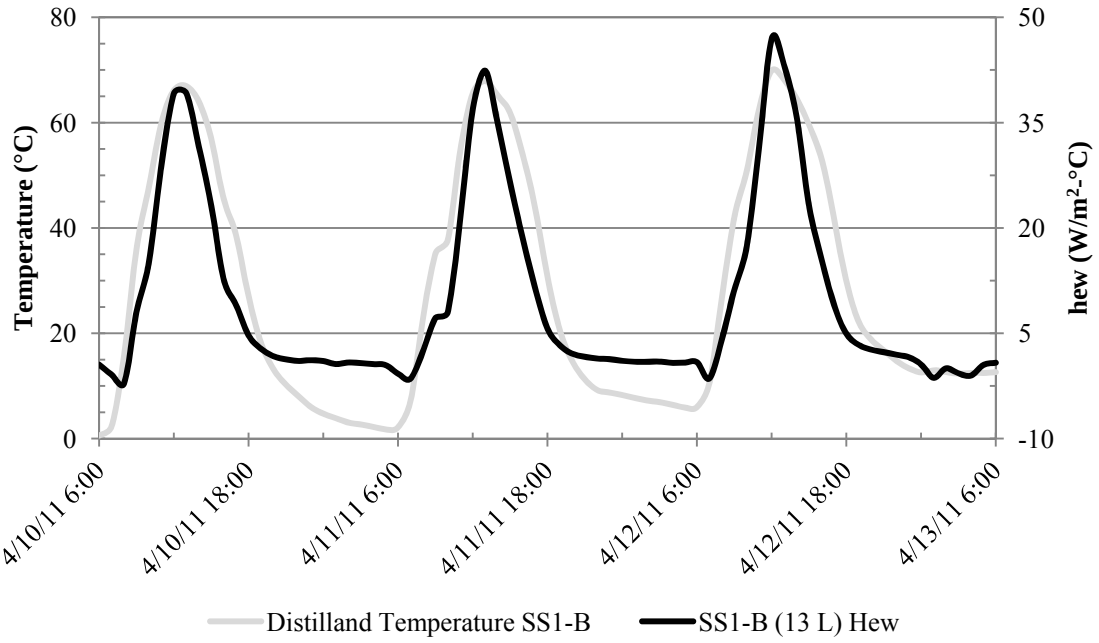


Figure 3.34 Hourly h_{ew} and distilland temperature for SS1-B (13 L) between 4/10/11 and 4/13/11

The calculated h_{cw} and h_{ew} both reach their maximum values near the same time the peak distilland temperature is reached. Figures 3.32 and 3.34 indicate that SS1-B (13 L) reaches the maximum value for h_{ew} sooner than SS1-A (39 L). Figures 3.31 and 3.33 also illustrate a one to two hour lag between the rise in distilland temperature and the initial increase for h_{cw} and h_{ew} for SS1-A (39 L). SS1-B (13 L) does not experience the lag that was noticed for SS1-A (39 L).

Figures 3.35 and 3.36 illustrate the h_{cw} and distilland temperature for SS1-A (13 L) and SS1-B (39 L) between 4/20/11 and 4/23/11, respectively. Figures 3.37 and 3.38 illustrate the h_{ew} and distilland temperature for SS1-A (13 L) and SS1-B (39 L) between 4/20/11 and 4/23/11.

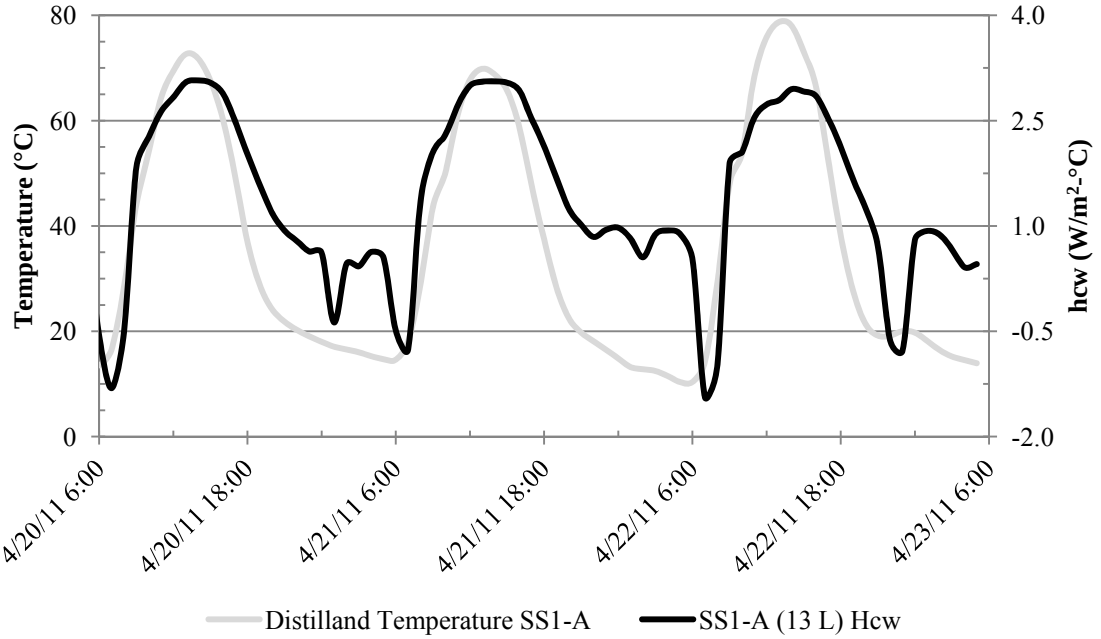


Figure 3.35 Hourly h_{cw} and distilland temperature for SS1-A (13 L) between 4/20/11 and 4/23/11

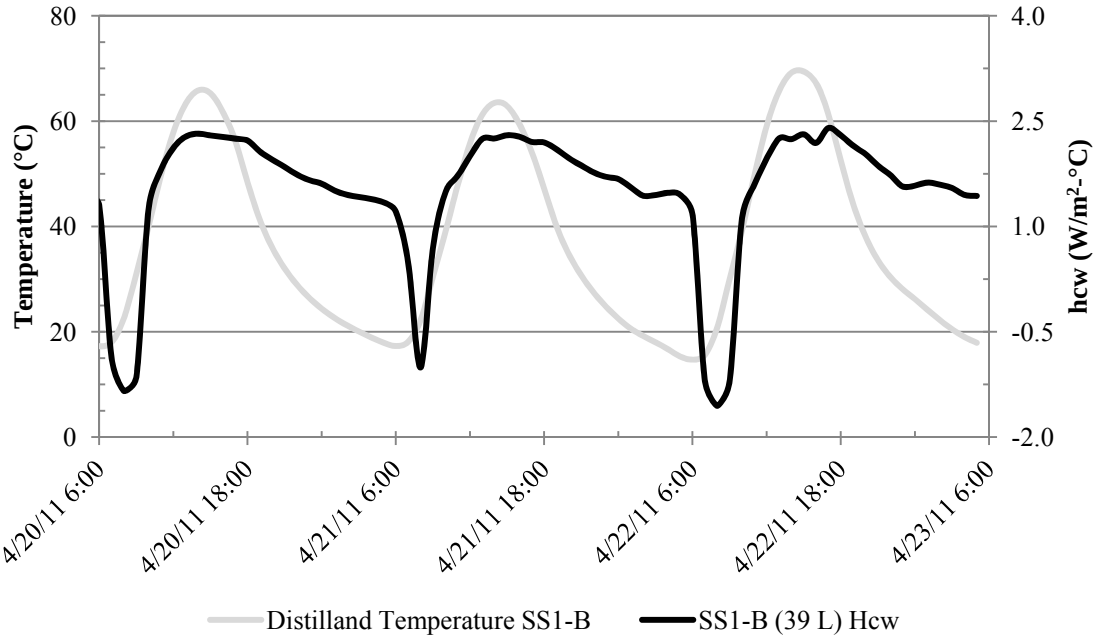


Figure 3.36 Hourly h_{cw} and distilland temperature for SS1-B (39 L) between 4/20/11 and 4/23/11

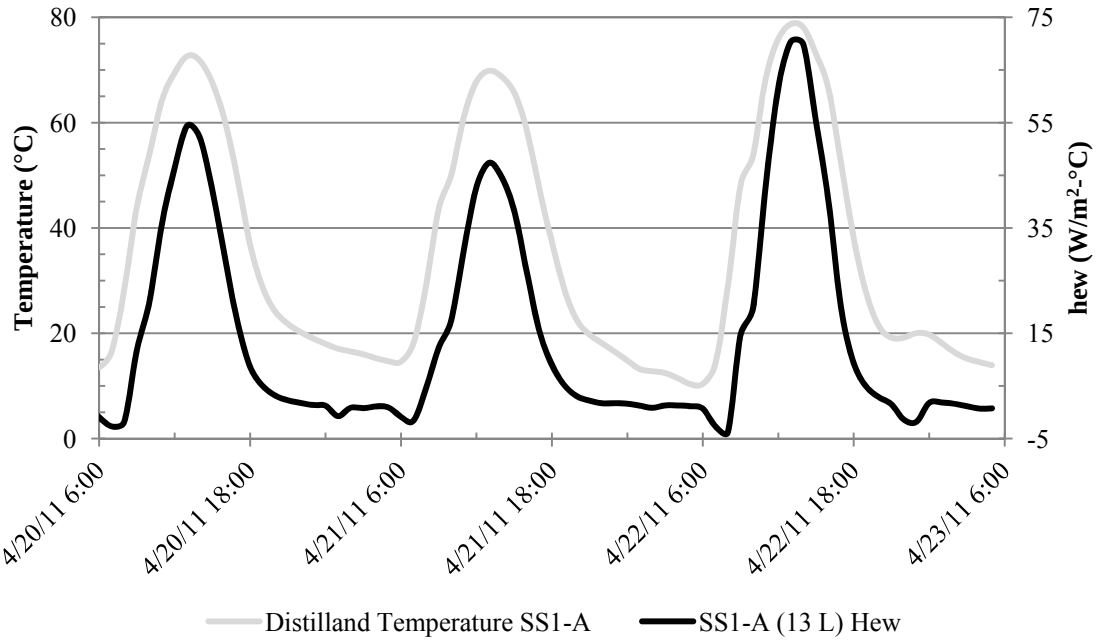


Figure 3.37 Hourly h_{ew} and distilland temperature for SS1-A (13 L) between 4/20/11 and 4/23/11

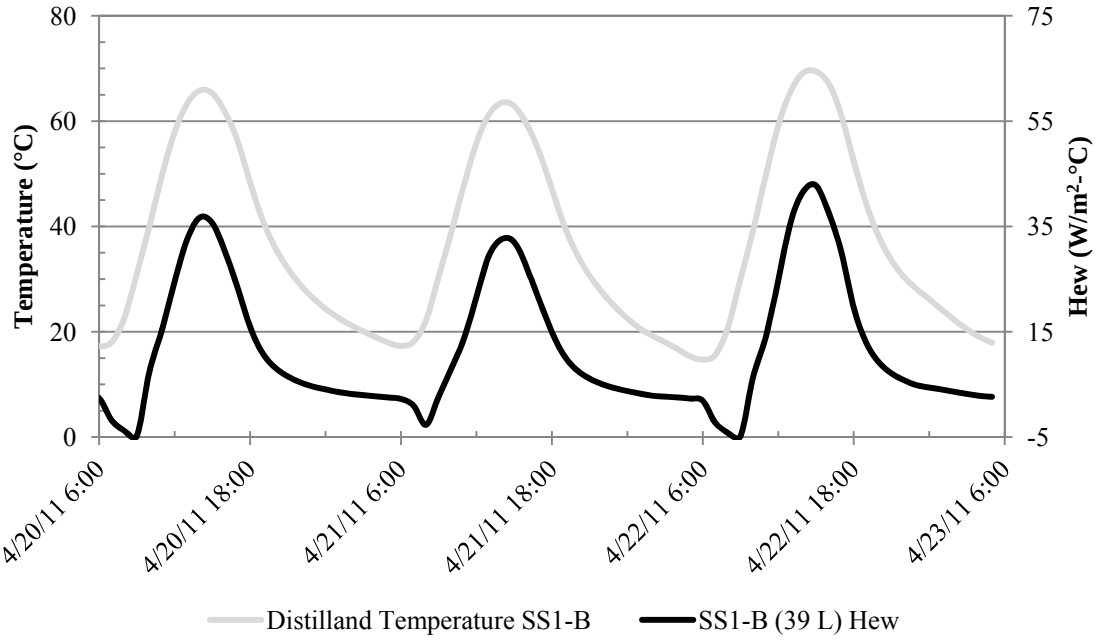


Figure 3.38 Hourly h_{ew} and distilland temperature for SS1-B (39 L) between 4/20/11 and 4/23/11

Figures 3.36 and 3.38 show the same trend regarding the lag in peak distilland temperature and maximum values for h_{cw} and h_{ew} as well as the lag between the rise in distilland temperature and the initial increase for h_{cw} for SS1-B (39 L) as was found for SS1-A (39 L). These trends are noticeable in the entire data set. Between 4/20-4/23, SS1-A's distilland temperature was higher (smaller distilland volume) than SS1-B's and therefore resulted in higher heat transfer coefficients.

3.5.2 Heat Transfer and Hourly Production

Figures 3.39 and 3.40 illustrate SS1-B (13 L)'s hourly distillate production graphed along with the convective and evaporative heat transfer coefficients, respectively, between 4/5/11 and 4/9/11. Figures 3.39 and 3.40 demonstrate how the distillate production and heat transfer coefficients track each other well throughout the day.

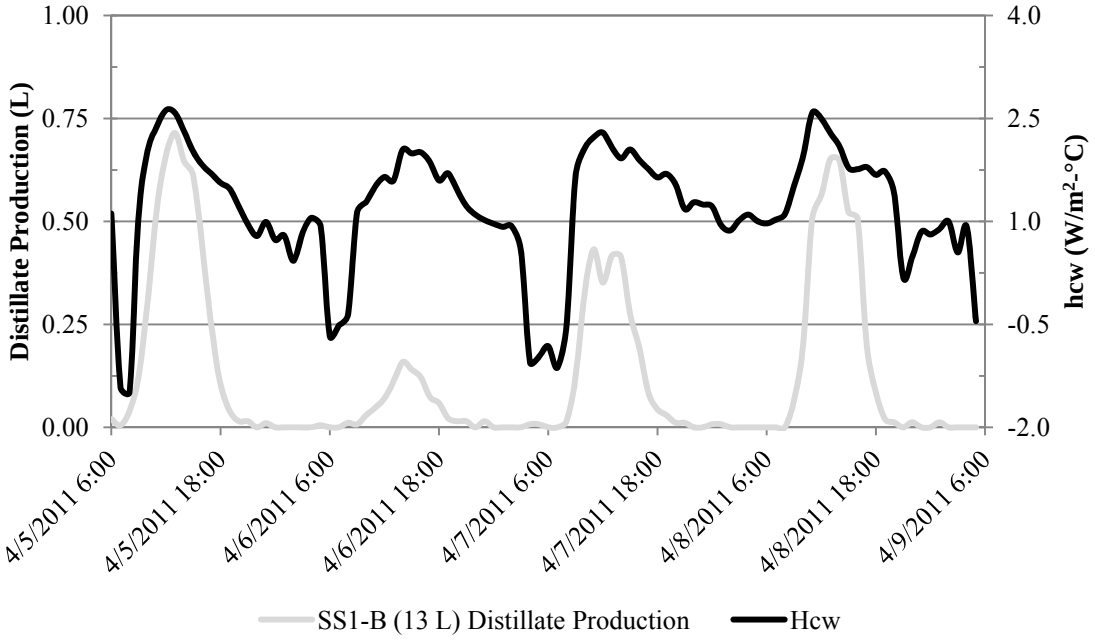


Figure 3.39 Hourly distillate production and h_{cw} for SS1-B (13 L) between 4/5/11 and 4/9/11

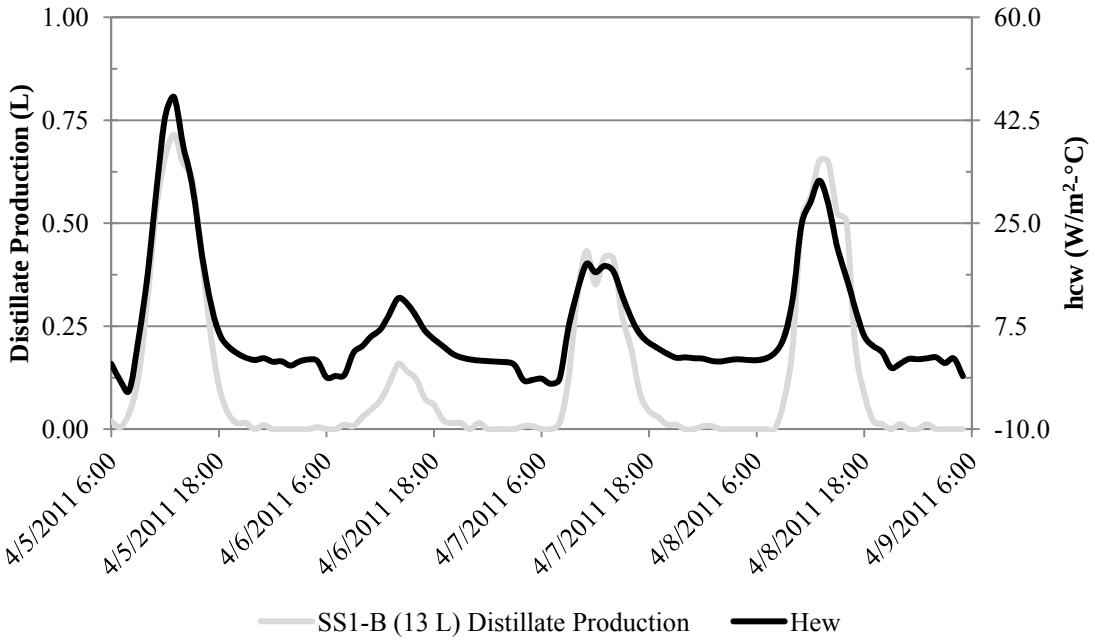
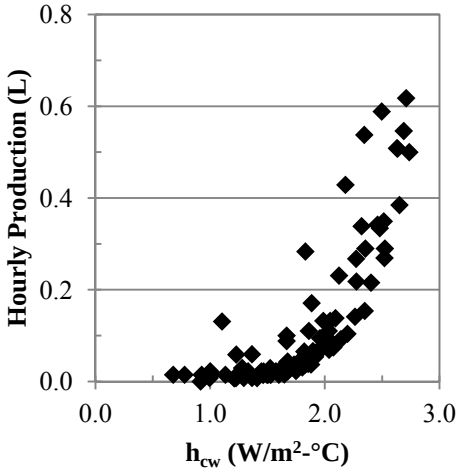


Figure 3.40 Hourly distillate production and h_{ew} for SS1-B (13 L) between 4/5/11 and 4/9/11

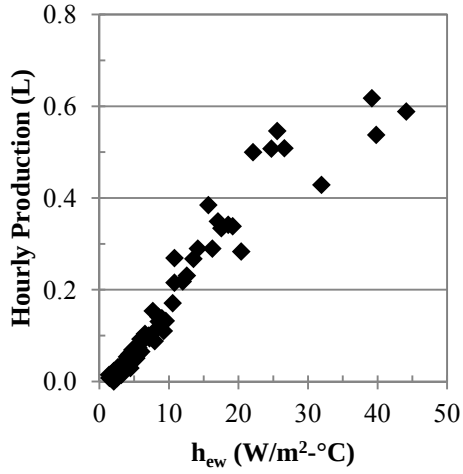
Pearson correlation coefficients (R values) were calculated to measure the linear dependence between the hourly distillate production and the h_{cw} and h_{ew}, respectively. The R value between hourly distillate production and h_{cw} was calculated to be 0.624 while the R value between hourly distillate production and h_{ew} was calculated to be 0.965. The stronger R value between h_{ew} and hourly distillate production can be attributed to the lack of lag between the increase/decrease between h_{ew} and distillate production.

Figure 3.41 illustrates various plots for the hourly production graphed against h_{cw} and h_{ew} for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11. The larger R value for production vs. h_{ew} compared to production vs. h_{cw} can be clearly identified when comparing Figures 3.41b and 3.41d to 3.41a and 3.41c. The graphs for production vs. h_{ew} for SS1-A (39 L) and SS1-B (13 L) are more linear in nature compared to the graphs for

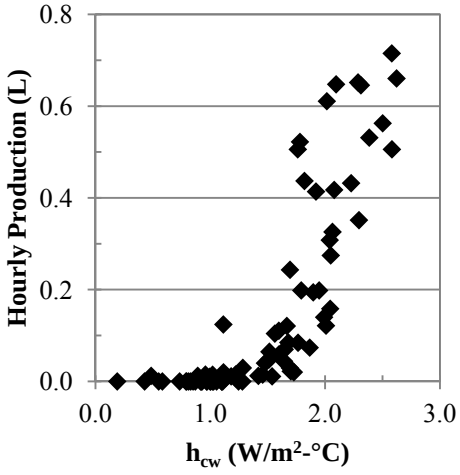
production vs. h_{cw} which illustrate some exponential behavior.



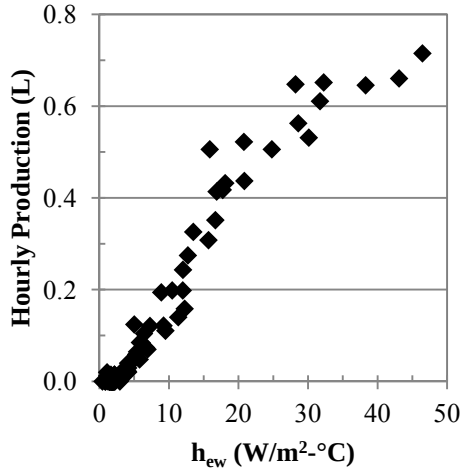
a. SS1-A (39 L) vs. h_{cw}



b. SS1-A (39 L) vs. h_{ew}



c. SS1-B (13 L) vs. h_{cw}



d. SS1-B (13 L) vs. h_{ew}

Figure 3.41 Hourly production vs. h_{cw} and h_{ew} for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11

Figures 3.42 and 3.43 illustrate SS1-B (13 L)'s hourly distillate production graphed

along with the convective and evaporative heat transfer coefficients, respectively, between 4/17/11 and 4/21/11.

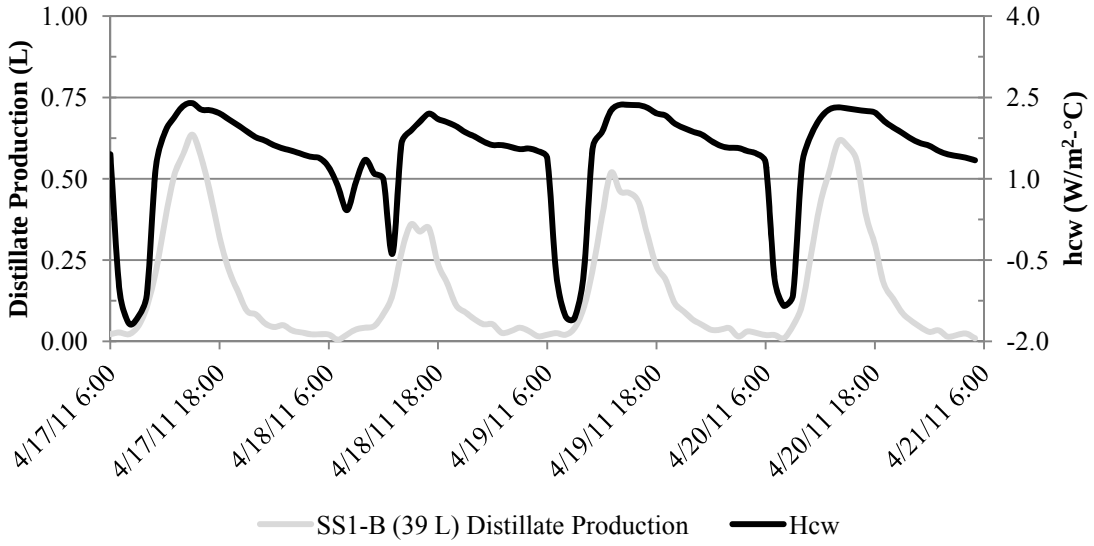


Figure 3.42 Hourly distillate production and h_{cw} for SS1-B (39 L) between 4/17/11 and 4/21/11

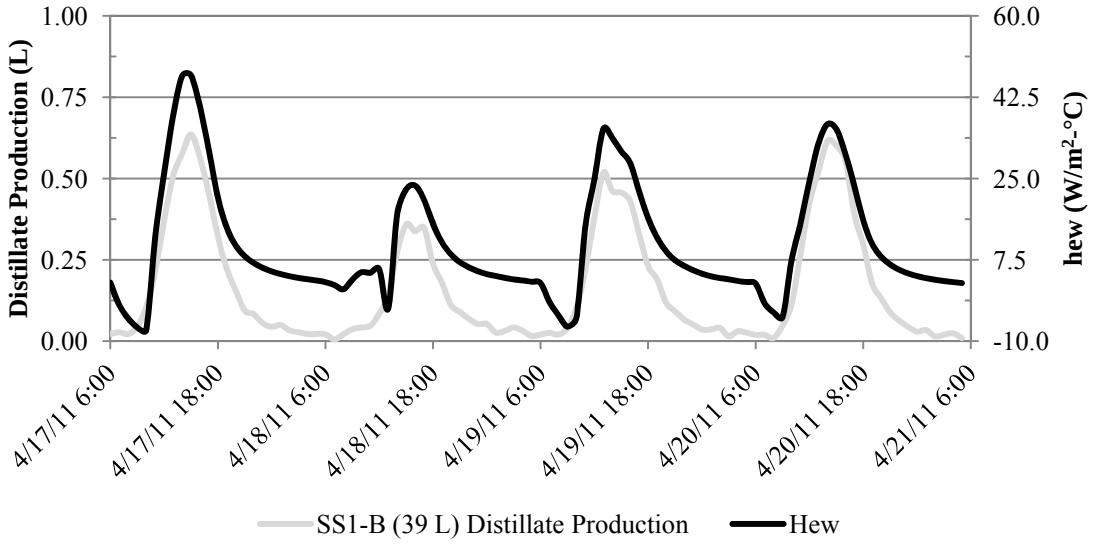
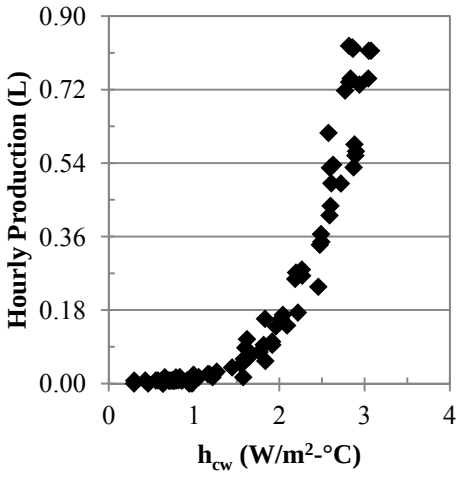
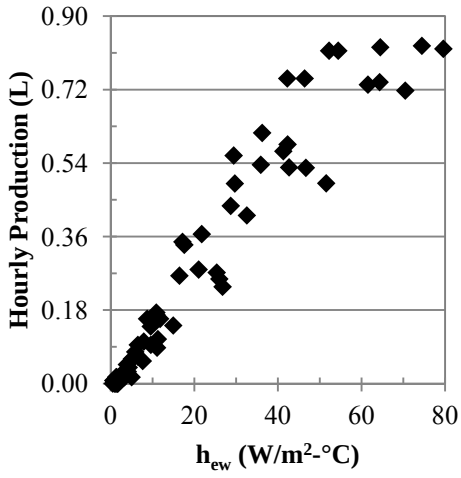


Figure 3.43 Hourly distillate production and h_{ew} for SS1-B (39 L) between 4/17/11 and 4/21/11

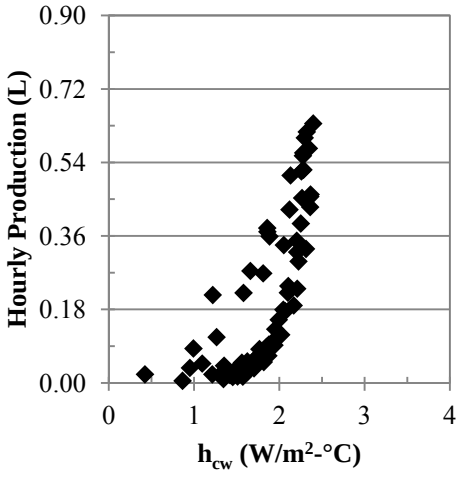
Figure 3.44 illustrates various plots for the hourly production graphed against h_{cw} and h_{ew} for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11.



