Degradation Linearity Determination and Temperature of

Soiled Photovoltaic Modules in the Field

by

Adit Patankar

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2017 by the
Graduate Supervisory Committee:

Govindasamy Tamizhmani, Co-chair
Liping Wang, Co-chair
Patrick Phelan

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

Photovoltaic modules degrade in the field. This thesis aims to answer two questions: 1. Do photovoltaic modules degrade linearly or not? 2. Do soiled modules operate at lower temperatures than clean modules? Answers to these questions are provided in part 1 and part 2 of this thesis respectively.

Part 1: Linearity determination in degradation: The electricity output from PV power plants degrades every year. Generally, a system's life is considered to last for 20-25 years and rate of degradation is commonly assumed as 1% per year. PV degradation can be found out using Performance Ratio (PR), Performance Index (PI) and raw kWh output. The rate of degradation is considered linear for simplicity of calculations. In this thesis, statistical methods are used to check whether systems in Arizona are degrading linearly or not. Time series modeling such as Winters' method and ARIMA are used to model the data. Winters' method and Seasonal ARIMA consider the seasonality component and perform well for small data sets of about 10 years. Rate of degradation is found out as linear for all the evaluated systems.

Part 2: Temperature analysis of clean and soiled modules: Soiling and temperature are important parameters in performance of PV modules. In this paper, an analysis is carried out on a soiling station located in Mesa, Arizona. The soiling station consists of 10 different c-Si coupons with tilt angles varying from 0° to 45° with the difference of 5°. These coupons are cut in half, one is cleaned periodically and the other is remained soiled naturally. The analysis involves data worth for 19 months. 6 dry spells in all four seasons within 19 months were analyzed. The temperature difference between a clean module and

a soiled module (ΔT) is compared with the soiling loss factor (SLF). The analysis concludes stating in which season a soiled module is hotter or cooler than a clean module.

To,
My parents, Satish Patankar and Madhuri Patankar; grandparents Bhalchandra Joshi and Anjali Joshi; and dear friends for always being there.

ACKNOWLEDGEMENTS

First, I would like to thank Dr. Govindasamy (Mani) Tamizhmani for accepting me in the lab and supporting me throughout my work at ASU-PRL. Working with ASU-PRL has provided me a lot of theoretical as well as hands on experience on various indoor and outdoor testing techniques. I would like to thank Dr. Kuitche for his constant help with statistical knowledge. I would like to thank Sai Tatapudi, technical manager of ASU-PRL for always being helpful and guiding me. I am grateful to Chris Raupp for his work in field degradation of modules and data availability for the same. I want to thank Prasanna Sundarajan for explaining statistical modeling of PV degradation. I especially want to thank Aravind Srinivasan for his friendly guidance in statistical concepts. I also thank all my fellow lab mates and former lab students Srinivas Mantha, Darshan Choudhary for their help. I thank all my friends and family for their support and love. I want to thank my thesis defense committee, Dr. Wang and Dr. Phelan for their support and help.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

PART 1: DEGRADATION LINEARITY DETERMINATION

1. INTRODUCTION

1.1.1 Background:

Commercial development of PV technology started in second half of the 20th century. Initially it was used in space applications. Later, from 1970s, PV was used to power residences and then it started growing rapidly on commercial scale. Now PV power plants are built at the capacity of over 500 MW. Performance of PV power plants depend upon many factors such as system efficiency, maintenance, reliability, and weather. Performance of a PV power plant degrades year by year in terms of output due to many reasons such as browning of encapsulant, leakage of current, mechanical damages and so on. The rate of degradation is an important factor that an owner or a developer of PV plant should know. A customer would purchase PV system or use PV electricity only if he knows how much it is worth. The degradation rate is helpful to predict or forecast performance. Also, it's a crucial parameter to perform lifecycle analysis and payback period.

There are two methods of calculating degradation rate of a power plant. One is to calculate I/V data of every module from time to time and then calculate the degradation rate using that data. The second method is to calculate rate of degradation using kWh output data from inverter. Performance Ratio (PR) is one common measure to calculate degradation. It's a ratio of output energy and expected input energy [11]. This calculation can be done for yearly data and to find out how much the PR has degraded. Another, a better measure of calculating degradation is Performance Index (PI). It considers system losses in the input energy and therefore the PI is slightly higher than the PR [11]. A simpler way to calculate the degradation rate is just to compare yearly kWh data and find the slope of degradation in output. Calculating degradation is important, but, finding nature of degradation is vital. It's been assumed that the degradation of PV systems is linear and is close to 1% per year for simplicity of calculation. However, a statistical proof whether it is really linear or non-linear is useful, because if it's non-linear, the assumption would be wrong and then the forecasts would be wrong. ASU-PRL have done extensive research on the degradation of PV power plants. Chris Raupp et al [11] used PR, PI and raw kWh data to find out degradation rates of PV systems located in the Arizona State University, Tempe campus. He used box plots on the median values of monthly data to find out if the degradation trend is linear or non-linear. Chris et al [11] stated in his thesis that, in hot and dry climatic condition of Arizona, PV systems degrade in output at higher rate in first three years and at a lower rate after three years. The rate of degradation could be logarithmic in nature. Prasanna Sundarajan et al [12] used Holt Winters' method to determine whether the systems degrade linearly or not. As mentioned earlier, the HoltWinters' method decomposes the data into three parts, trend, seasonality and level. In order to determine the nature of degradation, trend component was considered for analysis and seasonality was ignored. Ultimately, the analysis is a regression analysis on trend data.

1.1.2 Scope of work:

A basic method to determine whether the degradation is linear or not, is to make a regression analysis on the output data. However, only regression is not enough, because the output from a PV system has components of seasonality and trend. Seasonality is a phenomenon which shows that the output in summer is high when the angle of incidence is low and in winter, output is low when the angle of incidence is high. Time series is a specialized statistical method which can consider the effect of seasonality. There are various methods in time series such as moving average, exponential smoothing and their variations. A third order exponential smoothing is called "Holt Winters'" method which decomposes the data in three parts which are Level, Trend and Seasonality. An advanced method of moving average is called autoregressive method and their combination is called Autoregressive Moving Average (ARMA) and even more advanced technique is called Autoregressive Integrated Moving Average (ARIMA) which also considers seasonality effect. In this research, six PV systems located in hot dry climate of Arizona are evaluated and modeled using Holt Winters' method and ARIMA method. Difference between these two models is that the time series in Holt Winters' model is not stationary where ARIMA makes a time series stationary to fit the data and model it. ARIMA is an advanced method of Holt-Winters' method. In both methods, weights are used which have more importance to recent past data and less importance to older data. Both methods approximate linear models. Hence, in this research, PV systems are evaluated using these two methods and check whether they degrade linearly or not.

Seasonal ARIMA is an advanced method which makes a time series stationary meaning that the mean and variance is constant. The model has two integrated components, seasonal and non-seasonal. In this analysis, seasonal component gives linear correlation between particular months of consecutive years; and the non-seasonal component gives linear correlation between consecutive months of a year. Although systems degrade linearly, different systems have different correlations between months and years. For some systems, consecutive two months might be correlated, for some systems, consecutive three months might be correlated. A clear comparison between Winters' method and Seasonal ARIMA is that Winters' method could only tell you whether the system degradation is linear or not but seasonal ARIMA could tell you about linearity, model the data, draw correlations between years and months.

2. LITERATURE REVIEW

1.2.1 Degradation of PV modules

Reliability is probability of a product or a system to perform their function under certain conditions throughout their life. A PV system is said to be reliable if it performs above 80% of its rated power after 20 years of its life. Degradation rates depend heavily on climatic conditions. Regions with humid and windy weather experience more degradation. Regions with cold and snowy climate face lowest rates of degradation. On the other hand, degradation is observed high in hot and dry desert climates. In northern United States, degradation rate is found as low as 0.2% per year. In deserts like Arizona, the rate of degradation is observed as high as 1% per year. Determination of degradation rate is important for site assessment and lifetime prediction of energy production. Also, it is important in calculation of payback period. Degradation due to different factors can be different in nature either linear or non-linear. Degradation in different PV technologies is different in nature.

1.2.2 Statistical methods to determine linearity in degradation rate

Data consists of two components, signal and noise [12]. For a prediction or modeling of data, noise is removed and signal is predicted. Data smoothing is a statistical method which approximates a function of a signal and maintains it, removing or separating the noise from main signal. A smoother is a function which modifies signal which makes higher values reduce and points adjacent to these values are increased, to achieve a local mean . Seasonality in PV data is taken into account and worked upon using smoothers in Winters' method. Winters' method separates out seasonality component and predicts how much it deviates from local mean due to seasonality effect. ARIMA makes the time series stationary and then uses smoothers such as moving average and auto-regression. Seasonal ARIMA uses periodic smoothing. It smooths periodic data and modifies it. ARIMA estimates parameters in a linear model. The significance level of these parameters explains whether the model fits the data adequately or not.

3. METHODOLOGY

1.3.1 Preprocessing the data:

- Hourly raw kWh data of ASU systems was taken from previous research of ASU-PRL. Data for high sun hours (9:00 am – 3:00 pm) is selected.
- This data was converted to daily data using the pivot table tool in Microsoft Excel.
- The daily data was converted to median monthly data by taking median values for every month. Hence, 12 data points per year are obtained.
- Thus, a time series of monthly kWh data was created. The ASU systems have data available for about 6 years. Data available for Tempe warehouse is for 10 years.

1.3.2 Time series plot:

Following image shows time series plot of monthly median kWh data of PV system built on Hayden Library building.

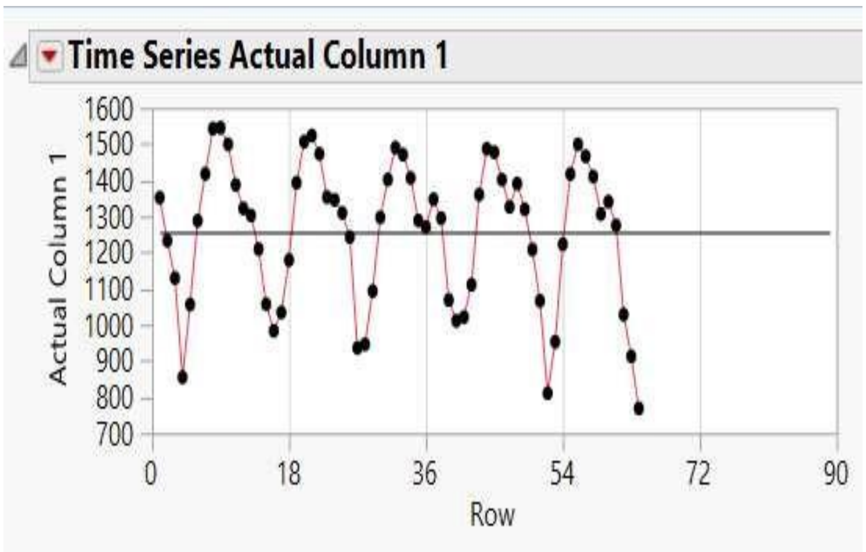


Figure 1.1: Time series plot of monthly median kWh

In the above plot, it is seen that kWh output is at its peak in summer months May and June. It is lowest in winter months December and January.

1.3.3 Time series methods:

Winters' method (Additive): Holt-Winters' method commonly known as Winters' method is a third order exponential smoothing method. It separates out three components from the data which are Level, Trend and Seasonality. Level is a local mean or "level" of data generating process at this time [5]. Trend component gives current trend of the data, or change in data at current time. Seasonality component estimates deviation from mean due to

seasonality effect.

Out of these three components, trend is useful to determine the degradation rate in the system output. A previous research by Prasanna Sundarajan et al [12] in ASU-PRL developed a method for this. It consisted of data filtration of trend data and then regression analysis on the filtered trend data. The degradation rate was found out using the Y-intercept of regression analysis. Determination whether the degradation rate is linear or not was worked out using hypothesis making. A null hypothesis was made that said the rate of degradation is not linear. The hypothesis was accepted or rejected based on the regression parameters. This method is applied in this research to the PV systems built in ASU campus.

JMP is a statistical tool which provides advanced time series modeling techniques. It facilitates to use Winters' additive method. Following image shows JMP output of Winters' method. It says the model is stable and not invertible. A model is invertible when past values can be predicted using current values. This phenomenon is crucial for forecasting.

Goal of this research is to model the existing data and find out whether the degradation rate is linear or not. Focus is not on forecasting. Therefore, this model of Winters' method works just fine.

Model: Winters Method (Additive)

Model Summary

DF	48	Stable	Yes
Sum of Squared Errors	175137.868	Invertible	No
Variance Estimate	3648.70558		
Standard Deviation	60.4045162		
Akaike's 'A' Information Criterion	582.494448		
Schwarz's Bayesian Criterion	588.289925		
RSquare	0.88272622		
RSquare Adj	0.87783981		
MAPE	4.62394946		
MAE	51.0375253		
-2LogLikelihood	576.494448		

Figure 1.2: JMP output window of Winters' method

A new approach using a different time series method is involved in this thesis. Autoregressive Integrated Moving Average (ARIMA) is an advanced version of moving average. The moving average technique uses averages of current value with past values with certain weights associated with them. Autoregressive component of the method consists of exponentially decreasing weights associated with the values in the same time series. Hence the name "autoregressive". Like Holt-Winters' method, ARIMA also deals with seasonality component. ARIMA has a seasonal component and a non-seasonal component.

1.3.4 Important facts about ARIMA: ARIMA is abbreviation for Autoregressive Integrated Moving Average. Both autoregressive (AR) and moving average (MA) terms estimate a linear model for a time series. ARIMA as a combination of both AR and MA terms also therefore estimates a linear model because one term is added to the another. Hence, the linearity of model is not altered. ARIMA is linear in its parameters.

1.3.5 Auto Correlation Function (ACF): In time series, current values of observation are dependent on previous values of observation. Hence there's a possible relation between current observation and previous observations.

If there's a time series $S_t = \{S_t, S_{t-1}, S_{t-2}, S_{t-3}, \ldots, S_{t-s}\}$. The values S_{t-1}, S_{t-2} and so on represent previous observations are also called "lags". They also represent different individual time series. To make it clear, S_{t-1} represents time series with lag 1. It is known that there's relation between S_t and its previous terms. Or, there's relation between S_t and its lags.

Autocovariance is covariance between two different time series.

$\gamma_s = E[(S_t - \mu_t)(St_{-s} - \mu_{t-s})]$ is autocovariance between S_t and S_{t-s}. γ function has a unit [2]. However, a unit-less function is desired. Hence, γ_s is divided by γ_0 where γ_0 is variance of S_t.

$\tau_s = \frac{\gamma s}{\gamma 0}$ Where τ_s autocorrelation function (ACF). It lies between -1 and +1. The term "auto" comes because both the time series S_t and S_{t-s} are part of same time series S_t. ACF plot suggests us MA process [2].

Sample ACF plot:

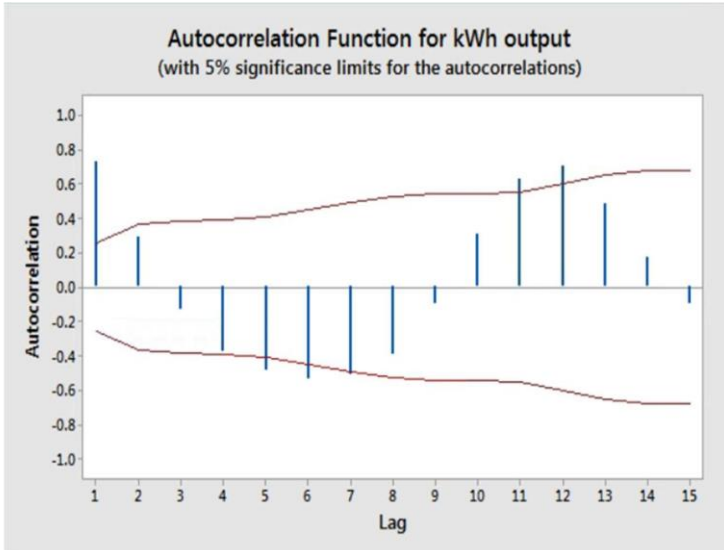


Figure 1.3: Sample ACF plot

In the above ACF plot on Hayden library kWh data, it is seen that there is one value which is cutting off the blue lines of "significance" or "limits". This value is called significant value. Here, the y – axis represents ACF and x – axis represents lag. This means that first lag has significant ACF value. It means, the series y_t is correlated with the series y_{t-1}. After lag 1, the ACF behaves a sinusoidal and exponentially decaying pattern. It does not conclude yet anything. Hence, to identify a process, partial autocorrelation function is used.

1.3.6 Partial Autocorrelation Function (PACF): If there are two series y_t and y_{t-s}, it is observed that there are intermediate series of y_{t-1}, y_{t-2},…, y_{t-s+1}. It is desired to take out the effect of these intermediate series on the correlation of y_t and y_{t-s} .

For example, there are two series y_t and y_{t-2}. The goal is to take out the effect of y_{t-1} on the correlation of y_t and y_{t-2}.

$$\text{PACF} = \frac{\text{Covariance }(yt, yt-2|yt-1)}{\text{Square root of}[\text{variance}(yt|yt-1)\text{variance}(yt-2|yt-1)]}$$

Sample PACF plot

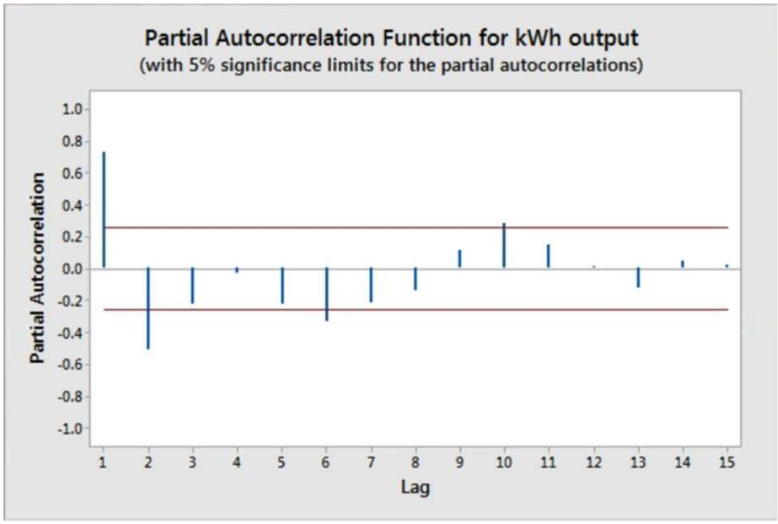


Figure 1.4: Sample PACF plot

Above figure shows PACF plot for the kWh data of PV system on Hayden Library. In the PACF plot, it is observed that first two lags have significant PACF values and further lags have values below significance limits. Which means, the series y_t is correlated independently with y_{t-1} and y_{t-2}, with removal of effect of intermediate series. This suggests for an AR (2) process.

1.3.7 Application of ARIMA: Now these concepts of ACF and PACF are applied to determine an ARIMA model. However, JMP makes this task easier. JMP has a tool in time series modeling named "ARIMA model group". In this tool, a time series column is input on which analysis is to be carried out. In ARIMA model group, input is maximum limits of AR, difference and MA terms for seasonal and non-seasonal parts. This, after running gives seasonal ARIMA models with all combinations of AR and MA terms. Following image shows input of parameters.

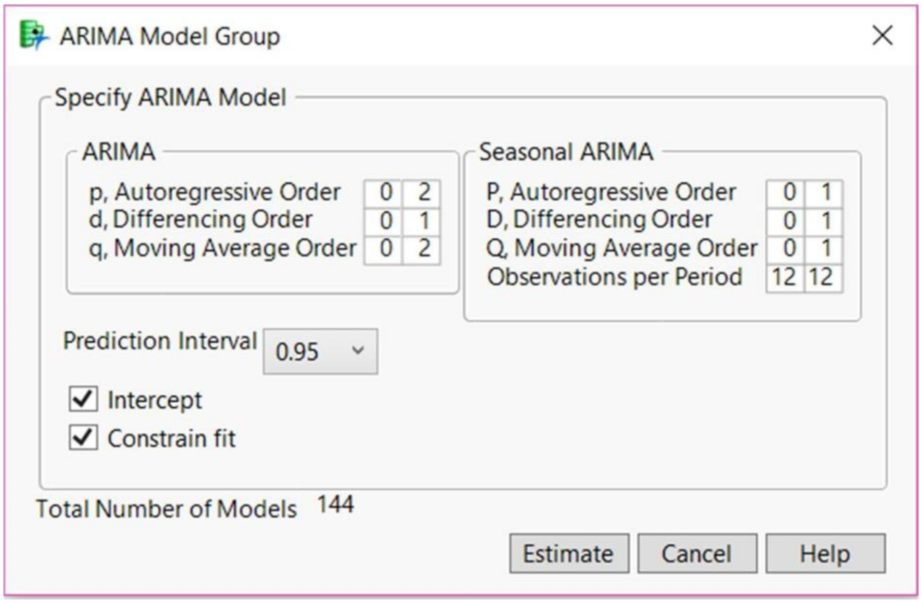


Figure 1.5: JMP input window for ARIMA model group

After clicking estimate, it gives a list of ARIMA models with all combinations of AR and MA terms and seasonal AR and MA terms.