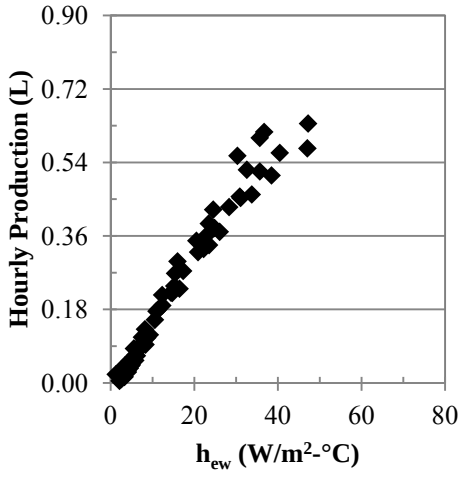
a. SS1-A (13 L) vs. h_{cw}



b. SS1-A (13 L) vs. h_{ew}



c. SS1-B (39 L) vs. h_{cw}



d. SS1-B (39 L) vs. h_{ew}

Figure 3.44 Hourly production vs. h_{cw} and h_{ew} for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11

The R value between the hourly production and h_{cw} and hourly production and h_{ew} was calculated to be 0.495 and 0.956, respectively. The larger R value for production vs. h_{ew} compared to production vs. h_{cw} can be clearly identified when comparing Figures 3.44b and 3.44d to 3.44a and 3.44c. Figures 3.42 and 3.43 illustrate the same behavior identified in Figures 3.39 and 3.40. Despite the different time of year and different operating distilland volume, SS1-B continues to perform with the hourly distillate production more closely matching the hourly h_{ew}.

Despite Figures 3.39, 3.40, 3.42, and 3.43 being examples for short time periods, the same behavior was found to occur for SS1-A at different operating distilland volumes. The correlation between hourly production, h_{cw}, and h_{ew} for SS1-A can be seen in Figures 3.41a/b and 3.44 a/b.

3.5.3 Modeling Hourly Production

The hourly production was estimated using the calculated h_{cw} for SS1-A. Equation 3.5 details the simplified equation used to estimate the distillate production (m_{ew}), in kilograms, based on the convective heat transfer coefficient by simplifying equations 1.6 and 1.7 into one equation

$$m_{ew} = \frac{0.01623 \cdot h_{cw} \cdot (P_w - \varphi \cdot P_{ci}) \cdot A_w \cdot t}{\Delta h_v}$$

Equation 3.5 Estimated distillate production as described in Tiwari and Tiwari (2006)

P_{ci}	Partial saturated vapor pressure of water at condensing cover temperature ($N/m^{2)}$	P_w	Partial saturated vapor pressure of water at water temperature (N/m^2)
φ	Relative humidity	Δh_v	Enthalpy of evaporation of water (J/kg)
t	Time (seconds)	A_w	Evaporative surface area (m^2)

Figures 3.45 and 3.46 illustrate the actual and estimated hourly production for SS1-A (39 L) and SS1-B (13 L) between 4/5/11 and 4/9/11, respectively.

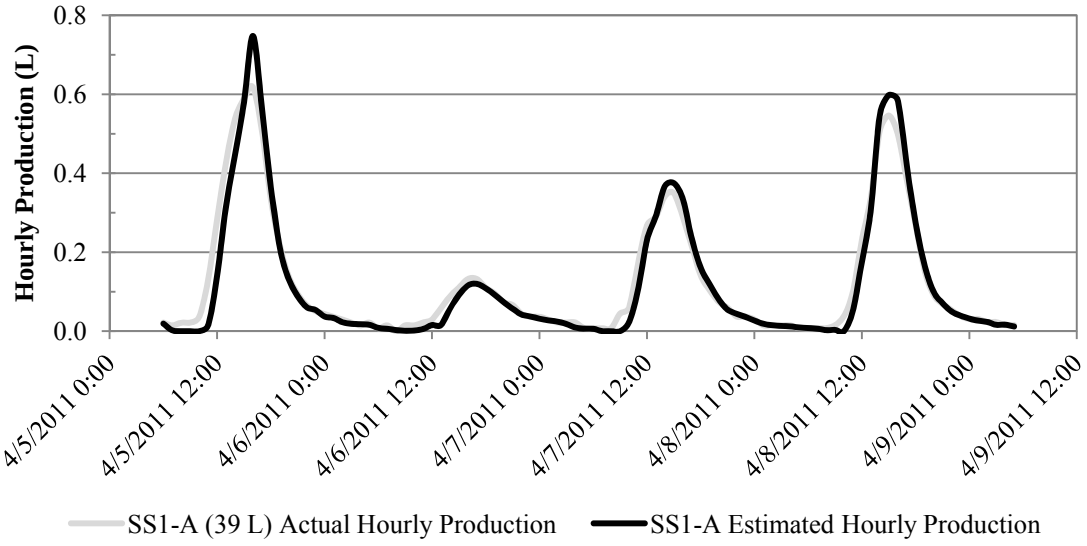


Figure 3.45 Actual and estimated hourly production for SS1-A (39 L) between 4/5/11 and 4/9/11

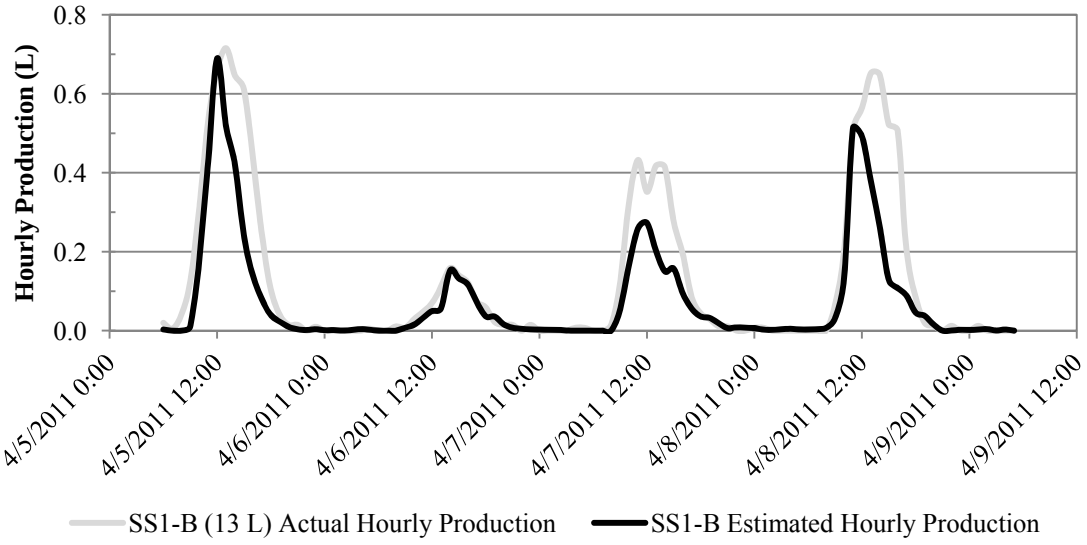


Figure 3.46 Actual and estimated hourly production for SS1-B (13 L) between 4/5/11 and 4/9/11

Figure 3.45 illustrates the estimated hourly production slightly over estimating for SS1-A (39 L) while Figure 3.46 illustrates the estimated hourly production under estimating the actual hourly production for SS1-B (13 L). Figures 3.45 and 3.46 both

show that estimated hourly production closely matched the actual measured production with little lag. The heat transfer method predicted hourly production with a mean absolute error of 26.9% and 49.7% for SS1-A (13 L) and SS1-B (39 L), respectively. The R^2 value was calculated to be 0.959 for SS1-A (13 L) and 0.813 for SS1-B (39 L).

Figures 3.47 and 3.48 illustrate the actual and estimated values graphed along the origin bisector for SS1-A (39 L) and SS1-B (13 L), respectively. Figures 3.47 and 3.48 illustrate the slight over estimation and under estimation in hourly production for SS1-A (39 L) and SS1-B (13 L), respectively.

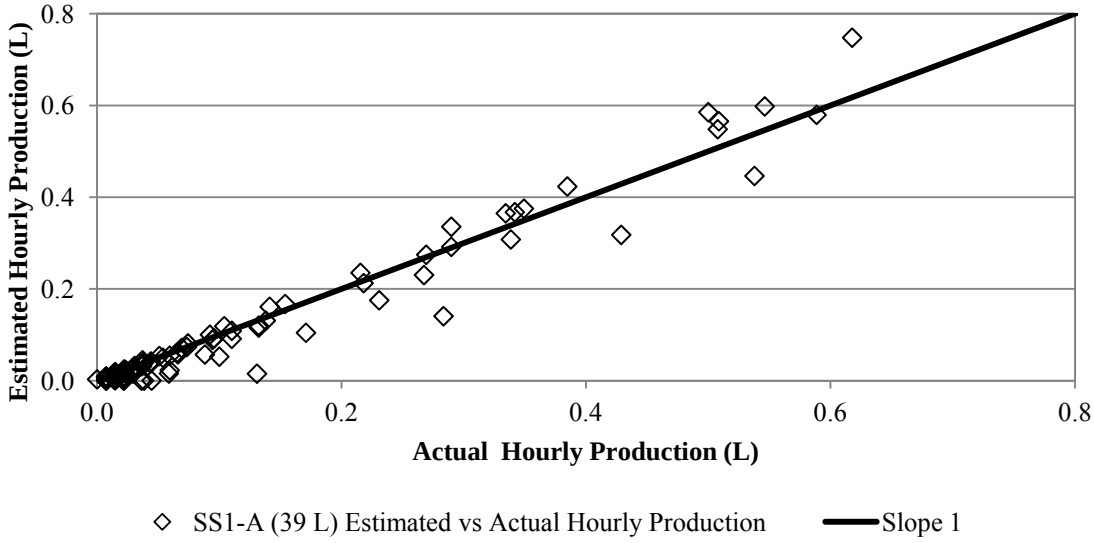


Figure 3.47 Estimated vs. actual hourly production for SS1-A (39 L)

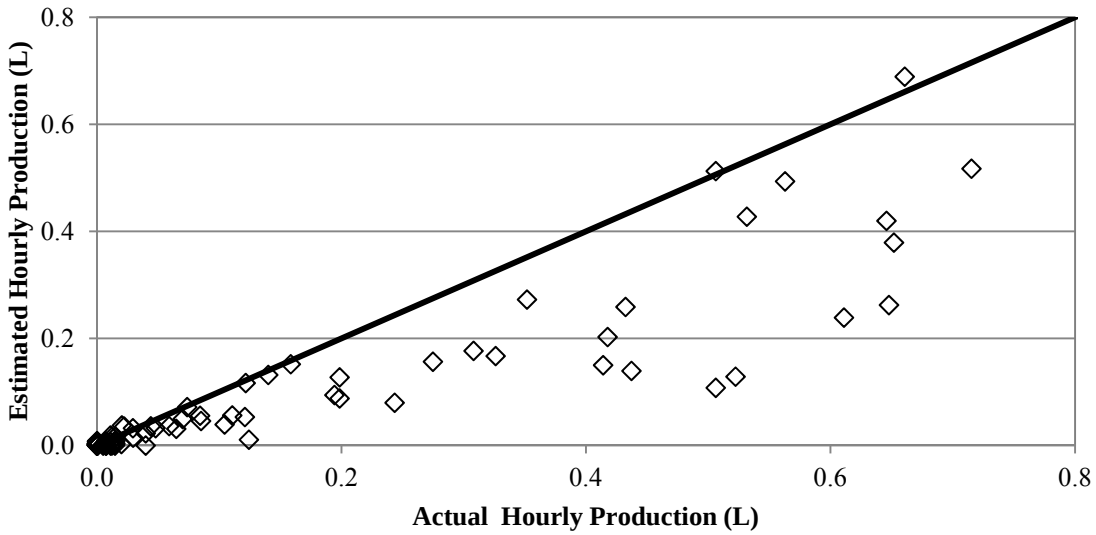


Figure 3.48 Estimated vs. actual hourly production for SS1-B (13 L)

Figures 3.49 and 3.50 illustrate the actual and estimated hourly production for SS1-A (13 L) and SS1-B (39 L) between 4/17/11 and 4/21/11, respectively.

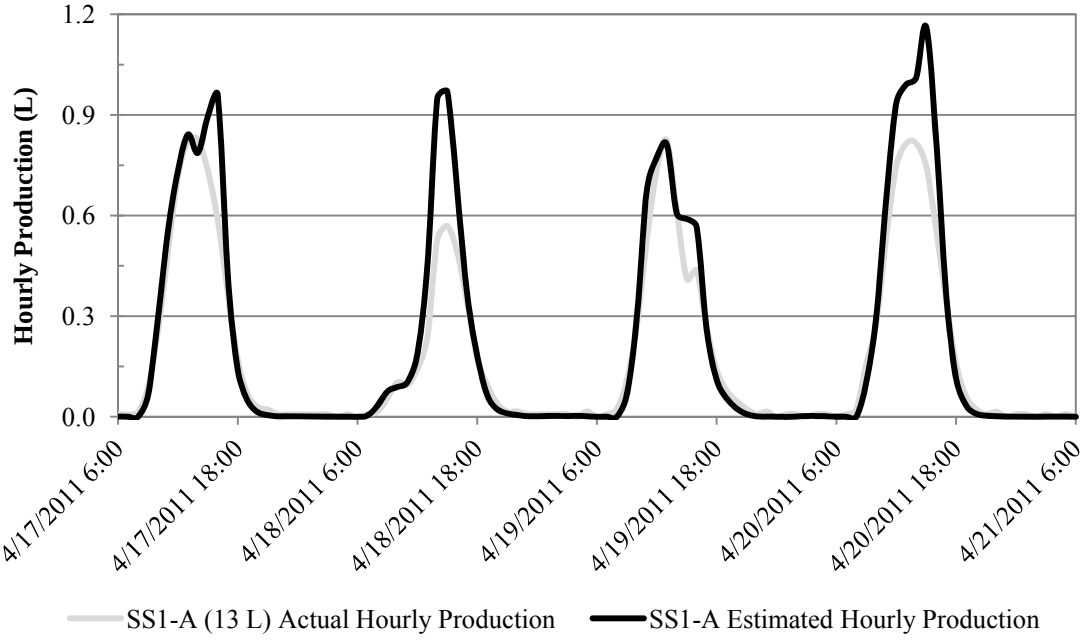


Figure 3.49 Actual and estimated hourly production for SS1-A (13 L) between 4/17/11 and 4/21/11

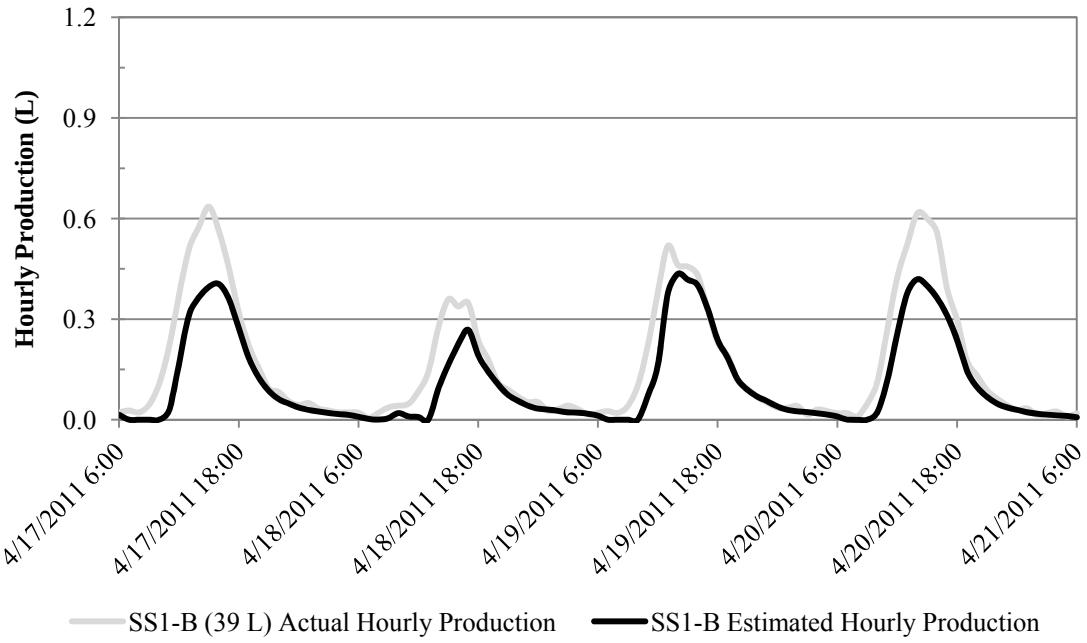


Figure 3.50 Actual and estimated hourly production for SS1-B (39 L) between 4/17/11 and 4/21/11

Figures 3.49 and 3.50 illustrate the same behavior shown in Figures 3.45 and 3.46 with regards to the over estimation of peak hourly production for SS1-A, under estimation of peak hourly production for SS1-B, and the little to no lag between the increase/decrease in estimated and actual hourly production. The over and under estimation for SS1-A and SS1-B occurs regardless of the distilland volume for each still. This particular behavior indicates that SS1-A and SS1-B are likely operating slightly different when examined at the hourly level. Figures 3.51 and 3.52 illustrate the actual and estimated values graphed along the origin bisector for SS1-A (13 L) and SS1-B (39 L), respectively.

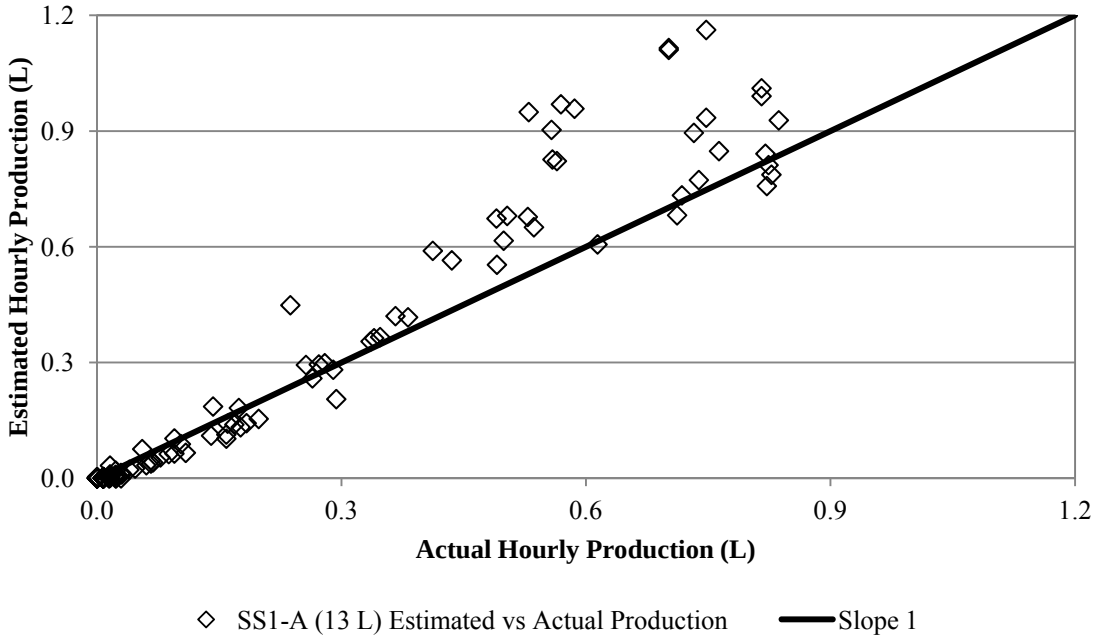


Figure 3.51 Estimated vs. actual hourly production for SS1-A (13 L)

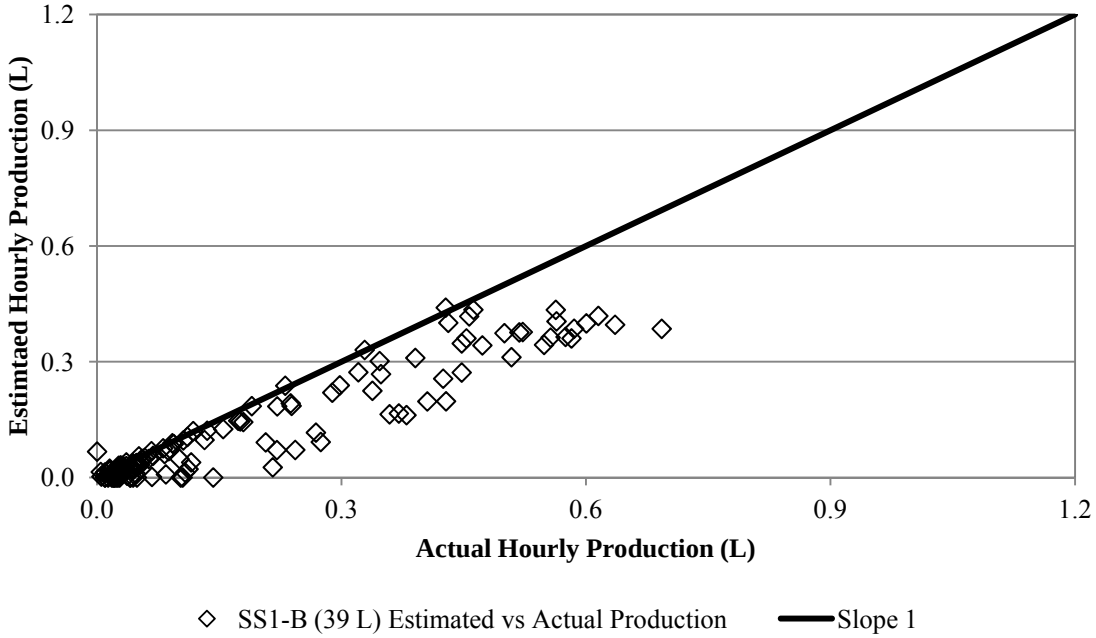


Figure 3.52 Estimated vs. actual hourly production for SS1-A (39 L)

Figures 3.51 and 3.52 illustrate the slight over estimation and under estimation in

hourly production for SS1-A (13 L) and SS1-B (39 L), respectively. The heat transfer method predicted hourly production with a mean absolute error of 53.2% and 40.8% for SS1-A (13 L) and SS1-B (39 L), respectively. The R^2 value was calculated to be 0.941 for SS1-A (39 L) and 0.904 for SS1-B (13 L).The R^2 values indicate the estimated hourly production from the Jakob (1949) correlation, as modified by Dunkle (1961), is able to account for a large portion of the variance for the actual hourly production. Furthermore, Figures 3.51 and 3.52 continue to illustrate the over estimation for SS1-A and under estimation for SS1-B for the peak hourly distillate production values.

3.6 Conclusions

Variability in daily production depends on the daily total insolation and other weather variables. Compared to spring and fall, fluctuations in production decrease as the weather conditions become less variable during the summer.

Recorded hourly data indicate a 1-2 hour lag between peak distilland/vapor/inner glass cover temperature and peak insolation. Operating solar stills with a larger distilland volume results in lower distillate production rates, cooler internal temperatures during the day, and warmer internal temperatures at night compared to operating solar stills with a smaller distilland volume.

Two approaches by Mathioulakis (1999) and Jakob (1949) as applied by Dunkle (1961) and translated by Tiwari and Tiwari (2006) that were presented in the literature review were tested using the data from the 2011 study. Mathioulakis et al. (1999) experimented with predicting daily production using daily temperature readings for the solar still and the ambient environment. Dunkle (1961) experimented with predicting hourly production by using heat transfer methods.

It is possible to use Mathioulakis et al.'s (1999) method to predict day/night distillate production provided daily total insolation and the temperature difference data between the distilland and the ambient temperature. Night time production is not as easy to calculate given the lack of an external energy source; however, nocturnal production prediction is improved when the distilland's stored energy is taken into account.

Hourly heat transfer coefficients also indicate a 1-2 hour lag between peak insolation and the maximum and minimum heat transfer coefficients. h_{cw} varied much less than h_{ew}. Heat and mass transfer modeling using correlations originally developed by McAdams (1954), Dunkle (1961), and A.K. Tiwari and G.N. Tiwari (2006) shows reasonable agreement with experimental measurements for hourly distillate production. The heat and mass transfer methods appeared to over and under predict hourly production depending on the type of solar still.

The validation of models to predict distillate production indicates the necessity of future studies to operate several solar stills in parallel at several different distilland depths. This would allow for a greater understanding of the magnitude of the effect of distilland volume on daily production and hourly performance characteristics.

CHAPTER 4

ARTIFICIAL NEURAL NETWORK MODELING

4.1 Overview

Solar still distillate yield varies greatly with the environmental conditions and the season in which the still is operated. In order to effectively implement passive solar distillation technology on a massive scale, it is necessary to develop accurate, worst case estimates for distillate production. As can be seen in Chapter 1, the current state of technology for solar still modeling generally lacks a simple and effective method to predict daily solar still production utilizing data that is commonly available worldwide. Modeling daily solar still production is limited within the existing field of solar distillation; Mathioulakis et al. (1999) performed one such study where daytime and nocturnal distillate production was predicted with low error by simply using insolation and distilland/ambient temperature data. Without an effective method to predict daily production, it will be difficult to determine the solar still basin area required to meet the purified water demand for either an individual, family, or large community in different parts of the world.

4.2 Artificial Neural Network Background

Multi-layer Perceptron (MLP) networks in artificial neural networks (ANNs) have been used in the past for engineering applications due to their ability to use non-linear transformations and to learn patterns of behavior between inputs and outputs (Haykin, 1994). Kalogiriou (2001) has reviewed multiple uses of ANNs for a wide range of fields for modeling and prediction in energy engineering systems. The architecture of a neural network helps to determine how a network transforms the inputs into an output

(Kalogirou, 2001). Furthermore, Kalogiriou (2001) states that it is essential to be able to identify the most important variables in a process. These networks are highly data driven and are capable of capturing complex behavior by learning from the user supplied input and target (output) data.

The MLP network can consist of input, output, and several hidden layers. Each layer can have many computational hidden nodes or neurons. The hidden layers' neurons connect the input and target layers by using a specified training function (Haykin, 1994). Each layer has units that are partially or fully connected to units in consecutive layers. Initially, the connections between consecutive units are assigned random weights to represent their strength or activity with regards to the noted patterns. The output from the first hidden layer is transferred to the second hidden layer whose outputs are then transferred to the subsequent hidden layers. This process is repeated for the rest of the network until reaching the final output layer which is the complete response of the ANN to the patterns and trends that were provided in the input layer. The training process ends once the validation data, supplied along with the training data, experiences a minimal change in error between the actual and predicted data.

ANNs are able to derive their predictive power through their parallel structure as well as their ability to learn and generalize. The generalization that occurs within a neural network allows for the prediction of reasonable outputs given inputs that were not originally included in the training dataset (Haykin, 1994). After an ANN architecture has been designed, the network must be trained in order to create the optimum set of weights for each connection until there is no more change in the synaptic weights. This results in a minimized difference between the actual and predicted target variable. ANNs have a

potential advantage over traditional empirical models and multi-variable regression analysis because they are able to account for the total interaction between input variables (Rumelhart, Hinton, and Williams, 1986). Figure 4.1 below displays an example ANN architecture as developed by Kalogirou (2001) (reprinted with permission).

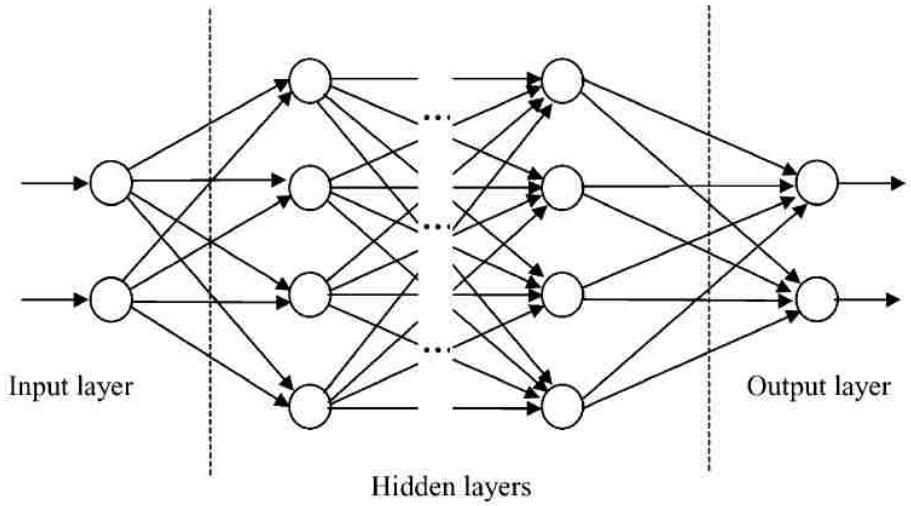


Figure 4.1 Example ANN architecture (Kalogirou, 2001)

4.3 Artificial Neural Network Modeling

The ANN for this study was first developed by using the data from the 2006 study using the SS1 and SS2 production data. The input layer consists of total global insolation, ambient temperature, wind speed, wind direction, cloud cover and distilland volume with the output layer consisting of the target (actual) distillate production. The hidden layer consisted of twenty processing units. The transfer function used for all processing units was the tangent sigmoid function due to its superior performance compared to alternative transfer functions (Haykin, 1994).

The ANN was created with a set of weather and distilland volume data as inputs and a set of daily distillate production as the target (output) variable. Prior to the training

process, the input and target variables were normalized between zero and 1 through dividing by the maximum value in each variable's range. This normalization accelerates the training process and enhances the network's generalization capabilities (Haykin, 1994). Eighty percent of the entire solar still performance dataset was used for training purposes and the remaining 20% was used to test/validate the network's predictive capabilities.

The training dataset for solar still SS1 included data from 2/1/2006 to 7/31/2007 and the validation and testing dataset included data from 1/3/2007 to 4/5/2007. The validation data are used to measure network generalization and to stop training when generalization stops improving. The testing dataset consists of data not previously included in training or validation and are used to provide an independent measure of network performance. SS1's dataset consisted of 453 data points with 360 points used for training, 26 points for validation, and 67 for testing the trained network. The training dataset for SS2 included data from 2/1/2006 to 1/17/2007 and the separate testing dataset included data from 1/18/2007 to 3/20/2007. SS2's dataset consisted of 312 data points with 250 points used for training, 19 points for validation, and 43 for testing. The ANN training/testing process was repeated with different combinations of weather variables until the best performing combination of input weather variables was found.

Combinations of selected input variables for the trial ANNs were varied in an attempt to find the optimum combination of data inputs that would produce the most accurate predicted results. From the literature, it is known that insolation and ambient temperature play an important role in the radiant and convective transfers of energy into the solar still. Because of their function as sources of energy, it was vital to incorporate the effects of

insolation and temperature in the ANN model.

The rate of heat removal from the glass cover is one of the limiting factors that influences the distillate condensation rate. Because of their ability to remove energy from a solar still, wind speed and wind direction were included as ANN input variables. Due to the presence of a 2.5 meter tall utility structure located 3 to 8 meters directly north of the solar still test site, the recorded data used in this modeling project was obtained from stills that were sheltered from northerly winds that prevail at the site in winter and spring. Because of the proximity of this structure, the solar stills likely experienced a reduced wind cooling effect if the wind direction was within 90º of geographic North. Wind direction was therefore included to study this variable's effect on the performance of the developed ANN models.

Distilland volume was included as an input variable because larger volumes (masses) of water would heat up more slowly than smaller volume at similar irradiances, as observed in Chapter 3 and as cited in the literature, and would therefore have lower vapor pressures and slower evaporation rates. The amount of cloud cover and its effect on the balance of diffuse and direct irradiance was also a focus of the study to determine cloud cover's effect on the ANN model's performance.

4.3.1 Artificial Neural Network Modeling Results

The results for the top ten ANN models for each solar still are presented in Tables 4.1 and 4.2; sorted according to performance by different combinations of input variables. Other input combinations were also tested with the ANN model, but did not perform as well. The primary criteria for evaluating the performance of different combinations of input data architectures were the percentages of model predictions that were within 10%,

20%, and 30% of actual daily distillate production. Tables 4.1 and 4.2 show that the top three performing architectures, for SS1 and SS2, used daily total insolation (I), average daily temperature (T), operating distilland volume (V), average daily wind velocity (W), and the corrected wind direction (D). Testing/validation data consisted of input data that had not been previously included in the trained network to demonstrate the ability of each trained ANN to predict results. The training scenarios are shown as non-highlighted while the testing scenarios are shown as highlighted in gray.

Table 4.1 SS1's ANN modeling results for 2006-2007 data

INPUTS	R^2	MEAN ABSOLUTE ERROR	% PREDICTIONS WITHIN 30% ERROR	% PREDICTIONS WITHIN 20% ERROR	% PREDICTIONS WITHIN 10% ERROR
ITVWD	0.935	19.6%	90.8%	87.8%	75.0%
	0.966	9.7%	93.5%	89.2%	76.3%
ITV	0.882	16.4%	91.7%	86.7%	73.3%
	0.945	10.0%	94.6%	87.1%	78.5%
ITVW	0.933	18.4%	91.9%	85.6%	70.0%
	0.966	8.4%	96.8%	89.2%	72.0%
ITVWC	0.925	21.2%	89.4%	85.6%	68.3%
	0.961	9.8%	93.5%	89.2%	75.3%
IVW	0.922	20.6%	90.0%	84.7%	63.9%
	0.916	15.4%	88.2%	78.5%	58.1%
IT	0.916	20.5%	90.6%	84.2%	67.5%
	0.930	13.2%	89.2%	83.9%	64.5%
ITVWDC	0.934	19.0%	91.1%	83.6%	68.9%
	0.961	9.8%	96.8%	89.2%	72.0%
I	0.920	21.9%	90.8%	82.2%	61.4%
	0.909	15.4%	86.0%	74.2%	60.2%
TV	0.813	39.2%	74.2%	63.3%	40.0%
	0.649	28.5%	72.0%	55.9%	38.7%
TVDC	0.784	44.3%	60.6%	46.9%	32.8%
	0.607	35.3%	58.1%	50.5%	31.2%

Table 4.2 SS2's ANN modeling results for 2006-2007 data

INPUTS	R^2	MEAN ABSOLUTE ERROR	% PREDICTIONS WITHIN 30% ERROR	% PREDICTIONS WITHIN 20% ERROR	% PREDICTIONS WITHIN 10% ERROR
ITV	0.969	9.4%	94.4%	92.4%	80.8%
	0.909	9.4%	95.2%	88.7%	72.6%
ITVW	0.975	8.9%	95.6%	92.4%	84.0%
	0.921	9.4%	93.5%	82.3%	69.4%
ITVWD	0.971	9.4%	95.2%	91.2%	83.2%
	0.908	10.7%	91.9%	82.3%	69.4%
ITVWDC	0.975	9.5%	95.2%	90.8%	80.0%
	0.921	10.7%	90.3%	83.9%	66.1%
IT	0.966	10.5%	94.0%	87.2%	73.2%
	0.768	23.6%	66.1%	54.8%	37.1%
I	0.954	13.0%	91.6%	86.0%	69.6%
	0.742	28.5%	58.1%	40.3%	24.2%
IVW	0.957	13.3%	90.8%	85.6%	68.8%
	0.833	14.9%	88.7%	69.4%	48.4%
ITVWC	0.955	13.9%	90.0%	85.2%	71.6%
	0.904	12.0%	90.3%	83.9%	66.1%
TVDC	0.916	18.6%	88.4%	80.0%	64.8%
	0.805	16.2%	88.7%	69.4%	58.1%
TV	0.803	35.1%	66.8%	53.6%	34.0%
	0.640	27.2%	75.8%	61.3%	37.1%

Table 4.1 shows the ANN modeling results for SS1's training and testing scenarios with regard to the R^2 value, mean absolute error, and the number of predicted results within 0-30% error for SS1. The four best performing models for SS1-A were ITV, ITVW, ITVWD, and ITVWDC in terms of mean absolute error. Six out of the ten presented scenarios in Table 4.1 show a higher percentage of model predictions falling within the 0-10% and 0-20% error categories for the testing scenario compared to the training scenario. This is possibly due to the smaller dataset size and lower variability of the testing dataset. The 0-10% and 0-20% error category shows the testing scenario having 0.4% to 3.6% more results within the 0-20% error category than the training

scenario for the top four performing input architectures. The 0-10% error category shows the testing scenario having 1.3% to 7% more results within the 0-10% error limit than the training scenarios for the top four performing input architectures.

The four best performing models for SS2 were ITV, ITVW, ITVWD, and ITVWDC. The results for SS2 in Table 4.2 show the testing scenarios were close to but did not exceed the training scenarios. The 0-20% error category shows the testing scenario having 3.7% to 10.1% less results within the 0-20% error category than the training category for the top four performing input architectures. The 0-10% error category shows the testing scenario having 8.2% to 13.9% less results within the 0-10% error category than the training scenarios for the top four performing input architectures. The lower performance for the SS2 testing categories could possibly be attributed to the smaller data set size which covered only one calendar year for SS2 compared to one year and a half of data for SS1.

Table 4.1 shows an increase in the number of results falling within the 0-10% and 0-20% model error categories for SS1 as the adjusted wind direction (D) is added to the ITVW (insolation, temperature, distilland volume, and wind speed) input architecture. This result is not shared with the solar still SS2's results in Table 4.2 where the results for both ITVW and ITVWD scenarios are extremely close to one another. The difference in reaction to the addition of the wind direction variable indicates that SS1 was more sensitive to wind direction than SS2. The addition of the cloud cover variable (C) to the ITVW input architecture scenario in Table 4.1 results in very little change in model performance. In contrast, there is a noticeable decrease in model performance for SS2 when the cloud cover variable is added to the ITVW input architecture in Table 4.2.

Since the addition of the cloud cover data results in a negligible effect on SS1's model performance, the variable can be seen as non-essential to its ANN model. Also, when using daily weather data, average cloud cover for conditions less than completely overcast may not accurately indicate the total amount of insolation actually received by the still. Additional research, at higher weather data collection frequencies may be needed in this area.

The ANN modeling results indicate eight out of ten and five out of ten input architectures having testing and training results with R^2 values greater than 0.9 for SS1 and SS2, respectively. The ITVWD input architecture performed the best for SS1 in terms of having low error, high R^2 value, and a high percentage of results in low total error categories for both training and testing scenario. The ITV input architecture performed the best for SS2 with performance characteristics similar to those described for SS1's best ANN architecture.

Overall, the results from the study show that ANN modeling can produce test results with up to 89% of the predictions being within 20% of the actual value. The ITVWD model for SS1 had a mean absolute error of 17.5% with 68.7% of predictions having less than 10% error. The ITV model for SS2 had a mean absolute error of 9.4% with 71.2% of predictions having less than 10% error. All residual plots typically had a slight right skew; indicating a slightly higher frequency of over predictions. An example testing/training residual histogram for SS2's ITV ANN model is shown in Figure 4.2.

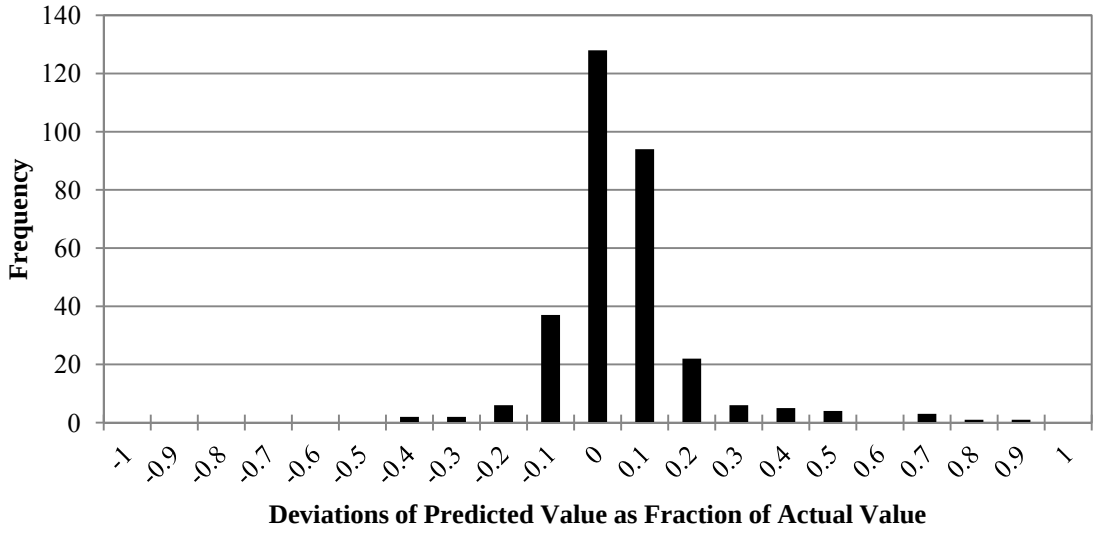


Figure 4.2 SS2's training/testing residuals histogram for ANN ITV model showing slight right skew

4.4 Correlation Coefficients for ANN modeling

Plots of the relationships between ANN predicted and actual daily distillate production, for each solar still under the best performing architecture, are shown in Figures 4.3 and 4.4. Although there are several outlying points due to over and under prediction, Figures 4.3 and 4.4 show tight distributions of both training and testing output data plotted against measured actual production, centered around the line of slope one through the origin.

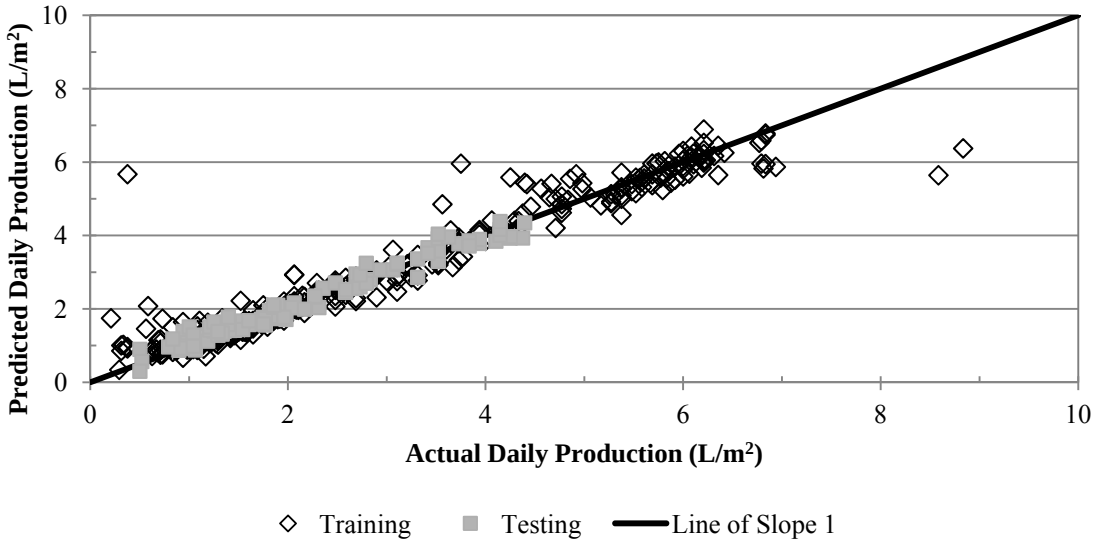


Figure 4.3 SS1's predicted vs. actual distillate production for ITVWD model

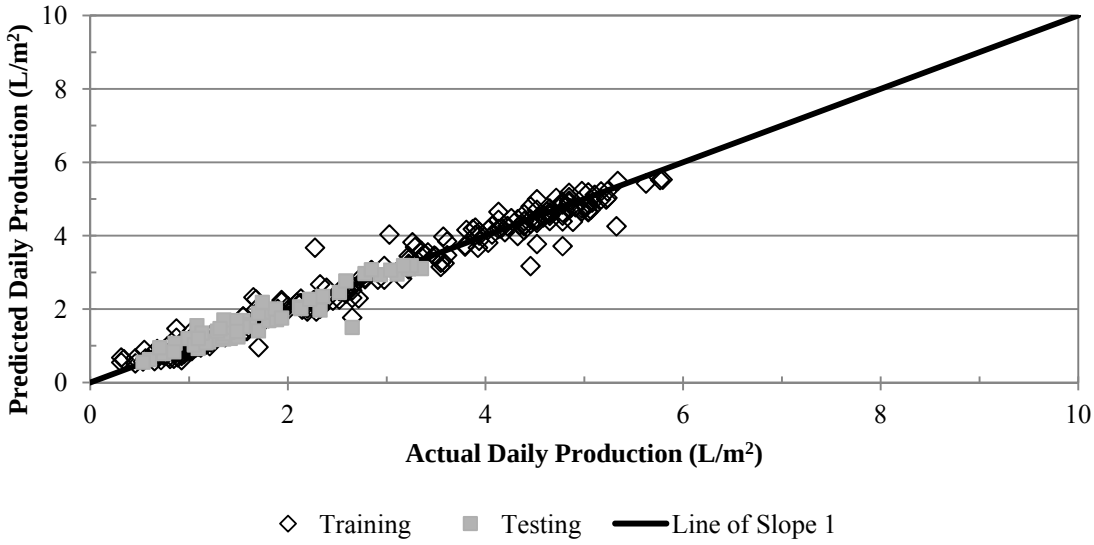


Figure 4.4 SS2's predicted vs. actual distillate production for ITV model

Coefficients of determination (R^2 values) were computed to determine the proportion of variance that is accounted for by the ANN model. The Pearson correlation coefficient (R value) was also calculated to measure the correlation between predicted and actual production. The R value provides a measure of how well future outcomes can be

predicted using the ANN model. Separate coefficients for the ANN models' training and testing data sets were calculated with respect to each set of predictions. The coefficients were also calculated for the entire data set combining training and testing values. Table 4.3 compares the different R and R^2 values for the training, testing, and the combined training/testing data sets for the best scenario for each still.

Table 4.3 Determination (R^2) and Pearson correlation (R) coefficients for the best scenarios for SS1 and SS2

	SS1 (ITVWD)		**SS2 (ITV)**	
	R^2	**R**	**R^2**	**R**
Training	0.935	0.967	0.969	0.984
Testing	0.966	0.983	0.909	0.953
Combined	0.939	0.969	0.974	0.987

The ANN predicted outputs for solar stills SS1 and SS2 exhibit high R^2 and R values for the training and testing data sets that all exceed 0.90. Similar R^2 and R values for both the training and testing data sets show that the ANNs were able to fairly well simulate results given data that were not originally used for the training process.

4.5 ANN's Data Requirements

Besides the number of inputs that are needed to create a reliable neural network, the amount of data needed to create reliable results also plays an important role. ANN model runs were evaluated by varying training input data set sizes to determine the effect of data set size on model accuracy for training and testing predictions. Figure 4.5 shows that the model for SS2 required approximately 75 data points for the ANN to attain a threshold of 60% of predicted results being within 0-10% of actual daily distillate production. Unlike the model for SS2, the model for SS1 required approximately 200 data points to achieve

60% of predicted results being within 0-10% of the actual daily distillate production. Additionally, SS1's ANN model does not quite reach the same performance level as SS2. These results show that to support the strong generalization capabilities found in neural networks, there is a need for large training data sets that incorporate weather conditions found throughout the year. Furthermore, the data requirement varies depending on solar still type.

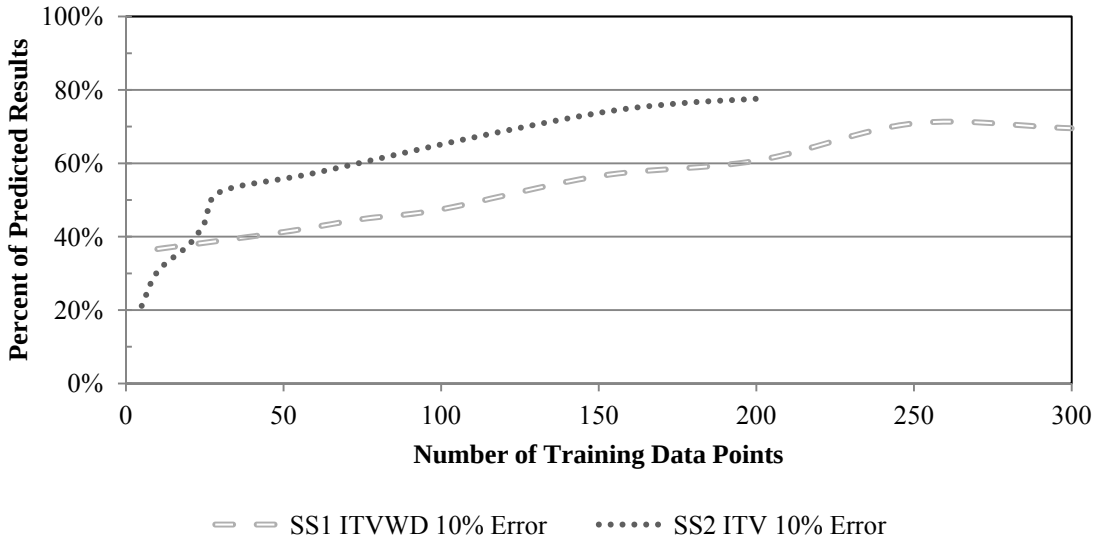


Figure 4.5 Percentage of predictions within 10% error as a function of data set size

4.6 Daily Production Reliability

As seen in Chapter 2, the July 2006-October 2006 production data had to be substituted from SS1-C to complete Figure 2.9. In order to complete the reliability analysis during the July 2006-October 2006 period, a separate ANN had to be created for SS1-C. The ANN for SS1-C performed similarly when compared to the ANN for SS1 and was built using insolation, temperature, wind speed, and distilland volume (ITVWD) as the input variables. SS1-C's ANN performed with 90.0% of training and 86.1% of

testing predictions being within 0-20% of actual daily production, respectively. The ANN also performed with 76.7% of training and 72.2% of testing predictions being within 0-10% of actual daily production, respectively. The ANN model for SS1-C performed with a mean absolute error of 10% for the training and 11% for the testing scenario.

As seen in Table 3.1, there is a substantial variation in daily experimental water production within each month due to varying weather conditions. By constructing still collector area to accommodate the lower 5^{th} percentile production value, a user could generally expect 95% of actual daily production quantities in each month to exceed the minimum water yield requirement.

Figures 4.6 and 4.7 show that the lower 5^{th} percentile predicted and actual monthly average daily distillate production for the winter (December-February) varied between 1.1 to 2.45 L/m^2 for SS1 and 0.8 to 1.9 L/m^2 for SS2. The lower 5^{th} percentile predicted and actual daily average distillate production for the summer (June-August) varied between 2.5 to 5.9 L/m^2 for SS1 and SS1-C and 3.4 to 4.2 L/m^2 for SS2. The July 2006-October 2006 data for SS1-C are indicated in Figure 4.6 with asterisk data markers instead of the squares and diamonds that were used for SS1.

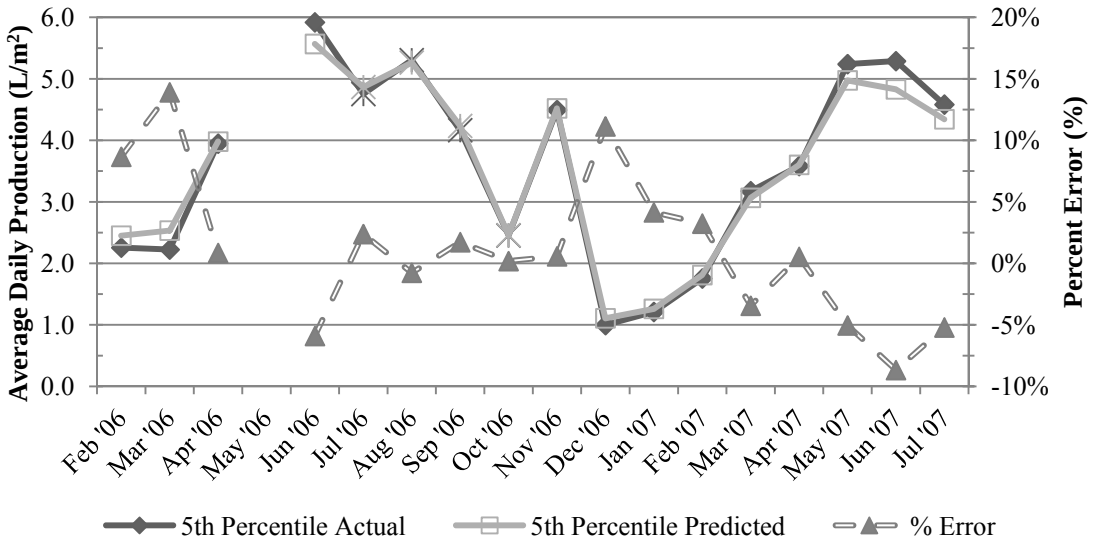


Figure 4.6 SS1's 5th percentile actual and predicted average daily production for ITVWD ANN model

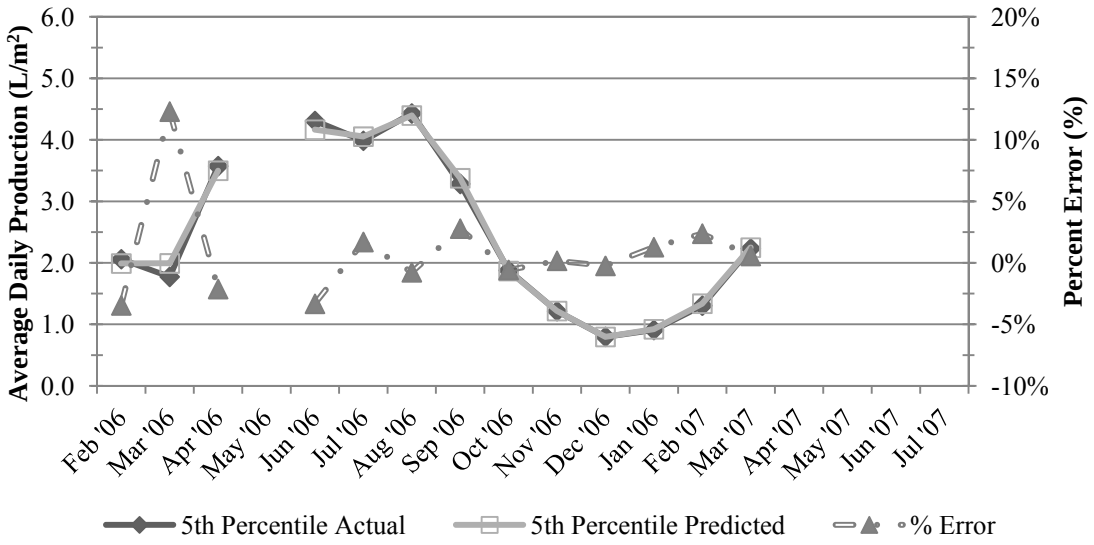


Figure 4.7 SS2's 5th percentile actual and predicted average daily production for ITV ANN model

Predicted solar still performance at lower 5th percentile yields were generally within 0-10% of the actual 5th percentile distillate production. As seen in Figure 4.6, the ANN 5th percentile predicted values for SS1 have a peak positive 13.9% difference for the

month of March 2006 and a peak negative 8.7% difference for the month of June 2007. The ANN model for solar still SS1 tended to over predict distillate production during the winter and under predict during the summer.