Model Comparison

Report	Graph	Model	DF	Variance	AIC	SBC	RSquare	-2LogLH	Weights	.2 .4 .6 .8	MAPE	MAE
		Seasonal ARIMA(0, 1, 2)(0, 1, 1)12	47	2786.3385	577.73873	585.46603	0.893	569.73873	0.171454		4.326479	47.757226
		Seasonal ARIMA(2, 1, 2)(0, 1, 1)12	45	2568.4353	578.27291	589.86386	0.898	566.27291	0.131266		4.286642	47.046064
		Seasonal ARIMA(1, 1, 1)(0, 1, 1)12	47	2834.9006	578.49620	586.22351	0.892	570.4962	0.117399		4.357141	48.368116
		Seasonal ARIMA(0, 1, 2)(1, 1, 1)12	46	2987.3321	579.14427	588.80340	0.894	569.14427	0.084906		4.319548	47.832085
		Seasonal ARIMA(2, 1, 1)(0, 1, 1)12	46	2828.1046	579.56283	589.22196	0.893	569.56283	0.068873		4.350125	47.954083
		Seasonal ARIMA(1, 1, 2)(0, 1, 1)12	46	2838.4617	579.67164	589.33077	0.894	569.67164	0.065226		4.323266	47.645304
		Seasonal ARIMA(1, 1, 1)(1, 1, 1)12	46	2996.7079	580.22096	589.88009	0.893	570.22096	0.049561		4.370520	48.642807
		Seasonal ARIMA(0, 1, 1)(0, 1, 1)12	48	2981.8943	580.70744	586.50292	0.886	574.70744	0.038860		4.554490	50.810113
		Seasonal ARIMA(2, 1, 1)(1, 1, 1)12	45	3044.5056	580.82850	592.41946	0.894	568.8285	0.036577		4.332894	47.847342
		Seasonal ARIMA(2, 0, 1)(0, 1, 1)12	47	2412.1538	581.81913	591.57535	0.912	571.81913	0.022290		3.968203	43.680150
		Seasonal ARIMA(2, 0, 1)(1, 1, 1)12	46	2632.465	582.02280	593.73026	0.913	570.0228	0.020131		3.975232	43.798114
		Seasonal ARIMA(2, 0, 2)(0, 1, 1)12	46	2562.8595	582.23773	593.94520	0.913	570.23773	0.018080		3.890272	42.836982
		Seasonal ARIMA(0, 0, 1)(0, 1, 1)12	49	2753.1334	582.50935	588.36308	0.903	576.50935	0.015784		4.111506	45.212838
		Seasonal ARIMA(0, 1, 1)(1, 1, 1)12	47	3092.9156	582.65899	590.38630	0.886	574.65899	0.014646		4.559396	50.893816
		Seasonal ARIMA(2, 1, 2)(1, 1, 1)12	44	3109.9493	582.98908	596.51186	0.894	568.98908	0.012418		4.298384	47.531946
		Seasonal ARIMA(2, 1, 2)(1, 1, 0)12	45	4123.5912	583.28957	594.88052	0.887	571.28957	0.010686		4.505747	49.346343
		Seasonal ARIMA(1, 0, 0)(0, 1, 1)12	49	2798.7809	583.32895	589.18268	0.902	577.32895	0.010477		4.149136	45.894846
		Seasonal ARIMA(0, 1, 2)(1, 1, 0)12	47	4558.5238	583.62453	591.35184	0.881	575.62453	0.009038		4.661167	51.623975
		Seasonal ARIMA(0, 0, 1)(1, 1, 1)12	48	2940.0983	583.78253	591.58750	0.904	575.78253	0.008351		4.093031	45.149087
		Seasonal ARIMA(1, 0, 2)(0, 1, 1)12	47	2552.9929	583.98350	593.73972	0.909	573.9835	0.007553		4.064699	44.732979
		Seasonal ARIMA(2, 1, 2)(0, 1, 0)12	46	4340.2116	583.99276	593.65189	0.881	573.99276	0.007518		4.513101	49.845952
		Seasonal ARIMA(2, 0, 0)(0, 1, 1)12	48	2798.0786	584.28275	592.08773	0.904	576.28275	0.006503		4.131984	45.418328
		Seasonal ARIMA(1, 1, 1)(1, 1, 0)12	47	4599.1645	584.35614	592.08344	0.879	576.35614	0.006269		4.709706	52.453912
		Seasonal ARIMA(1, 0, 2)(1, 1, 1)12	46	2781.1232	584.43079	596.13825	0.910	572.43079	0.006039		4.045892	44.601088
		Seasonal ARIMA(1, 0, 1)(0, 1, 1)12	48	2805.7571	584.43124	592.23622	0.903	576.43124	0.006038		4.105131	45.073003
		Seasonal ARIMA(0, 0, 2)(0, 1, 1)12	48	2806.463	584.43985	592.24483	0.903	576.43985	0.006012		4.108433	45.116556
		Seasonal ARIMA(1, 0, 0)(1, 1, 1)12	48	2955.7938	584.96642	592.77140	0.902	576.96642	0.004620		4.149495	45.992639
		Seasonal ARIMA(2, 1, 1)(1, 1, 0)12	46	4631.3754	585.36399	595.02312	0.881	575.36399	0.003787		4.650413	51.358923
		Seasonal ARIMA(2, 0, 0)(1, 1, 1)12	47	3004.4144	585.36478	595.12099	0.904	575.36478	0.003786		4.115836	45.348938

Figure 1.6: JMP output of Seasonal ARIMA model comparison

Out of these models, best model is selected based on R-square value and simplicity of the model. Seasonal ARIMA (0,0,0) (0,1,1) 12 is selected because it's the simplest model with good value of R-square.

Model: Seasonal ARIMA(0, 0, 0)(0, 1, 1)12

Model Summary

DF	50	Stable	Yes
Sum of Squared Errors	149263.165	Invertible	Yes
Variance Estimate	2985.26331		
Standard Deviation	54.6375632		
Akaike's 'A' Information Criterion	585.647643		
Schwarz's Bayesian Criterion	589.55013		
RSquare	0.89550373		
RSquare Adj	0.8934138		
MAPE	4.34056402		
MAE	48.3196907		
-2LogLikelihood	581.647643		

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
MA2,12	2	12	1.00000	0.305429	3.27	0.0019*
Intercept	1	0	-17.69085	4.374456	-4.04	0.0002*

Constant Estimate	Mu
-17.690853	-17.690853

Figure 1.7: JMP output of selected model of seasonal ARIMA

A time series is stable when the residuals are normal and random, variance is constant. A model is stable means that the data fits the model.

Invertibility: A model is invertible if the previous values can be predicted using current values. Invertibility is important for forecasting. Here, the focus is not on forecasting but an invertible model is always a better fit.

1.3.8 Parameter analysis: The model only consists of one term: seasonal moving average. The estimated weight of this MA term is 1. The significance of this value is shown in the same table. The "prob" value is shown as 0.0019. Since, the probability of the parameter is less than the threshold value 0.05 (0.0019<0.05), it's chance of being greater than the statistical “t” ratio 3.27 is almost negligible. Therefore, the estimated weight constant “1” is a good fit and model fits a linear pattern. This means the degradation is linear.

Calculation of degradation rate:

Since, it is proved that the model fits adequately and the degradation rate is linear, the rate of degradation is calculated using the slope of modelled kWh data.

Degradation rate is calculated for four selected months in summer which are April, May, June and July. Reason behind this is that insolation in rest of the months is variable.

The slope is calculated using Excel function “slope”, where y-values are degradation values and x-values are in the year numbers.

The degradation rates and slopes are showed in Chapter 4.

4. RESULTS AND DISCUSSION

1.4.1 Time series plot:

Following plot is a time series plot of kWh output data from inverter of a PV system built on Hayden Library in ASU Tempe campus.

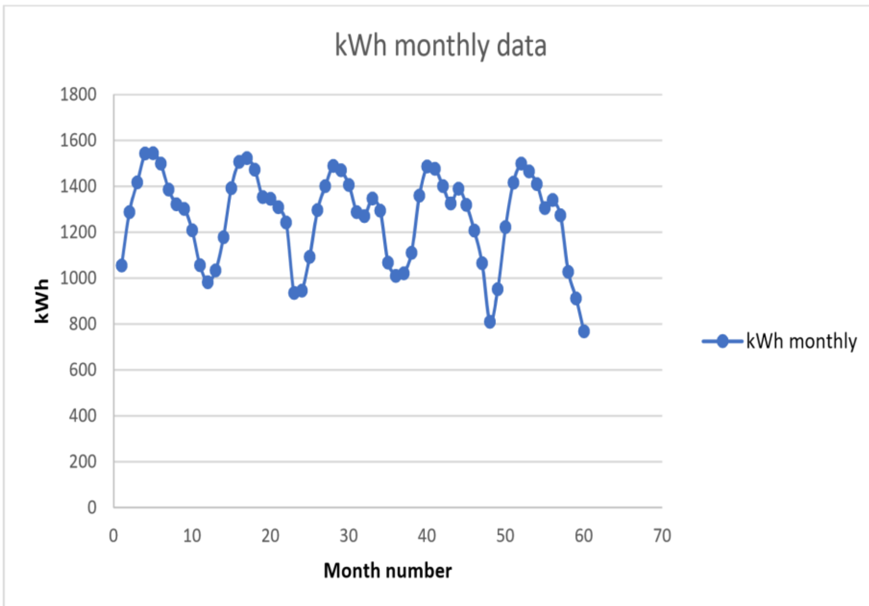


Figure 1.8: Time series plot of kWh monthly data

In this plot, it is observed that kWh output on Y axis and month number on the X axis. There's a seasonal pattern seen in this plot. In winter months, the kWh output is low and in summer months, the kWh output is high. The reason for this is that the angle of incidence in winter months is high and therefore direct normal irradiance (DNI) is low in winter. However, the angle of incidence in summer months on the on the PV modules in winter months is low and DNI is high.

Another thing which is observed is that there is some trend in reducing the power output, which is same as the trend in degradation of output. The kWh output data is separated by equal number of intervals in time. One data point represents one month. Thus, this is a time series data. KWh output data is available for 5 years and the aim is to check whether this degradation of output is linear in nature or not. Another aim is to compare results obtained from Winters' method and ARIMA method.

1.4.2 Simple regression of kWh output:

ASU-PRL have done research in finding out degradation rates of power output in hot-dry climate of Arizona. In that research, Chris et al [11] used different irradiance models, temperature models and Performance Index (PI) ratio. PI is the best method of calculating degradation rates because it is most accurate. Now, the task is to check that this PV output from the systems is degrading linearly or not.

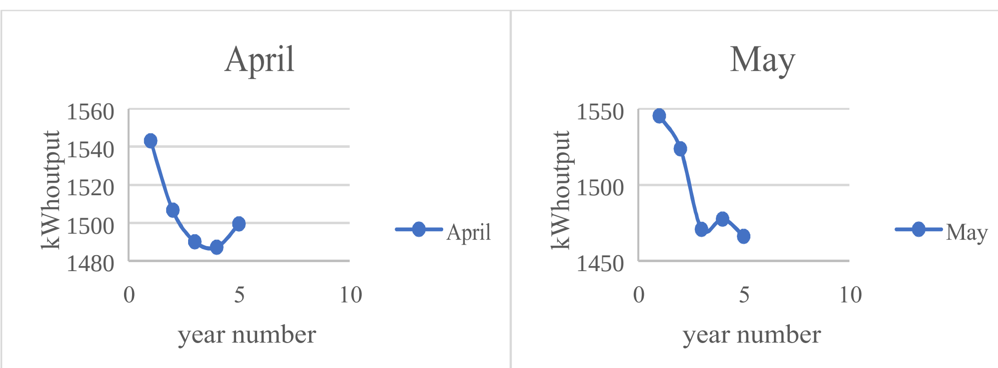


Figure 1.9: kWh degradation for April and May

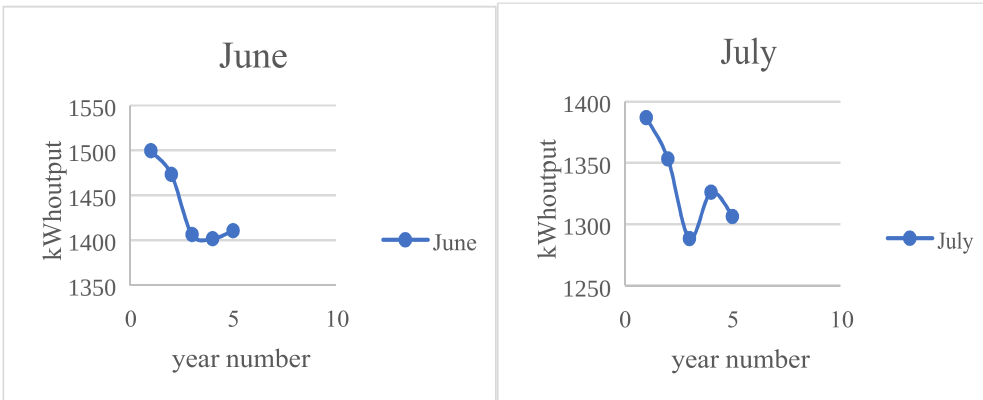


Figure 1.10: kWh degradation for June and July

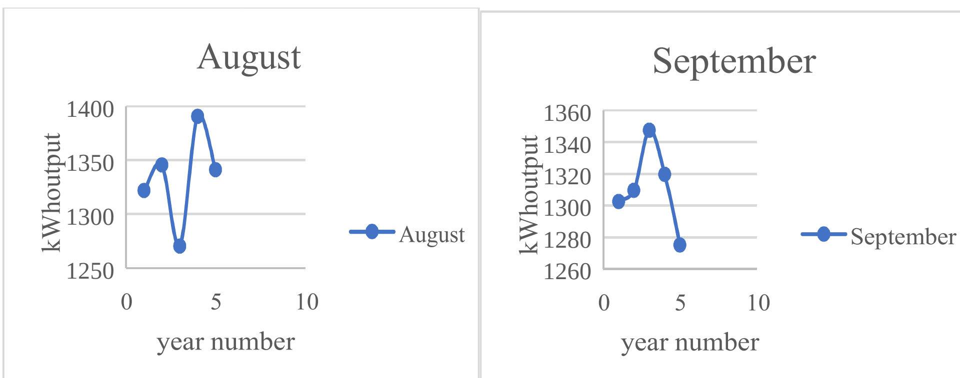


Figure 1.11: kWh degradation for August and September

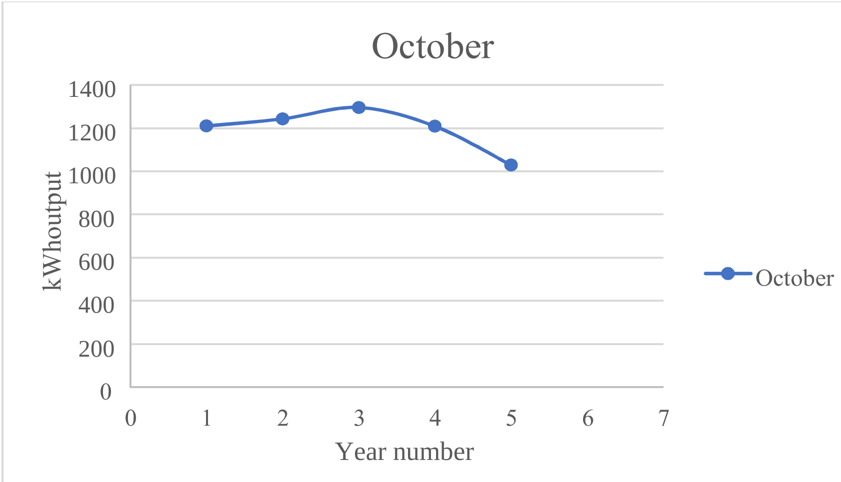


Figure 1.12: kWh degradation for October

The above plots show kWh output over five years from 2011 to 2015 for seven selected months which are April, May, June, July, August, September and October. These are the major sunny months for which kWh output is high. It is observed that for months May, June and July, the plots are almost linear. For other months, the plots do not show a linear trend. Hence, regression is performed on plots of May, June and July to check if the regression model fits.

1.4.3 Regression Analysis:

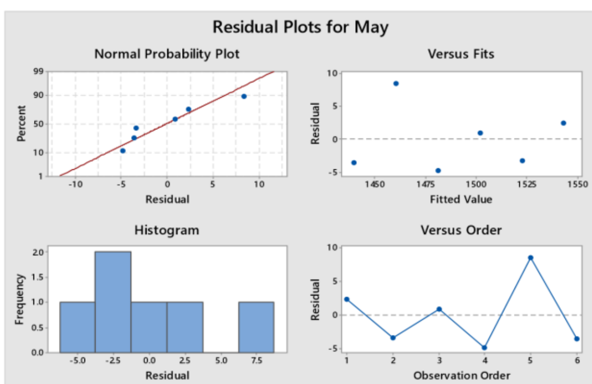


Figure 1.13: Residual plots for May

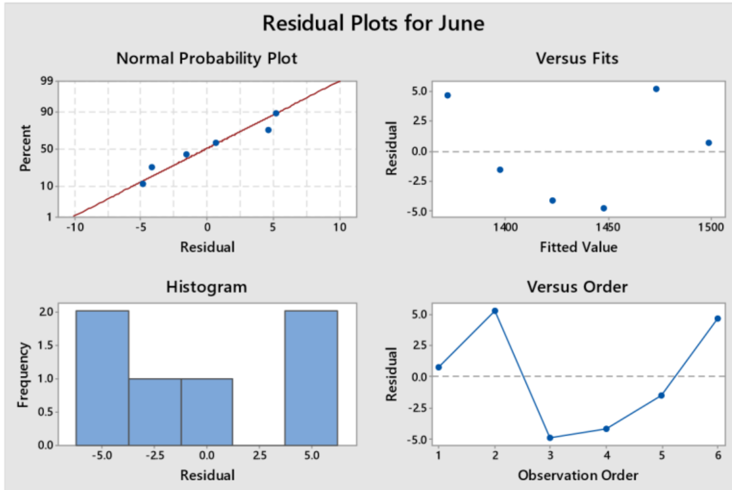


Figure 1.14: Residual plots for June

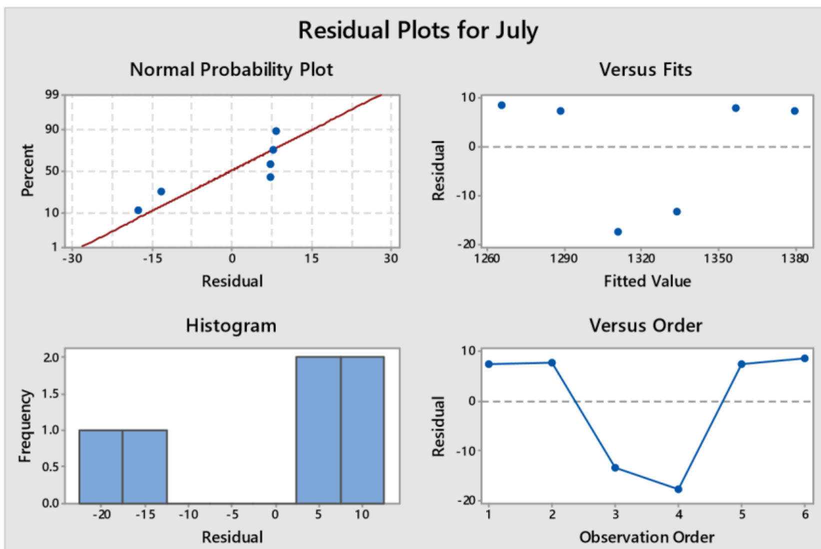


Figure 1.15: Residual plots for July

Above plots are generated by the software Minitab after performing Regression. Following results are observed from regression analysis.

- Normal probability plot: For the data to be normal, it must be along the straight line in the normal probability plot. Except for June, the other two months poorly satisfy this condition.

- Histogram: An ideal histogram is a bell-shaped curve which means that the residuals are normally distributed. In this case, all three months have poor histograms.
- Versus fits: This plot shows how scattered or skewed the residuals are. For the data to be normal, residuals should be randomly distributed. In this case, the number of data point is small, however any trend is not seen. So, the plots are adequately
  fitting.
- Versus order: This plot shows if the variance remains constant through time. All three plots have changing variance. Hence, the regression model is not appropriate or it can be said this kWh degradation is non-linear.

It is occurred that these systems have a non-linear power degradation. However, it's not a conclusion based on regression model, since only 5 years of data is worked on. Further analysis is performed using time series methods. ASU-PRL have evaluated PV systems for degradation using Winters' method, which is a third order exponential smoothing method. In this research, Winters' method and ARIMA method are performed on kWh data of PV systems built in ASU Tempe campus.

1.4.4 Application of time series methods on kWh pre-processed data:

First, ARIMA method is performed on these systems.

Model: Seasonal ARIMA(0, 0, 0)(0, 1, 1)12

Model Summary

DF	57	Stable	Yes
Sum of Squared Errors	54205.5197	Invertible	Yes
Variance Estimate	950.97403		
Standard Deviation	30.8378668		
Akaike's 'A' Information Criterion	578.044908		
Schwarz's Bayesian Criterion	582.199983		
RSquare	0.97183025		
RSquare Adj	0.97133605		
MAPE	2.00886978		
MAE	23.0502316		
-2LogLikelihood	574.044908		

Figure 1.16: JMP output window of selected model of seasonal ARIMA

This model is invertible and fits the data. The R-square value is 0.971 which is highly favorable.

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
MA2,12	2	12	0.53583	0.161213	3.32	0.0016*
Intercept	1	0	-19.13607	2.449237	-7.81	<.0001*

Constant Estimate	Mu
-19.136074	-19.136074

Forecast

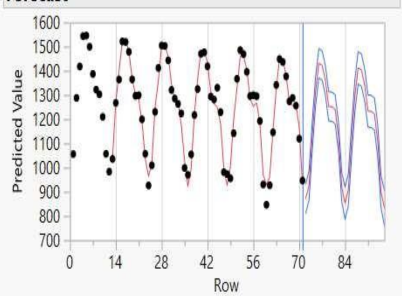


Figure 1.17: JMP output showing parameter estimates of selected model

In the above figure, output window shows parameter estimates. This model has only seasonal component of Moving Average term. The p value of moving average parameter and the intercept is less than 0.05 which is ideal. These values are shown in red and marked with asterisks. It means these values are highly acceptable.

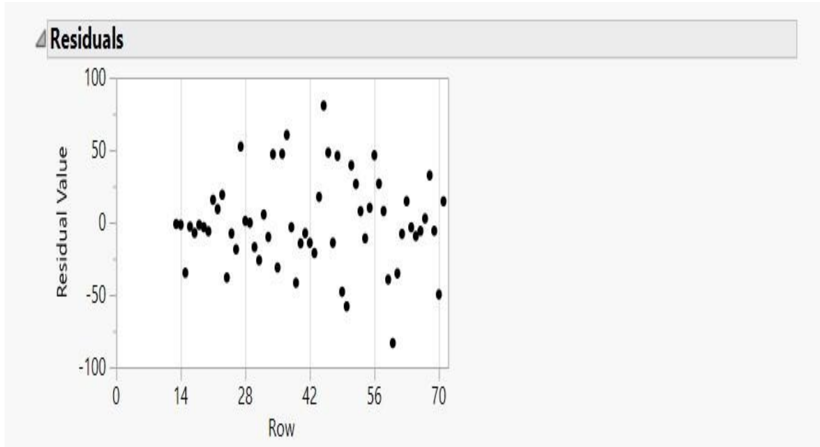


Figure 1.18: JMP output showing residuals

The above plot shows residual values of the model seasonal ARIMA (0,0,0) (0,1,1)12. A model is good if it has scattered residuals. This plot shows that the residuals are scattered and they do not show any trend.

All the systems evaluated are located in ASU Tempe campus. All those systems are new, about 5-6 years old. For these systems, the ARIMA has worked well. However, it is essential to evaluate performance of older PV systems to check if they are degrading with the same rate and if they are still degrading in a linear manner.

A PV system was built in Tempe warehouse, which is in Tempe, Arizona. It started producing power in November 2004 and researchers have collected data available till December 2013. It’s data worth for more than 9 years, precisely 110 months. Now, this data is analyzed using Winters’ and ARIMA. First, Seasonal ARIMA is performed.