The ANN modeled results for SS2 were less erratic than SS1. Figure 4.7 illustrates that the ANN 5th percentile predicted values for SS2 have a peak positive 12.3% difference for the month of March 2006 and a peak negative 3.3% difference for the month of June 2006.

Tables 4.4 and 4.5 show the summary for the number of months that were over and under predicted with the neural network model. Table 4.4 shows that over a 17 month period the ANN ITVWD model for SS1 under predicted by an average 4.8%, 35% of the time and over predicted by 4.3%, 65% of the time. Table 4.5 shows that the ANN model for SS2 under predicted by 1.8%, 46% of the time and over predicted by 3.0%, 54% of the time.

Table 4.4 SS1 summary of lower 5th percentile average daily actual and predicted distillate production with ANN ITVWD model

Number of Months	17	**Percent of Time**	**Average Error**	**Standard Deviation of Error**
ANN Monthly Predictions > Actual	6	35%	-4.8%	2.6%
ANN Monthly Predictions < Actual	11	65%	4.3%	4.8%

Table 4.5 SS2 summary of lower 5th percentile average daily actual and predicted distillate production with ANN ITV model

Number of Months	13	**Percent of Time**	**Average Error**	**Standard Deviation of Error**
ANN Monthly Predictions > Actual	6	46%	-1.8%	1.4%
ANN Monthly Predictions < Actual	7	54%	3.0%	4.2%

Figures 4.8 and 4.9 graphically summarize the actual and predicted values for SS1 and SS2, respectively.

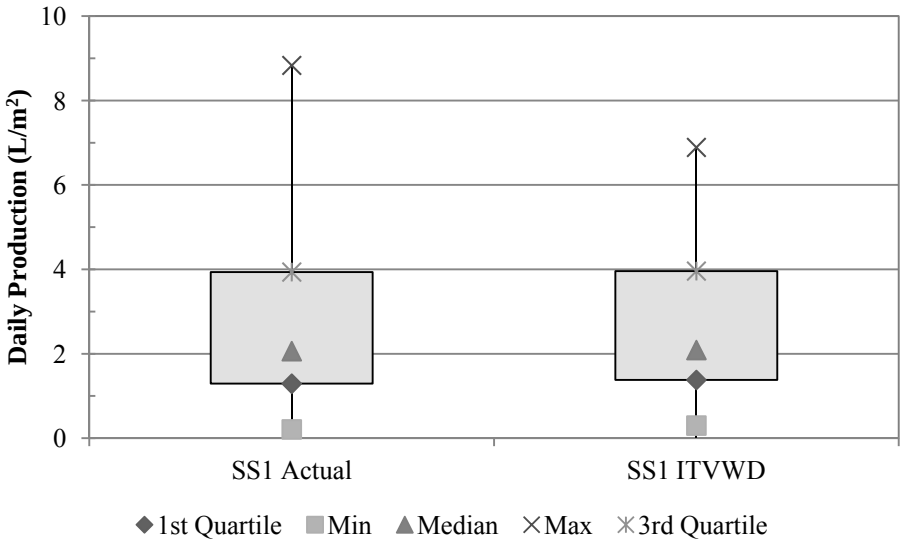


Figure 4.8 Descriptive statistics box plot for SS1's actual and ITVWD predicted production

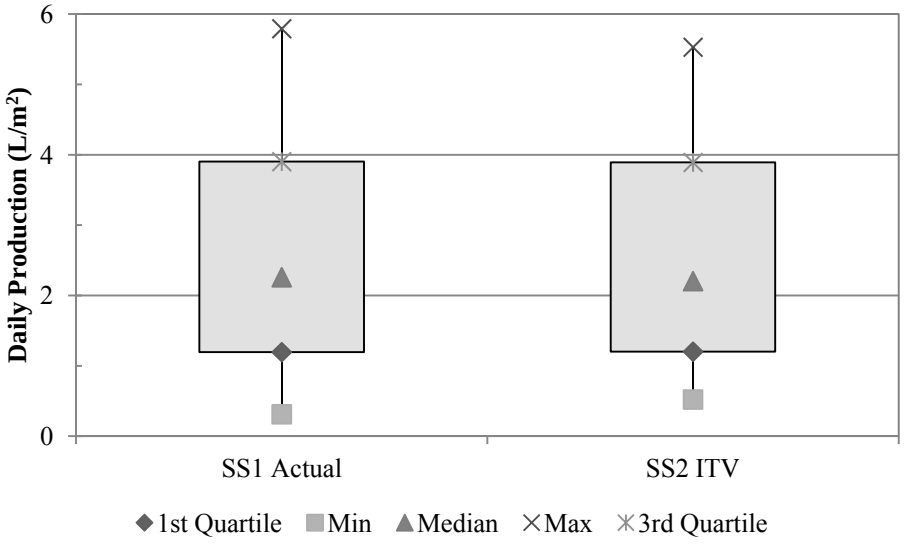


Figure 4.9 Descriptive statistics box plot for SS2's actual and ITV predicted production

Figure 4.8 illustrates how the ITVWD ANN model predictions for SS1 are close to the actual daily production values; however, the ANN model is not able to predict the more extreme maximum values that occurred. The maximum value for SS1 might be attributed to an original data entry error by Venkatesh (2007) since the maximum value was higher than any other recorded yield in the entire data set. Figure 4.9 illustrates similar results to Figure 4.8; however, the ANN ITV model for SS2 is able to predict the more extreme minimum and maximum values more closely than the ANN model for SS1.

The percent difference for the 1st quartile, minimum, median, maximum, and 3rd quartile between the actual daily production and the ANN ITVWD model for SS1 was 6.6%, 36.1%, 1.2%, 24.7%, and 0.6%, respectively. The highest percent difference between the predicted and actual values for SS1 occurred for the minimum and maximum production values.

The percent difference for the 1st quartile, minimum, median, maximum, and 3rd quartile between the actual daily production and the ANN ITV model for SS2 was 0.6%, 50.3%, 2.5%, 4.6%, and 0.3%, respectively. The highest percent difference between the predicted and actual values for SS2 occurred for the minimum production value.

4.7 Parametric Study

Systematic evaluations of the predicted yields, depending on the combined effects of several weather variables and solar still operating depths, were developed by conducting runs of the developed ITVWD ANN model for SS1 with 30 simulated input data points. The domain for each simulated input combination of variables was limited to the actual range of values that were observed during the original testing conditions. The predicted contributions of simulated interactions between several weather variables were evaluated

by plotting the results of parametric studies calculated by the developed ANN.

4.7.1 Effects of Insolation and Wind Speed

The parametric study was performed using SS1's ITWVD ANN model by systematically increasing daily total insolation at different levels of average wind speed while maintaining the distilland depth constant at 1 cm and a low average daily temperature of 9ºC. This study was performed because the estimated total amount of produced distillate greatly depends on both the total amount of energy received by the solar still and how heat is convectively removed from all faces of the solar still. Figure 4.10 illustrates how the insolation and wind speed variables are predicted to affect estimated daily distillate production.

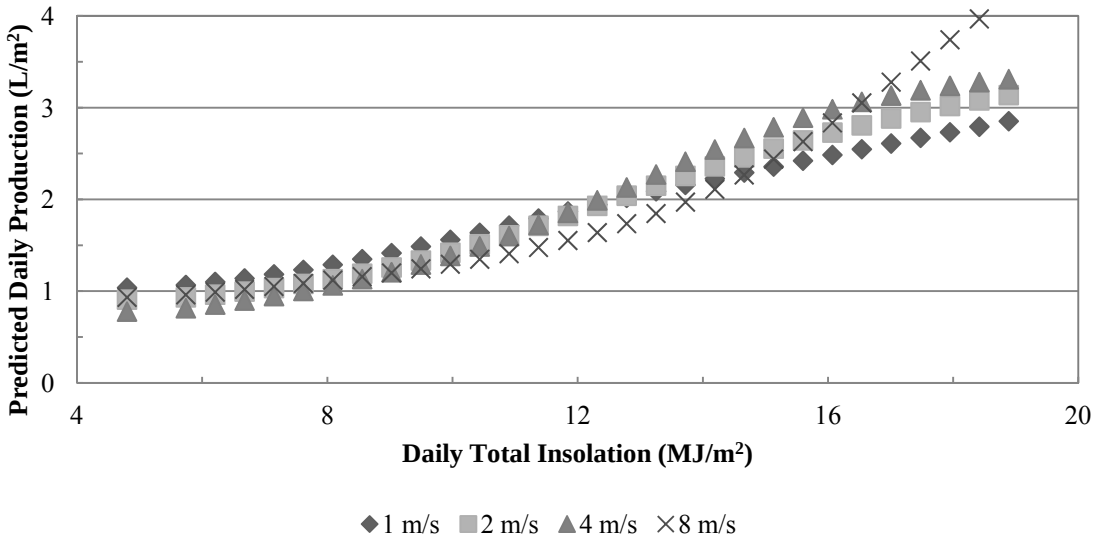


Figure 4.10 SS1 parametric study showing predicted production vs. total daily insolation at different levels of daily average wind speed

The parametric study predicts that maximum production occurs for combinations of high insolation and high wind speed. The 35% increase in predicted total daily production

at maximum insolation and wind speed at 8 m/s compared to the 1 m/s scenario could have been due to increased heat removal over the glass cover due to higher wind speed. Furthermore, a slight decline in yield, with increasing wind speed at low insolation, suggests more rapid heat removal and a need to keep solar stills covered or shielded from high winds during cool, windy weather.

4.7.2 Effect of Insolation and Distilland Depth

A second parametric study was performed using SS1's ITVWD ANN model by systematically increasing daily total insolation at different levels of distilland depth while maintaining constant the daily average temperature at 9ºC and wind velocity at 3.2 m/s. Figure 4.11 shows the results for this scenario.

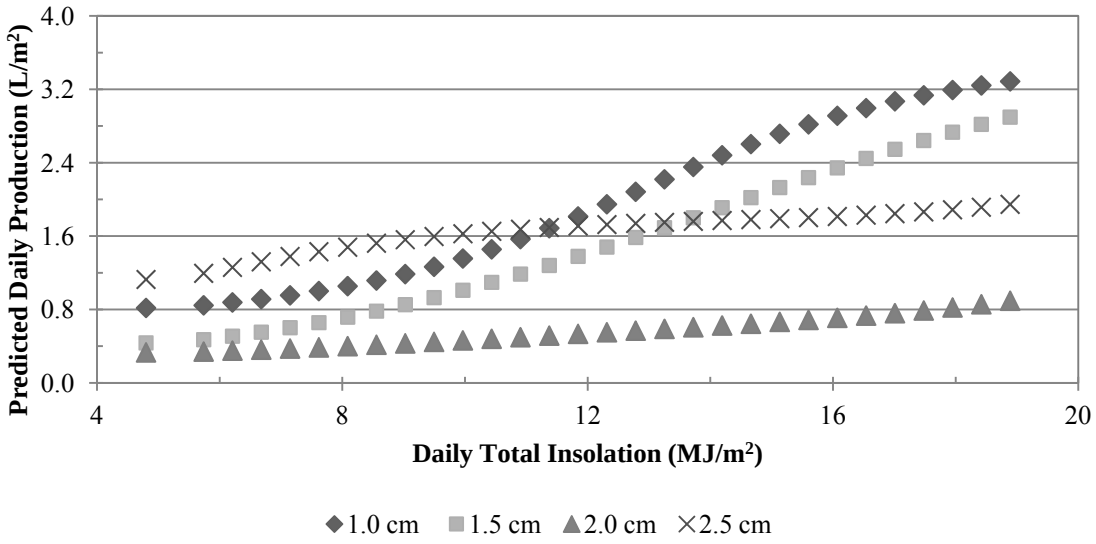


Figure 4.11 SS1 parametric study showing predicted production vs. daily total insolation at different levels of distilland depth

Higher total daily distillate production is predicted for high depths and low insolation when compared to low depths and low insolation. In contrast, Figure 4.11 also shows

higher total daily distillate production for low depth and high insolation compared to high depth and high insolation. Predicted distillate production for maximum insolation at 1.0 cm depth is nearly 60% higher than the predicted production at a depth of 2.5 cm. However, at the lowest insolation, the 2.5 cm depth is predicted to produce more distillate than at other depths. It is hypothesized that this occurs because the distilland stores more thermal energy at a depth of 2.5 cm than at shallower depths. The monitoring data showed that the deeper distilland also stayed warmer for a longer period of time into the evening hours producing more distillate even when daytime insolation was low.

4.8 ANN Validation

In order to validate the developed ANN models, the new data from the 2011 study was used as additional test data for the neural networks that were developed for SS1 using the 2006-2007 data. Table 4.6 summarizes the domain of the training and testing data from the 2006 and 2011 study.

Table 4.6 shows how the input variables for the ANN for the 2011 study were all within the minimum and maximum values for the 2006-2007 study with the exception of distilland volume. The stills were operated at a higher distilland volume during the summer phase of the 2011 study compared to the 2006-2007 study.

Once the 2011 daily production data, for SS1-A and SS1-B, were paired with the respective daily weather data and normalized, the 2011 data set was run through various ANN models that were built using the 2006-2007 data for SS1. Figure 4.12 shows the results of the simulated predictions for the 2011 data with respect to the coefficient of determination (R^2) and the mean absolute error.

Table 4.6 Daily weather data domain for 2006-2007 and 2011 experiments

Variable	Study Year	Maximum	Minimum
Insolation (MJ/m^2)	06-07	33.3	3.4
	11	33.3	4.8
Temperature (°C)	06-07	38.3	-0.6
	11	37.2	1.1
Wind Speed (m/s)	06-07	9.83	0.89
	11	9.83	1.79
Wind Direction (From North)	06-07	180	24
	11	169	25
Cloud Cover (Fraction)	06-07	0.83	0.00
	11	0.74	0.00
Distilland Volume (L)	06-07	26.7	10.0
	11	40.0	13.0
Daily Production (L/m^2)	06-07	8.83	0.21
	11 (SS1-A)	6.81	0.47
	11 (SS1-B)	6.83	0.57

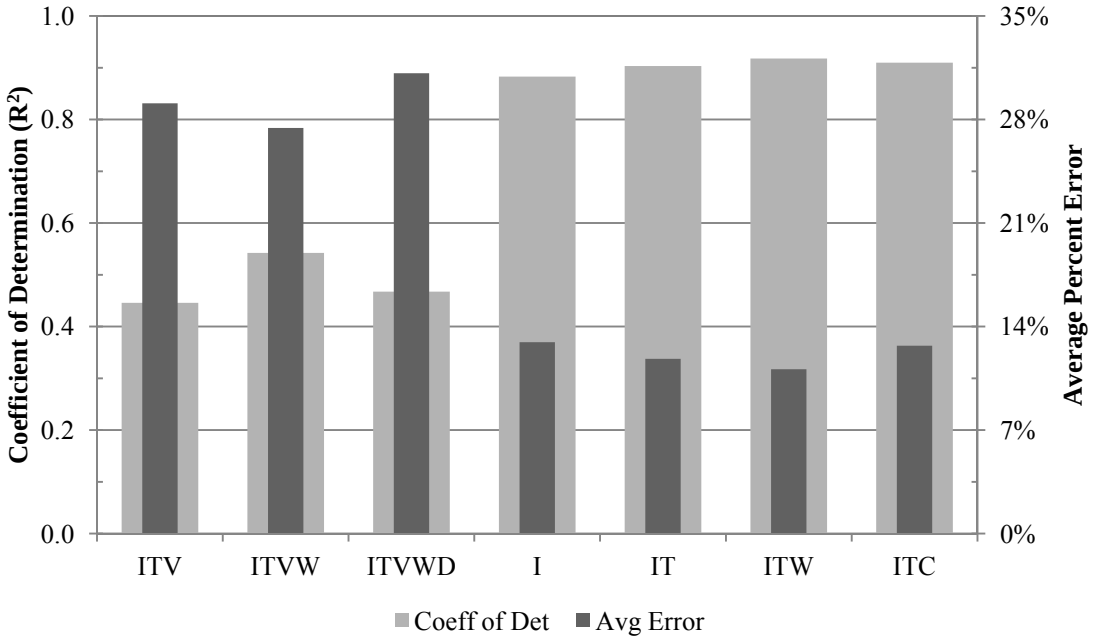


Figure 4.12 R^2 and mean absolute error for 2011 SS1-A/B data using models developed with data from 2006-2007, without recalibration

The results presented in Figure 4.12 indicate the 2006-2007 ANN models' ability to predict daily production with less than 15% mean absolute error for four out of seven models by using data collected during 2011. The three models performing with error greater than 15% all included the distilland volume variable.

The ANN architecture featuring insolation, temperature, and wind speed (ITW) performed the best out of all the scenarios with a mean absolute error of 11.1%, an R^2 value of 0.918, and with 88.9% of predictions within 0-20% error. The lower performance for the scenarios involving distilland volume is hypothesized to occur due to scenarios where the operating distilland volume was higher for the 2011 study compared to the 2006-2007 study; and therefore, out of range for the original ANN model.

Due to the high calculated error in scenarios involving distilland volume, the ANN model was re-calibrated by combining the 2006-2007 dataset with 50% of the 2011 dataset. The data from the 2011 study chosen for the recalibration consisted of distilland volume scenarios involving 40 L, 39 L, 30 L, and 26 L. Figures 4.13 and 4.14 show the results of the new training and testing scenarios for the recalibrated ANN model.

The recalibration for the SS1 ANN models results in the IT architecture performing the best with regards to the mean absolute error and the coefficient of determination. While the ITV architecture performs slightly better than the IT architecture for the training scenario, the IT architecture outperforms the ITV architecture for the testing scenario. Tables 4.7 and 4.8 detail the performance of each trained SS1 ANN model, following the recalibration, for the testing and training scenarios. The ITV and IT architectures are highlighted.

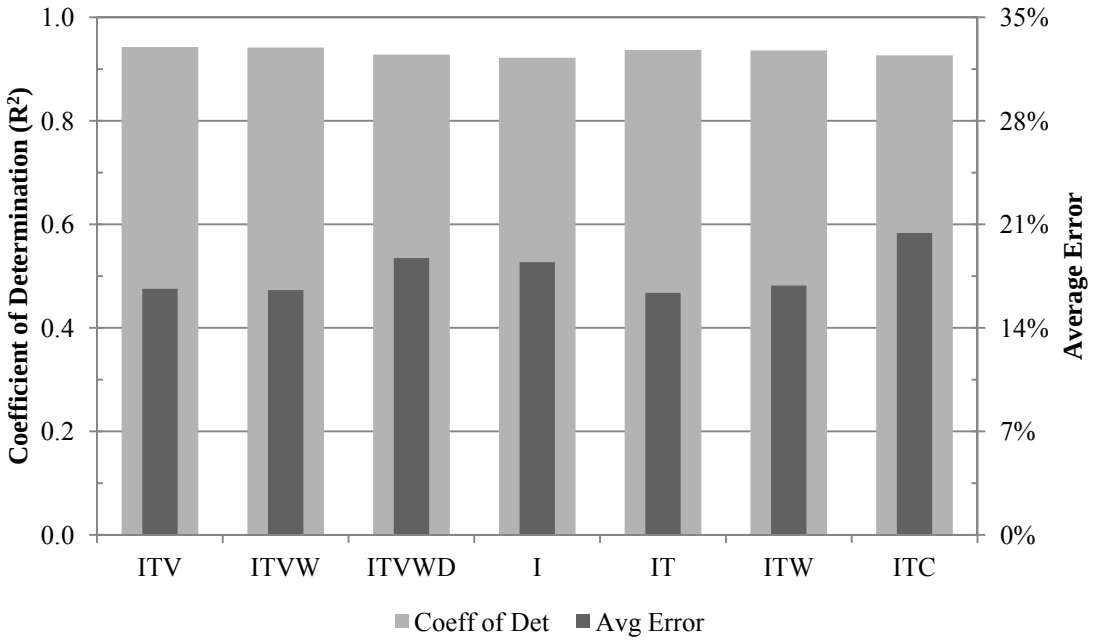


Figure 4.13 SS1 ANN training results for the 2006-2007 models recalibrated with 50% of SS1-A/B data from 2011

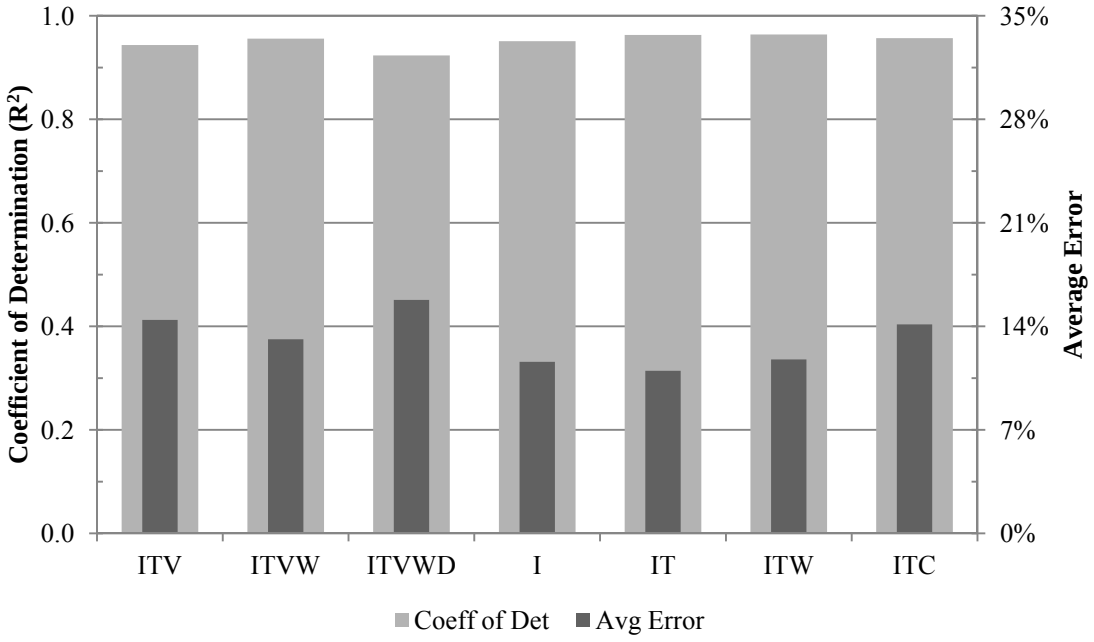


Figure 4.14 SS1 ANN testing results for the 2006-2007 models recalibrated with 50% of SS1-A/B data from 2011

Tables 4.7 and 4.8 indicate that it is possible to create an ANN model with low error by using training data that has a domain capable of representing as many operational conditions as possible. Furthermore, the testing results for the ANN models generally had lower error and higher R^2 values than the training scenario. This behavior indicates the ANN models' ability to predict results with low error for a large amount of data that wasn't originally included in the training process.

Table 4.7 SS1 ANN training results for the 2006-2007 models recalibrated with 50% of SS1-A/B data from 2011

	Mean Absolute Error	**R^2**	**0-20% Model Error Training**	**0-10% Model Error Training**
ITV	16.6%	0.942	82.9%	64.9%
ITVW	16.6%	0.942	82.7%	61.9%
ITVWD	18.7%	0.928	80.4%	59.2%
I	18.4%	0.922	82.5%	54.9%
IT	16.4%	0.937	84.5%	65.2%
ITW	16.9%	0.936	83.0%	64.6%
ITC	20.4%	0.926	79.7%	63.2%

Table 4.8 SS1 ANN testing results for the 2006-2007 models recalibrated with 50% of SS1-A/B data from 2011

	Mean Absolute Error	**R^2**	**0-20% Model Error Testing**	**0-10% Model Error Testing**
ITV	14.4%	0.944	76.3%	48.7%
ITVW	13.1%	0.956	84.3%	51.6%
ITVWD	15.8%	0.923	77.7%	47.8%
I	11.6%	0.951	86.9%	59.3%
IT	11.0%	0.963	89.9%	54.9%
ITW	11.8%	0.964	87.8%	55.2%
ITC	14.1%	0.957	80.4%	54.0%

Tables 4.9 and 4.10 compare the before and after test results for the separate recalibrations of the ANN model for SS1-A and SS1-B, respectively, using the 2011 dataset. Tables 4.9 and 4.10 exhibit low error and high R^2 values for ANN architectures that didn't include the distilland volume in the pre-calibration scenario. Moreover, Tables 4.9 and 4.10 illustrate a high mean absolute error and low R^2 value for ANN architectures that included the distilland volume in the pre-calibration scenario. This is attributed to a higher operating distilland volume being used during the 2011 study. Once the ANNs were recalibrated with new test data from the 2011 study, the ANN architectures that included distilland volume performed just as well as the ANN architectures that did not include distilland volume. However, it is possible to avoid recalibration by using an ANN architecture that avoids the distilland volume as an input. Tables 4.9 and 4.10 reflect the results for the 2006-2007 ANN model for SS1 with test data from 2011 prior to recalibration and post-recalibration. The test data for the post-recalibration scenario included data that was not included in the re-calibration training data.

Table 4.9 SS1-A ANN model performance before and after recalibration for the testing scenario

	No Recalibration		**After Recalibration**	
Model	**Mean Absolute Error**	**R^2**	**Mean Absolute Error**	**R^2**
ITV	17.7%	0.693	13.6%	0.935
ITVW	22.1%	0.649	12.6%	0.949
ITVWD	27.6%	0.421	15.0%	0.872
I	12.0%	0.899	9.0%	0.951
IT	10.5%	0.925	9.9%	0.958
ITW	10.2%	0.936	10.7%	0.956
ITC	12.9%	0.918	14.1%	0.947

Table 4.10 SS1-B ANN model performance before and after recalibration for the testing scenario

	No Recalibration		After Recalibration	
Model	**Mean Absolute Error**	**R^2**	**Mean Absolute Error**	**R^2**
ITV	40.5%	0.269	9.9%	0.960
ITVW	32.8%	0.460	9.6%	0.962
ITVWD	34.7%	0.513	11.7%	0.958
I	13.9%	0.870	9.4%	0.947
IT	13.1%	0.886	9.6%	0.961
ITW	12.1%	0.904	9.5%	0.965
ITC	12.5%	0.905	11.3%	0.964

Tables 4.9 and 4.10 illustrate a considerable decrease in the mean absolute error and increase in the R^2 value once the 2006-2007 ANN models are recalibrated with data from 2011 for four out of seven ANN models for SS1-A and three out of seven ANN models for SS1-B. The remaining SS1 recalibrated scenarios perform nearly as well as the non-recalibrated scenarios.

4.9 ANN Modeling Conclusions

The ANN modeling method has shown it is possible to create a model to accurately predict daily solar still production for commercially designed solar stills. Furthermore, the modeling methods prove that local environmental factors can be used to accurately predict solar still production with minimal error. ANN modeling yielded results with up to 78% of the predictions being within 0-10%, and up to 89% of the predictions being within 0-20% of actual distillate production. The use of historical insolation, temperature, distilland depth, and wind speed data (ITVWD) produced the best results for SS1 while insolation, temperature, and wind velocity data (ITV) produced the best results for solar still SS2.

Recalibration had to be performed for scenarios involving 2011 input data that exceeded the original domain of the 2006-2007 training data. Once the recalibration was performed, the resulting ANN is able to perform as well as the original ANN model. It is recommended that the original ANN training datasets include a wide array of historical conditions to produce the best simulated results without having to recalibrate the model.

CHAPTER 5

EQUATION BASED MODELING

5.1 Genetic Algorithms Overview

This study utilized Genetic Algorithms (GAs) as the second method to model daily solar still production using local weather data. Genetic algorithms serve as optimization techniques for problems with complex or non-linear relationships and use biological theories to execute the optimization processes. Natural selection and biological evolution serves as the basis of their computational power. Genetic algorithms are able to provide optimum solutions for what would otherwise be approximately formulated problems (Holland, 1975; Goldberg, 1989). The success of the genetic algorithm approach depends greatly on the appropriate selection of certain parameters that control reproduction, mutation, and the diversity of the individuals/chromosomes (Nehdi, El Chabib, & Said, 2007).

Genetic algorithms were first developed by John Holland and his students/colleagues in the 1960s and 1970s to understand adaptation as it naturally occurs throughout nature and to apply the natural mechanisms into computer systems (Mitchell, 1995). The ability to search through a large array of possible solutions, the ability to adapt to any environment or the user's demands, to innovate and construct something new to accomplish a task, and the ability to solve complex problems that are too difficult to program individually are just some of the computational problems required of computer programs (Mitchell, 1995). The mechanisms involved in adaptation and evolution are well suited for the wide array of computational problems in many scientific fields (Mitchell, 1995). In biological terms, the set of possibilities in an environment is the set

of possible genetic sequences while the "solutions," organisms able to survive and reproduce in the environment, are high fitness organisms (Mitchell, 1995). The fitness of each organism depends on factors such as the ability to withstand the physical properties of the environment, how well it can cooperate with other organisms in the environment, and the ability to reproduce offspring that are in turn also highly fit organisms (Mitchell, 1995).

Genetic algorithms are capable of solving constrained and unconstrained optimization problems based on processes that drive biological evolution (MathWorks, 2011; Mitchell, 1995; Coello, Lamont, & Van Veldhuizen, 2007; Pazos, Sierra, & Buceta, 2009). GAs are capable of searching for individuals/chromosomes in a space to find the best candidate solution to a given problem (Mitchell, 1995). Most GA methods have elements such as populations of chromosomes, methods to select chromosomes according to fitness, crossover to produce new offspring, and mutation that takes place during reproduction (Mitchell, 1995).

The chromosomes in a GA population consist of bit strings, strings of 1s and 0s where each bit position has two possible values, either 1 or 0 (Mitchell, 1995). The search for the best chromosome takes place by processing populations of chromosomes resulting in new and different populations from one iteration to the next (Mitchell, 1995).

Selection rules are used to select individuals/chromosomes to become parents and populate the next generation. Crossover rules combine the two parent individuals to create children for the next generation. Mutation rules apply changes to random individual parents to form mutated children (Mitchell, 1995; MathWorks, 2011). General optimization algorithm methods such as enumerative, deterministic, and stochastic

methods (Coello et al., 2007) generate single points, at each iteration, which are then sequenced until an optimal solution is reached. Progressive points in a sequence are then selected by using deterministic computation (Coello et al., 2007). Genetic algorithms, on the other hand, generate a population of points for every iteration. Each progressive population is selected by computation utilizing random number generators. The best individual in each population then approaches the optimal solution (Coello et al., 2007; MathWorks, 2011).

Genetic algorithms require a fitness or objective function with which to optimize by finding the minimum of the objective function (Coello et al., 2007; MathWorks, 2011). The value of the objective function for any chromosome/individual is known as the "score" or "fitness." The GA performs a series of computations to create successive new populations until the minimum of the objective function is found (Mitchell et al., 1995; MathWorks, 2011).

According to Mitchell (1995), a simple genetic algorithm works in the following way:

1. Begin with a randomly generated population of size "n" "m"-bit chromosomes
2. Fitness calculation for each chromosome "x" in the random population
3. Repeat steps a-c, below, until "y" offspring have been created
    a. Select a pair of parent chromosomes, with the probability of selection being an increasing function of fitness. This selection is done "with replacement" so that the same chromosome can be selected more than once to reproduce.
    b. A selected pair is crossed over at a random point, chosen with uniform probability, to form two offspring. If no crossover takes place, decided

according to the crossover probability, two offspring are created that are exact copies of the parents.

c. Offspring are mutated, according to the mutation probability, and are placed among the other resulting chromosomes in the new population.

4. Replace the starting population with the new population.
5. Repeat steps 2-4 until the best individual/chromosome is found according to the stopping criteria (number of generations, time, fitness level tolerance).

This study attempted to create GA models using objective functions to calculate daily distillate production. Multiple objective functions were created to estimate daily production with minimum error. The score for each individual was the GA models' estimate for daily production. The "fitness" for each individual is usually determined by the objective function as the score; however, this study utilized the percent error between the score for each objective function's prediction and the actual value for daily production to measure fitness. This method allowed for the GA to solve the various coefficients and exponents designated in each objective function while simultaneously minimizing the error for each progressive individual score.

Previous uses of GAs, in solar energy research, include Kalogirou's (2004) study to optimize a solar energy system in order to maximize its economic benefits, Varun's (2010) study to optimize the thermal performance a flat plate solar air heater, Cabello, Cejudo, Luque, Ruiz, Deb, and Tewari's (2011) study to optimize the size of a solar thermal electricity plant, and Loomans and Visser's (2002) study to optimize large solar hot water systems. These are a limited set of examples from a wider variety of genetic algorithm modeling attempts. Genetic algorithms have been successfully implemented in

the past to optimize large and small systems alike in terms of performance and cost.

5.2 Genetic Algorithm Modeling

The GA for this study was first developed using the data from the 2006-2007 study using the SS1 and SS2 production data. The input data for the GA optimization included total global insolation, ambient temperature, wind speed, wind direction, cloud cover and distilland volume. The detailed description and organization of the datasets for SS1 and SS2 can be found in sections 2.1, 4.3, and 4.6. The input variables were used in order to optimize an objective function to predict solar still production. Unlike ANNs, GAs do not require any normalization of the input data to enhance prediction performance. Similar to the ANN modeling methodology, 80% of each data set for SS1 and SS2 was used to create the optimized production function and the remaining 20% was used to validate the function's effectiveness.

The calibration and validation data sets for the GA modeling method contained the same data from the ANN modeling method as described in section 4.3. One difference between the two data sets is that the GA method does not need to separate validation and testing data. The GA validation data set contains the combined data from the validation and testing data set for the ANN method. Similar to the ANN method, the GA method varied the combination of input variables to find the optimum combination of inputs that would yield the best predicted results.

For the GA method, the optimization process calculates the best coefficients and exponents to minimize the error of the fitness function. The fitness function was designed to correlate daily production with daily weather data and the distilland volume. Linear, power, exponential, and sinusoidal functions were used to approximate daily production.

The fitness of the function was measured by calculating the mean absolute error between predicted and actual production at every iteration of the optimization process. The GA modeling was carried out by setting the stopping criteria to include a maximum iteration of 100 and a maximum change in percent error of $1x10^{-9}$. This study used the global optimization toolbox found in MathWorks's MATLAB® software and was computed using a 32 bit, 2.40 GHz Intel processor operating Microsoft's Windows XP ®.

5.2.1 Genetic Algorithm Modeling Results

The top ten GA models for SS1 and SS2 are presented in Tables 5.1 and 5.2, respectively. Linear (L), exponential (E), power (P), and sinusoidal (S) combinations were used with the historical weather data to predict daily production. The weather variables that were used included insolation (I) measured in MJ/m^2, temperature (T) measured in °C, distilland volume (V) measured in liters, wind speed (W) measured in m/s, wind direction (D) measured in degrees from north, and cloud cover (C) measured as a fraction of the total sky area. The GA models used daily insolation measured as MJ/m^2 as opposed to J/m^2 to render the magnitude of the insolation data to be similar to other input variables. The remaining variables used the same units as those used for the ANN models.

Coefficients of determination (R^2 value) were computed to determine the proportion of variance that is accounted for by the GA model. The Pearson correlation coefficient (R value) was also calculated to measure the correlation between predicted and actual production. The R value also provides a measure of how well future outcomes can be predicted using the GA model.

Table 5.1 Top ten developed fitness functions for SS1

Model	Fitness Function
L-I	$0.211 \cdot I - 1.069$
L-IT	$0.111 \cdot I + 0.045 \cdot T - 0.380$
L-ITV	$0.136 \cdot I + 0.066 \cdot T - 7.61x10^{-4} \cdot V - 0.434$
E-I	$0.707 \cdot e^{0.074 \cdot I}$
P-I	$0.010 \cdot I^{1.747} + 1.00$
P-IT	$0.043 \cdot I^{1.124} + 0.019 \cdot T^{1.480} + 0.046$
LS-I	$8.26 \cdot \sin(I \cdot \frac{\pi}{180}) + 0.044$
PS-I	$14.7 \cdot (\sin\left(I \cdot \frac{\pi}{180}\right))^{1.43} + 0.223$
LS-IT	$2.98 \cdot \sin\left(I \cdot \frac{\pi}{180}\right) + 3.49 \cdot \sin\left(T \cdot \frac{\pi}{180}\right) + 0.242$
PS-IT	$10.1 \cdot (\sin\left(I \cdot \frac{\pi}{180}\right))^{1.542} + 8.35 \cdot (\sin\left(T + \frac{\pi}{180}\right))^{1.813} + 0.149$

Table 5.2 Top ten developed fitness functions for SS2

Model	Fitness Function
L-I	$0.125 \cdot I + 0.033$
L-IT	$0.124 \cdot I + 0.055 \cdot T - 1.049$
L-ITV	$0.170 \cdot I + 0.049 \cdot T - 0.052 \cdot V - 0.199$
L-ITVW	$0.180 \cdot I + 0.035 \cdot T - 0.064 \cdot V + 0.033 \cdot W + 0.076$
L-ITVWD	$0.162 \cdot I + 0.041 \cdot T - 0.061 \cdot V - 0.021 \cdot W + 0.001 \cdot D + 0.354$
L-ITVWDC	$0.167 \cdot I + 0.023 \cdot T - 0.064 \cdot V - 0.038 \cdot W + 0.002 \cdot D - 0.381 \cdot C + 0.212$
E-I	$0.500 \cdot e^{0.080 \cdot I}$
P-I	$0.011 \cdot I^{1.788} + 0.008$
P-IT	$0.006 \cdot I^{1.245} + 0.039 \cdot T^{1.321} + 0.031$
PS-I	$17.0 \cdot (\sin\left(I \cdot \frac{\pi}{180}\right))^{1.793} + 0.030$

The top ten GA model results for SS1 and SS2 are presented in Tables 5.3 and 5.4, respectively. The criteria for evaluating the performance of different input combinations were the mean absolute error, percentage of model predictions within 30%, 20% and 10%

of the actual daily production, and the coefficient of determination (R^2).

Table 5.3 shows the calibration and validation results for the GA for SS1. The shaded values indicate the results for the validation scenario while the non-shaded values indicate the results for the calibration scenario. The power based fitness function based on the sine function using insolation data (PS-I) performed the best with respect to the R^2 value for both the calibration and validation scenario. Moreover, the linear based production function using insolation data (L-I) performs the best with respect to minimum error for both the calibration and validation scenario. The L-I scenario also proves to perform the best with regards to the error distribution for the 30%, 20%, and 10% categories. Despite having a low R^2 value for the validation data, L-I performs the best in 4 out of 5 performance categories.

Table 5.3 shows four out of ten, three out of ten, and five out of ten fitness functions performing with more results in the calibration scenario for the 0-30%, 0-20%, and 0-10% error categories, respectively. The calibration scenario can have as much as 30% (as seen in L-IT) and as little as 5% (as seen in P-IT) more results than the validation results. Table 5.3 indicates that the performance of the genetic algorithm does not increase as more variables are added. A high R^2 value is obtained by solely using insolation data for the L-I fitness function.

Table 5.4 shows L-ITV performing the best with respect to the R^2 value and the mean absolute error for SS2. The L-ITV scenario also performs the best with regards to the error distribution for the 10% and 20% categories. The L-IT function performs the best with respect to the 0-30% error category. Overall L-ITV performs the best in 4 out of 5 performance categories.

Table 5.4 shows the calibration results generally having a higher amount of predictions within 0-30% than the validation scenario. The calibration results can have as much as 12% (as seen in L-I) and as little as 0.1% (as seen in L-ITV) more results than the validation results. The inclusion of any more variables does not increase the performance of the fitness function.

The results presented in Table 5.3 show that it is possible to predict SS1's daily production relying solely on insolation data. This differs from the results presented in Table 4.1 since the ANN model relied on insolation, temperature, distilland volume, wind speed, and wind direction to produce the best model for SS1. The results presented in Table 5.4 align with the results shown in Table 4.2 for SS2. The two different modeling methods required insolation, temperature, and distilland volume to produce the best predictions for SS2.

The results in Tables 5.3 and 5.4 show that GA modeling can produce results with up to 89% and 92% of the testing predictions within 20% of the actual values for SS1 and SS2, respectively. Simple linear combinations of input variables tended to produce the best results. Tables 5.3 and 5.4 show that the best GA model for SS1 and SS2 featured higher R^2 values for the calibration data set than for the validation data set. Despite the lower R^2 values for the validation data set, the absolute difference in error between the calibration and validation data was 10.8% for SS1 (L-I) and 3.5% for SS2 (L-ITV).

Table 5.3 Top ten GA modeling results for SS1

MODEL	R^2	MEAN ABSOLUTE ERROR	% PREDICTIONS WITHIN 30% ERROR	% PREDICTIONS WITHIN 20% ERROR	% PREDICTIONS WITHIN 10% ERROR
L-I	0.911	21.4%	87.3%	78.7%	56.6%
	0.630	10.6%	93.4%	89.0%	70.3%
L-IT	0.921	21.1%	85.1%	67.7%	49.7%
	0.642	20.3%	89.0%	39.6%	17.6%
L-ITV	0.914	39.0%	70.7%	53.0%	33.7%
	0.629	11.8%	94.5%	86.8%	76.9%
E-I	0.892	33.3%	74.9%	64.4%	45.6%
	0.599	22.5%	90.1%	76.9%	45.1%
P-I	0.895	35.5%	75.3%	49.6%	18.4%
	0.931	27.1%	74.8%	52.2%	21.7%
P-IT	0.848	28.7%	82.9%	66.0%	33.8%
	0.852	25.3%	80.9%	61.7%	29.6%
LS-I	0.908	35.6%	68.6%	36.1%	14.1%
	0.922	26.1%	70.4%	43.5%	21.7%
PS-I	0.907	28.9%	77.5%	68.2%	42.9%
	0.934	21.2%	74.8%	60.0%	42.6%
LS-IT	0.871	36.6%	52.8%	33.1%	20.3%
	0.877	30.6%	60.9%	44.3%	20.9%
PS-IT	0.889	26.5%	84.6%	74.5%	46.5%
	0.908	20.8%	81.7%	74.8%	43.5%

I	Insolation (MJ/m^2)	C	Cloud Cover
T	Temperature (°C)	L	Linear Combination
V	Distilland Volume (m^3)	P	Power Combination
W	Wind Speed (m/s)	E	Exponential (e) Function
D	Wind Direction (Degrees from north)	S	Sine Function

Table 5.4 Top ten GA modeling results for SS2

MODEL	R^2	MEAN ABSOLUTE ERROR	% PREDICTIONS WITHIN 30% ERROR	% PREDICTIONS WITHIN 20% ERROR	% PREDICTIONS WITHIN 10% ERROR
L-I	0.938	29.8%	78.4%	44.8%	24.4%
	0.763	31.9%	66.1%	56.5%	29.0%
L-IT	0.933	15.7%	90.0%	74.8%	46.0%
	0.895	11.8%	96.8%	91.9%	69.4%
L-ITV	0.956	12.9%	90.4%	85.2%	74.8%
	0.881	16.5%	90.3%	80.6%	69.4%
L-ITVW	0.952	13.8%	90.0%	85.2%	74.8%
	0.841	19.4%	87.1%	74.2%	64.5%
L-ITVWD	0.956	13.2%	93.6%	88.0%	75.6%
	0.882	20.6%	85.5%	79.0%	64.5%
L-ITVWDC	0.941	16.8%	90.4%	81.6%	48.4%
	0.799	20.5%	85.5%	75.8%	58.1%
E-I	0.894	21.1%	84.4%	76.8%	49.6%
	0.821	20.5%	87.1%	72.6%	51.6%
P-I	0.930	17.9%	85.2%	64.4%	37.2%
	0.913	18.3%	80.6%	69.4%	41.9%
P-IT	0.723	29.7%	63.6%	50.0%	28.0%
	0.661	33.3%	64.5%	56.5%	37.1%
PS-I	0.933	16.4%	86.4%	71.2%	42.0%
	0.919	17.2%	85.5%	77.4%	53.2%

I	Insolation (MJ/m^2)	C	Cloud Cover
T	Temperature (°C)	L	Linear Combination
V	Distilland Volume (m^3)	P	Power Combination
W	Wind Speed (m/s)	E	Exponential (e) Function
D	Wind Direction (Degrees from north)	S	Sine Function

The residual plot for L-I for SS1 exhibited a right skew (skewness = 13.5) indicating a higher frequency of over predicting production. The residual plot for L-ITV for SS2 exhibited a slight right skew (skewness = 2.24) indicating a slightly higher frequency of over predicting production. Figure 5.1 illustrates the residual histogram for the L-I scenario for SS1 and Figure 5.2 illustrates the residual histogram for the L-ITV scenario for SS2.

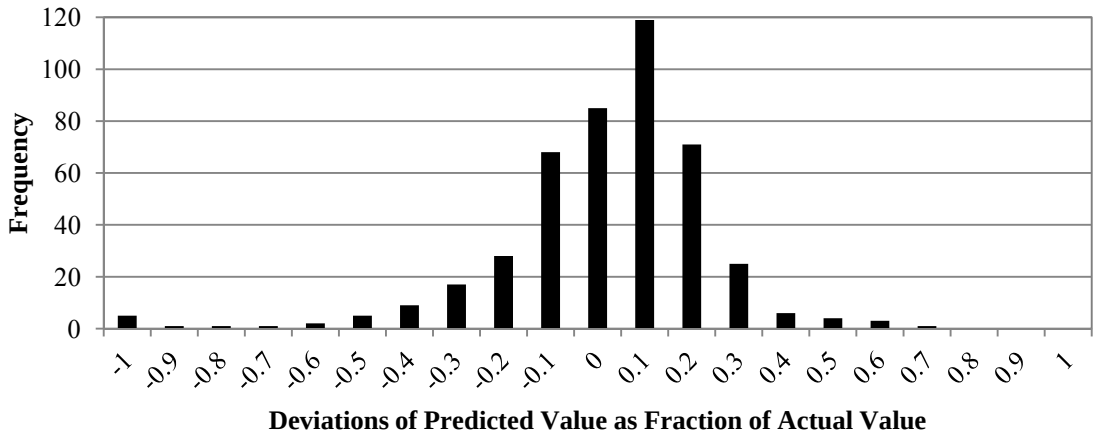


Figure 5.1 SS1's residual histogram for L-I model exhibiting right skew

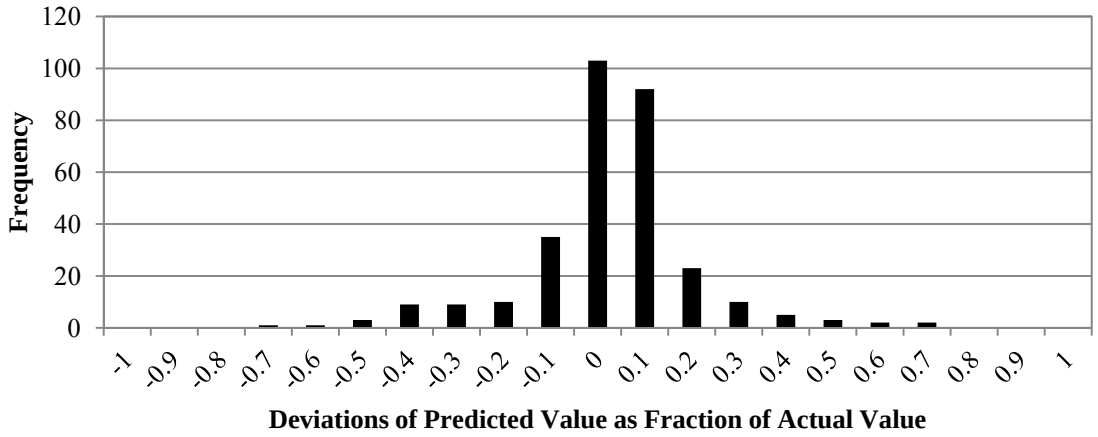


Figure 5.2 SS2's residual histogram for L-ITV model exhibiting slight right skew

5.2.2 Correlation Coefficients for GA Modeling

Plots of the relationships between GA predicted and actual daily production for each still, under the best performing input scenario, are shown in Figures 5.3 and 5.4.

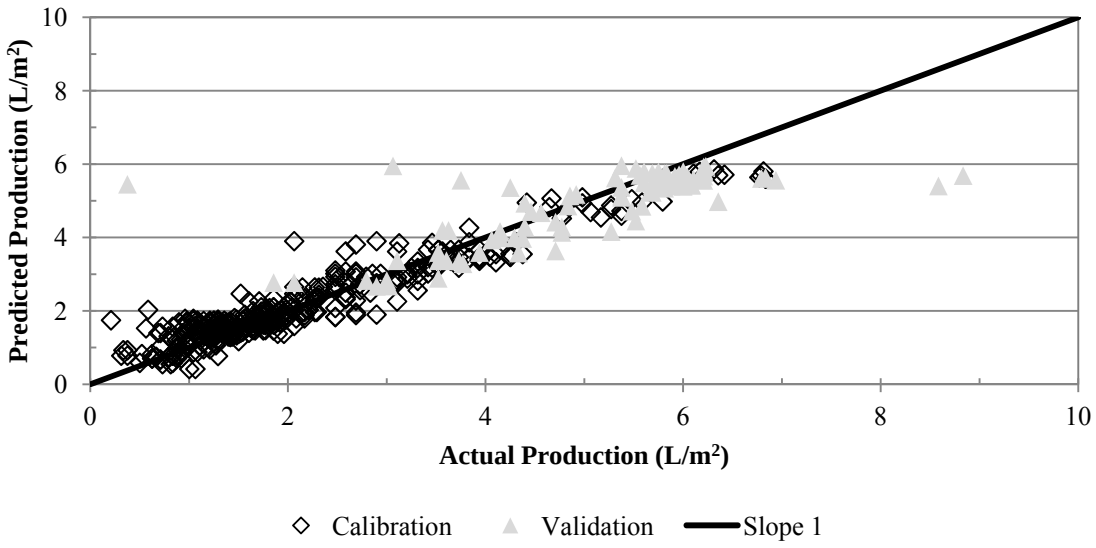


Figure 5.3 Predicted vs. actual distillate production for SS1 using L-I GA model

Figure 5.3 indicates a tight distribution of predicted vs. actual production data points centered on a bisecting line through the origin with a slope of one. Figure 5.3 shows SS1 having a couple of outlying points that could not be accounted for by the GA model.

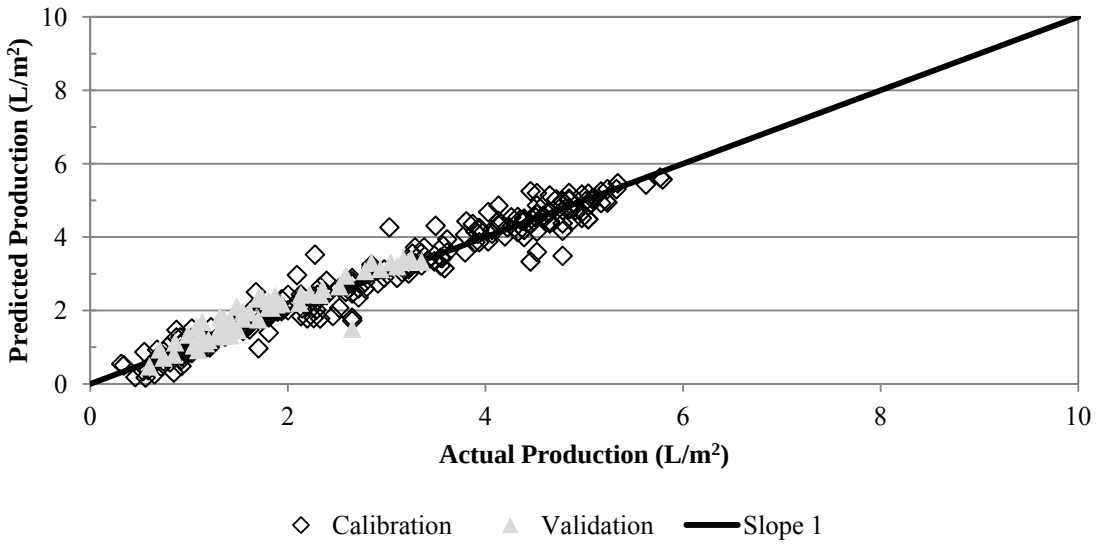


Figure 5.4 Predicted vs. actual distillate production for SS2 using L-ITV GA model

Figure 5.4 indicates a tight distribution of predicted vs. actual production data points centered on a bisecting line through the origin with a slope of one. Moreover, Figure 5.4 shows the model for SS2 performing better than the model for SS1 with fewer points lying away from the origin bisector.

Separate coefficients of determination for the GA models' training and testing data sets were calculated with respect to each set of experimental predictions. The coefficients were also calculated for the entire data set combining calibration and validation actual/predicted values. Table 5.5 shows the different coefficient values for the calibration, validation, and the combined calibration/validation data sets for the best input data scenario for SS1 and SS2. Table 5.5 illustrates the ability of the best GA models, for SS1 and SS2, to predict daily production with high correlation values.

Table 5.5 Determination (R^2) and Pearson correlation (R) coefficients for the best scenarios for SS1 and SS2 GA models

	SS1 (L-I)		**SS2 (L-ITV)**	
	R^2	**R**	**R^2**	**R**
Calibration	0.911	0.955	0.956	0.978
Validation	0.630	0.794	0.881	0.938
Combined	0.925	0.961	0.952	0.976

5.3 Multivariable Least Squares Regression

Regression is a frequently used method to analyze the relationship between dependent and independent variables. Statistical evaluations of regression analysis can also be used to evaluate the significance of each independent variable with regards to the dependent variable. Multivariable least squares techniques were used to minimize the error between the observed data and the predicted responses developed by the linear regression model.

5.3.1 Multivariable Least Squares Regression Methods

The least squares regression method was carried out using the data from Venkatesh (2007) for SS1 and SS2. The weather variables that were used for this study included daily total insolation (I) measured in MJ/m^2, temperature (T) measured in °C, distilland volume (V) measured in liters, wind speed (W) measured in m/s, wind direction (D) measured in degrees from north, and cloud cover (C) measured as a fraction of total sky area. The data was organized into calibration and validation data sets as was described in Section 5.2. The least squares regression was carried out using Microsoft's Excel® spreadsheet software using a 64 bit, 2.00 GHz Intel processor operating Microsoft's Windows 7®.

5.3.2 Multivariable Least Squares Regression Results

Tables 5.6 and 5.7 illustrate the top performing regression models in terms independent variable significance, minimum error, and error distribution. The significance of the independent variables were determined with an F-Test at a significance level of $\alpha = 0.05$. Insolation, temperature, wind speed, and cloud cover were the most significant variables for SS1 while insolation, temperature, distilland volume, and cloud cover were the most significant for SS2. The units for each input variable were the same as those used for the GA model method. Coefficients of determination (R^2 value) and Pearson correlation coefficients (R) were calculated for SS1 and SS2 as they had been previously done for the GA and ANN modeling methods.