Model: Seasonal ARIMA(1, 0, 0)(0, 1, 1)12

Model Summary

DF	95	Stable	Yes
Sum of Squared Errors	26218.5026	Invertible	Yes
Variance Estimate	275.984238		
Standard Deviation	16.6127733		
Akaike's 'A' Information Criterion	849.839734		
Schwarz's Bayesian Criterion	857.594637		
RSquare	0.96066463		
RSquare Adj	0.95983652		
MAPE	3.72566251		
MAE	13.9018345		
-2LogLikelihood	843.839734		

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR1,1	1	1	0.403315	0.0983212	4.10	<.0001*
MA2,12	2	12	0.896660	0.2646604	3.39	0.0010*
Intercept	1	0	-5.072274	0.9817088	-5.17	<.0001*

Constant Estimate	Mu
-3.026549	-5.0722743

Figure 1.19: JMP output showing seasonal ARIMA results

It is seen in the above output window that the model seasonal ARIMA (1,0,0) (0,1,1)12 is stable and invertible. Also, the ‘p’ values of AR and MA parameters and the intercept are very good. These values are shown in red in the output window and marked with asterisk. It means those values are just perfect and the model fits the data properly.

1.4.5 Summary of all the systems:

Following table shows degradation rates, ARIMA models and linearity in degradation for all analyzed systems.

System name	Degradation rate (% per year)	ARIMA model fit	Linearity of system degradation
Hayden Library	-1.55	(0,0,0) (0,1,1)12	Yes
Packard parking	-2.9	(0,0,0) (0,1,1)12	Yes
Wrigley Hall	-3	(0,0,2) (1,1,0)12 (low R square)	Yes
Barret Honors College	-1.3	(0,0,0) (1,0,0)12	Yes
Weathercup Center	-1.64	(0,0,0) (0,1,1)12	Yes
Tempe Warehouse	-1.69	(1,0,0) (0,1,1)12	Yes

Table 1.1: Summary of analysis

Above tabulated models are analyzed further in the research. This analysis involves correlating data using the ARIMA model numbers.

1.4.6 Analysis of seasonal ARIMA models:

Seasonal ARIMA (0,0,0) (0,1,1) 12: First pair of parentheses represent non-seasonal part of the model and second pair of parentheses represent seasonal part of the model. “12” is the seasonal period in the model, which is number of months. Seasonal input is 12 because one month repeats after next 12 months. Which also means, this model correlates data of every month of consecutive years. For example, January of 1st month and January of 2nd year, January of 2nd year and January of 3rd year and so on. It does this for all months.

Now let’s understand Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) and their relation with autoregressive and moving average terms.

Seasonal part (0,0,1): AR term is 0, Difference is 0 and MA term is 1. It means, 1st difference is taken to make the time series stationary; the ACF plot has 1 lag significant. This suggests that there's a seasonal correlation between months of consecutive years. To be more clear, there's a correlation between January of year1 and January of year2, January of year2 and January of year3 and so on.

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
MA2,12	2	12	0.53583	0.161213	3.32	0.0016*
Intercept	1	0	-19.13607	2.449237	-7.81	<.0001*

Constant Estimate	Mu
-19.136074	-19.136074

Forecast

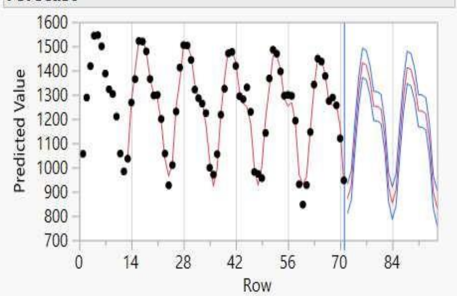


Figure 1.20: Parameter estimates of selected ARIMA model

The parameters show a probability of t- statistics as 0.0016 and 0.0001 for MA term and intercept. These parameters are smaller than the threshold value of 0.05. Probability of the parameter values being greater than the t ratio is very low which means that the parameters have significant values, residuals are normally distributed and model fits the data adequately.

Similarly, inferences are made with observation of results of other models.

Wrigley hall [(0,0,2) (1,1,0)12]: Seasonal part: AR term is 1 which means seasonal every month of one year is correlated with corresponding month of next consecutive year. In the non-seasonal part, MA term is 2 which means in each year, every month is correlated with next two consecutive months. It suggests that output increases or decreases linearly for three consecutive months. For example, Jan, Feb and March; Feb, March and April; and so on.

Barret Honors college [(0,0,0) (1,0,0) 12]: This model suggests that the system behaves similarly as Hayden Library, only the linear degradation behavior is represented by autoregressive term instead of moving average term.

Tempe warehouse [(1,0,0) (0,1,1)12]: Similar behavior as previous system, but also for every year, two consecutive months are linearly correlated.

From above table, it is seen that Hayden Library, Packard Parking and WeatherCup Center systems behave exactly in the same manner. From above analysis, it is seen that ARIMA models are consistent in its parameters. ARIMA approximates a linear model. Hence, it can be concluded that all above systems which are located in hot and dry weather of Arizona degrade in a linear manner.

1.4.7 Application of Winters' method:

ARIMA fits the degradation in kWh output data from PV systems in hot dry climate of Arizona in linear manner. Further, Winters' method is used to evaluate these systems.

Model: Winters Method (Additive)

Model Summary

DF	94	Stable	Yes
Sum of Squared Errors	27957.779	Invertible	Yes
Variance Estimate	297.423181		
Standard Deviation	17.2459613		
Akaike's 'A' Information Criterion	857.407197		
Schwarz's Bayesian Criterion	865.13133		
RSquare	0.95318715		
RSquare Adj	0.95219113		
MAPE	4.20441806		
MAE	15.7607574		
-2LogLikelihood	851.407197		

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Level Smoothing Weight	0.50682653	0.1232548	4.11	<.0001*
Trend Smoothing Weight	1.5159e-10	5.8734e-9	0.03	0.9795
Seasonal Smoothing Weight	0.00000025	0.0000097	0.03	0.9795

Figure 1.21: JMP output showing Winters' method results

This is JMP output of Winters' method performed on the kWh data of Tempe Warehouse system. It is the model is stable, invertible, the R square value is 0.953 which means the model fits good to the data. In parameter estimates and their significance values, it is seen that the P value for level smoothing weight is significant and the P values or trend smoothing weight and seasonal smoothing weight are not significant.

Model: Winters Method (Additive)

Model Summary

DF	48	Stable	Yes
Sum of Squared Errors	175137.868	Invertible	No
Variance Estimate	3648.70558		
Standard Deviation	60.4045162		
Akaike's 'A' Information Criterion	582.494448		
Schwarz's Bayesian Criterion	588.289925		
RSquare	0.88272622		
RSquare Adj	0.87783981		
MAPE	4.62394946		
MAE	51.0375253		
-2LogLikelihood	576.494448		

Failed: Cannot Decrease Objective Function Hessian is not positive definite.

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Level Smoothing Weight	0.10722998	0.091840	1.17	0.2487
Trend Smoothing Weight	0.00000000	.	.	.
Seasonal Smoothing Weight	0.12681490	2.973304	0.04	0.9662

Figure 1.22: JMP output of Winters' method

The above image shows output for Winters method performed on kWh data of Hayden Library. It says the model failed for forecasting. The goal is to determine whether the system is degraded linearly or not based on available data and not forecasting. Therefore, a comparative analysis between predicted kWh outputs of Holt-Winters' and ARIMA with the raw kWh values is performed further in this report.

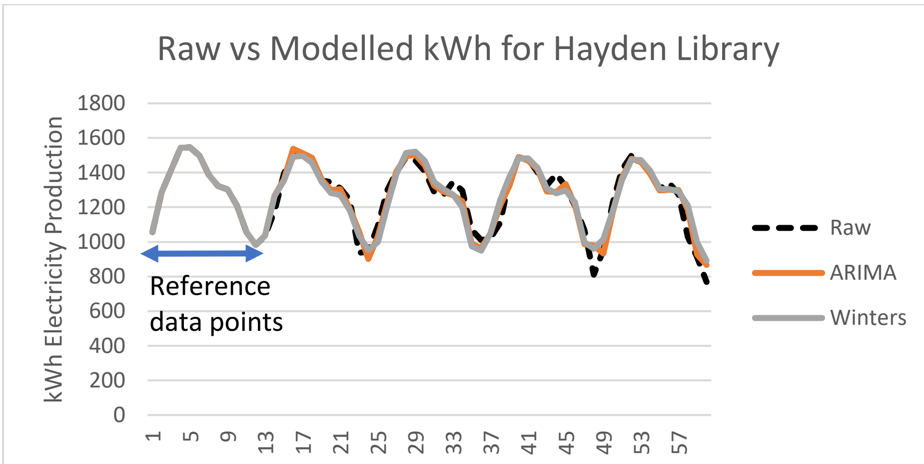


Figure 1.23: Comparison of ARIMA and Winters' predicted kWh values with raw kWh values for Hayden Library

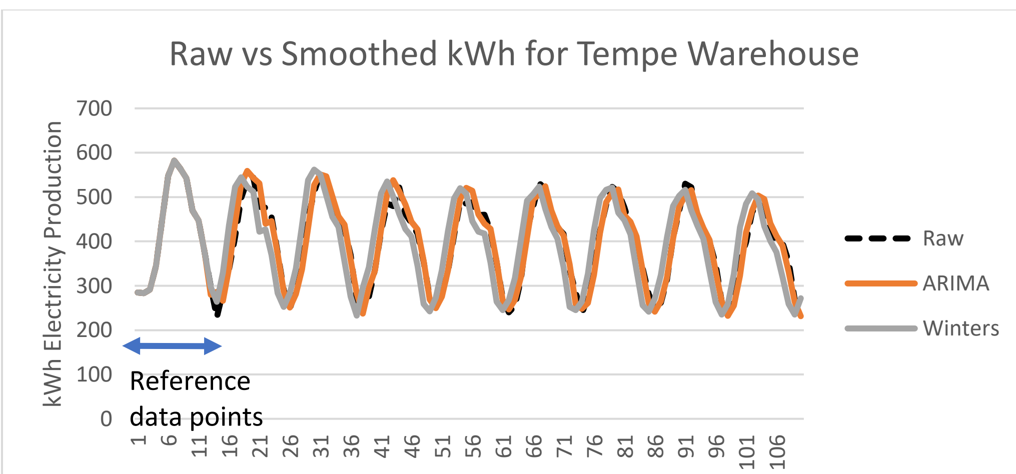


Figure 1.24: Comparison of ARIMA and Winters' predicted kWh values with raw kWh values for Tempe Warehouse

In the above plots, we can see that the models have smoothed the data and removed any peaks present. First 12 data points are for reference used by both models to take seasonality into account.

1.4.8 Degradation slopes:

Following plot shows degradation slopes for selected summer months. The degradation rates for April, May, June and July are -1.12%, -1.34%, -1.86%, -1.8% respectively.

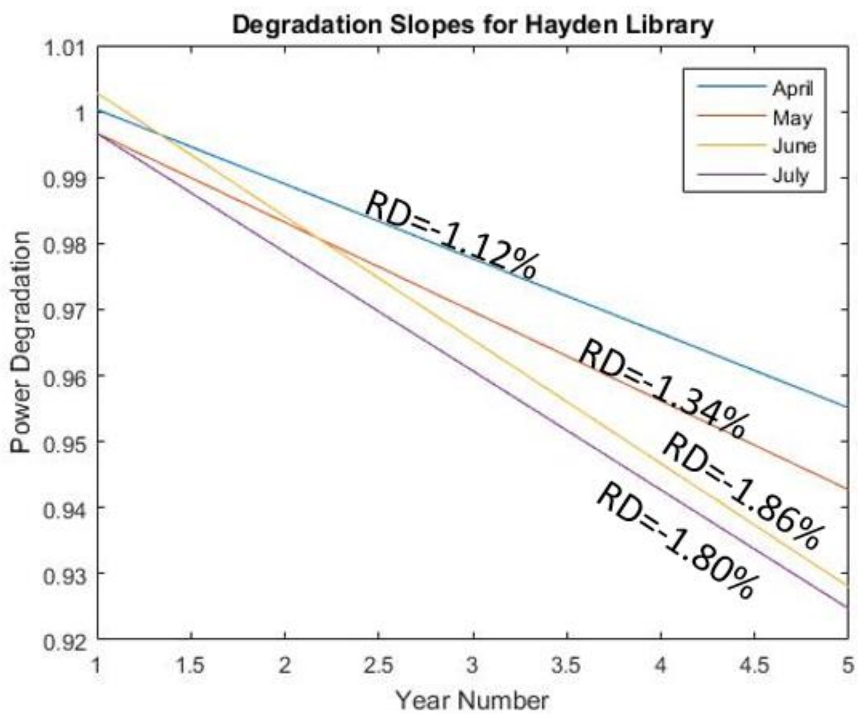


Figure 1.25: Degradation slope lines for summer months for Hayden Library

The following plot shows an average degradation for Hayden library system, which is -1.53% per year.

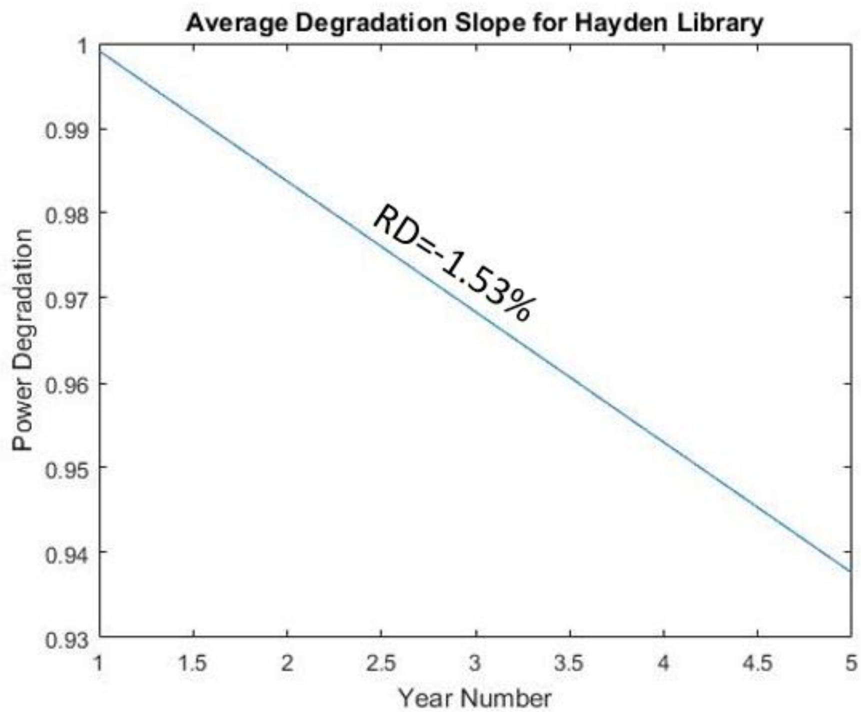


Figure 1.26: Average degradation slope for Hayden library PV system

Degradation rate can be calculated more accurately by using PR and PI ratio. PR is Performance Ratio which corrects for hourly irradiance hence monthly energy. PI is Performance Index which corrects for hourly irradiance as well as temperature and other losses. PI is a better measure than PR.

5. CONCLUSION

In the first part of this thesis, a statistical analysis is performed to model performance of PV power plants and determine whether the degradation rate is linear or not for those plants located in hot-dry climate of Arizona, consisting of c-Si modules.

This part of thesis is an extended research of ASU-PRL's previous work on climate specific degradation rate and linearity analysis using I/V, PR, PI and kWh data. This analysis is performed on the raw kWh data of the plants. The analyzed power plants of different ages are built in Arizona State University. Previously, it was stated that these power plants have tendency to degrade at different rates for first three years from the years after. This conclusion was based on standard deviation and box plots. In this research a statistical analysis using time series methods is performed. These methods are Winters' method and ARIMA method which are exponential smoothing and an advanced version of moving average respectively.

A measured value consists of two components, signal (a desired value) and white noise (random errors). Statistical methods such as Winters' method and ARIMA separate out noise from data. Winters' method, which is a third order exponential smoothing method uses smoothing weights and averages current data with previous data. ARIMA is an advanced method of a moving average which integrates moving average terms with autoregressive terms. The autoregressive weights are exponentially decreasing associated with historical data. Both methods estimate linear models. Winters' method separates out data in three components level, trend and seasonality. A regression analysis on trend component helps us tell whether the model represents a linear degradation or not. ARIMA has two components seasonal and non-seasonal. Seasonal part of ARIMA explains correlation between yearly data, i.e. value of this year January can be predicted from value(s) of previous year(s) January. Non-seasonal ARIMA explains trends in the same year, i.e. value of August can be predicted from value(s) of previous month(s). Winters' method and ARIMA both estimate a linear fit, however, the difference between two methods is that an ARIMA (p,d,q) (P,D,Q)S model is defined but There's no definition for Winters' model. Knowing an ARIMA (p,d,q) (P,D,Q)S model one can predict the data based on both seasonality and trend. Only trend component of Winters' method is used to make analysis on linearity. Therefore, ARIMA could be more accurately modeling the data. In the output window from JMP shows significance values of parameters. These parameters are smoothing weights in case of Winters' method and moving average or autoregressive weights in case of ARIMA. Winters' method is not very sensitive to small trend or seasonality. Therefore, the smoothing weights are not significant. For small amounts of data of five-ten years both Winters' and ARIMA are good fits and the fitted values for both methods are almost equal.

In the seasonal ARIMA models, it is clearly observed that seasonal dependence and current monthly data is mainly depending upon previous year monthly data by a fixed constant. It means that the output is degrading linearly in hot and dry climate of Arizona. A future work using these methods in determining linearity in degradation of PV modules at different climates for example, cold and humid or hot and humid can be done which would help estimating behavior of c-Si modules in various climatic conditions.

Future work in this research may consist of calculating degradation rates using PR and PI ratio and analyzing time series of PR and PI using ARIMA. It would give a more accurate degradation rate.

PART 2: TEMPERATURE OF SOILED MODULES

1. INTRODUCTION

2.1.1 Background:

Performance of PV modules is affected by different environmental parameters such as irradiance levels, ambient temperature, wind, soil, humidity, precipitation and so on [3]. Soil factor is mainly observed in hot and dry desert climatic conditions like in Arizona, Texas, New Mexico, Nevada and some parts of California. Dust storms are frequently observed in such deserts. After a dust storm, soil gets deposited on the PV modules and forms a layer. Besides in hot and dry desert climatic conditions, pollution also causes soil/dirt deposition of PV modules. This mainly occurs in highly dense cities such as Beijing, Hong Kong, Mumbai and New Delhi.

Soil deposition reduces performance of PV in terms of power output by variable amounts. Typically, in Arizona, power loss due to soiling is considered as 3-5%. The soiling loss varies with different angles of tilt. Modules with 0° angle of tilt are considered to have maximum soil deposition and modules with increasing tilt angles would have lesser amount of soil because of gravity. A soiling station constructed by Sandia national labs with ASU-PRL is located at ASU-PRL, Mesa, Arizona. This soiling station consists of 10 PV cell coupons tilted at different angles, ranging from 0° to 45° with difference of 5° in each coupon. All coupons are divided in two cells. One cell is cleaned periodically and the other is remained soiled naturally.

2.1.2 Scope of work:

The temperature data of coupons is stored in a data logger system. Data for 19 months is collected and used for analysis. This research includes analysis of temperature difference between cleaned cell and soiled cell. It involves analysis of rain data, correlation of soiling loss factor (SLF) with the temperature difference and variations of temperature difference with seasons. A probability analysis is made to determine if clean module is hotter than a soiled module. Dry spells cause negative slope in SLF plots. Temperature analysis is performed on these periods. An analysis of temperature difference with different levels of irradiance in one day is made which represents dependability of temperature difference on irradiance as well as ambient temperature.

2. LITERATURE REVIEW

2.2.1 Soiling loss analysis

Soiling loss factor (SLF) is crucial in determining power loss due to soiling. It depends on dust frequency, intensity of site, humidity, wind speed, height of module installation and rain frequency [3]. Sandia National Laboratories in collaboration with Arizona State University (ASU) constructed five soiling stations around the country to analyze SLF. One of these soiling stations is located in ASU's Photovoltaic Reliability Lab [1]. This soiling station consists of 10 c-Si coupons tilted at 10 different angles 0°, 5°, 10°, 15°, 20°, 25°, 30°, 35°, 40°, 45° which are cut in half. One half is cleaned periodically and the other half is remained soiled naturally. It was observed that Arizona soiling station had the highest soiling loss factor amongst the other sites. The loss depends significantly on tilt angle.

2.2.2 Temperature analysis

Performance of PV is affected by irradiance, ambient temperature. For a crystalline Silicon module, voltage decreases by 1% for every 2.5° C rise in temperature and power decreases by 1% for every 2.2° C rise in temperature [10]. Also, temperature increases associated with thermal expansion; therefore, temperature is a significant factor affecting the performance of PV modules [10].

Previously, a PV system in Phoenix was inspected for its performance and it was found that, at different locations in the same system, modules were behaving differently. Overall performance of system was 40% of its rated capacity, however, performance of a group of modules at a particular end of the system, where trees were present, was about 60% of the rated capacity [9]. The reason suspected for this anomaly was air turbulence due to trees and hence, a cooler environment in that area. Therefore, temperature plays a major role in performance of PV modules.

3. METHODOLOGY

2.3.1 Outdoor Soiling Station:

Experiments and analysis were carried out on an outdoor soiling station located at ASUPRL, Mesa.

Description: The soiling station is constructed such that coupons are facing south. There are 10 coupons in total. Each coupon is divided in two parts. Coupons are tilted at tilt angles 0°, 5°, 10°, 15°, 20°, 25°, 30°, 35°, 40° and 45°. These coupons are irradiance sensors. Thermocouples are also connected at the back of each sensor. All sensors are connected to data logger through metallic pipes to protect connections from weather.

Fig 2.1: SANDIA soiling station located in outdoor testing field at ASU-PRL

A special plastic structure was constructed on the soiling station to avoid bird drops.