Table 5.6 Developed multivariable regression models for SS1

Model	Regression Function
L-I	$0.229 \cdot \mathrm{I} - 1.361$
L-IT	$0.194 \cdot \mathrm{I} + 0.038 \cdot \mathrm{T} - 1.361$
L-ITW	$0.197 \cdot \mathrm{I} + 0.036 \cdot \mathrm{T} - 0.023 \cdot \mathrm{W} - 1.309$
L-ITC	$0.183 \cdot \mathrm{I} + 0.045 \cdot \mathrm{T} - 0.444 \cdot \mathrm{C} - 1.152$
LN-I	$3.41 \cdot \mathrm{LN(I)} - 6.93$
E-I	$0.414 \cdot \mathrm{e}^{0.096 \cdot \mathrm{I}}$
P-I	$0.029 \cdot \mathrm{I}^{1.541}$
P-T	$0.271 \cdot \mathrm{T}^{0.746}$

Table 5.7 Developed multivariable regression models for SS2

Model	Regression Function
L-I	$0.202 \cdot \mathrm{I} - 1.119$
L-IT	$0.173 \cdot \mathrm{I} + 0.028 \cdot \mathrm{T} - 1.141$
L-ITV	$0.171 \cdot \mathrm{I} + 0.040 \cdot \mathrm{T} - 0.049 \cdot \mathrm{V} - 0.084$
L-ITVC	$0.167 \cdot \mathrm{I} + 0.042 \cdot \mathrm{T} - 0.049 \cdot \mathrm{V} - 0.228 \cdot \mathrm{C}$
LN-I	$3.11 \cdot \mathrm{LN(I)} - 6.17$
E-I	$0.421 \cdot \mathrm{e}^{0.087}$
P-I	$0.038 \cdot \mathrm{I}^{1.43}$
P-T	$0.166 \cdot \mathrm{T}^{0.904}$

The results for the top performing regression models are summarized in Tables 5.8 and 5.9 for SS1 and SS2, respectively. The criteria for evaluating the performance of different input combinations were the mean absolute error, percentage of model predictions within 30%, 20%, and 10% of the actual daily production, and the coefficient of determination (R^2).

Table 5.8 Top eight MVR modeling results for SS1

INPUTS	R^2	MEAN ABSOLUTE ERROR	% PREDICTIONS WITHIN 30% ERROR	% PREDICTIONS WITHIN 20% ERROR	% PREDICTIONS WITHIN 10% ERROR
L-I	0.911	23.2%	86.2%	77.9%	54.4%
	0.630	25.3%	93.4%	90.1%	74.7%
L-IT	0.930	18.8%	87.8%	83.1%	68.8%
	0.680	25.9%	94.5%	91.2%	82.4%
L-ITW	0.927	19.8%	85.4%	77.4%	54.3%
	0.539	26.9%	93.4%	85.7%	72.5%
L-ITC	0.931	18.4%	86.0%	80.2%	56.7%
	0.536	25.9%	94.5%	87.9%	72.5%
LN-I	0.793	34.5%	80.9%	69.6%	47.2%
	0.635	29.2%	92.3%	70.3%	30.8%
E-I	0.862	22.9%	86.2%	69.6%	41.4%
	0.587	56.0%	63.7%	48.4%	22.0%
P-I	0.922	20.1%	88.1%	79.8%	48.3%
	0.624	26.8%	94.5%	87.9%	67.0%
P-T	0.595	44.2%	59.9%	42.0%	25.4%
	0.251	42.6%	37.4%	22.0%	13.2%

Table 5.9 Top eight MVR modeling results for SS2

INPUTS	R^2	MEAN ABSOLUTE ERROR	% PREDICTIONS WITHIN 30% ERROR	% PREDICTIONS WITHIN 20% ERROR	% PREDICTIONS WITHIN 10% ERROR
L-I	0.938	18.2%	86.4%	80.8%	66.4%
	0.763	30.0%	71.0%	51.6%	41.9%
L-IT	0.953	13.7%	91.2%	86.0%	74.0%
	0.848	21.7%	83.9%	75.8%	58.1%
L-ITV	0.957	12.7%	92.0%	87.6%	77.2%
	0.869	19.6%	83.9%	77.4%	67.7%
L-ITVC	0.957	12.6%	89.6%	82.4%	68.4%
	0.865	19.4%	77.4%	62.9%	41.9%
LN-I	0.819	34.2%	81.2%	73.2%	48.0%
	0.648	50.7%	38.7%	30.6%	17.7%
E-I	0.885	20.3%	85.6%	72.4%	44.8%
	0.825	16.2%	91.9%	82.3%	59.7%
P-I	0.939	17.5%	87.2%	78.8%	58.8%
	0.789	23.6%	80.6%	67.7%	46.8%
P-T	0.652	37.0%	65.6%	48.0%	30.8%
	0.562	30.4%	77.4%	59.7%	37.1%

Table 5.8 shows the calibration and validation results for the MVR model for SS1. The highlighted values indicate the results for the validation scenario while the non-highlighted values indicate the results for the calibration scenario. The multiple regression model using insolation and temperature (L-IT) performs the best with regards to the R^2 value for the calibration and validation scenario. In terms of mean absolute error, both L-IT and L-ITC perform nearly identically with L-IT having 0.4% more error for the calibration scenario. L-IT also performs the best with regards to the amount of predictions within 10% and 20% error. L-IT and P-I are nearly identical with regards to the 30% error category. L-IT has 0.3% less results within 30% compared to P-I. Overall, the L-IT model performs the best in 3 out of the 5 performance categories; however, the results for L-IT are extremely close to the best model in the remaining 2 performance categories.

Table 5.8 shows the validation results generally having a higher amount of predictions within 0-30% than the calibration scenario. The validation scenario can have as much as 20% more results, in a particular prediction category, than the calibration scenario (as seen in L-I). Moreover, the calibration scenario can have as much as 22% more results than the validation scenario (as seen in E-I and P-T). Table 5.8 shows that the performance of the MVR models does not increase by solely adding new variables. Instead, the proper combination of variables yields the best results.

Table 5.9 shows the L-ITV model performing the best in terms of the R^2 value, average error, and the error distribution for SS2. The L-ITVC performs closely to L-ITV in terms of the R^2 value and the mean absolute error. Overall, L-ITV performs the best in 5 out of 5 performance categories. The validation scenario can have as much as 15%

more results, in a particular prediction category, than the calibration scenario (as seen in E-I) for the 0-30% error categories. Moreover, the calibration scenario can have as much as 29% more results than the validation scenario (as seen in L-I). Tables 4.2, 5.4, and 5.9 indicate that the performance of SS2 can be best modeled by using insolation, temperature, and distilland volume data.

The results shown in Tables 5.8 and 5.9 indicate that MVR modeling can produce results with up to 87% of predictions being within 20% of the actual values. Figures 5.5 and 5.6 illustrate the residual histogram for each still's best performing model. SS1's residual histogram for L-IT exhibits a strong left skew (skewness = -13.9) indicating a higher frequency of under predicting production. SS2's residual histogram for L-ITV exhibits a slight right skew (skewness = 1.47) indicating a slightly higher frequency of over predicting production.

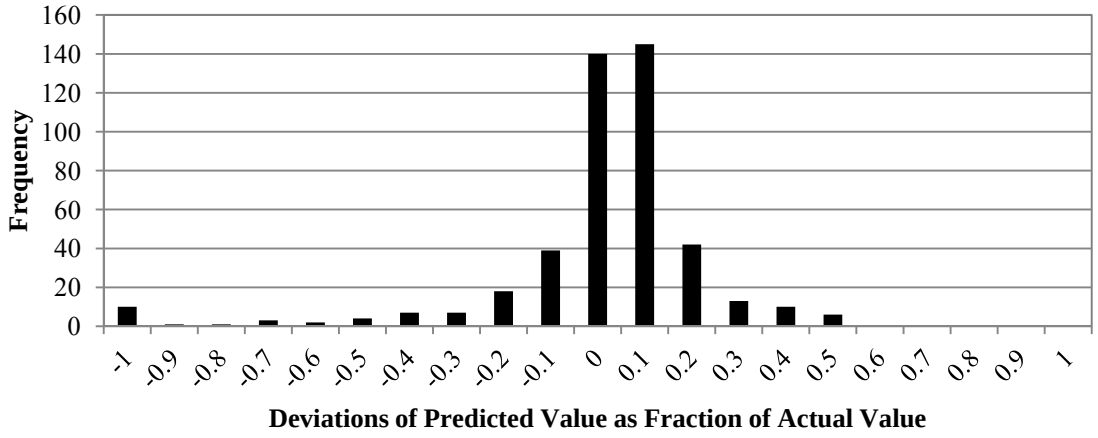


Figure 5.5 SS1's residual histogram for L-IT model exhibiting a left skew

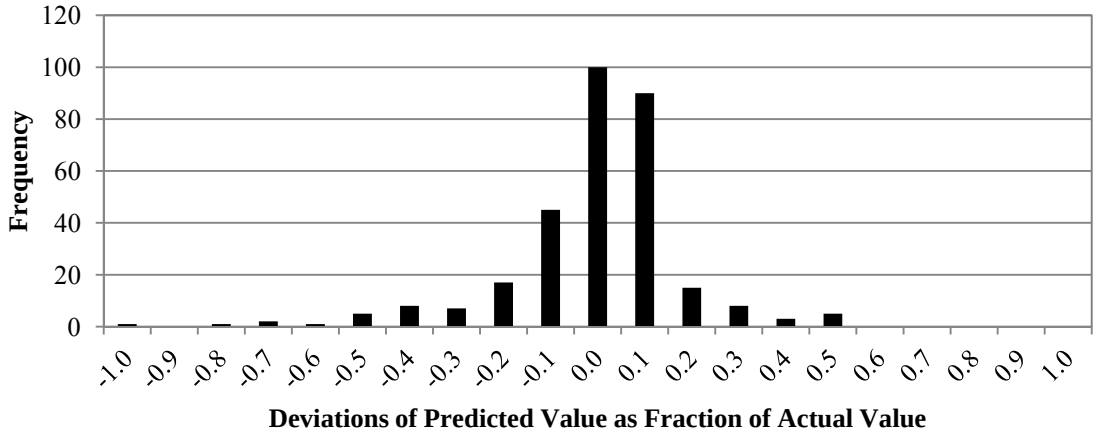


Figure 5.6 SS2's residual histogram for L-ITV model exhibiting a slight right skew

5.3.3 Correlation Coefficients for MVR Modeling

Plots of the relationships between MVR predicted and actual daily production for each still, under the best performing input scenario, are shown in Figures 5.7 and 5.8.

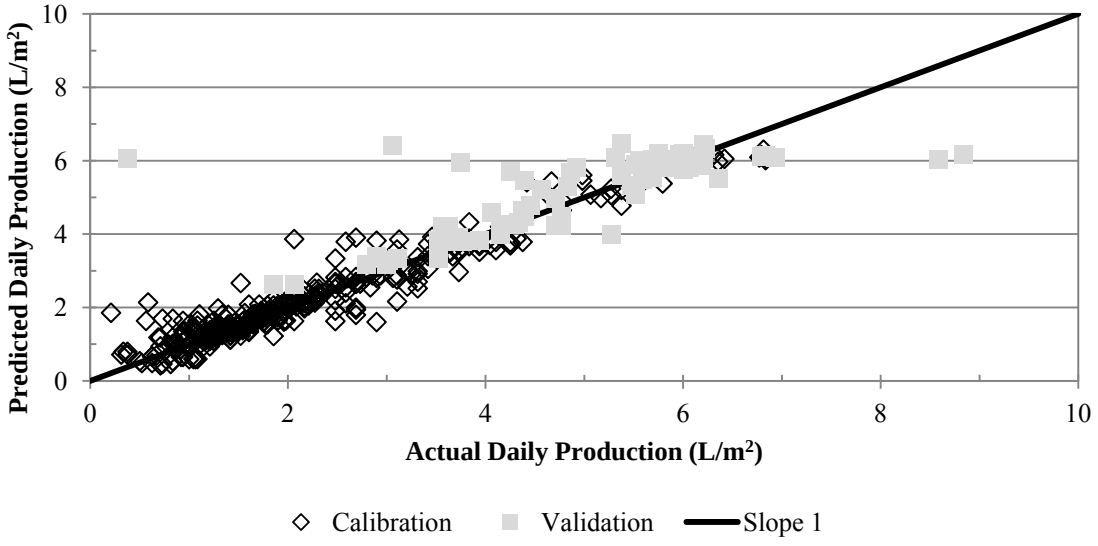


Figure 5.7 Predicted vs. actual daily production for SS1 using L-IT MVR model

Figure 5.7 shows the distribution of model predictions centered on a bisecting line

through the origin with a slope of one. Similar to Figure 5.3, Figure 5.7 shows SS1 having a couple of outlying points that could not be accounted for by the MVR model. Besides the outlying data points, a majority of the results show that predicted and actual values for daily production are clustered close together along the origin bisector line.

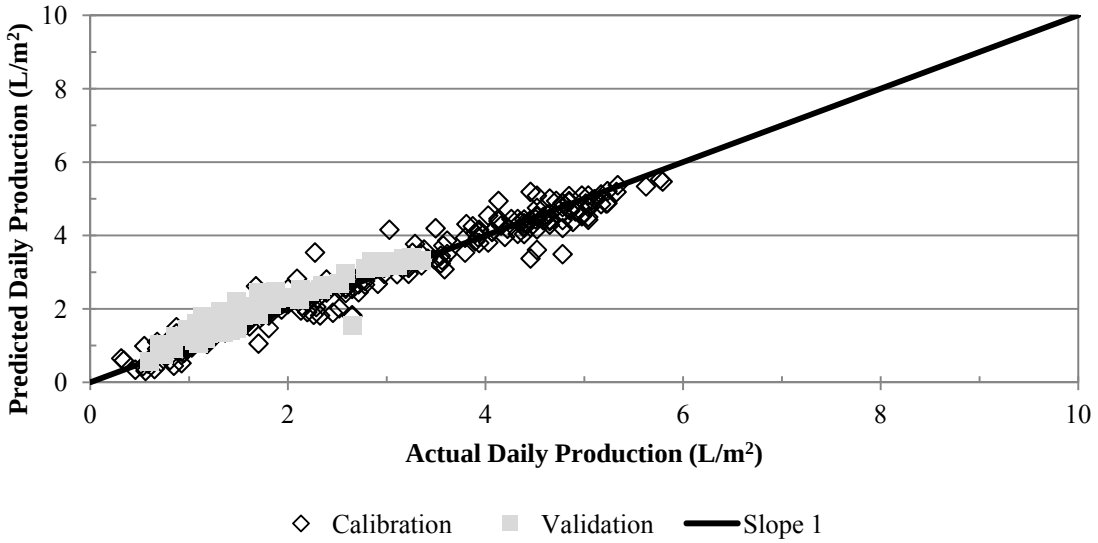


Figure 5.8 Predicted vs. actual daily production for SS2 using L-ITV MVR model

Figure 5.8 indicates a tight distribution of data points centered on a bisecting line through the origin with a slope of one. Figure 5.8 shows the model for SS2 performing better than the model for SS1 with less points lying away from the origin bisector.

Table 5.10 shows the different coefficient values for the calibration, validation, and the combined calibration/validation data sets for the best input scenario for SS1 and SS2. The MVR predictions for SS1 and SS2 exhibit high R^2 and R values for the calibration data and lower values for the validation data. The MVR model for SS1 produced a much smaller validation R^2 value compared to the validation R^2 value for SS2.

Table 5.10 Determination (R^2) and Pearson correlation (R) coefficients for the best scenarios for SS1 and SS2 regression models

	SS1 (L-IT)		SS2 (L-ITV)	
	R^2	**R**	**R^2**	**R**
Calibration	0.930	0.964	0.967	0.978
Validation	0.681	0.825	0.869	0.932
Combined	0.936	0.968	0.951	0.975

5.4 Reliability

The 5th percentile production for each still was calculated using the L-I and L-ITV GA model for SS1 and SS2, respectively. As discussed in section 4.6, data from SS1-C had to be incorporated into the reliability calculations to complete the summer trend for SS1. The L-I GA model for SS1 was used to predict values for the missing Summer 2006 data. The predicted values for SS1-C had an average 8% error and an R^2 value of 0.879. Figures 5.9 and 5.10 illustrate the results for the predicted and actual lower 5th percentile production for SS1 and SS2, respectively.

Figure 5.9 shows that SS1's lower 5th percentile predicted and actual monthly average daily production for the winter (December – February) varied between 0.38 L/m^2 to 1.56 L/m^2. Lower 5th percentile predicted and actual monthly average daily production for the summer (June – August) varied between 2.46 L/m^2 and 4.63 L/m^2. The predicted monthly average daily production values were within 30-80% of the actual daily distillate production.

The GA model for SS1 has a peak positive 35.3% difference for the month of October 2006 and a peak negative 79.5% difference for the month of December 2006. A positive percent difference indicates the predicted value was less than the actual value while a negative percent difference indicates the predicted value was greater than the actual vale.

Figure 5.9 indicates the L-I GA model for SS1 tended to under predict average daily production throughout the year except for the months of March ’06, April ’06, December ’06, January ’06, May ’07 and June ’07.

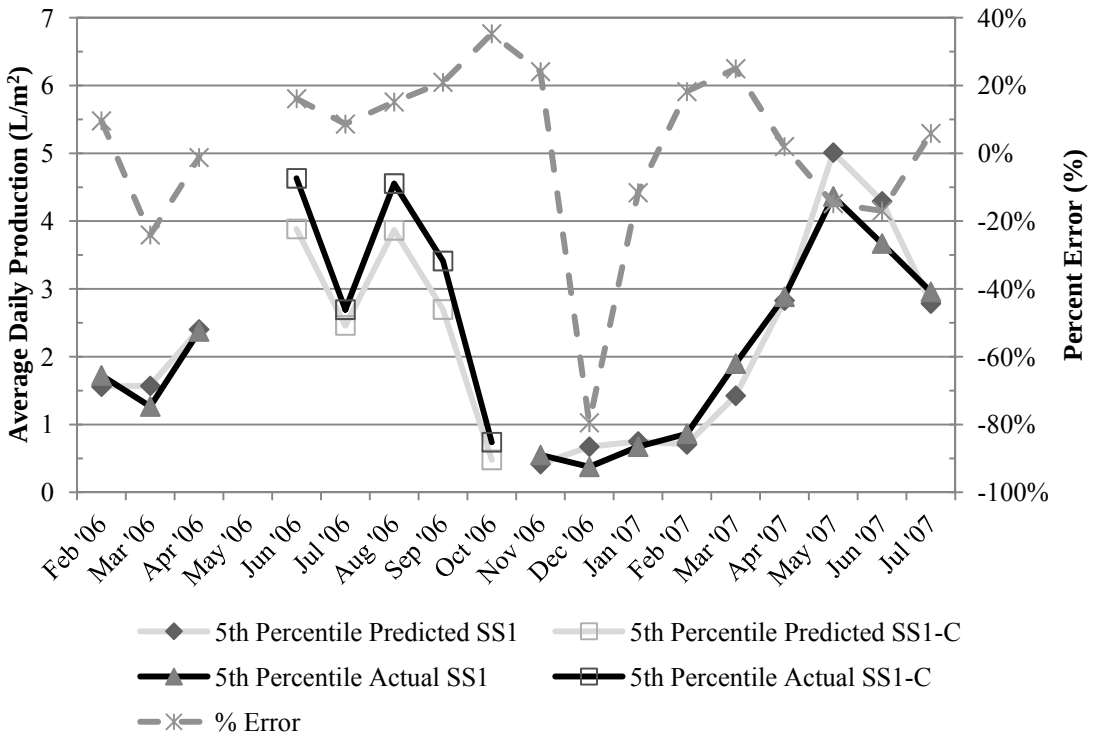


Figure 5.9 SS1’s 5th percentile actual and predicted average daily production for L-I GA model

Figure 5.10 shows that the lower 5th percentile predicted and actual monthly average daily production for the winter (December – February) varied between 0.22 L/m^2 to 1.70 L/m^2. Lower 5th percentile predicted and actual monthly average daily production for the summer (June – August) varied between 2.55 L/m^2 and 4.48 L/m^2. The predicted monthly average daily production values were within 50% of the actual monthly average daily distillate production.

The GA model for SS2 has a peak positive 35.9% difference for the month of

December 2006 and a peak negative 60.7% difference for the month of October 2006. Figure 5.10 indicates the L-ITV GA model for SS2 tended to over predict average daily production throughout the year except for the months of February ’06, November ’06, December ’06, and January ’07. Tables 5.11 and 5.12 summarize the results of the 5th percentile production throughout the 2006-2007 study.

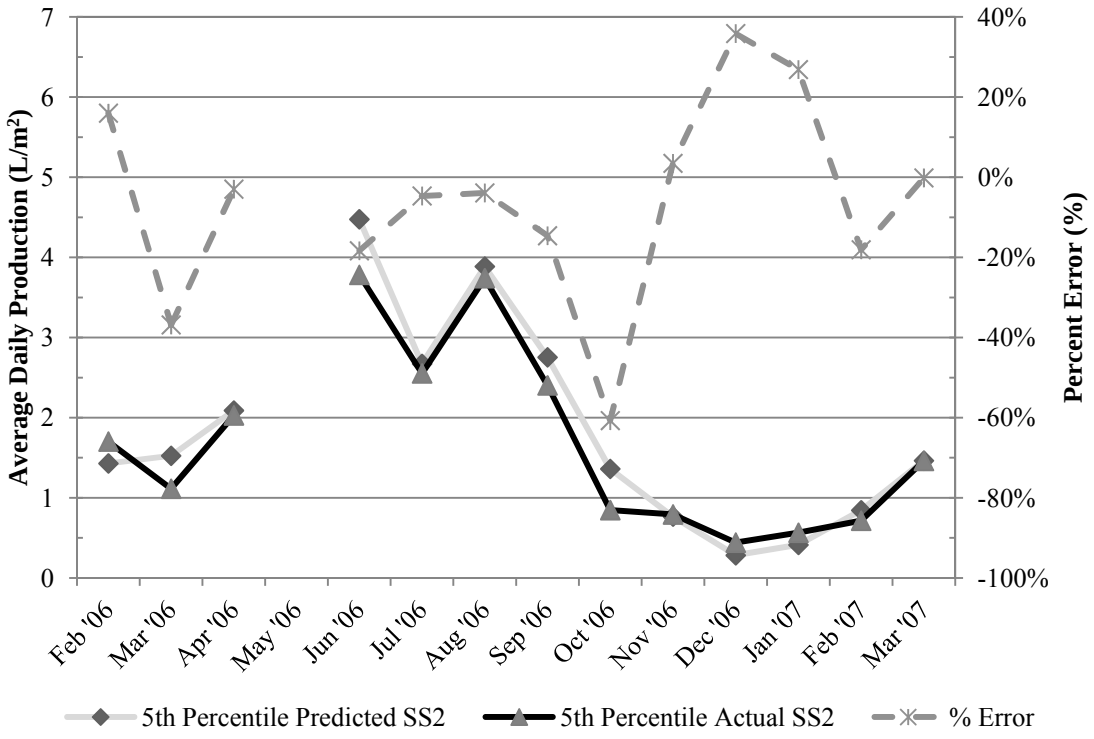


Figure 5.10 SS2’s 5th percentile actual and predicted average daily production for L-ITV GA model

Table 5.11 SS1 summary of lower 5th percentile average daily actual and predicted distillate production with L-I GA model

Number of Months	17	Percent of Time	Average Error	Standard Deviation of Error
Monthly Predictions > Actual	6	35%	-24.7%	27.9%
Monthly Predictions < Actual	11	65%	16.4%	9.7%

Table 5.12 SS2 summary of lower 5th percentile average daily actual and predicted distillate production with L-ITV GA model

Number of Months	13	**Percent of Time**	**Average Error**	**Standard Deviation of Error**
Monthly Predictions > Actual	9	69%	-9.7%	11.8%
Monthly Predictions < Actual	4	31%	29.6%	14.8%

Table 5.11 shows that the L-I GA model for SS1 over predicted by 24.7%, 35% of the time and under predicted by 16.4%, 65% of the time. The L-I GA model for SS1 tends to underestimate 5^{th} percentile values which indicates that a design based on these data would be conservative. Table 5.12 shows that the L-ITV GA model for SS2 over predicted by 9.7%, 69% of the time and under predicted by 29.6%, 31% of the time. The L-ITV GA model for SS2 tends to over predict 5^{th} percentile daily distillate production. A solar still system design might not be conservative for SS2 and could be under sized thus preventing the proper supply of water to be produced. Figures 5.11 and 5.12 illustrate a comparison between actual and predicted values for SS1 and SS2, respectively.

Figure 5.11 illustrates how the GA model predictions for SS1 are close to the actual daily production values; however, the GA model is not able to predict the more extreme minimum and maximum values that occurred. Figure 5.12 illustrates similar results to Figure 5.11; however, the GA model for SS2 is better able to predict the more extreme minimum and maximum values more closely than the GA model for SS1.

The percent difference for the 1^{st} quartile, minimum, median, maximum, and 3^{rd} quartile between the actual daily production and the L-I GA model for SS1 were 13.8%, 66.8%, 5.2%, 38.3%, and 8.5%, respectively. The highest percent difference between the predicted and actual values for SS1 occurred for the minimum and maximum daily production values.

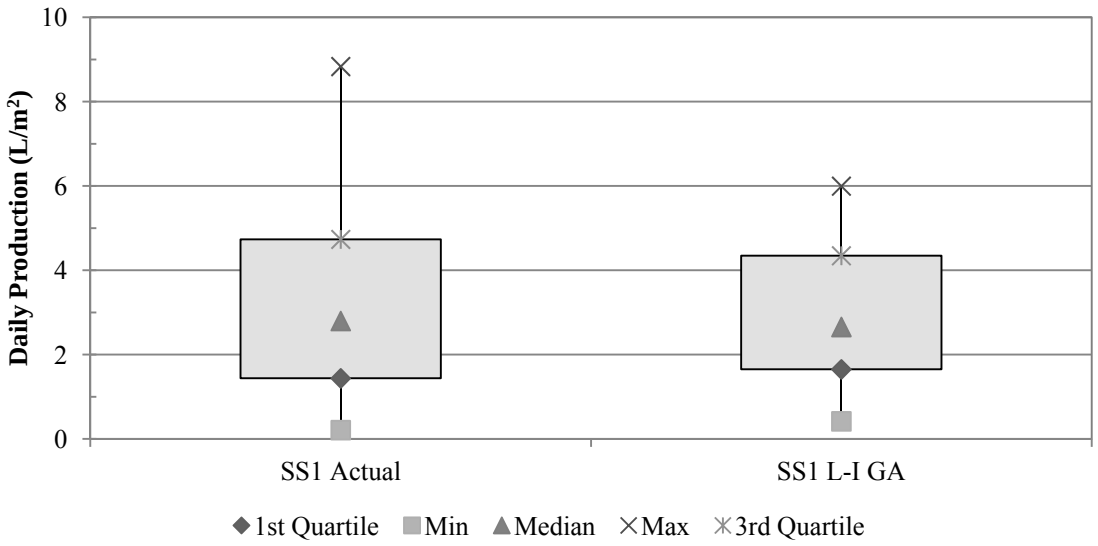


Figure 5.11 Descriptive statistics box plot for SS1's actual and L-I GA predicted production

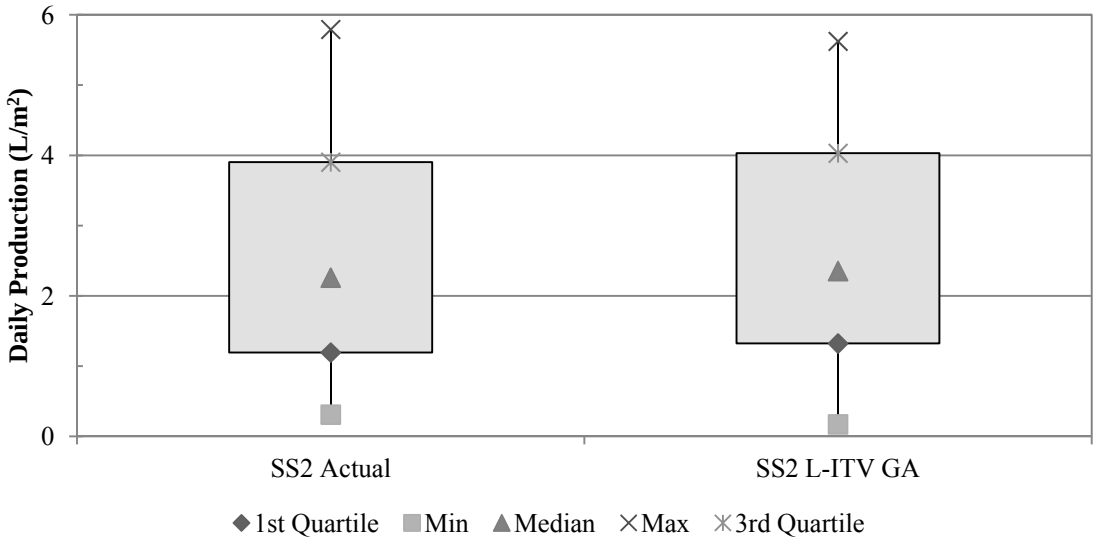


Figure 5.12 Descriptive statistics box plot for SS2's actual and L-ITV GA predicted production

The percent differences for the 1st quartile, minimum, median, maximum, and 3rd quartile between the actual daily production and the L-ITV GA model for SS2 was 10.2%, 57.4%, 4.1%, 3.0%, and 3.2%, respectively. The highest percent difference

between the predicted and actual values for SS2 occurred for the minimum and maximum daily production values.

5.5 GA Validation

In order to validate the developed GA models, the new data from the 2011 study was used as additional validation data for the GA models that were developed for SS1 using the 2006-2007 data. Table 4.6 previously summarized the domain of the calibration and validation data from the 2006-2007 and the 2011 study.

The input scenarios for the recalibration were chosen based on the significance of each variable as seen during the regression modeling and the performance of each model prior to recalibration. The distilland volume was included as a variable to determine if high error occurs for the L-ITV scenario given new data from 2011. This was done to compare with the ANN model that performed with high error for the ITV architectures being tested with new 2011 data prior to recalibration.

Figure 5.13 illustrates the results of the predictions for the 2011 data with respect to the R^2 value and the mean absolute error. The results presented in Figure 5.13 indicate the 2006-2007 GA models' ability to predict daily production with less than 15% error for four out of five models given data collected during 2011. The L-I model performed the best with regards to the mean absolute error (12.4%) while the L-ITC model performed the best with regards to the R^2 value (0.954). The L-IT model featured the highest mean absolute error (23.4%) however the R^2 value (0.941) was the 3rd highest out of the presented models. The L-ITV GA model validation, before recalibration, performed with a mean absolute error of 13.8% and featured an R^2 value of 0.930 which is much better compared to the ANN ITV model validation, before recalibration, with a

mean absolute error of 29.1% and an R^2 value of 0.446.

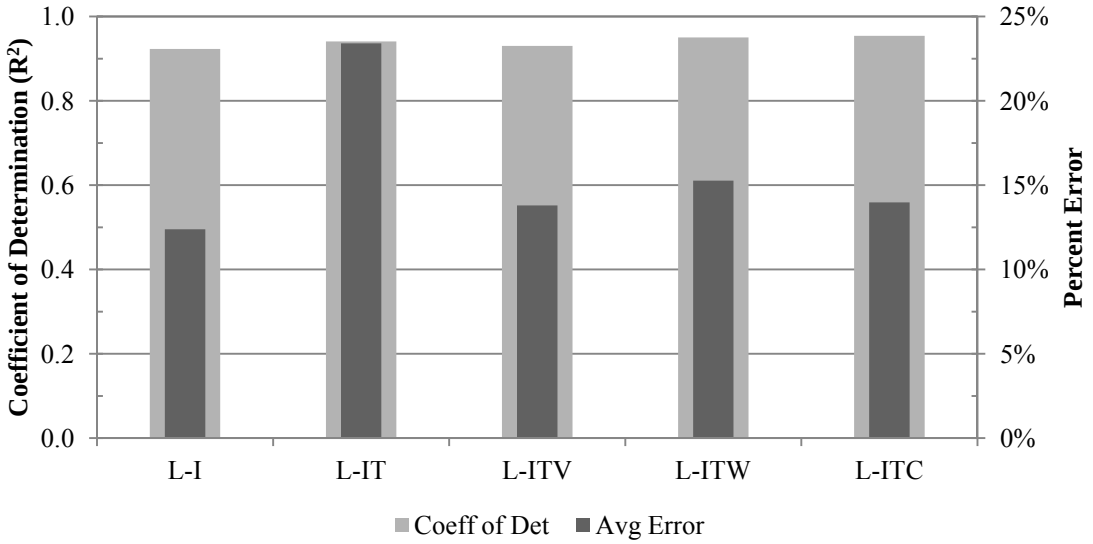


Figure 5.13 R^2 and mean absolute error for 2011 SS1-A/B using GA models developed with data from 2006-2007, without recalibration

The GA models were recalibrated using the 2011 data following the same procedure presented in section 4.8 for the ANN models. The results of recalibration are shown in Figures 5.14 and 5.15. The recalibration for the GA models results with the L-ITW model performing the best with regards to mean absolute error and the R^2 value for the calibration as well as the validation results. Tables 5.13 and 5.14 detail the performance of each GA model, following the recalibration, for the calibration and validation scenarios.

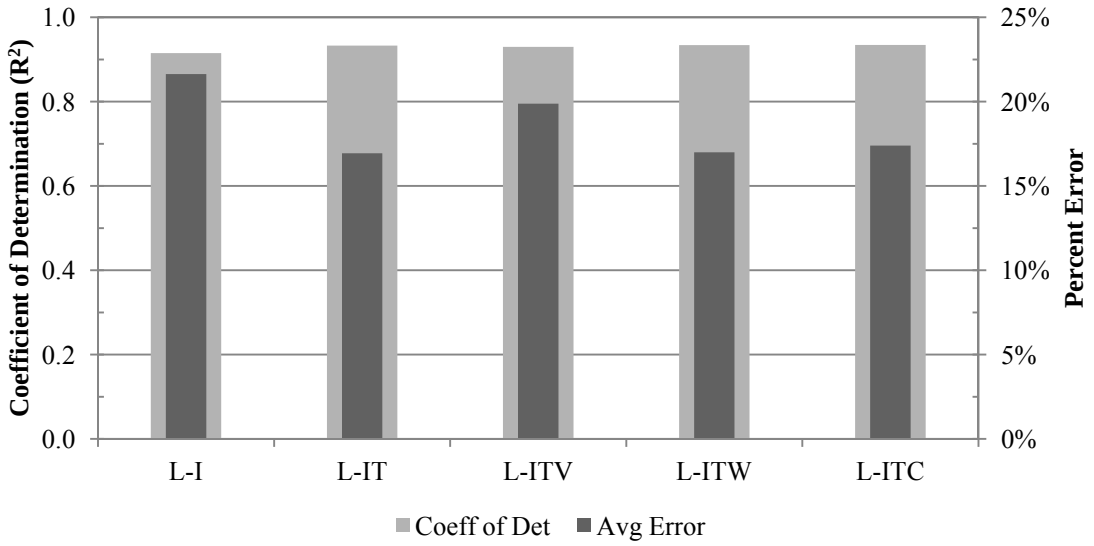


Figure 5.14 SS1 GA calibration results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

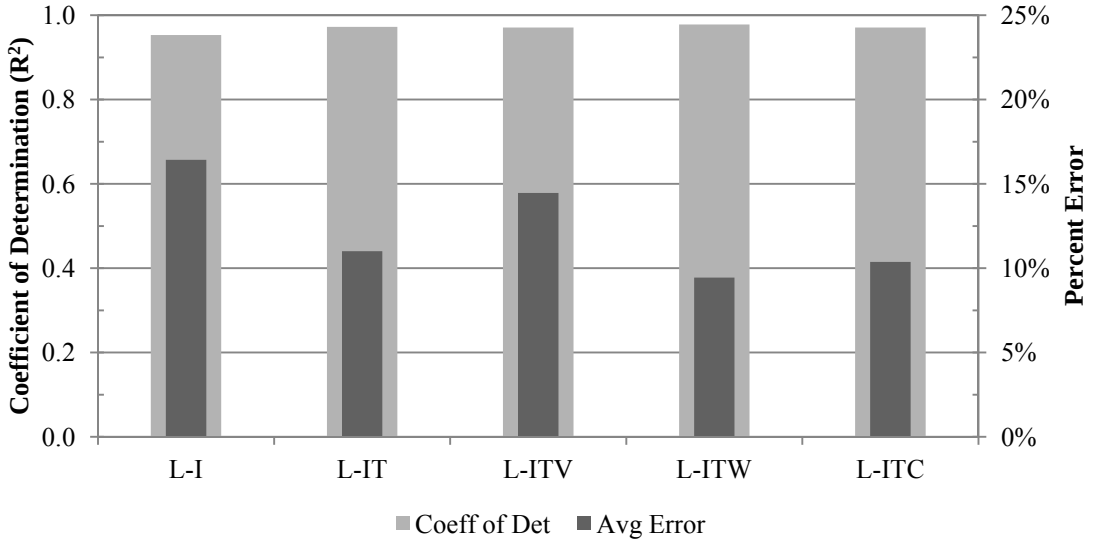


Figure 5.15 SS1 GA validation results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

Table 5.13 SS1 GA calibration results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

	Mean Absolute Error	R^2	0-20% Model Error Calibration	0-10% Model Error Calibration
L-I	21.6%	0.915	73.2%	29.1%
L-IT	16.9%	0.933	83.7%	64.1%
L-ITV	19.9%	0.930	79.7%	37.8%
L-ITW	17.0%	0.934	84.7%	68.6%
L-ITC	17.4%	0.934	82.7%	62.7%

Table 5.14 SS1 GA validation results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

	Mean Absolute Error	R^2	0-20% Model Error Calibration	0-10% Model Error Calibration
L-I	16.4%	0.953	71.8%	29.7%
L-IT	11.0%	0.972	91.7%	53.1%
L-ITV	14.5%	0.971	85.2%	30.9%
L-ITW	9.4%	0.978	94.1%	67.7%
L-ITC	10.4%	0.971	89.0%	66.8%

Tables 5.13 and 5.14 indicate that it is possible to generate a GA model with lower error by using an extended calibration data that has a domain capable of representing as many operational conditions as possible. Furthermore, the validation results for the GA models generally had lower error and higher R^2 values than the calibration scenario. This behavior indicates the GA models' ability to predict results with low error for input data that wasn't originally part of the calibration process.

Tables 5.15 and 5.16 illustrate the before and after results for the recalibration of the GA models for the validation data of SS1-A and SS1-B, respectively. Tables 5.15 and 5.16 reflect the validation results for the 2011 dataset for SS1-A and SS1-B, respectively.

Table 5.15 GA model validation performance for SS1-A before and after recalibration

	No Recalibration		After Recalibration	
Model	**Mean Absolute Error**	**R^2**	**Mean Absolute Error**	**R^2**
L-I	11.6%	0.949	18.0%	0.958
L-IT	23.9%	0.946	11.5%	0.966
L-ITV	12.2%	0.932	16.0%	0.974
L-ITW	16.2%	0.971	9.9%	0.978
L-ITC	14.1%	0.973	9.2%	0.973

Table 5.16 GA model validation performance for SS1-B before and after calibration

	No Recalibration		After Recalibration	
Model	**Mean Absolute Error**	**R^2**	**Mean Absolute Error**	**R^2**
L-I	13.1%	0.902	14.5%	0.950
L-IT	22.9%	0.940	10.5%	0.971
L-ITV	15.4%	0.932	12.8%	0.970
L-ITW	14.4%	0.935	8.0%	0.980
L-ITC	13.9%	0.940	7.6%	0.971

Tables 5.15 and 5.16 indicate a high average error for the L-IT GA models for the pre-calibration scenario. Despite the fact that the temperature data for the 2011 study were within the minimum and maximum range for the 2006-2007 study, the inclusion of the temperature data along with insolation results in a GA model performing with high mean absolute error. The inclusion of temperature along with other variables does not yield as high an error as it did when paired only with insolation.

Tables 5.15 and 5.16 illustrate a decrease in the mean absolute error and an increase in the R^2 value once the 2006-2007 GA models were recalibrated with data from 2011 for six out of seven GA models for SS1-A and SS1-B. The L-I GA model was the only scenario that did not improve with regards to the mean absolute error following the

recalibration.

5.6 Regression Validation

In order to validate the developed regression models, the new data from the 2011 study was used as additional validation data for the regression models that were developed using 2006-2007 data for SS1. Table 4.6 previously summarized the domain of the calibration and validation data from the 2006-2007 and the 2011 study.

Figure 5.16 illustrates the results of the predictions for the 2011 data with respect to the R^2 value and the mean absolute error for the top five performing regression models. Four out of the five presented input scenarios display error near or below 10% and R^2 values greater than 0.90.

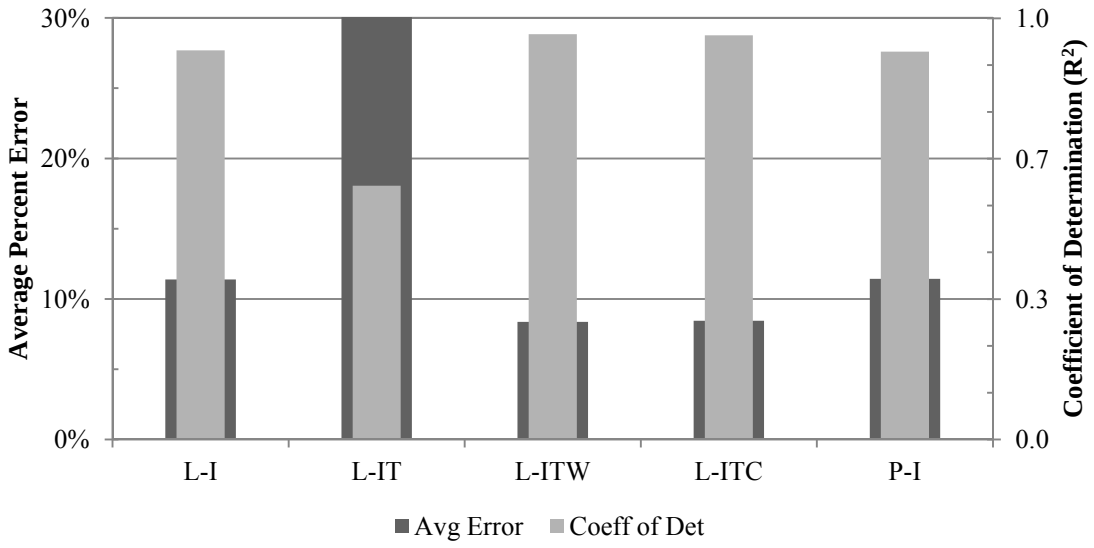


Figure 5.16 R^2 and mean absolute error for 2011 SS1-A/B using regression models developed with data from 2006-2007, without recalibration

The results presented in Figure 5.16 indicate the 2006-2007 regression models' ability to predict daily production with less than 15% error for four out of the five models by

using data collected during 2011. The L-IT regression model performed with the highest mean absolute error and lowest R^2 value out of the top five performing regression models. Table 4.6 previously illustrated the minimum and maximum values for the historical ambient temperature conditions. The 2011 study was performed well within the minimum and maximum ambient temperature conditions that were experienced during the 2006-2007 study. Since different window seals were used for the 2011 study, there could have been a different relationship between the daily production and ambient temperature for the 2006-2007 and the 2011 study.

The top five regression models were recalibrated using data from the 2011 study to examine if any change in model performance occurs. Each regression model was recalibrated by combining the 2006-2007 data with 50% of the 2011 dataset. Figure 5.17 and 5.18 show the results of the new recalibrated and validation scenarios for the regression models.

The recalibration of the regression models results with the L-ITW scenario performing the best with regards to the mean absolute error and the coefficient of determination. The results of the recalibrated regression models indicate a better overall performance for each data input scenario compared to the original 2006-2007 regression models. Tables 5.17 and 5.18 detail the performance of each regression model following the recalibration.

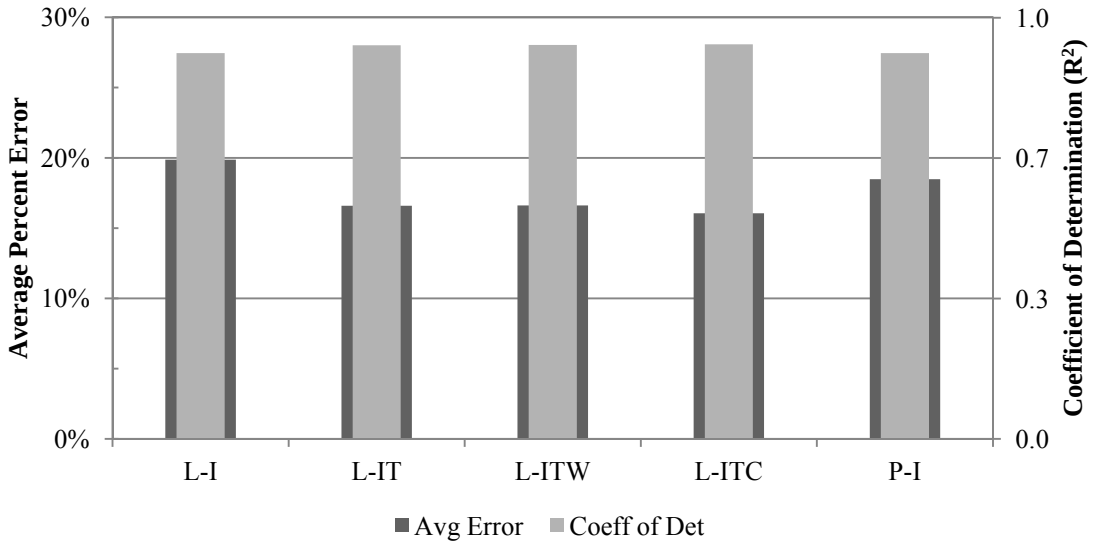


Figure 5.17 Regression calibration results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

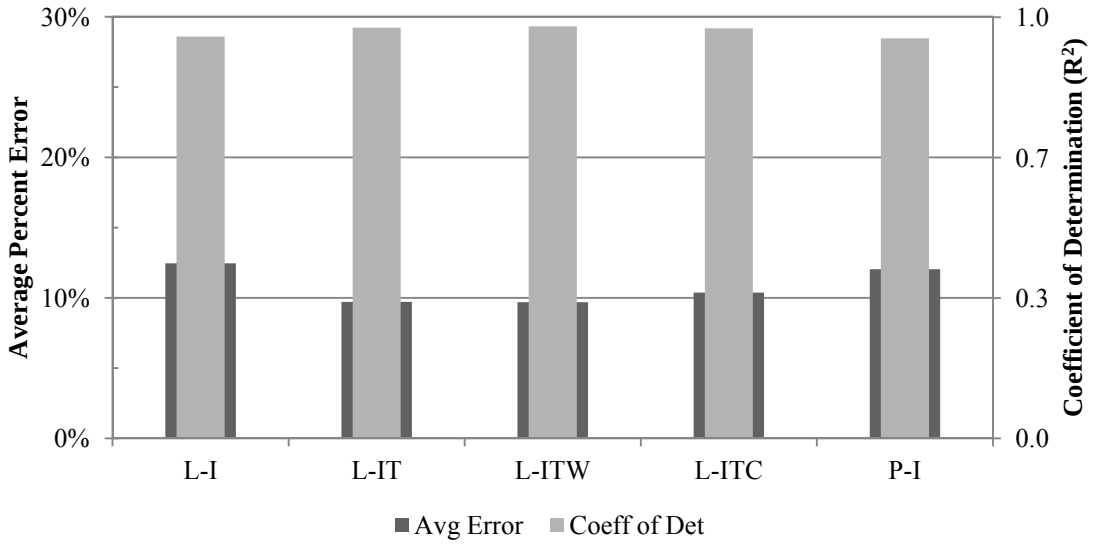


Figure 5.18 Regression validation results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

Table 5.17 Regression calibration results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

	Mean Absolute Error	R^2	0-20% Model Error Calibration	0-10% Model Error Calibration
L-I	19.9%	0.915	79.7%	55.1%
L-IT	16.6%	0.934	85.4%	68.1%
L-ITW	16.6%	0.934	84.9%	70.4%
L-ITC	16.1%	0.936	86.4%	66.9%
P-I	18.5%	0.915	83.0%	48.8%

Table 5.18 Regression validation results for the 2006-2007 models recalibrated with 50% of the data from 2011 for SS1-A/B

	Mean Absolute Error	R^2	0-20% Model Error Validation	0-10% Model Error Validation
L-I	12.5%	0.953	87.5%	59.9%
L-IT	9.7%	0.975	94.1%	62.9%
L-ITW	9.7%	0.977	94.1%	62.0%
L-ITC	10.4%	0.973	93.5%	58.8%
P-I	12.0%	0.949	86.4%	51.0%

Tables 5.17 and 5.18 indicate that it is possible to achieve a low error for regression models by using calibration data that has a domain capable of representing as many operational conditions as possible.

Tables 5.19 and 5.20 illustrate the before and after results for the recalibration of the regression models for SS1-A and SS1-B, respectively. Half of the 2011 data from SS1-A and SS1-B were added to the 2006-2007 dataset. The recalibrated scenario for SS1-A included only data from SS1-A with the 2006-2007 and 2011 SS1-A data to perform the recalibration. Likewise, the recalibrated scenario for SS1-B included only data from SS1-B with the 2006-2007 and 2011 SS1-B data to perform the recalibration.

Table 5.19 Regression model performance for SS1-A before and after recalibration

	No Recalibration		After Recalibration	
Model	**Mean Absolute Error**	**R^2**	**Average Error**	**R^2**
L-I	10.4%	0.949	10.4%	0.958
L-IT	31.4%	0.579	9.6%	0.973
L-ITW	7.8%	0.975	10.1%	0.977
L-ITC	7.8%	0.972	10.1%	0.970
P-I	10.9%	0.943	11.2%	0.948

Table 5.20 Regression model performance for SS1-B before and after recalibration

	No Recalibration		After Recalibration	
Model	**Mean Absolute Error**	**R^2**	**Average Error**	**R^2**
L-I	12.3%	0.902	9.1%	0.950
L-IT	29.8%	0.627	8.7%	0.974
L-ITW	8.9%	0.952	8.6%	0.978
L-ITC	9.1%	0.949	8.9%	0.973
P-I	12.0%	0.901	9.5%	0.947

Tables 5.19 indicate a high mean absolute error and low R^2 value for the L-IT model prior to recalibration. Once the recalibration was performed, the L-IT model performed similar to the other regression models and feature low mean absolute error and high R^2 values. Table 5.19 indicates that the recalibration for SS1-A's L-ITW, L-ITC, and P-I models slightly increases the mean absolute error for each model.

Table 5.20 illustrates a decrease in mean absolute error and an increase in the R^2 value once the 2006-2007 regression models are recalibrated with data from 2011 for five out of the five regression models.

5.7 A System Dynamics Model for a Solar Still System

The application of system dynamics allows for an approach to understand the

behavior of complex systems over time (Sterman, 2000). A traditional system dynamics model features internal feedback loops and time delays to replicate the behavior of a system. Unlike ANNs, GA, and regression methods, system dynamics allows the incorporation of public or private policies (Saleh, Oliva, Kampmann, and Davidsen, 2009). The purpose of system dynamics is to identify how different policies affect system behavior that could be problematic. In turn, new structural or policy based solutions could then be implemented to correct problematic behavior (Sterman, 2000).

Since system dynamics modeling is a problem driven method, it also takes on a functional perspective in that validation is an iterative process (Saleh et al., 2009). Confidence in the developed system dynamics model is gradually built as the model becomes a useful representation of the actual problem at hand (Saleh et al., 2009). A system dynamics application for solar stills would allow for the analysis of the effects of supply, demand, and storage capabilities for an actual solar still system.

5.7.1 Solar Still System Dynamics Model Organization

Figure 5.19 illustrates the developed System Dynamics (SD) model for an example solar still system using STELLA modeling software from ISEE Systems.

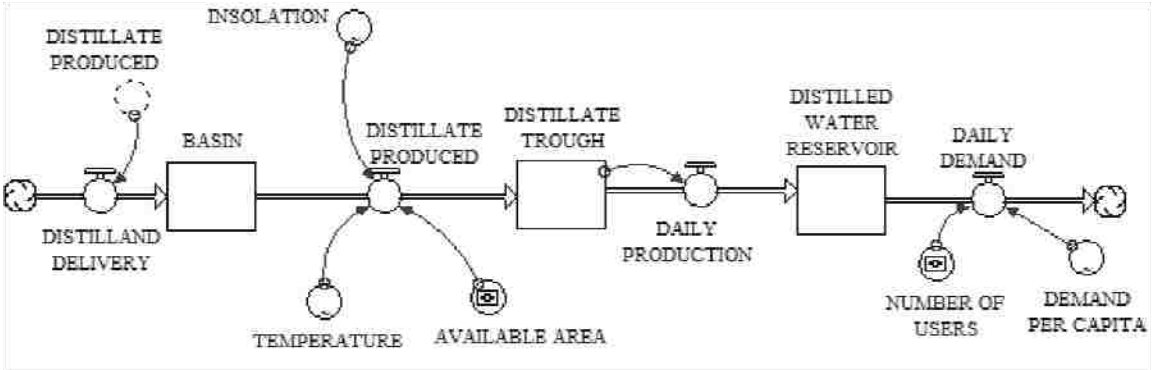


Figure 5.19 System dynamics model for a solar still system

The system dynamics model is composed of stocks (represented by rectangles) which suggest a container holding the contents of a stock, inflows represented by a pipe with an arrow pointing into a stock which suggests the addition of content into a stock, and outflows represented by a pipe with an arrow point out of a stock which suggests the removal of content from a stock. The stocks for the solar still SD model include the distilland basin, the distillate trough, and the distilled water reservoir. Each stock represents where water is stored and is later removed.

The inflow into the distilland basin represents the delivery of the water that is to be purified. The distilland basin has an outflow that represents the distillate produced as a result of evaporation which in turn is an inflow for the distillate trough. The distillate trough collects water for a short period of time before the water leaves the solar still as an outflow which in turn becomes an inflow for the distilled water reservoir. Water is then removed from the distilled water reservoir from an outflow caused by daily demand from the end user.

The circular units are known as converters or transforming variables. The converters add new information to a stock or flow and can be used to relate the effects of multiple variables on a stock or flow. Figure 5.19 illustrates the effect of the various converters on the amount of distillate produced and the total amount of water demanded by the end user. The produced distillate appears as a ghost (converter with dashed lines) to replicate the flow from the produced distillate. The produced distillate is replicated to mirror the real behavior of an actual system where the solar still basin is refilled with water to replenish the water that evaporated as a result of the distillation process. The replication ensures that the solar still basin operates with a constant distilland volume over time.

5.7.2 System Dynamics Model Preparation

The converters that are connected to the produced distillate flow are variables that have been shown to affect the amount of water produced by a solar still. The recalibrated SS1 L-IT multivariable regression model was used as the foundation for the relationship between the outflow and connectors for the produced distillate due to its low error. The insolation and temperature variables both vary depending on the time of year and have been shown to affect daily production. The distilland volume was not included since the variable was considered non-significant by the regression model's F-Test.