Fig 2.2: Plastic structures to avoid bird drops

Fig 2.3: Data logger

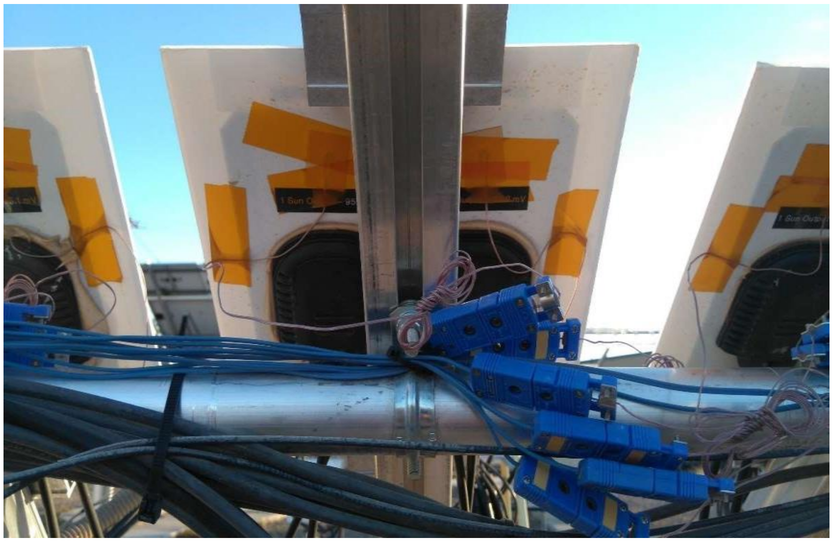


Fig 2.4: Back side of sensor showing attached thermocouples

2.3.2 Experiment description:

This soiling station was initially built to find out soiling loss in hot dry weather of Arizona. The right half part of every sensor was cleaned every Tuesday and Friday at 9:00 am. It was cleaned with distilled water and was dried with a dry paper towel. The left half part of the coupon is kept unclean or soiled. Following figure shows a coupon with left sensor soiled and right sensor cleaned.

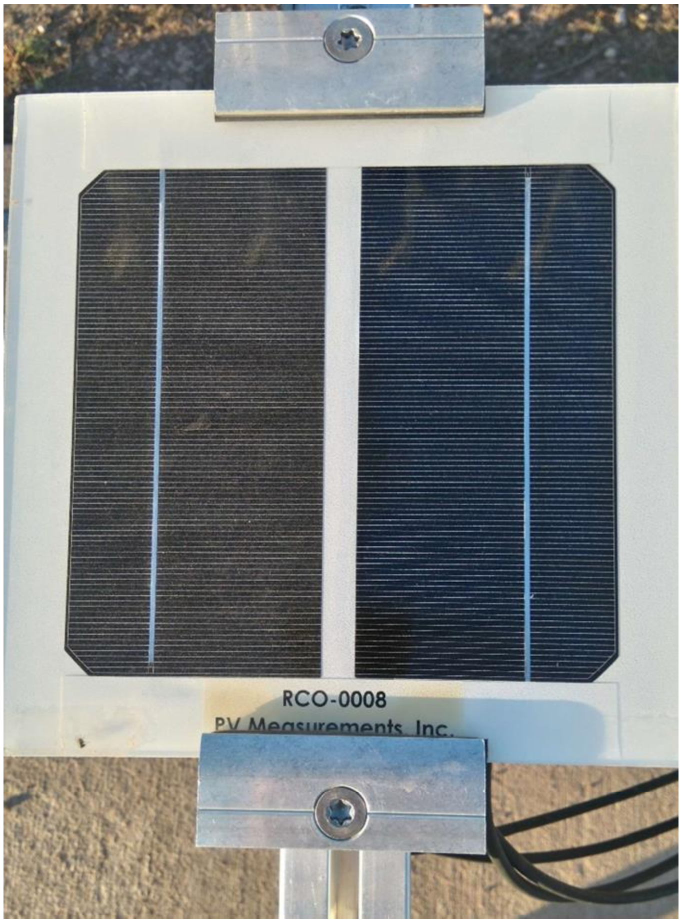


Fig 2.5: Sensor showing left half soiled and right half clean

Data collection and processing: Irradiance and temperature data are collected and stored in data logger. Irradiance and temperature are stored for every minute. These data are extracted and stored in server time to time. For this analysis, data are extracted in following manner:

- Note down date for First Friday in December 2014
- Select a time in morning (9-9:30) am (low irradiance and high angle of incidence)
- Take average temperatures of these time spans and store them.
- Do this for both soiled sensor and clean sensor for all 20 sensors

- Find ΔT = Tclean - Tsoiled
- Now go to next Friday and repeat the whole process

This process is followed from December 2014 to July 2016. The same process is repeated for afternoon (high irradiance and low angle of incidence) time and late afternoon (high irradiance and high angle of irradiance).

2.3.3 Time series plot:

$\Delta T = T_{clean} - T_{soiled}$ is observed to be varying in nature. Following time series plot shows that the ΔT varies with a polynomial curve with order of 5.

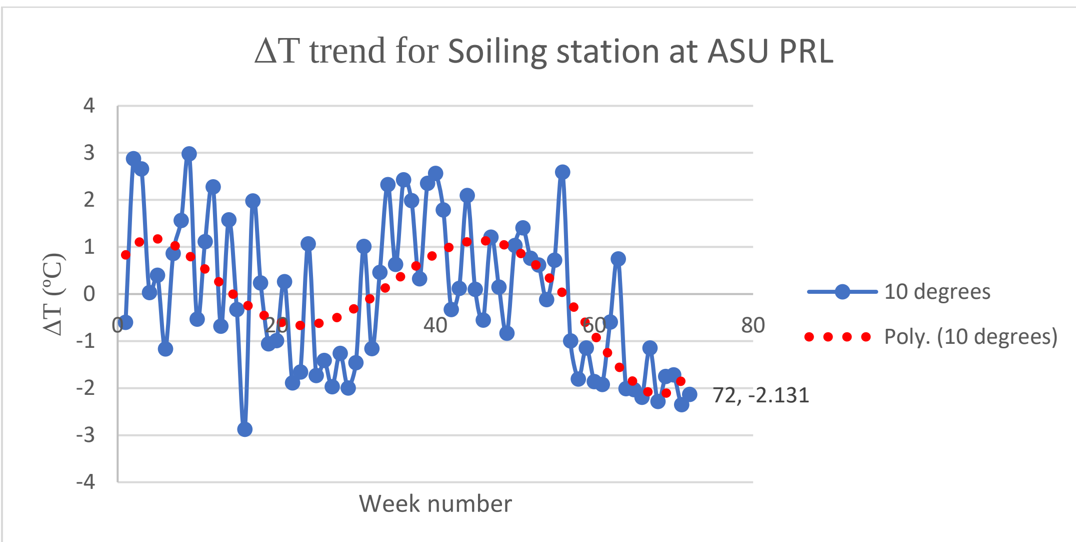


Fig 2.6: Time series plot of ΔT showing 5- degree variation

The plot also shows some seasonal variation as shown in the following figure:

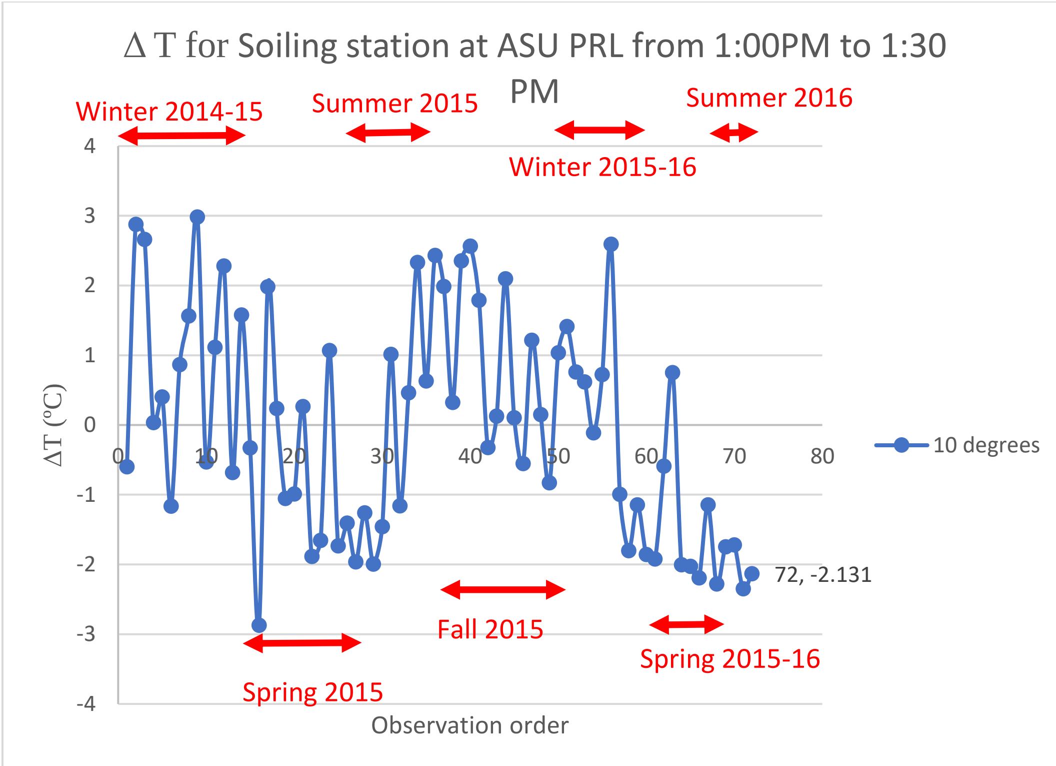


Fig 2.7: Time series plot showing seasonal variation

2.3.4 Seasonal Analysis:

It was observed that the ΔT has some influence of seasons on it. Therefore, the data were sorted seasonally as follows:

- Winter 2014: December 23rd 2014 to March 20th 2015
- Winter 2015: December 29th 2015 to March 15th 2016
- Spring 2015: March 27th 2015 to June 19th 2015
- Spring 2016: March 22nd 2016 to June 14th 2016
- Summer 2015: June 26th 2015 to September 25th 2015
- Summer 2016: June 21st 2016 to July 12th 2016

- Fall 2015: October 23rd to December 18th 2015

For simplicity of analysis, seasons were combined as:

- Combination of winter 2014 and 2015
- Combination of spring 2015 and 2016
- Combination of Summer 2015 and 2016
- Fall 2015

2.3.5 Collection of weather data:

Weather data are important in this analysis to study impact of weather on ΔT. The data are collected by a weather station near Phoenix Sky Harbor Airport. These data are available at NOAA website.

2.3.6 Calculating median ΔT:

To know how much a clean cell is hotter or cooler than a soiled cell, median values of ΔT were calculated for every season for every tilt angle.

2.3.7 Calculation of Angle of Incidence:

Angle of incidence varies with different seasons as well as time. Arizona modules are popularly placed at an angle about 10° of tilt, facing south. Considering this, angle of incidence in winter is high and in summer it is low. Also, Angle of incidence in the morning is low and in the afternoon, it is high. This is important because it influences irradiance and in turn the temperature. Solar elevation is calculated using the calculator available on NOAA website. Angle of incidence is a function of solar elevation angle.

Angle of incidence (AOI) = 90° - (angle of elevation) - (angle of tilt)

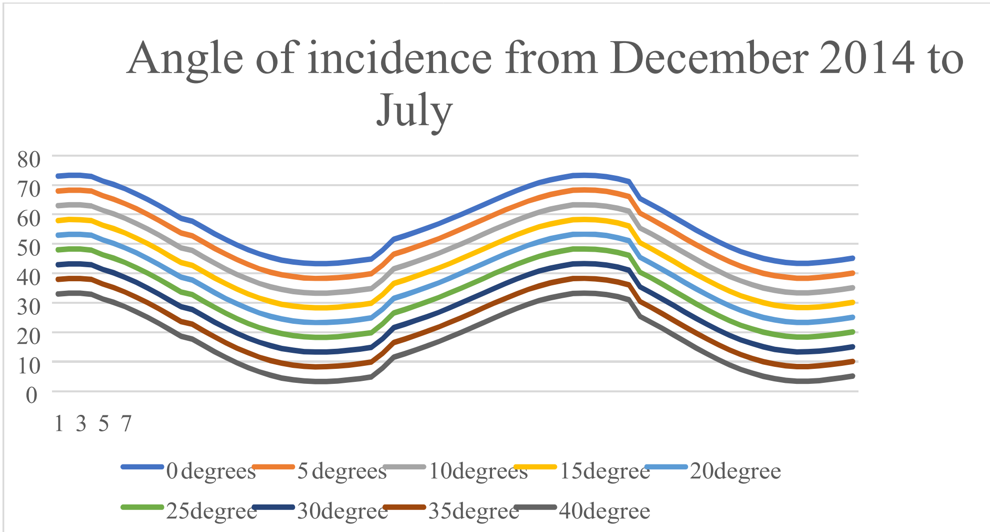


Fig 2.8: Angle of incidence at 9:15 am from December 2014 to July 2016

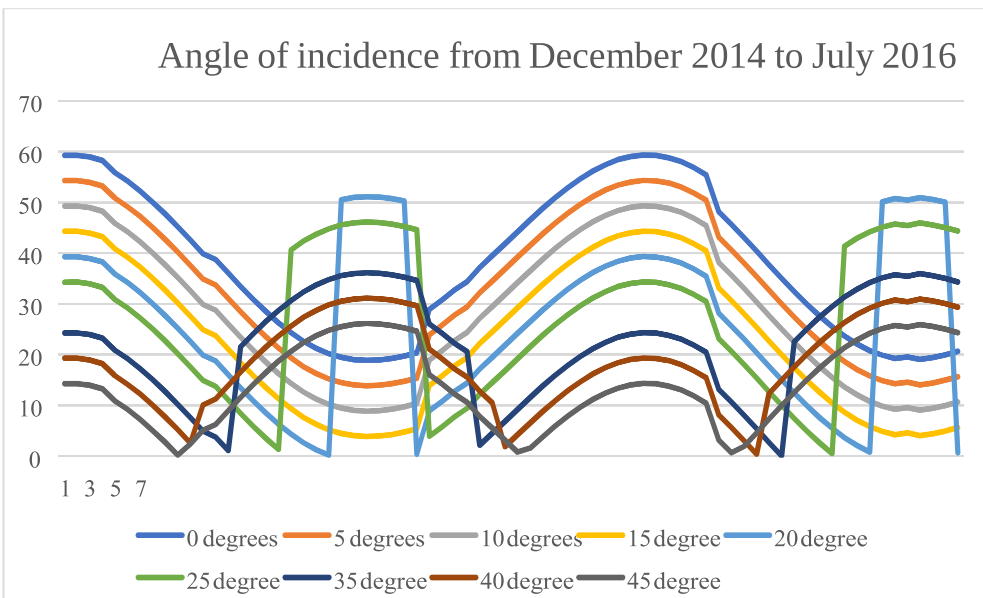


Fig 2.9: Angle of incidence at 1:15 pm from December 2014 to July 2016

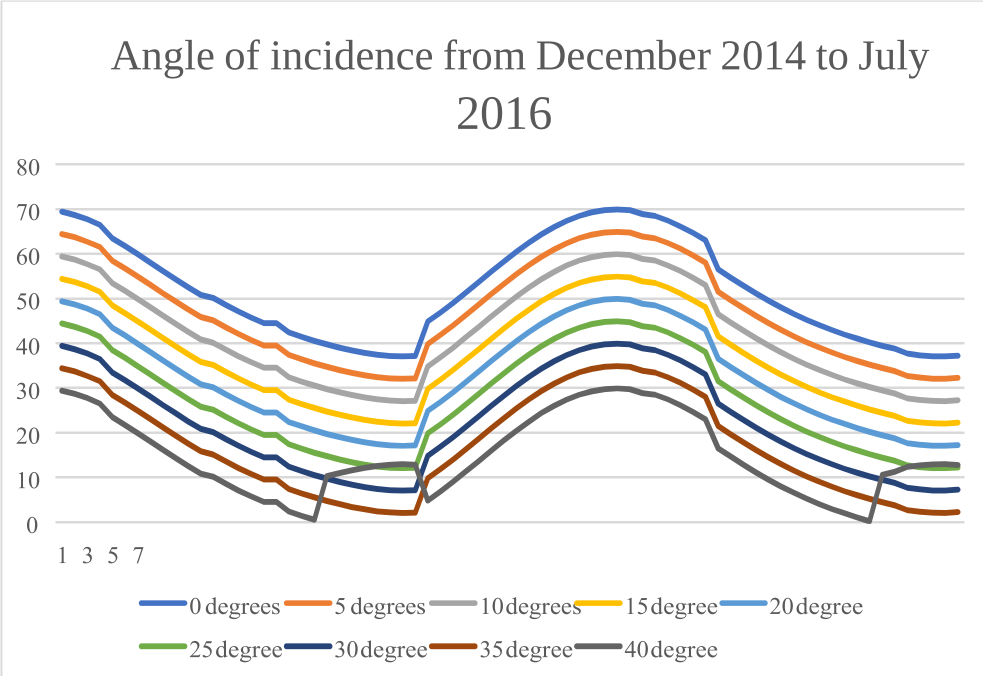


Fig 2.10: Angle of incidence at 3:15 pm from December 2014 to July 2016

2.3.8 Calculating probability:

The main objective of this analysis is to state whether a clean module would be hotter or cooler. Calculating probability is one of the best measures to state or prove this mathematically.

Probability of clean cell being hotter = $\frac{Number\ of\ positive\ \Delta T}{Total\ number\ of\ \Delta T}$

All positive ΔT's are numbered as "1" and all negative ΔT's are numbered as "0". Number of positive ΔT's are calculated using the "countif" function of Excel. Total number of ΔT's are calculated using "count" function of Excel. Probabilities are calculated seasonally.

Calculating Uncertainty in measurement of temperature: Uncertainty is the best estimate of how far the measured value can be from the actual value.

Accuracy of T type thermocouple: ±1°C

Accuracy of CR 1000 data logger in temperature: ± (0.055 + 0.0057 x temperature) °C

Accuracy of the data logger is of the order of 10^{-2} which is very small as compared to accuracy of T- type thermocouple. Hence it is justifiable to ignore the accuracy of data logger. Hence, uncertainty consists of only accuracy of thermocouple which is ±1°C.

RMS value of uncertainty:

Two uncertainties are present in measuring temperature. Hence, RMS (Root Mean Square) is more accurate measure of calculating uncertainty. The RMS value is: 1.05~1 (Order of 10^{-2} is not significant for temperature difference and its effect on power drop).

4. RESULTS AND DISCUSSION

2.4.1 Time series plot:

Time series is a sequence of time-indexed observations, equally spaced in time. It depicts chronological relation of output. In this analysis, the data are in weekly order. This plot is for 10° angle of tilt, which is most commonly used in Arizona. Every point in the plot represents ΔT of Friday, as calculated in methodology chapter. Many crests and troughs are seen in this plot. X-axis shows week number. Some points are above the zero line and some points are below the zero line. A seasonal pattern is seen. ΔT is positive in certain period of year and it is negative in some period of year.

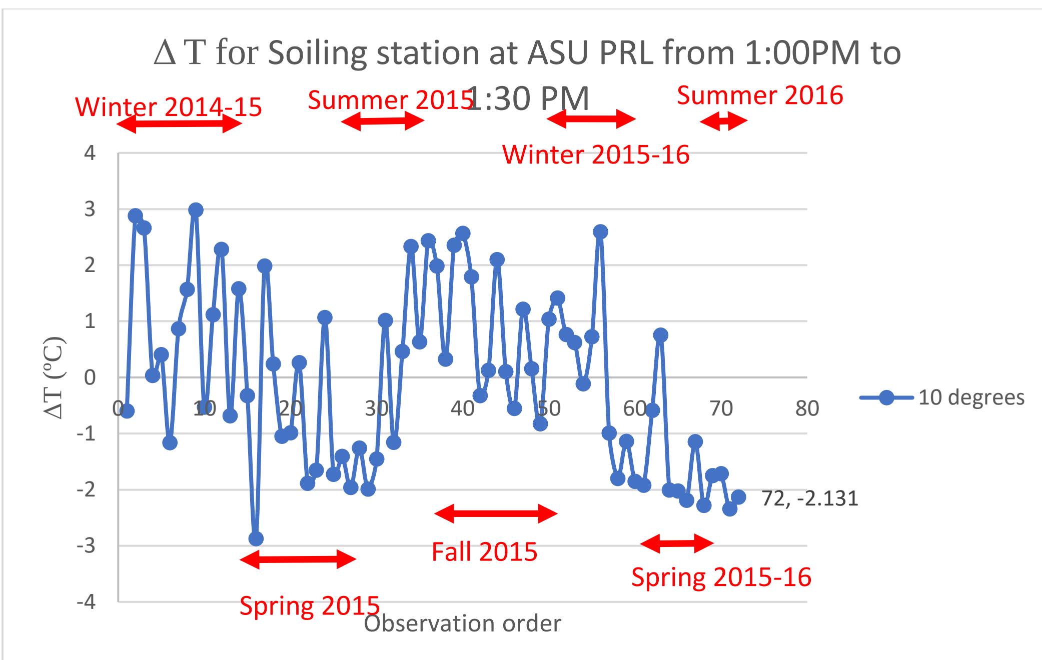


Figure 2.11: Time series plot of ΔT

It is observed that ΔT in winter and fall is positive and ΔT in spring and summer is negative. If a trendline is fit in this plot, it shows a 5th order polynomial curve.

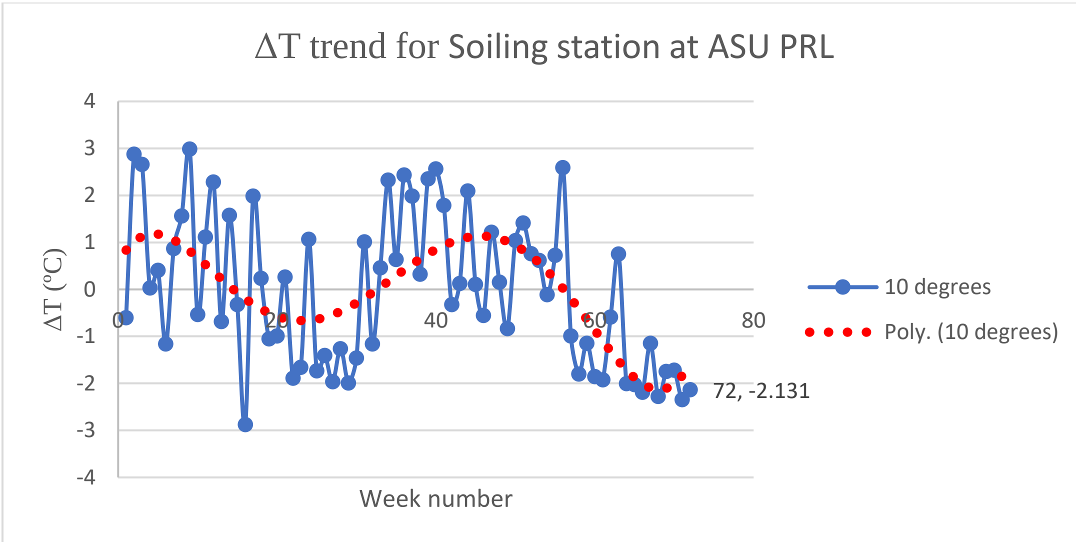


Figure 2.12: Time series plot with trendline

Following figures show time series plots of three tilt angles 0°, 20° and 35°.