The goal of this phase of the study was to examine the effect that basin area and population have on the water demand and supply for a simple solar still system. The data from SS1 and SS1-C were used since the combined dataset contains 547 data points and contains insolation and temperature data between February 2006 and July 2007. Since the system dynamics model can handle up to 1,500 data points, the remaining data points consisted of insolation and temperature data between August 2007 and March 2010. The L-IT regression model was used by the SD model to project daily production between August 2007 and March 2010.

The minimum water demand per capita was varied depending on the season of operation. A 5 L/Day per capita demand was used during the summer period, a 4 L/Day per capita demand was used during the spring/fall period, and a 3 L/Day per capita demand was used during the winter period (World Health Organization, 2005). The graphical user interface (GUI) was set up to vary the total amount of solar still basin area and the population based on the choice of the user. The model was run using a 64 bit, 2 GHz processor and utilized a 4^{th} order Runge-Kutta integration method with a time step

of 1/4 of a day.

5.7.3 System Dynamics Model Results

In general, surplus water is produced during the spring and summer seasons when the stills produce more water than is actually demanded. A deficit in water production generally occurs during the fall and winter when production alone cannot sustain the minimum water demand. In order to meet water demand year round, the use of a distilled water reservoir allows for the constant supply of water regardless of the season. The water reservoir will gain supply when the production exceeds demand; furthermore, the reservoir will lose supply when the demand exceeds production. Figure 5.20 illustrates the distilled water reservoir volume, over time, for a solar still system operated with a user population of two and a total basin area of 2.50 m^2. The initial volume of the reservoir was set to 400 L.

Figure 5.20 illustrates the seasonal pattern of daily production and the volume of the distilled water reservoir. Over the period of 4 years, there is a maximum storage volume of 1123 L, minimum storage volume of 70 L, maximum daily production of 16 L and a minimum daily production of 0.07 L. Figure 5.20 illustrates the need to have a storage capacity of 1,200 L to meet the long term demands of two people.

Figure 5.21 illustrates the solar still system for a population of 10 and a total basin area of 12.4 m^2. The initial volume of the reservoir was set to 2,000 L. Figure 5.21 illustrates a maximum storage volume of 5,280 L, minimum storage volume of 140 L, maximum daily production of 79 L, and a minimum daily production of 0.35 L. The 10 person scenario illustrates a high demand during the first year which requires a large basin area and a higher initial stored water volume in order to prevent running out of

potable water during the summer of 2007 and 2008.

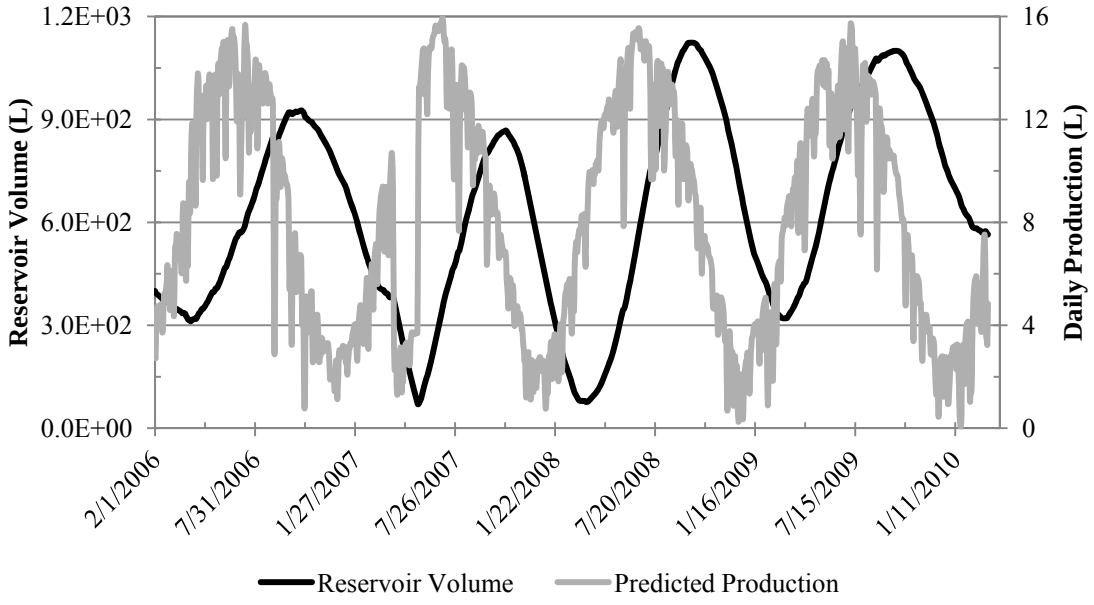


Figure 5.20 System dynamics modeled production and reservoir volume requirement for two people

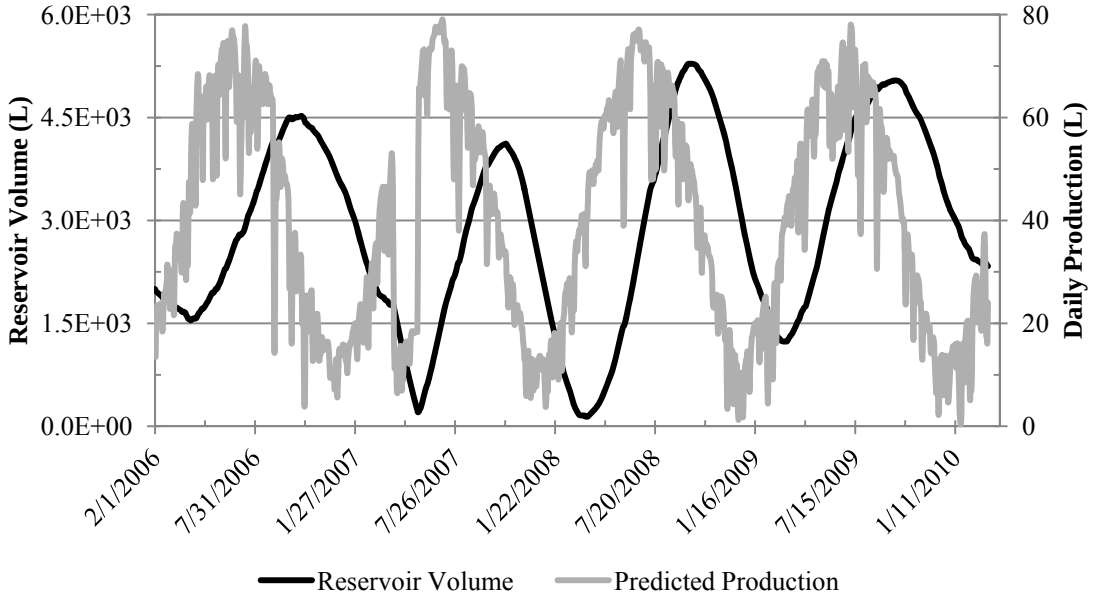


Figure 5.21 System dynamics modeled production and reservoir volume requirement for ten people

Figures 5.20 and 5.21 illustrate a 3 month lag between the maximum/minimum daily production and the maximum/minimum reservoir storage volume. The lag occurs due to seasonally varying production quantities supplying more than the required amount of water well past the peak production date. Once the daily production can no longer maintain the minimum demand, the reservoir volume is depleted to satisfy demand. The same behavior occurs when the minimum production is observed. Despite increasing daily production after the minimum production date, the required daily production to meet demand is not met until three months later. Once the minimum daily production is met, the reservoir's stored volume begins to increase.

Figures 5.20 and 5.21 also illustrate a difference in response to seasonal conditions between the daily production and the reservoir volume. Figures 5.20 and 5.21 indicate sporadic daily production values that fluctuate as a result of the insolation and ambient temperature. There is no smooth line connecting daily estimated production points throughout the four years represented by the data. On the other hand, the reservoir volume is shown to behave smoothly over time. This behavior indicates that the long term storage volume is more dependent on the seasonally varying average daily production values than on the short term daily fluctuations caused by sudden weather changes.

5.8 Equation Based Modeling Conclusions

Genetic algorithms and multivariable regressions are capable of predicting daily solar still production using local weather data. Insolation, ambient temperature, wind speed, and cloud cover were the most significant weather variables in terms of their contribution to daily solar still production.

The original genetic algorithm and regression models developed with the 2006-2007 solar still data were tested with new data collected in 2011. Despite having used data from a different time period, the genetic algorithm and regression models predicted daily production with a mean absolute error as low as 12% and 8.4%, respectively, for the 2011 data. Recalibrating the developed genetic algorithm and regression models with data from 2011 resulted in daily production predictions with a mean absolute error as low as 9.4% and 9.7%, respectively.

The developed models for solar still production were used along with a system dynamics model to forecast solar still production between August 2007 and March 2010. The system dynamics model was also developed to calculate the required distilled water storage volume to provide a sufficient amount of water to individuals. The system dynamics model illustrated the storage requirement of 1,100 L to maintain a group of two people and 5,200 L to maintain a group of 10 people over a period of four years. Furthermore, the maximum/minimum storage volume lagged the maximum/minimum daily production by three months. The system dynamics model also illustrated how the storage volume was more dependent on seasonally varying production than short term fluctuations in production.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

6.1 General Conclusions

This study has shown that passive, single basin solar stills are capable of producing as much as 6.8 L/m^2 of distilled water in the summer and as little as 0.21 L/m^2 of distilled water in the winter. The seasonally varying production of solar stills could be seen in the datasets from both the 2006-2007 and the 2011 study.

The ability to model daily solar still production was made possible by utilizing local weather data and by implementing artificial neural networks, genetic algorithms, and multivariable regression. While each method had a very different approach to developing a predictive model, the models with the simplest inputs generally worked the best. Even though datasets for input variables that included wind speed, wind direction, and cloud cover were acquired and tested for each different modeling method, the total insolation and ambient temperature data proved to be the most important inputs needed to obtain a model with low error.

Once each model type was completed, the determination of the lower 5^{th} percentile production for average daily production by month was calculated to determine the reliability of solar still production year round. A system dynamics model was also created using a developed regression model to project daily solar still production for 1,500 days. The system dynamics model was also used to determine the required basin area and storage volume to fulfill the daily potable water demands year round. Two and a half square meters of basin area and a storage capacity of 1,100 L were found to be necessary to support 2 people while 12.4 m^2 of basin area and a storage capacity of 5,280 L were

found to be necessary to support 10 people year round.

The analysis of hourly solar still production and hourly temperature readings illustrated how the distilland volume affects the timing of the peak production and distilland/vapor temperatures as well as the proportion of total production made during the day and night. A lag of 1-2 hours was observed between peak insolation and the peak hourly distilland, vapor, and inner glass cover temperatures. Furthermore, the recorded data illustrated how larger distilland volume scenarios experienced a longer lag between peak insolation and peak distilland, vapor, and inner glass cover temperatures.

The evaporative and convective heat transfer coefficients were calculated based on methods developed by Jakob (1949) and later implemented by Dunkle (1961) for solar stills. The heat transfer coefficients illustrated a stronger correlation between hourly production and the evaporative heat transfer coefficient than for the convective heat transfer coefficient. Furthermore, the predictions for hourly production based on the heat transfer model performed with R^2 values greater than 0.8 and with a mean absolute error between 26% and 53%.

The validation of the Mathioulakis et al. (1999) model for day and night time production produced results with a day production mean absolute error of 9.4% and a night production mean absolute error of 86.1%. The Mathioulakis model when modified to include the distilland's stored energy for the night production model and resulted with a mean absolute error of 34.9%, a 51.2% decrease. Overall, the original Mathioulakis et al. (1999) model performed the best for total daily production (day plus night production) with a mean absolute error of 7.5% compared to 8.8% for the modified Mathioulakis et al. (1999) model.

6.2 Modeling Results Summary

Tables 6.1 and 6.2 summarize the daily production modeling results for SS1 (Sunwater) and SS2 (SolAqua) according to the various model types and performance characteristics. The highlighted values indicate the results for the testing/validation scenarios.

Table 6.1 Model summary for 2006-2007 SS1 following recalibration with 50% of data from 2011 for SS1-A/B

	R^2		**Mean Absolute Error**		**Percent of Predictions Within 20%**		**Percent of Predictions Within 10%**	
ANN (IT)	0.937	0.963	16.4%	11.0%	84.5%	89.9%	65.2%	54.9%
GA (L-IT)	0.933	0.946	16.9%	11.0%	83.7%	91.7%	64.1%	53.1%
MVR (L-IT)	0.934	0.975	16.6%	9.7%	85.4%	94.1%	68.1%	62.9%

Table 6.2 Model summary for 2006-2007 SS2

	R^2		**Mean Absolute Error**		**Percent of Predictions Within 20%**		**Percent of Predictions Within 10%**	
ANN (ITV)	0.969	0.909	9.4%	9.4%	92.4%	88.7%	80.8%	72.6%
GA (L-ITV)	0.956	0.881	12.9%	16.5%	85.2%	80.6%	74.8%	69.4%
MVR (L-ITV)	0.957	0.869	12.7%	19.6%	87.6%	77.4%	77.2%	67.7%

The ANN, GA, and MVR modeling techniques produced comparable daily production predictions for SS1 with mean absolute error less than 17% and R^2 values around 0.930 for the training scenario; furthermore, the distribution of error between ANN, GA, and MVR methods was also comparable. The ANN modeling technique produced the best results for SS2 with a mean absolute error of 9.4% and an R^2 of 0.969

for the training scenario.

6.3 Recommendations for Future Studies

The initial investigations of this study on the stored energy and heat transfer produced results that could be more fully explored in the future by utilizing the artificial intelligence modeling methods used for daily production. Furthermore, genetic algorithms could be employed to improve the accuracy of heat and mass transfer and Mathoulakis's (1999) models by calculating convective and evaporative heat transfer coefficients using measured temperature differences and hourly production values. The applications of these models could be implemented in the following areas:

1. Optimize the coefficients used in the Mathioulakis et al. (1999) model by utilizing genetic algorithms
    a. Develop a method to predict temperature differences using weather data to avoid intensive data logging
    b. Analyze the effect of seasons on the coefficients for the Mathioulakis et al. (1999) model
    c. Evaluate the effect of stored energy and its role in night time production
2. Optimize the size of the distilland volume throughout different seasons to maximize daily production
3. Optimize the coefficients and exponents used to calculate the Nusselt number and the convective heat transfer coefficients by utilizing genetic algorithms
    a. Analyze the effect of seasons and distilland volume on the constants used in the heat transfer method for solar still modeling

4. Implement the use of the Schmidt number to model hourly production in terms of mass transfer instead of heat transfer

The 2011 study gathered a massive quantity of sub-hourly temperature and production data which could all be employed to complete the recommended studies above.

REFERENCES

Abu-Hijleh, B.A.K. (1996) Enhanced solar still performance using water film cooling of the glass cover. *Desalination*, 107, 235-244.

Adhikari, R.S., & Kumar, A. (1990). Estimation of mass transfer rates in solar stills. *International Journal of Energy Research*, 14, 737-744.

Akash, B.A., Mohsen, M.S., Nayfeh, W. (2000). Experimental study of the basin type solar still under local climate conditions. *Energy Conversion & Management*, 41, 883-890.

Al-Karaghouli, A.A., & Alnaser, W.E. (2004). Performances of single and double basin solar stills. *Applied Energy*, 78, 347-354.

Cabello, J.M., Cejudo, J.M., Luque, M., Ruiz, F., Deb, K., and Tewari, R. (2011). Optimization of the size of a solar thermal electricity plant by means of genetic algorithms. *Renewable Energy*, 36, 3146-3153.

Chen, Z., Ge, X., Bar, L., & Miao Y.X. (1984). Natural convection heat transfer across air layers at various angles of inclination. *Engineering Thermophysics*, 211-220.

Coello, C. A., Lamont, G. B., & Van Veldhuizen, D. A. (2007).*Evolutionary algorithms for solving multi-objective problems*. (pp. 21-57). New York: Springer.

Cooper, P.I. (1969). Digital simulation of transient solar still processes. *Solar Energy*, 12, 313-331.

Dhiman, N.K., & Tiwari, G.N. (1990). Effect of water flowing over the glass cover of a multi wick solar still. *Energy Converstion and Management,* 30(3), 245-250.

Dunkle, R.V. Solar water distillation, the roof type still and a multiple effect diffusion still. (1961). *International Developments in Heat Transfer*, ASME, Proc.

International Heat Transfer, Part V, University of Colorado, p. 895.

Dutt, D.K., Kumar, A., Anand, J.D., & G.N. Tiwari. (1989). Performance of a double basin solar still in the presence of dye. *Applied Energy*, 32, 207-223.

El-Sebaii, A.A., Yaghmour, S.J., Al-Hazmi, F.S., Faidah, A.S., Al-Marzouki, F.M., & Al-Ghamdi, A.A. (2009). Active single basin solar still with a sensible storage medium. *Desalination*, 249, 699-706

Esfahani, J.A., Rahbar, N., & Lavvaf, M. (2011). Utilization of thermoelectric cooling in a portable active solar still – an experimental study on winter days. *Desalination,* 269, 198-205.

Fath, H.E.S. (1998). Solar distillation: A promising alternative for water provision with free energy, simple technology and a clean environment. *Desalination*, 116, 45-56.

Fath, H.E.S., & Elsherbiny, S.M. (1993). Effect of adding a passive condenser on solar still performance. *Energy Conversion & Management*, 34, 63-72.

Fath, H.E.S., Elsherbiny, S.M., & Ghazy, A. (2004). A naturally circulated humidifying/dehumidifying solar still with a built-in passive condenser. *Desalination*, 169, 129-149.

Fishenden, M. (1957). *An introduction to heat transfer*. (pp. 89-108). Oxford: Clarendon Press.

Frick, B. Some new considerations about solar stills. (1970). *Proceedings of International Solar Energy Congress*, Melbourne, p. 395.

Foster, R. E., Eby, S., & Amost, W. (2005). Ten years of solar distillation application along the u.s.-mexico border. Paper presented at: *Solar World Congress:*

International Solar Energy Society, Orlando.

Garg, H.P., & Mann, H.S. (1976). Effect of climactic, operational and design parameters on the year round performance of single sloped and double sloped solar still under indian arid zone conditions. *Solar Energy*, 18, 169-164.

Goldberg, D.E. (1989). Genetic algorithm in search, optimization, and machine learning, Addison-Wesley, Reading, Mass.

Haykin, S. (1994). Neural Networks: A Comprehensive Foundation. Macmillan, New York, p. 842.

Holland, J. H. (1975). Adaptation in natural and artificial systems, University of Michigan Press, Ann Arbor, Mich.

Hongfei, Z., Zhang, X., Jing, Z., & Yuyuan, W. (2002). A group of improved heat and mass transfer correlations in solar stills. *Energy Conversion & Management*, 43, 2469-2478.

Jakob, M. (1949). *Heat transfer*. (pp. 534-542), New York: Wiley.

Kabeel, A.E., & El-Agouz, S.A. (2011). Review of researches and developments on solar stills. *Desalination*, 276, 1-12.

Kalogirou, S.A. (2001). Artificial neural networks in renewable energy system applications: A review. *Renewable & Sustainable Energy Reviews*, 5, 373-401.

Kalogirou, S.A. (2004). Optimization of solar systems using artificial neural networks and genetic algorithms. *Applied Energy*, 77, 383-405.

Kalogirou, S.A. (2005). Seawater desalination using renewable energy sources. *Progress in Energy and Combustion Science*, 31, 242-281.

Khalifa, A.J. (2011). On the effect of cover tilt angle of the simple solar still on its

productivity in different seasons and latitudes. *Energy Conversion and Management*, 52, 431-436.

Kumar, S., & Sinha, S. (1996). Transient model and comparative study of concentrator coupled regenerative solar still in forced circulation mode. *Energy Conversion & Management*, 37(5), 629-636.

Las Vegas Valley Water District. (2011). *Lvvwd water quality summary report*. Retrieved from http://www.lvvwd.com/assets/pdf/wq_summary_lvvwd.pdf

Lof, G., Eibling, J.A., & Blowemer, J.W. (1961). Energy balances in solar distillation. *Journal of the American Institute of Chemical Engineers*, 7, 641.

Loomans, M., & Visser, H. (2002). Application of the genetic algorithm for optimization of large solar hot water systems. *Solar Energy*, 72(5), 427-439.

Malik, M., & Tran V. (1972). A simplified mathematical model for predicting the nocturnal output of a solar still. *Solar Energy,* 14, 371-385.

Mathioulakis, E., Voropoulous, K., & Belessiotis, V. (1999). Modeling and prediction of long term performance of solar stills. *Desalination*, 122, 85-93.

MathWorks. (2011). *What is the genetic algorithm*? Retrieved from http://www.mathworks.com/help/toolbox/gads/f6636.html

McAdams, W. H. (1954). *Heat transmission*. (pp. 165-183). New York: McGraw-Hill.

Mitchell, M. (1995). Genetic algorithms: An overview. *Complexity*, 1(1), 31-39.

Morcos, V.H. (1994). Some experimental and theoretical studies of a single basin solar still. *Renewable Energy,* 4(4), 401-407.

McCracken, H. (1985).*Understanding solar stills*. Virginia: PACT Pub. Retrieved from http://pdf.usaid.gov/pdf_docs/PNABC961.pdf

Mohamad, M.A., Soliman, S.H., Abdel-Salam, M.S., & Hussein, H.M.S. (1995). Experimental and financial investigation of asymmetrical solar stills with different insulation. *Applied Energy*, 52, 265-271.

Murugavel, K.K., Sivakumar, S., Ahamed, J.R., Chockalingam, K., & Srithar, K. (2010). Single basin double slope solar still with minimum basin depth and energy storing materials. *Applied Energy*, 87, 514-523.

Murugavel, K.K., & Srithar, K. (2011). Performance study on basin type double slope solar still with different wick materials and minimum mass of water. *Renewable Energy,* 36, 612-620.

Nafey, A.S., Abdelkader, M., Abdelmotalip, A., & Mabrouk, A.A. (2002). Enhancement of solar still productivity using floating perforated black plate. *Energy Conversion & Management*, 43, 937-946.

Nehdi, M., El Chabib, H., & Said, A. (2007). Proposed shear design equations for frp-reinforced concrete beams based on genetic algorithms approach. *Journal of Materials in Civil Engineering*, 19, 1033-1042.

NREL. (2011). *University of nevada, las vegas*. Retrieved from http://www.nrel.gov/midc/unlv/

Pazos, A. B., Sierra, A. P., & Buceta, W. B. (2009). *Advancing artificial intelligence through biological process applications*. (pp. 231-249). Hershey: Information Science Reference.

Rubio, E., Porta, M.A., & Fernandez, J.L. (2000). Cavity geometry influence on mass flow rate for single and double slope solar stills. *Applied Thermal Engineering*, 20, 1105-1111.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, D. E. Rumelhart and J. L. McClelland, Eds. MIT Press Computational Models of Cognition and Perception Series. MIT Press, Cambridge, MA, 318-362.

Saleh, M., Oliva R., Kampmann, C.E., & Davidsen, P.I. (2009). A comprehensive analytical approach for policy analysis of system dynamics models. *European Journal of Operational Research*, 203, 673-683.

Sakthivel, M., Shanmugasundaram, S., & Alwarsamy, T. (2010). An experimental study on a regenerative solar still with energy storage medium – jute cloth. *Desalination*, 264, 24-31.

Sartori, E. (1987) On the nocturnal production of a conventional solar still using solar pre-heated water. *Advances in Solar Energy Technology*, 2, 1427.

Setoodeh, N., Rahimi, R., & Ameri, A. (2011). Modeling and determination of heat transfer coefficients in a basin solar still using CFD. *Desalination*, 268, 103-110.

Sodha, M.S., Kumar, A., Tiwari, G.N., & Tyagi, R.C. (1981) Simple multiple wick solar still: Analysis and performance. *Solar Energy*, 26, 127-131.

Sodha, M.S., Nayak, J.K., Tiwari, G.N., Kumar, A. (1980). Double basin solar still. *Energy Conversion & Management*, 20, 23-32.

Solaqua. (2011). Retrieved from http://www.solaqua.com/

State of Nevada Department of Conservation and Natural Resources, Division of Water Resources. (2009). *Nevada water facts: Water Resource Issues in Nevada*. Retrieved from http://water.nv.gov/WaterPlanning/wat-fact/issues.cfm.

Sterman, J. D. (2000). *Business dynamics, systems thinking and modeling for a complex world*. Irwin McGraw-Hill, Boston.

Tanaka, H. (2011) Tilted wick solar still with flat plate bottom reflector. *Desalination*, 273, 405-413.

Tiwari, A.K., & Tiwari, G.N. (2006). Effect of water depths on heat and mass transfer in a passive solar still: In summer climactic condition. *Desalination*. 195, 78-94.

Tiwari, G.N., Gupta, S.P., & Lawrence, S.A. (1989). Transient analysis of solar still in the presence of dye. *Energy Conversion & Management*, 29, 59-62.

Tiwari, G.N., & Rao, V. (1984). Transient performance of single basin solar still with water flowing over the glass cover. *Desalination*, 49, 231-241.

Tiwari, G.N., Singh H.N., and Tripathi, R. (2003). Present status of solar distillation. *Solar Energy*, 75, 367-373.

Toure, S., & Meukam, P. (1997). A numerical model and experimental investigation for a solar still in climactic conditions in Abidjan (Côte d'Ivoire*). Renewable Energy*, 11, 319-330.

Varun, S. (2010). Thermal performance optimization of a flat plate solar air heater using genetic algorithm. *Applied Energy*, 87, 1793-1799.

Venkatesh, N.H (2007). Performance evaluation of single and double-basin solar stills in Las Vegas, Nevada. M.S.E. Dissertation, University of Nevada, Las Vegas, United States.

Weather Underground. (2011).*Almanac for las vegas, nv*. Retrieved from http://www.wunderground.com/history/airport/KLAS/2011/10/9/DailyHistory.html?req_city=Las Vegas&req_state=NV&req_statename=Nevada

World Health Organization. (2005). *Water for life: making it happen*. Retrieved from http://www.who.int/water_sanitation_health/monitoring/jmp2005/en/index.html

VITA

Graduate College
University of Nevada, Las Vegas

Noe I. Santos

Degrees:

Bachelor of Science in Engineering, Civil Engineering, 2009
University of Nevada, Las Vegas

Publications:

Santos, N.I., Said, A.M., James, D.E., & Venkatesh, N.H. (2011). Modeling solar still production using local weather data and artificial neural networks. *Renewable Energy*, 40(1), 71-79.

Santos, N.I., Said, A.M., James, D.E., & Venkatesh, N.H. (2011). Comparing multivariate regression and artificial neural networks to model solar still production. Proceedings from American Solar Energy Society: *Solar 2011*, Raleigh, NC.

Santos, N.I., Said, A.M., James, D.E., & Venkatesh, N.H. (2012). Modeling solar still production with genetic algorithms and multivariate regression. *Renewable Energy,* [In Progress].

Santos, N.I., James, D.E., Venkatesh, N.H., & Said, A.M. (2012). Pursuing solar distillation as a reliable water resource. *Desalination,* [In Progress].

Santos, N.I., James, D.E., and Said, A.M. (2012). Evaluating heat and mass transfer relationships to predict solar still production. *Renewable Energy*, [In Progress].

Thesis Title: Modeling Passive Solar Still Production in Las Vegas, Nevada

Thesis Examination Committee:

Chairperson, Dr. Aly Said, Ph. D., P.E.
Committee Member, Dr. Dave James, Ph. D., P.E.
Committee Member, Dr. Sajjad Ahmad, Ph. D., P.E.
Graduate Faculty Representative, Dr. Robert Boehm, Ph. D., P.E.