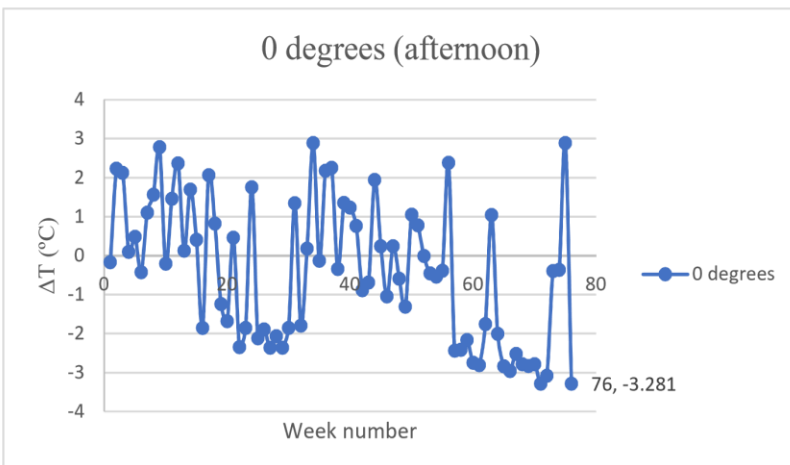


Figure 2.13: Time series plot for 0°

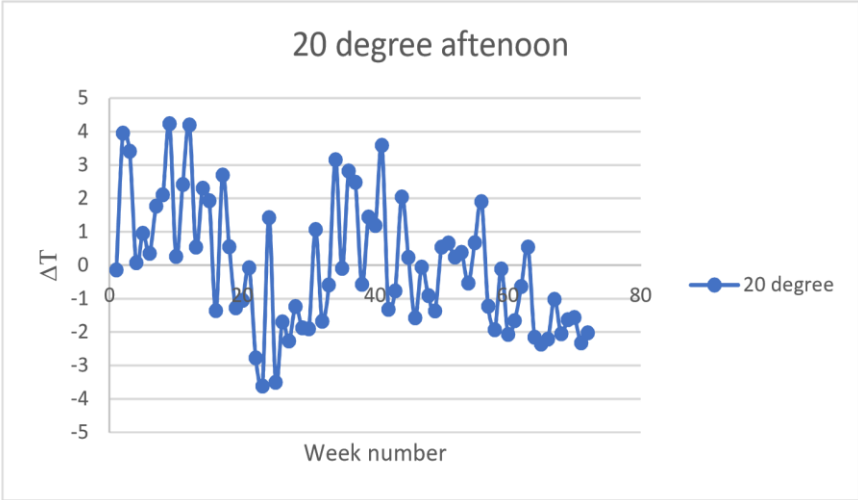


Figure 2.14: Time series plot of 20˚

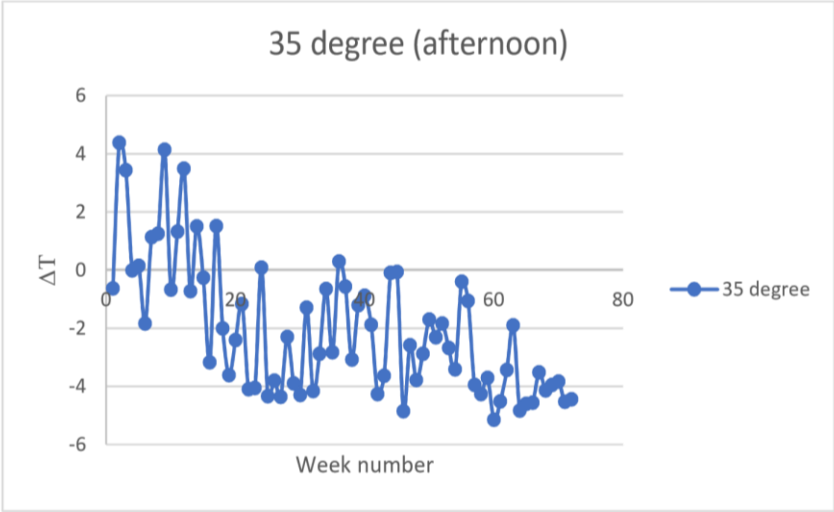


Figure 2.15: Time series plot for 35°

In above figures, it is seen that the time series plots of different angles have similar pattern for seasons. It is also see that 35° tilt angle has more negative values than 0° and 10° and 20°.

2.4.2 Median values of ΔT:

The ΔT varies with time and season. Further analysis shows how ΔT is influenced by certain parameters. Median is a good measure of statistical data. For every tilt angle, median value of ΔT in morning, afternoon and late afternoon is calculated.

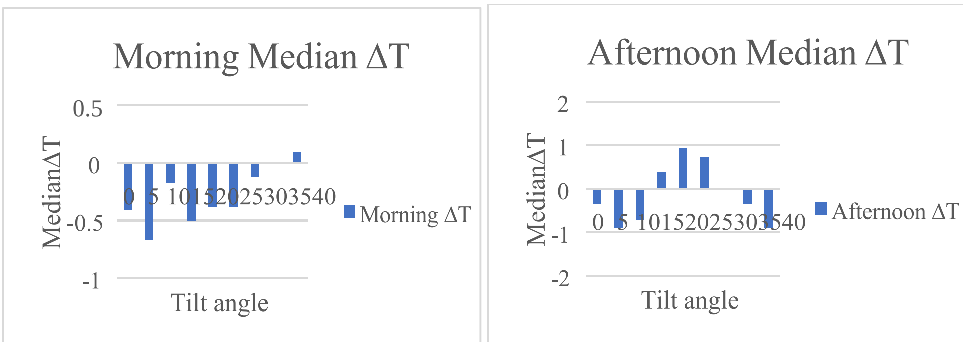


Figure 2.16: Median ΔT in morning and afternoon

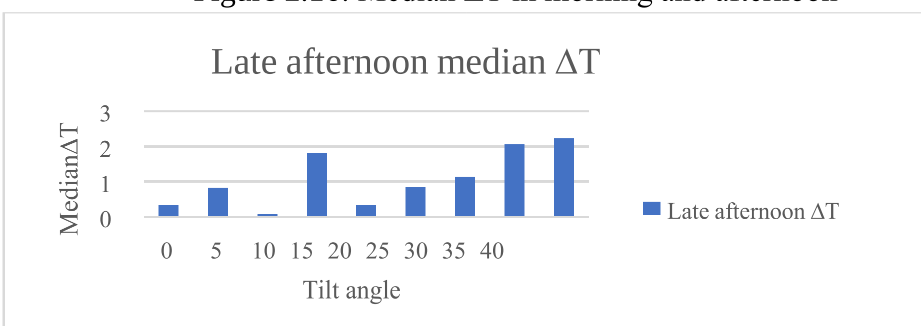


Figure 2.17: Median ΔT late in the afternoon

Median ΔT in morning is negative for all angles. Median ΔT in afternoon is negative for 0,5,10 degrees; positive for 15,20 and 25 degrees; negative for higher angles. Median ΔT is positive for all angles in the late afternoon. Hence, irradiance is one of the parameters which influences ΔT.

Superimposing all these values in one plot helps us compare median ΔT values in the morning, afternoon and late afternoon. It is clearly seen that all values in the late afternoon are higher than the morning and afternoon.

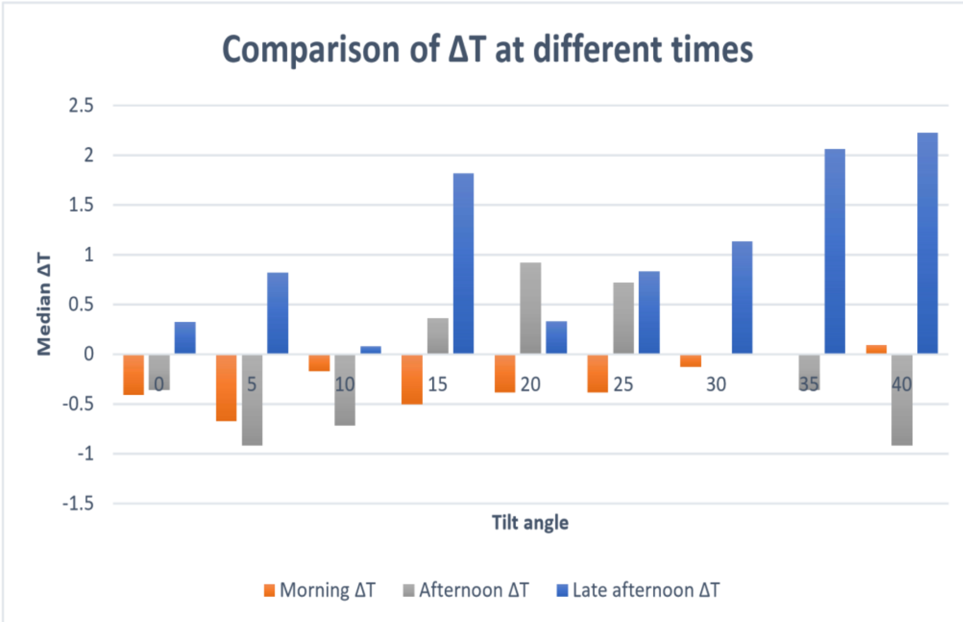


Figure 2.18: Comparison of median ΔT at different time spans

This is overall comparison throughout 2 years. However, it was initially seen that there's a seasonal influence on ΔT values. Therefore, a seasonal median ΔT comparison was made. Seasonal values were combined for two years as described in methodology. Median values of combined years were calculated and plotted.

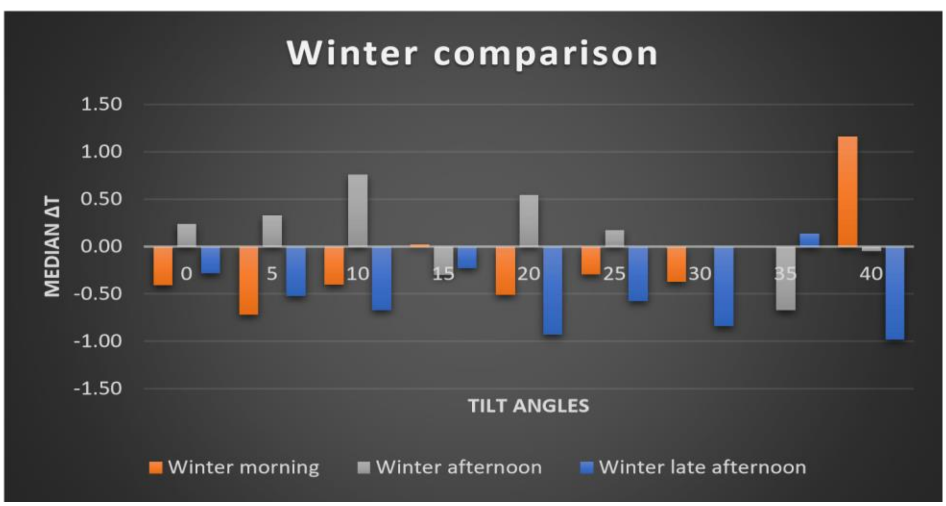


Figure 2.19: Comparison of median ΔT in Winter

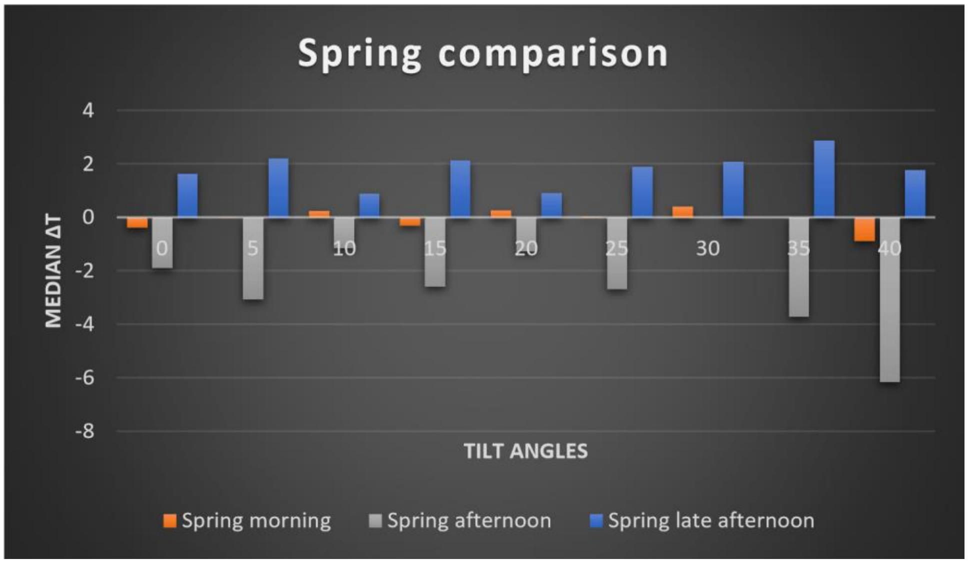


Figure 2.20: Comparison of median ΔT in spring

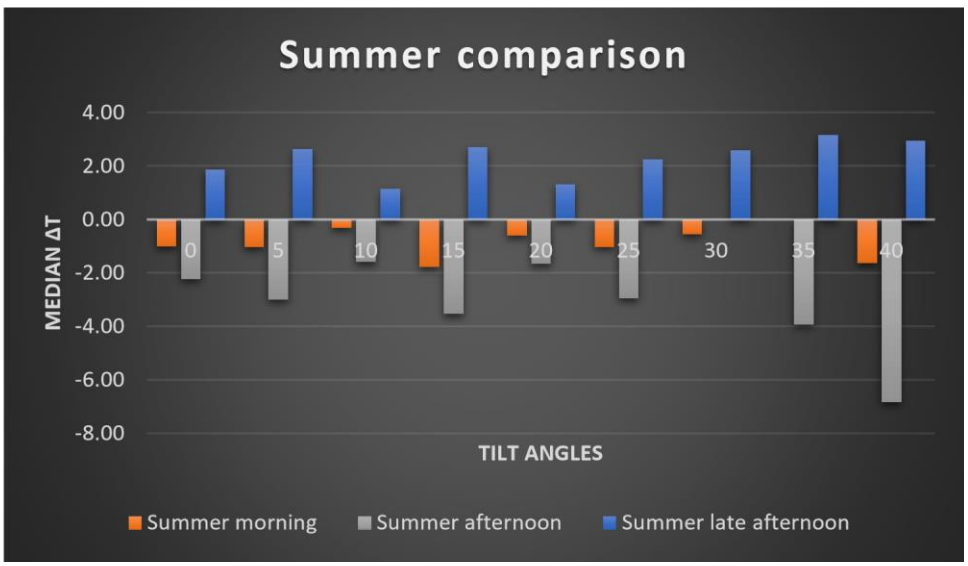


Figure 2.21: Comparison of median ΔT in summer

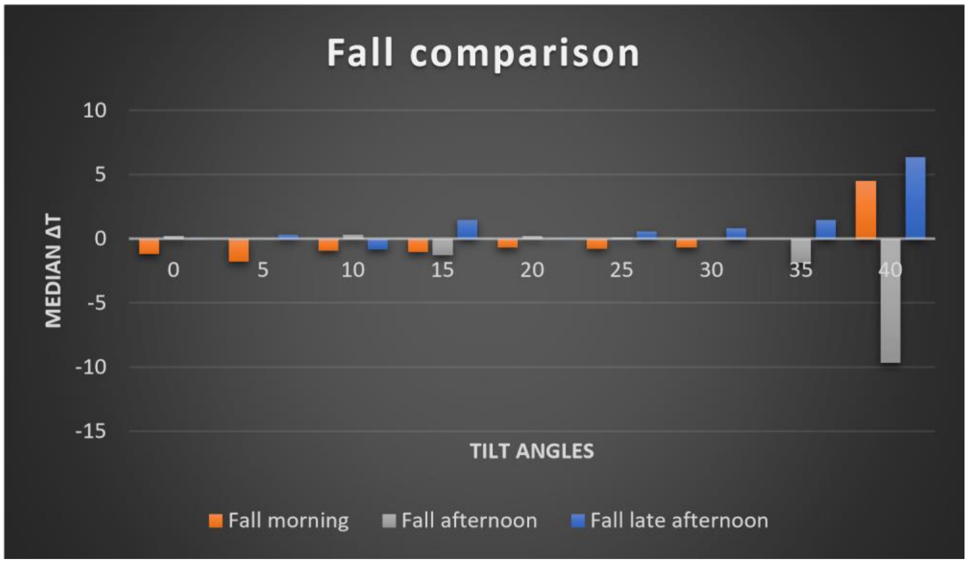


Figure 2.22: Comparison of median ΔT in Fall

In above plots showing median values of ΔT, it is observed that there's a seasonal effect. Fig (a) shows winter values of median ΔT. The morning and late afternoon values are mainly negative and afternoon values are positive. Maximum absolute value of ΔT is 1. Fig (b) shows median ΔT in spring. Morning values are close to 0. Afternoon values are significantly negative. Some of them exceed -4°C and the tilt angle 40° has the greatest negative value of -6°C. Late afternoon values are all positive significantly. Highest median ΔT value is about 3. Fig (c) shows median ΔT values in summer. Spring and summer values appear similar. Though the morning values are mostly negative about -1°C. Fig (d) shows median ΔT values in Fall. All values are small in magnitude about 1°C. Except for 40° tilt angle. This could be an exception or an outlier.

2.4.3 Probability plots:

Goal of this research is to find out whether a clean module is hotter than a soiled module or is cooler than a soiled module. To achieve this, a probability analysis is performed.

Further results show probability that a clean module is hotter than a soiled module.

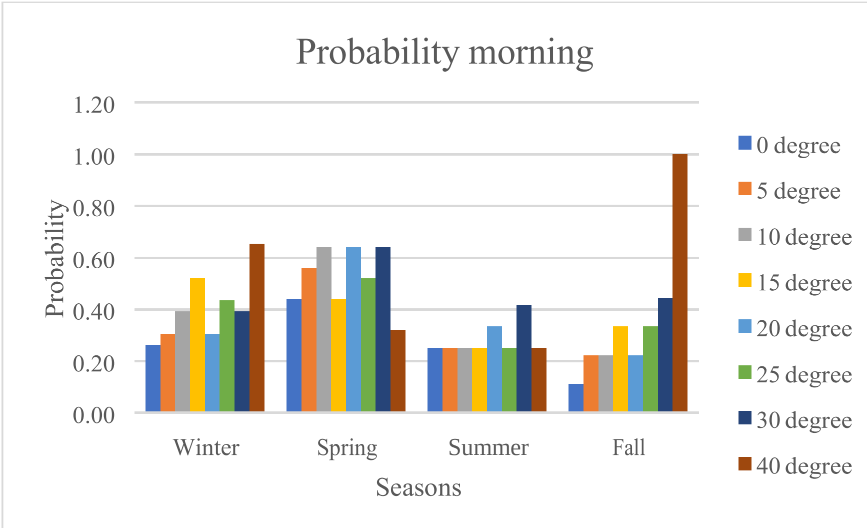


Figure 2.23: Probability of $T_{clean} > T_{soiled}$ in morning time

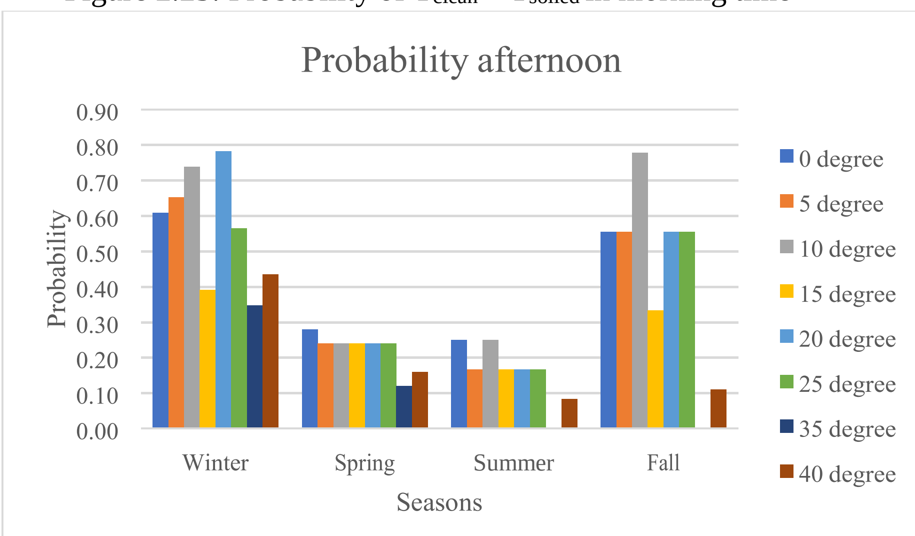


Figure 2.24: Probability of $T_{clean} > T_{soiled}$ in afternoon time

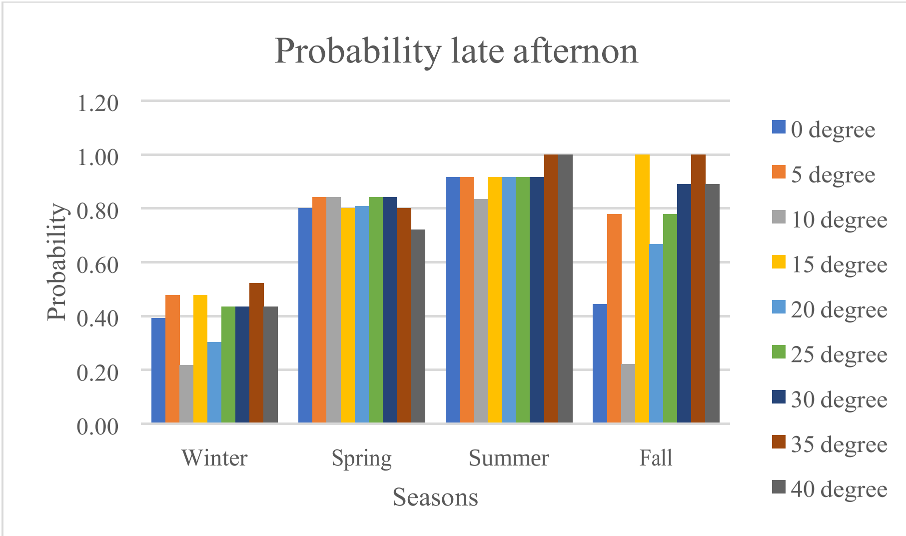


Figure 2.25: Probability of $T_{clean} > T_{soiled}$ in late afternoon time

From above plots, it is inferred that probability of clean module being hotter is low in morning time, it is moderate in the afternoon and it's high late in the afternoon. This suggests a possibility of influence of ambient temperature and angle of incidence on ΔT.

2.4.4 Soiling loss factor (SLF):

Soling loss factor (SLF) is a measure to determine power loss due to soiling. Generally, in Arizona's hot and dry climate, average soiling loss is about 3%. SLF is the ratio of short circuit current for a soiled module to the short circuit current of a clean module. Following plot shows weekly SLF for 10° tilt angle for 72 weeks of analysis. It is seen that average SLF is about 0.99. Some trends of SLF drops which are indicated using red lines.

These drops are called 'dry spells'.

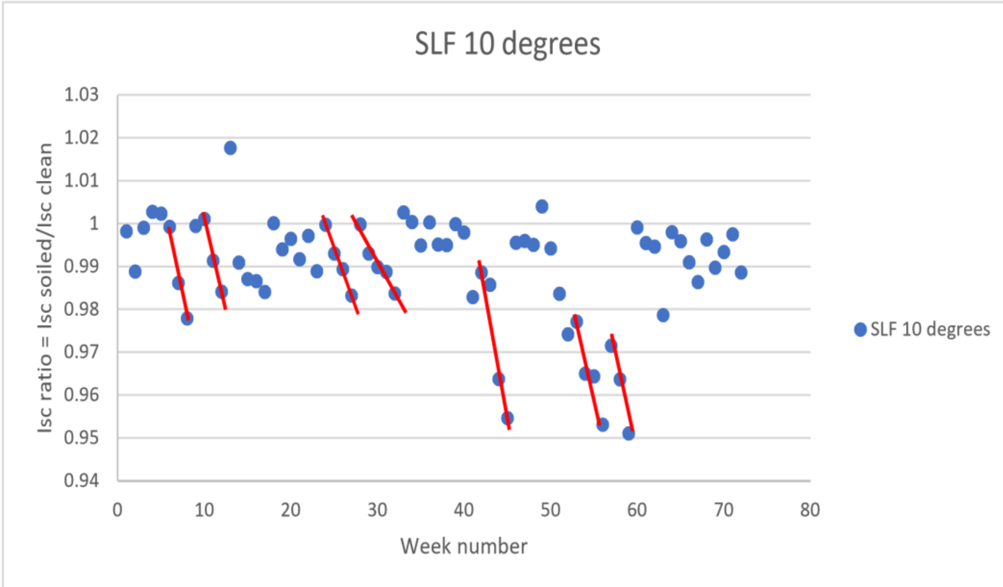


Figure 2.26: Soiling factor plot showing dry spells

Dry spell: In hot and dry weather of Arizona, soiling is an important issue in solar industry. Soiling loss factor (SLF) or I_{sc} ratio is a measure to find out how much power is lost due to soiling. The SLF is about 0.99 which means, power loss is about 1%. However, sometimes, the SLF drops to 0.97 or even 0.95 in consecutive weeks as shown in the plot. More soiling has been recorded in these dry spell periods. In this research, particularly these dry spells are considered for analysis.

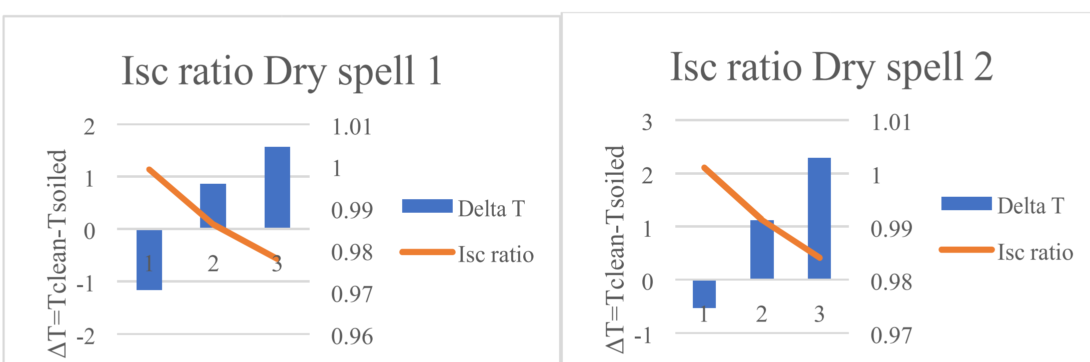


Figure 2.27: Soiling factor v/s ΔT

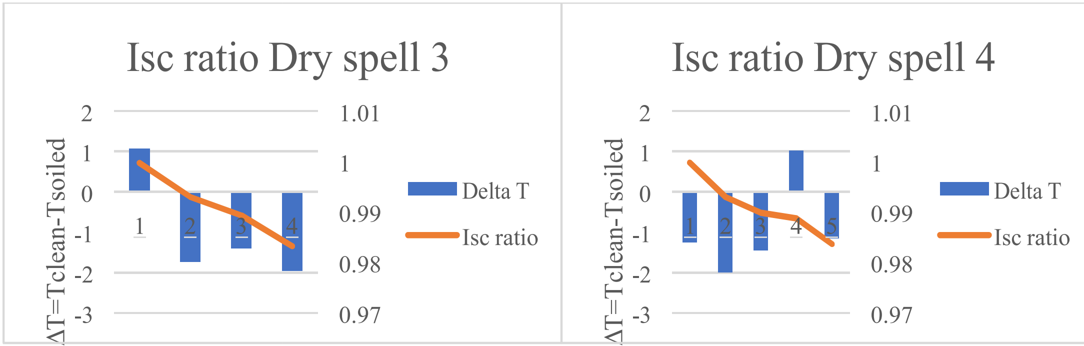


Figure 2.28: Soiling factor v/s ΔT

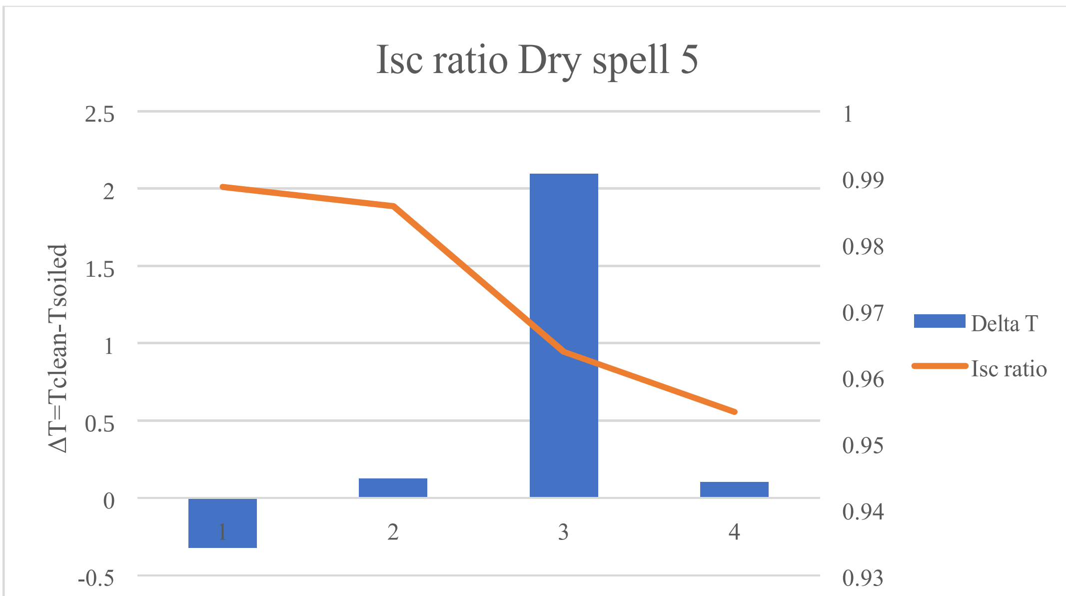


Figure 2.29: Soiling factor v/s ΔT

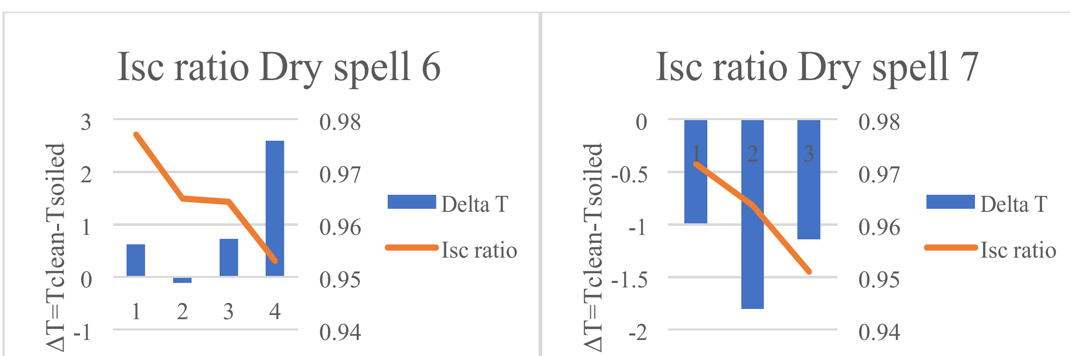


Figure 2.30: Soiling factor v/s ΔT

Discussion on above plots:

- Dry spells 1,2 and 6: These dry spells occurred in winter and ΔT tends to increase as soiling factor goes down. Soiling factor goes down when soiling density increases. Thus, in winter, a clean module is hotter than a soiled module.
- Dry spell 3 is occurred in late spring and early summer. ΔT seems going down as soiling factor decreases. Thus, in late spring and summer, clean module is cooler than a soiled module.
- Dry spell 4 is summer and ΔT is about same but negative. Clean module is cooler than a soiled module.
- Dry spell 5 is in fall. Soiling factor decreases and ΔT seems constant. Clean module and soiled module seem to have about equal temperatures.
- Dry spell 7 occurs in spring. As soiling factor decreases, ΔT is about same but negative. Thus, clean module is cooler than soiled module.

2.4.5 Plotting ΔT against day hours:

These plots give a more determining idea of relation of ΔT with irradiance levels as well as dry spells in different seasons.

2.4.5 Plotting ΔT against day hours

Following plots show temperature variation for clean and soiled modules in a day starting from 9:00 AM to 3:30 PM. Y axis represents time in minutes and X axis represents temperature in °C.

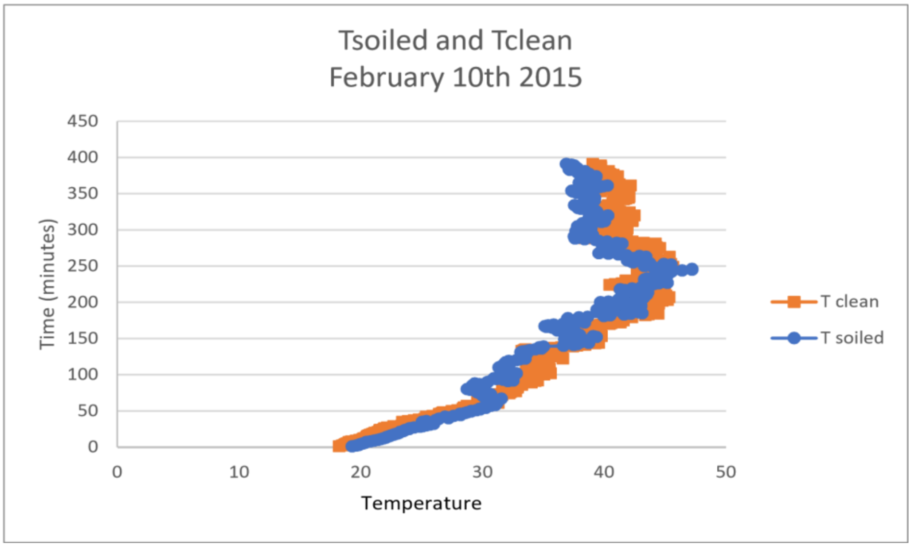


Figure 2.31: Temperature comparison with time

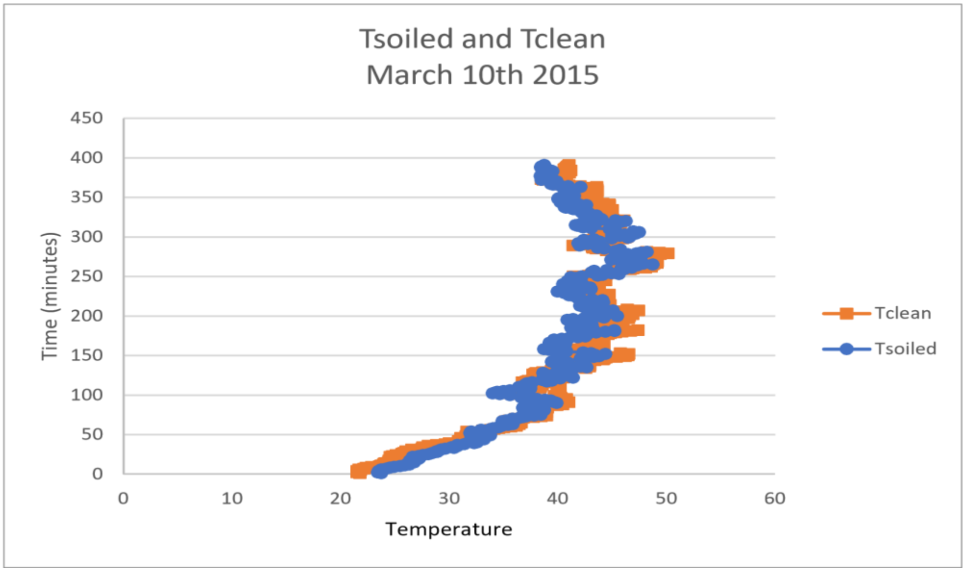


Figure 2.32: Temperature comparison with time

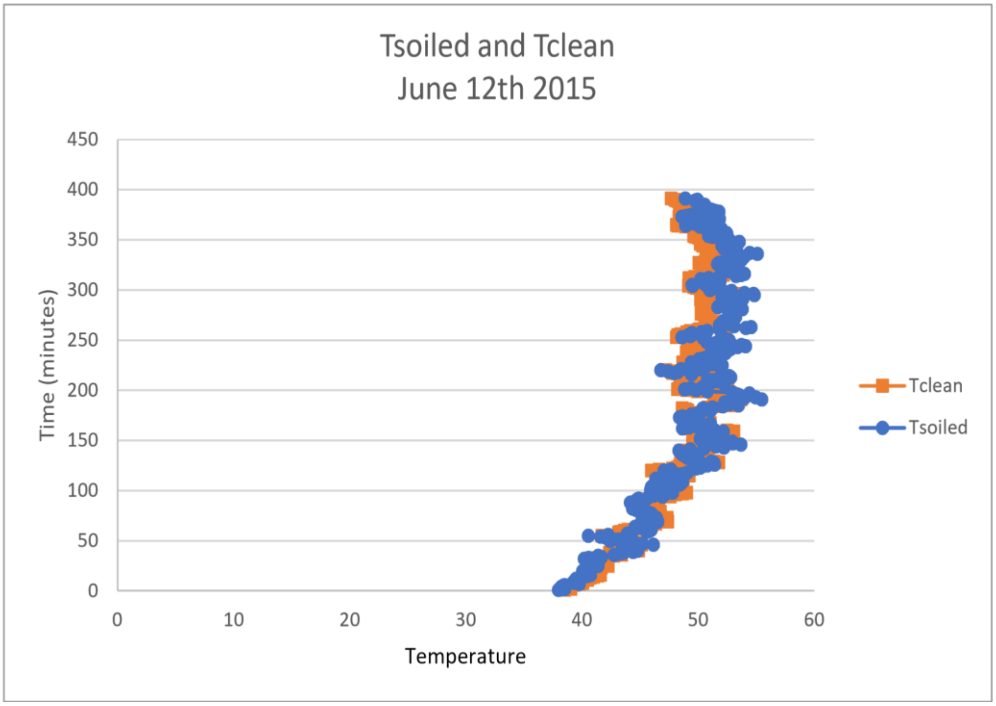


Figure 2.33: Temperature comparison with time

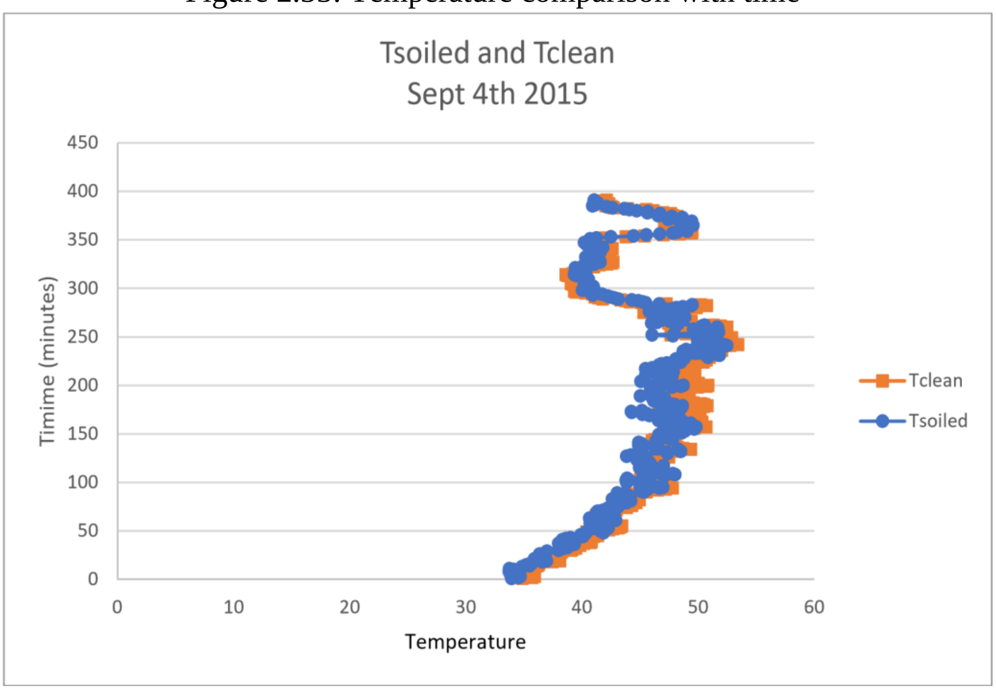


Figure 2.34: Temperature comparison with time

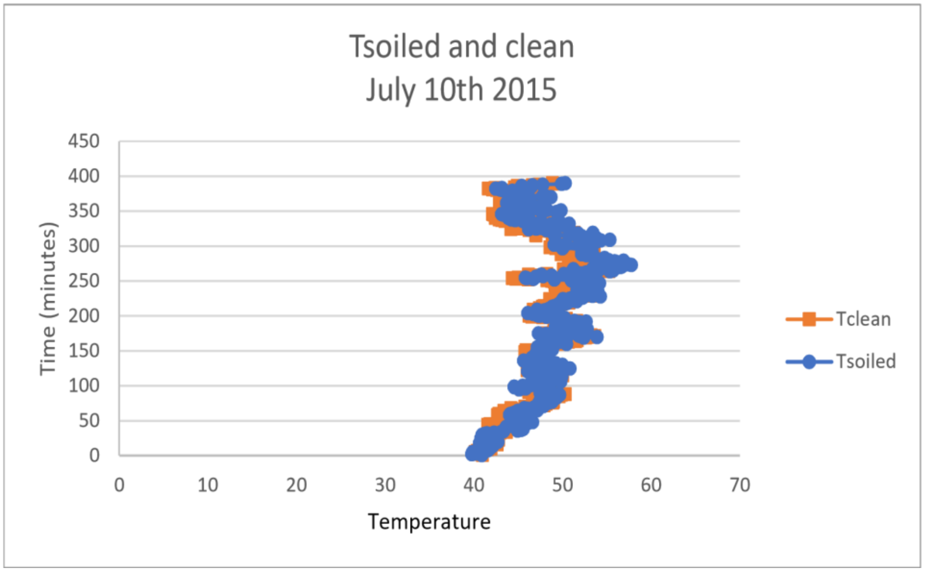


Figure 2.35: Temperature comparison with time

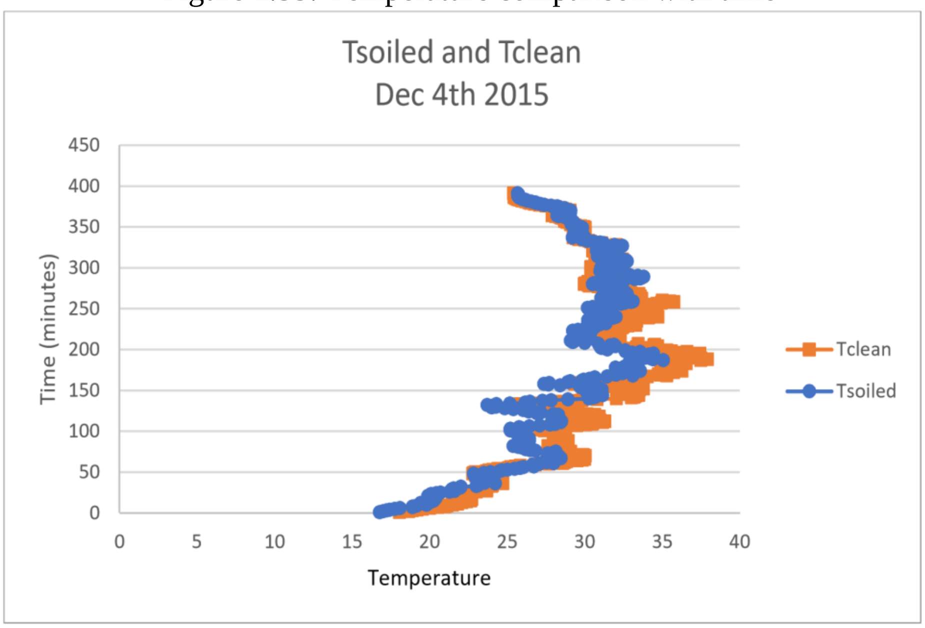


Figure 2.36: Temperature comparison with time

In above plots, Y-axis represents time in minutes starting from 9:00 AM in the morning till 3:30 PM in the afternoon. X-axis represents Temperature in °C. Orange curve represents temperature of clean module. Blue curve represents temperature of soiled module.

Generally, temperature of module increases till 250 minutes (till about 1:00 PM) and then decreases till 3:30 PM.

Inferences:

- Clean module is hotter in December, February, March and September which means clean module is hotter in winter and fall.
- Clean module is cooler in June and July, which means clean module is cooler in Summer.

5. CONCLUSION

In the second part, an analysis on temperature of soiled modules is performed to determine whether a soiled module is hotter than a clean module or it is cooler than a clean module.

This analysis was performed on the data from winter 2014 to July 2016. The experiment was carried out on a soiling station consisting of 10 different PV cell coupons, mounted at different tilt angles. A time series plot of temperature difference ΔT between a clean cell and a soiled cell against week number is generated. The trendline of this plot goes on decreasing as the tilt angle goes on increasing. It can be inferred that ΔT is low when angle of tilt is low and ΔT is higher when angle of tilt is high. This can be explained; a module with 0° tilt angle holds the most amount of soil and amount of soil deposition goes on decreasing as the tilt angle goes higher because the soil slides down due to gravity. This observation suggests that a module is cooler when soil deposition is high. A module is cooler when soil deposition is low.

In addition to above observations, ΔT varies with seasons and irradiance. ΔT is low in the morning (9:00-9:30am), higher in the afternoon (1:00-1:30pm) and highest late in the afternoon (3:00-3:30pm). ΔT is positive in fall and winter and it is negative in spring and summer.

In a dry spell, soil deposition goes on increasing and soiling loss factor (SLF) goes on decreasing. A temperature analysis was carried in dry spells. In winter, dry spell, a soiled module is cooler than a clean module. However, in late spring and summer, soiled module is hotter than a clean module. In further analysis of temperature comparison with time from morning to afternoon, the same observations are seen. In Arizona, typically, the modules are installed at about 10° of tilt angle. For this condition, a clean module is hotter by (1 ± 1° C) than a soiled module in winter. However, a clean module is cooler by (2 ± 1° C) than a soiled module in summer.

REFERENCES

[1] B. King1, G. T. (n.d.). *Regional Soiling Stations for PV: Design, Calibration and Installation.*

[2] Douglas C. Montgomery, C. L. (n.d.). *Introduction to Time Series Analysis and Forecasting.*

[3] G. TamizhMani1, B. K. (n.d.). *Regional Soiling Stations for PV: Soling Loss Analysis.*

[4] http://robjhyndman.com/talks/RevolutionR/10-Seasonal-ARIMA.pdf. (n.d.).

[5] http://stattrek.com/regression/slope-test.aspx?Tutorial=AP. (n.d.).

[6] http://www.nrel.gov/docs/fy12osti/51664.pdf. (n.d.).

[7] http://www.nrel.gov/docs/fy12osti/51664.pdf. (n.d.).

[8] (n.d.). *https://en.wikipedia.org/wiki/Smoothing.*

[9] https://onlinecourses.science.psu.edu/stat510/node/67. (n.d.).

[10] (n.d.). *https://www.youtube.com/watch?v=yBOqUVQ5zzA.*

[11] J. Belmont, K. O. (n.d.). *Degradation Rate Evaluation of 26-Year-Old 200kW Power Plant in a Hot-Dry Desert Climate.*

[12] Pavgi, A. (2016). *Temperature Coefficients and Thermal Uniformity Mapping of PV Modules and Plants.*

[13] Raupp, C. (2016). *Climate-Specific Degradation Rate and Linearity Analysis of Photovoltaic Power Plants.*

[14] Sundarajan, P. (n.d.). *Defects and Statistical Degradation Analysis of Photovoltaic Power Plants.*

[15] TamizhMani, G. (n.d.). *Reliability and standards of Photovoltaics.*

[16] Umachandran, N. (n.d.). *Spatial Temperature Uniformity and Statistical Degradation of Dominant Degradation Modes in PV Modules.*

APPENDIX A

SEASONAL ARIMA MODELS FOR OTHER SYSTEMS

Model: Seasonal ARIMA(0, 0, 0)(1, 0, 0)12

Model Summary

DF	66	Stable	Yes
Sum of Squared Errors	476110.486	Invertible	Yes
Variance Estimate	7213.79525		
Standard Deviation	84.9340641		
Akaike's 'A' Information Criterion	810.134932		
Schwarz's Bayesian Criterion	814.573947		
RSquare	0.50896569		
RSquare Adj	0.50152578		
MAPE	10.473287		
MAE	72.4549773		
-2LogLikelihood	806.134932		

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|	Constant	Mu
AR2,12	2	12	0.77671	0.06650	11.68	<.0001*	Estimate	774.462598
Intercept	1	0	774.46260	30.52681	25.37	<.0001*	172.929534	

JMP output of seasonal ARIMA for PV system on building of Barret Honors College

Model: Seasonal ARIMA(0, 0, 0)(0, 1, 1)12

Model Summary

DF	46	Stable	Yes
Sum of Squared Errors	190156.425	Invertible	Yes
Variance Estimate	4133.83532		
Standard Deviation	64.2949089		
Akaike's 'A' Information Criterion	557.176195		
Schwarz's Bayesian Criterion	560.918597		
RSquare	0.88821553		
RSquare Adj	0.88578543		
MAPE	4.23143689		
MAE	57.912402		
-2LogLikelihood	553.176195		

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|	Constant	Mu
MA2,12	2	12	0.99987	0.315442	3.17	0.0027*	Estimate	-51.452848
Intercept	1	0	-51.45285	5.745322	-8.96	<.0001*	-51.452848	

JMP output of seasonal ARIMA for PV system on Packard Parking lot

Model: Seasonal ARIMA(0, 0, 0)(0, 1, 1)12

Model Summary

DF	46	Stable	Yes
Sum of Squared Errors	62455.5659	Invertible	Yes
Variance Estimate	1357.72969		
Standard Deviation	36.8473838		
Akaike's 'A' Information Criterion	492.603422		
Schwarz's Bayesian Criterion	496.345824		
RSquare	0.8589229		
RSquare Adj	0.855856		
MAPE	4.3556149		
MAE	29.27634		
-2LogLikelihood	488.603422		

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
MA2,12	2	12	0.715452	0.293398	2.44	0.0187*
Intercept	1	0	-6.728296	3.171201	-2.12	0.0393*

Constant Estimate	Mu
-6.728296	-6.728296

JMP output of seasonal ARIMA for PV system on building of Weathercup center

Model: Seasonal ARIMA(0, 0, 2)(1, 1, 0)12

Model Summary

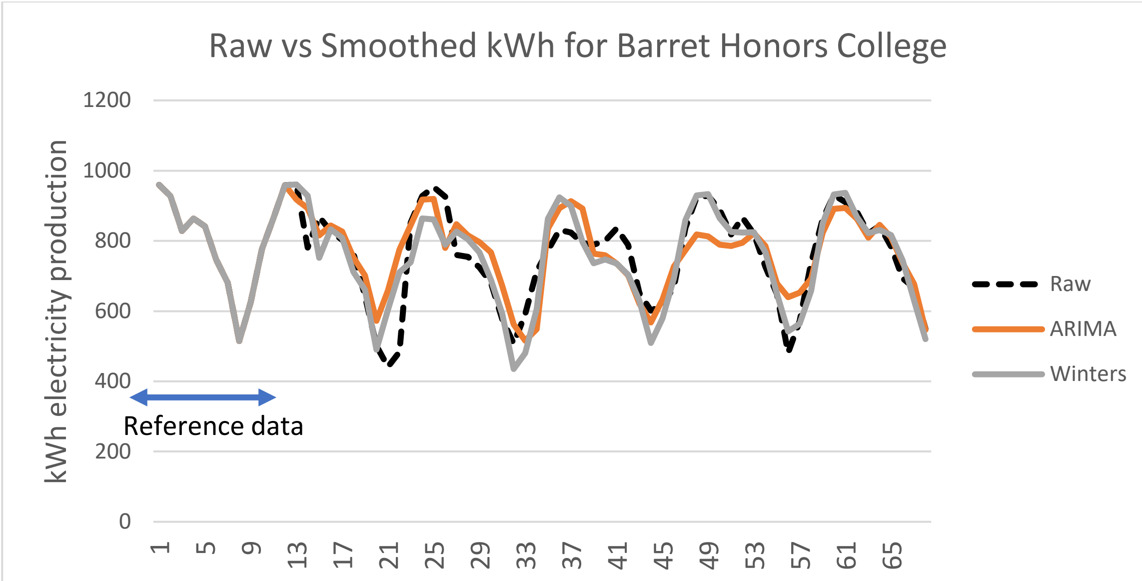
DF	56	Stable	Yes
Sum of Squared Errors	8561.89088	Invertible	Yes
Variance Estimate	152.890909		
Standard Deviation	12.3649063		
Akaike's 'A' Information Criterion	481.963616		
Schwarz's Bayesian Criterion	490.340994		
RSquare	0.62403503		
RSquare Adj	0.60389405		
MAPE	9.45475201		
MAE	9.64871561		
-2LogLikelihood	473.963616		

Parameter Estimates

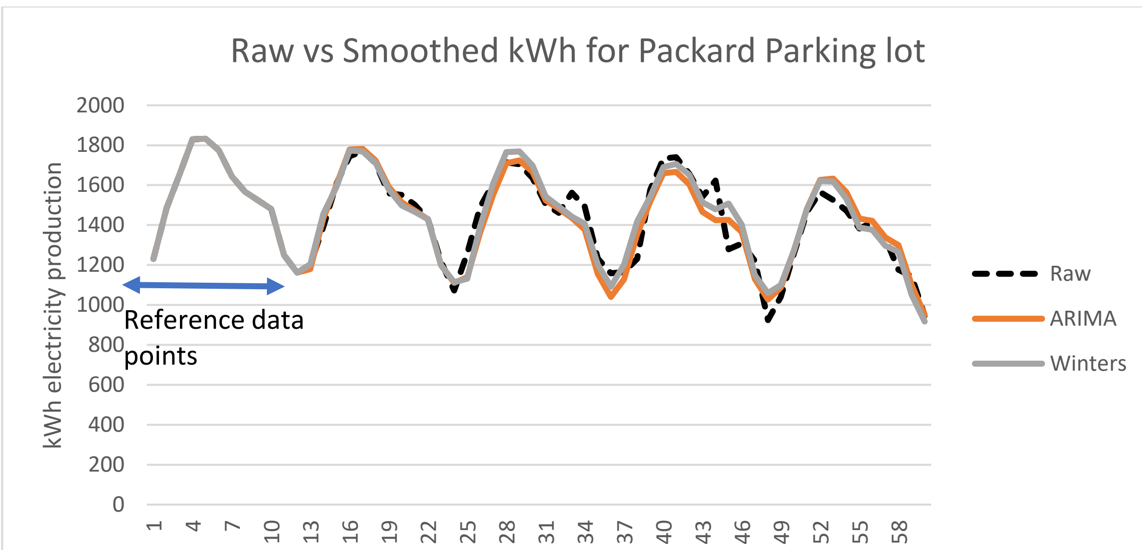
Term	Factor	Lag	Estimate	Std Error	t Ratio	Prob>|t|
AR2,12	2	12	-0.585994	0.122118	-4.80	<.0001*
MA1,1	1	1	-0.427393	0.118505	-3.61	0.0007*
MA1,2	1	2	-0.654406	0.103058	-6.35	<.0001*
Intercept	1	0	-5.140206	2.163637	-2.38	0.0210*

Constant Estimate	Mu
-8.1523356	-5.1402063

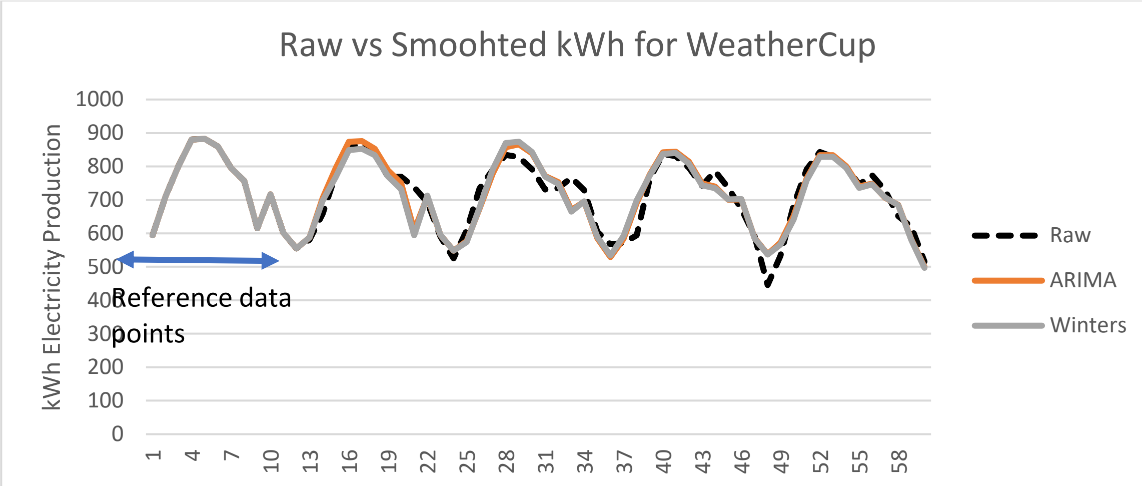
JMP output of seasonal ARIMA for PV system on building of Wrigley hall



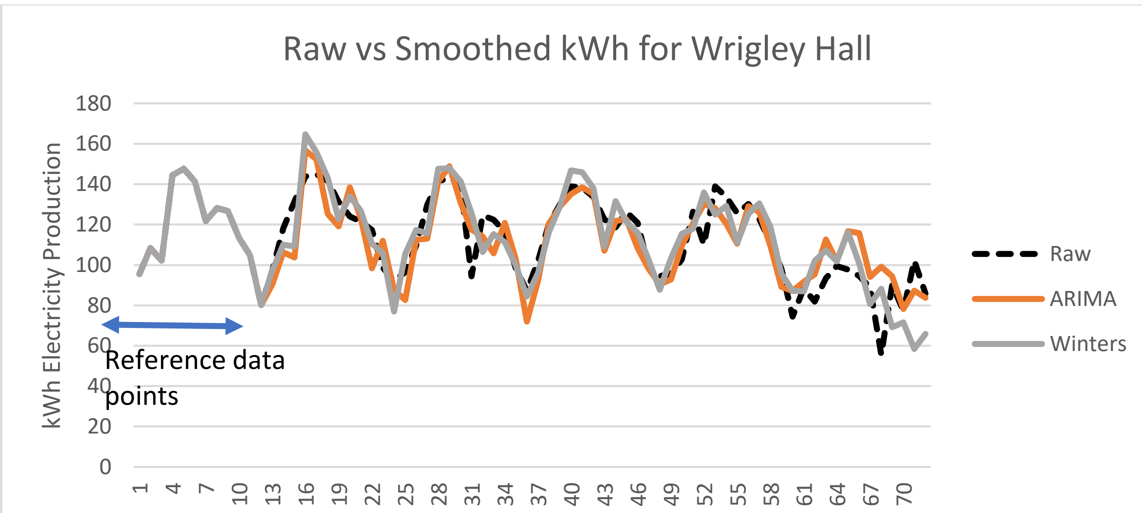
Comparison of raw kWh, modelled ARIMA and Winters for Barret Honors College system



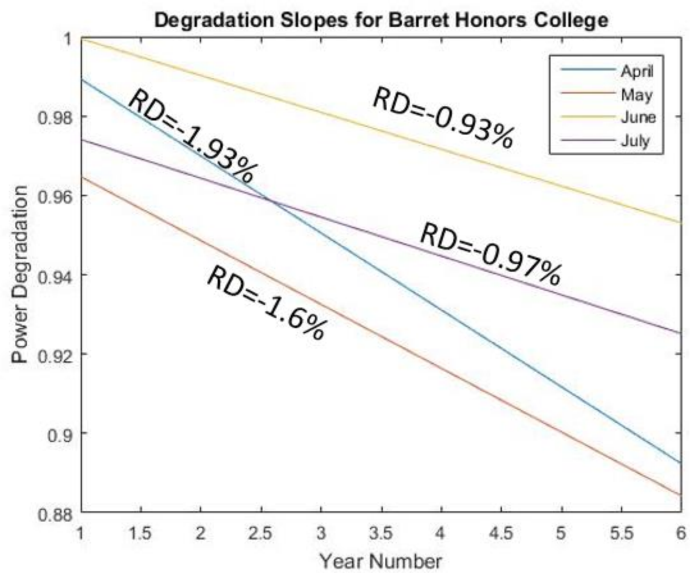
Comparison of raw kWh, modelled ARIMA and Winters for Packard Parking lot system



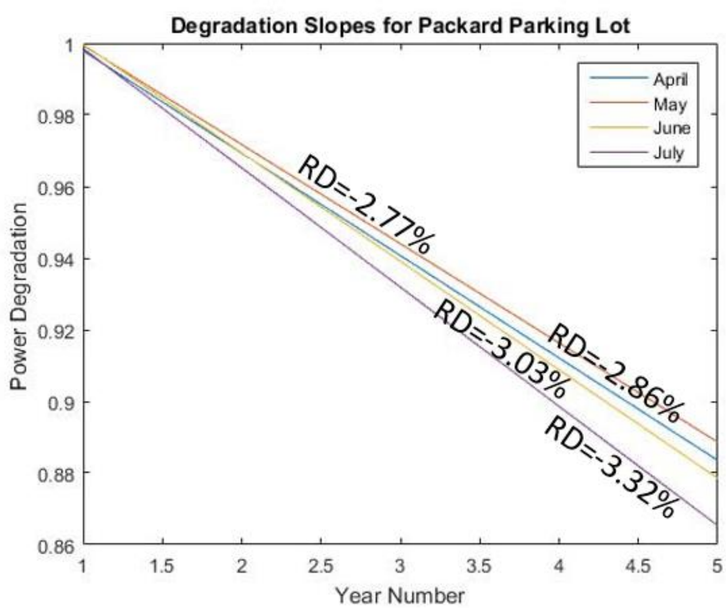
Comparison of raw kWh, modelled ARIMA and Winters for WeatherCup center system



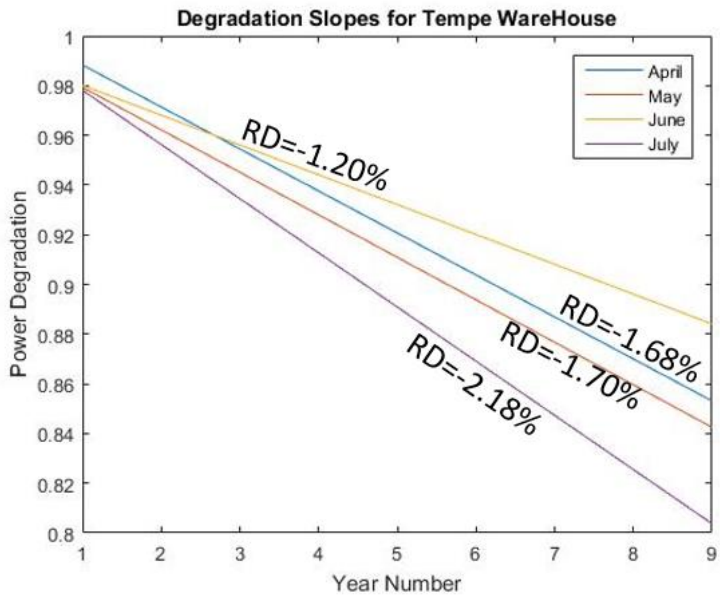
Comparison of raw kWh, modelled ARIMA and Winters for WeatherCup center system



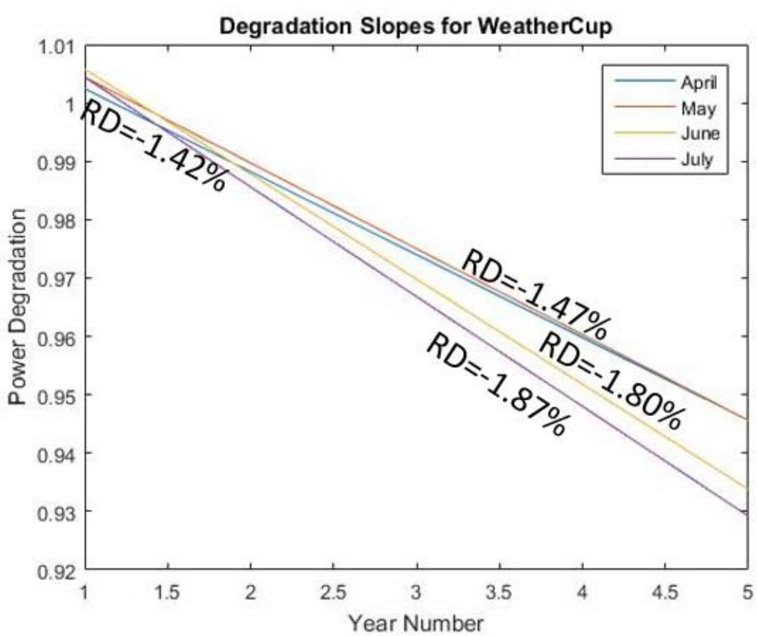
Degradation slopes for Barret Honors college



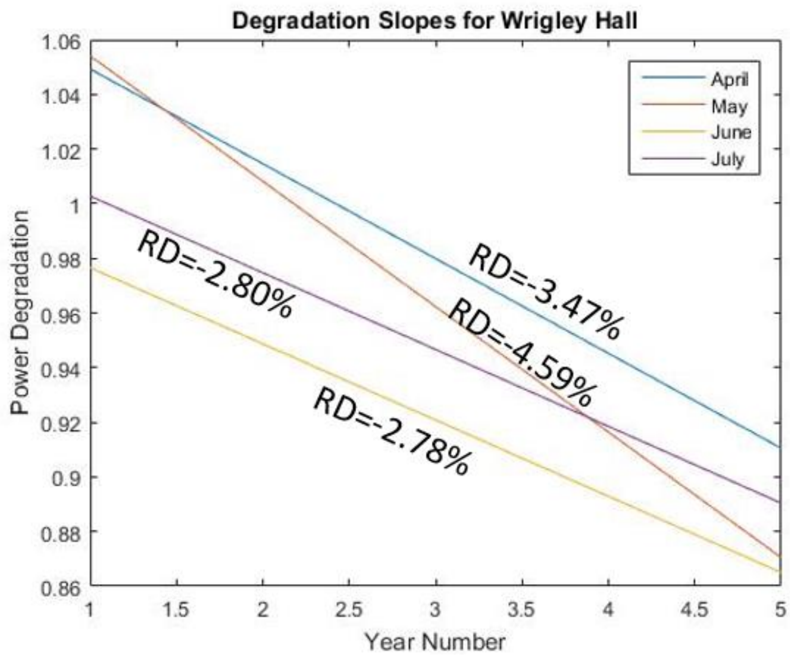
Degradation slopes for Packard Parking lot



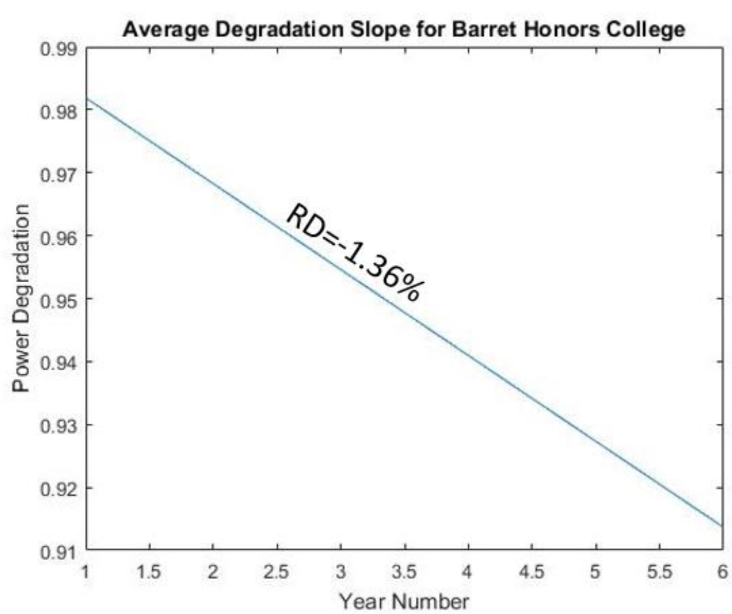
Degradation slopes for Tempe Warehouse



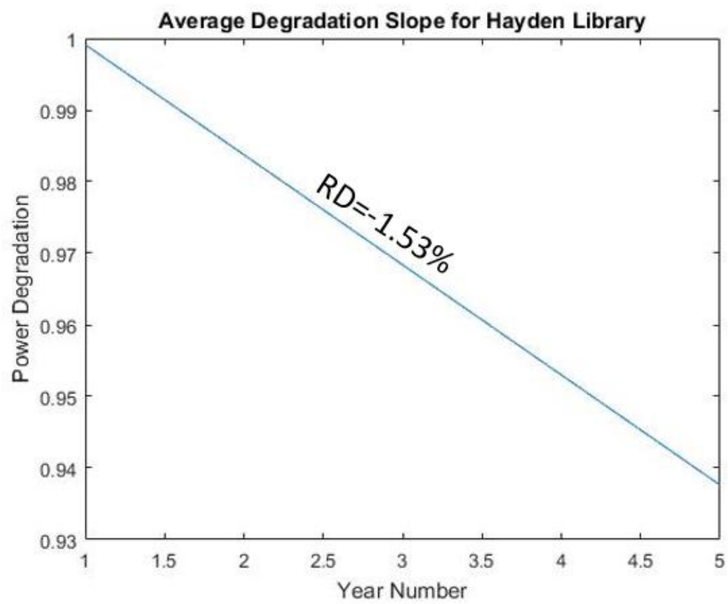
Degradation slopes for WeatherCup center



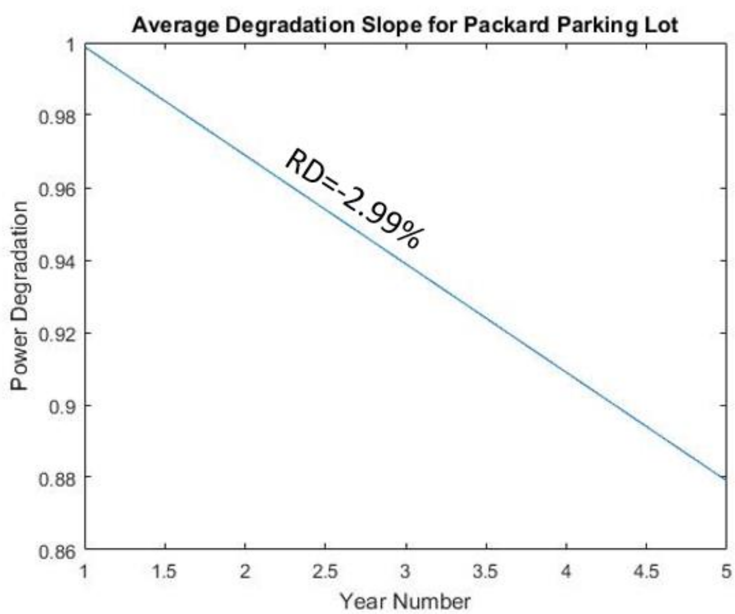
Degradation slopes for Wrigley Hall



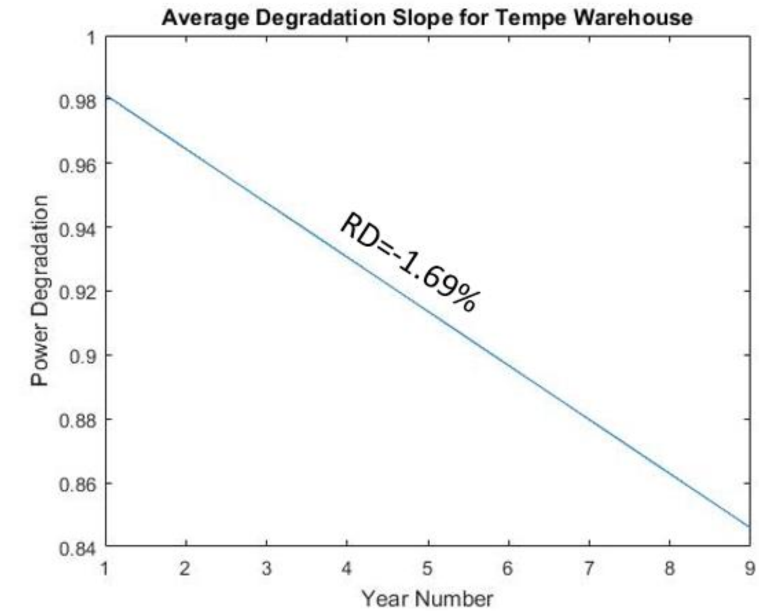
Average yearly degradation for Barret Honors college



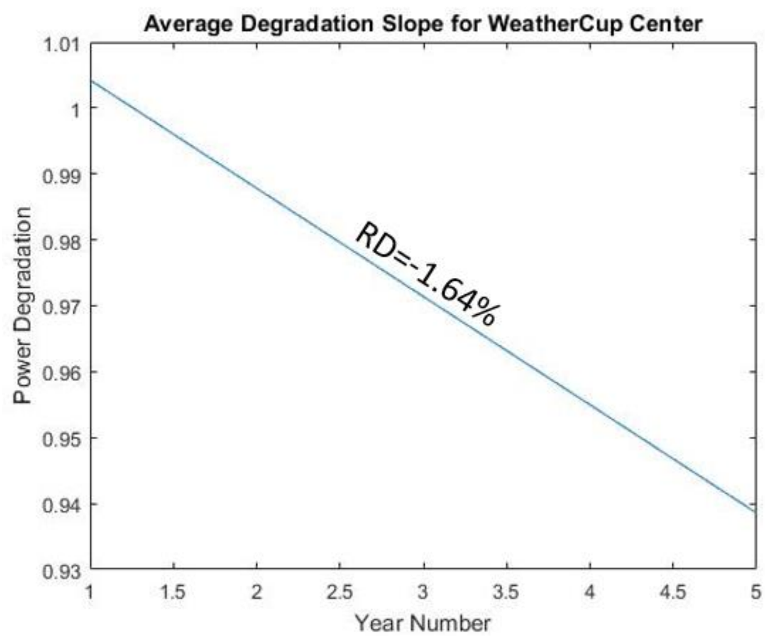
Average yearly degradation for Hayden Library



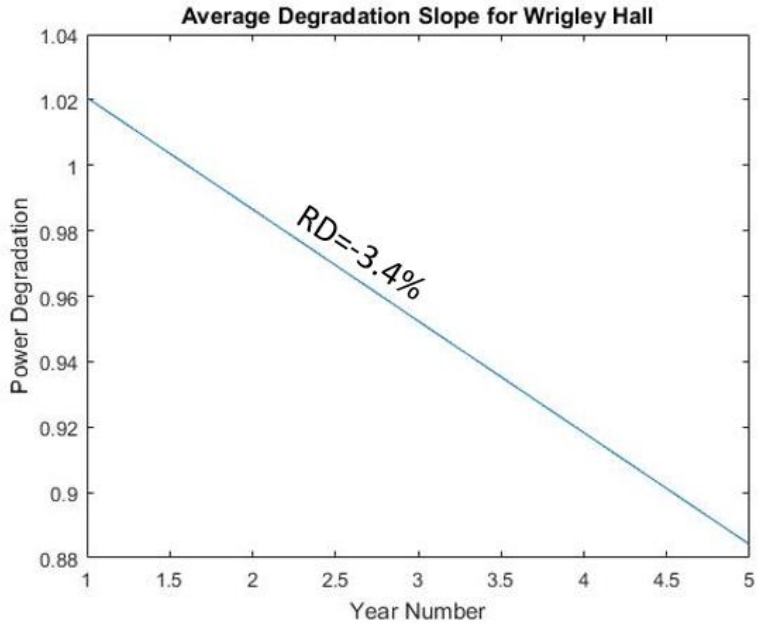
Average yearly degradation for Packard Parking lot



Average yearly degradation for Tempe warehouse



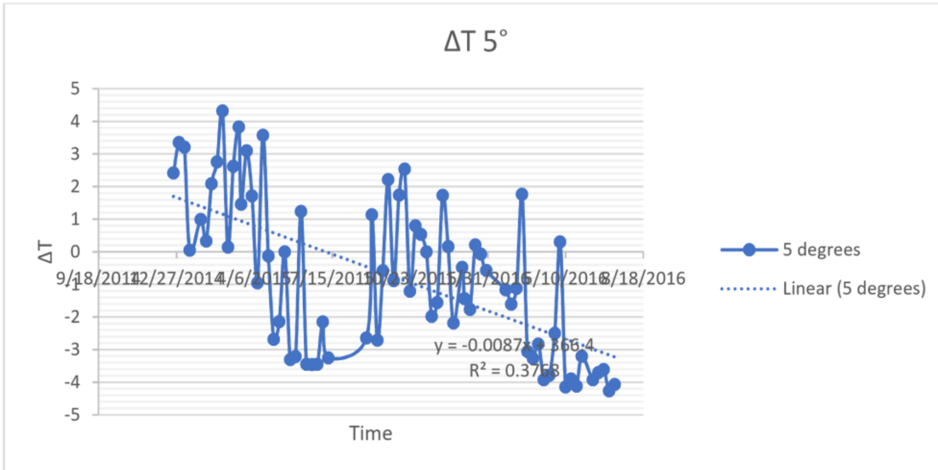
Average degradation for WeatherCup center



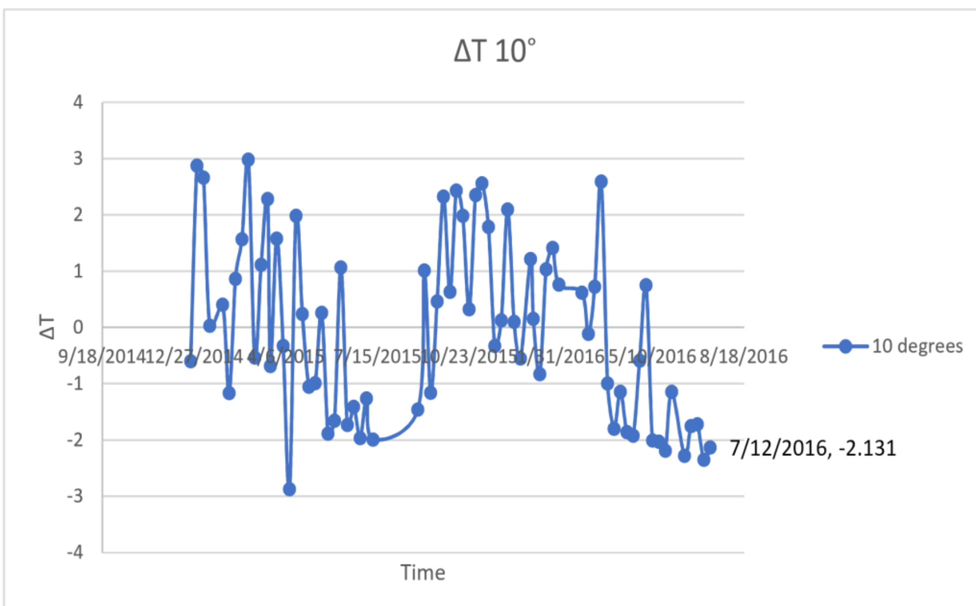
Average degradation for Wrigley Hall

APPENDIX B

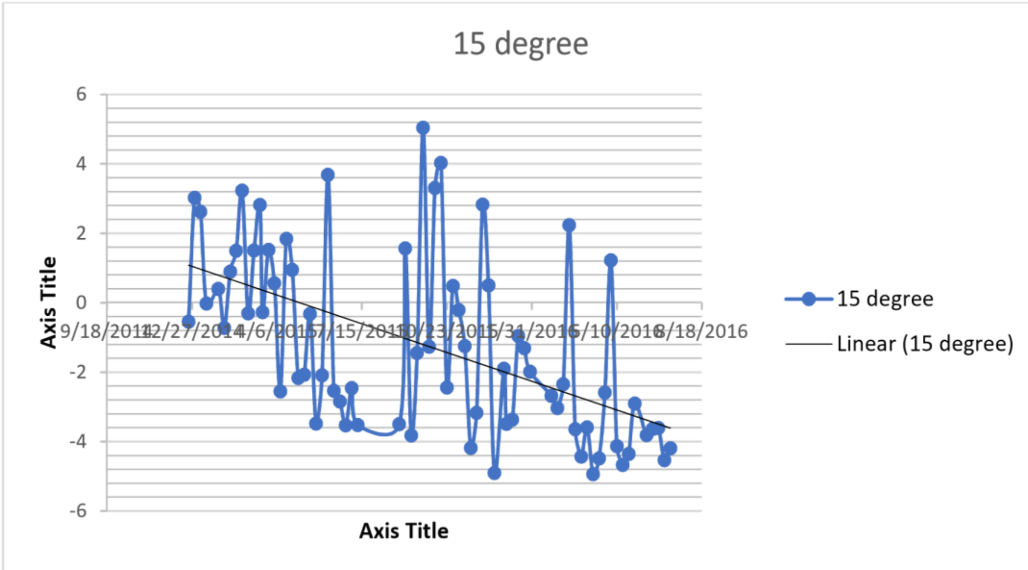
ΔT, SLF AND SEASONAL PLOTS OF MEDIAN ΔT AND ANGLE OF INCIDENCE



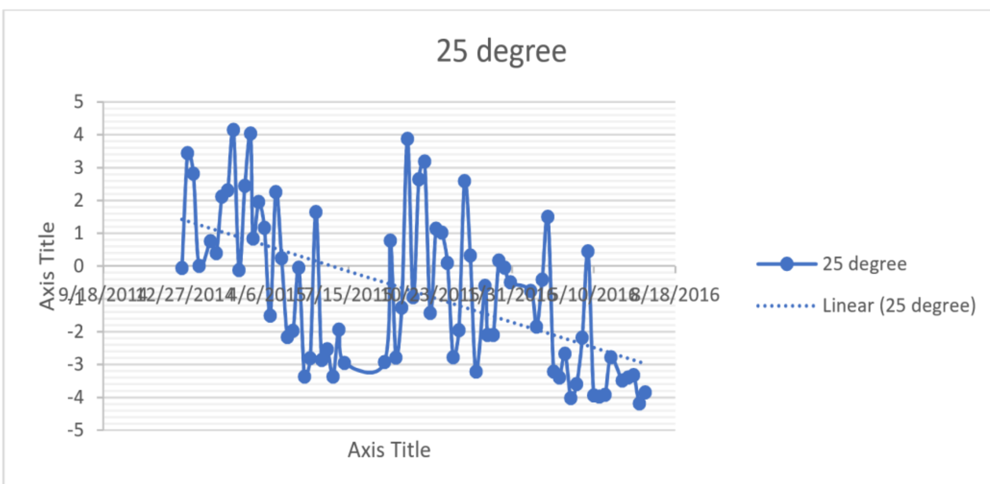
Time series plot of ΔT for 5° of tilt angle



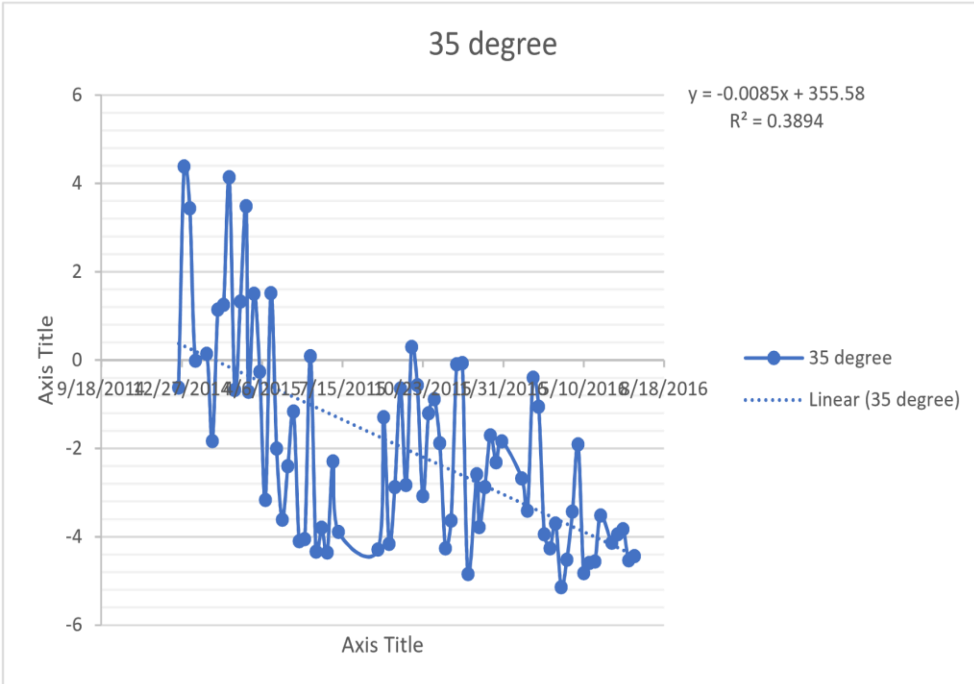
Time series plot of ΔT for 10° of tilt angle



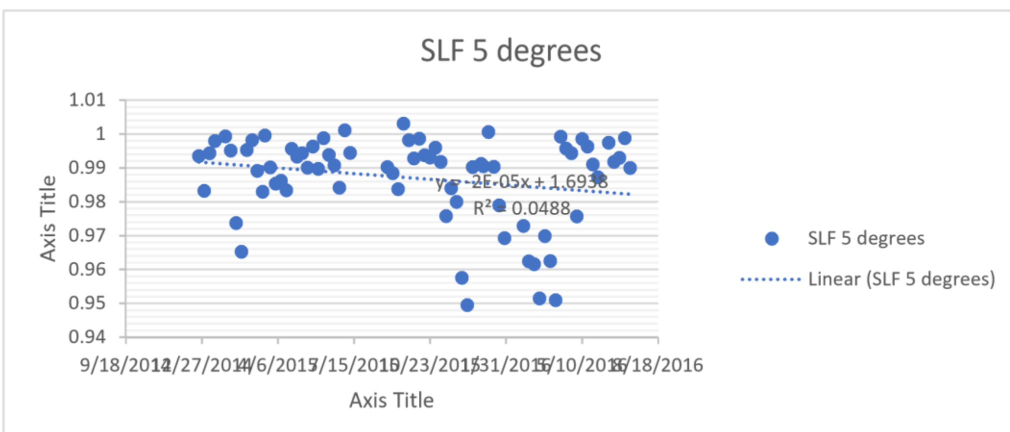
Time series plot of ΔT for 15° of tilt angle



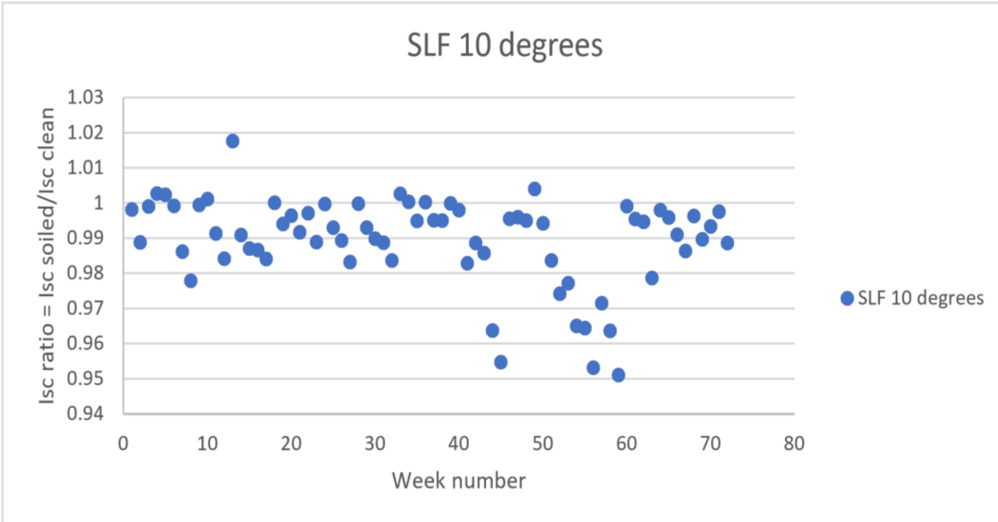
Time series plot of ΔT for 20° of tilt angle



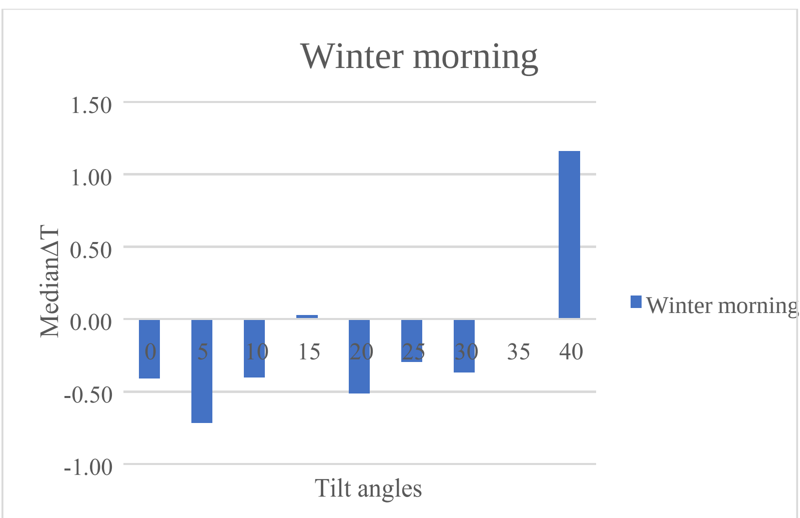
Time series plot of ΔT for 35° of tilt angle



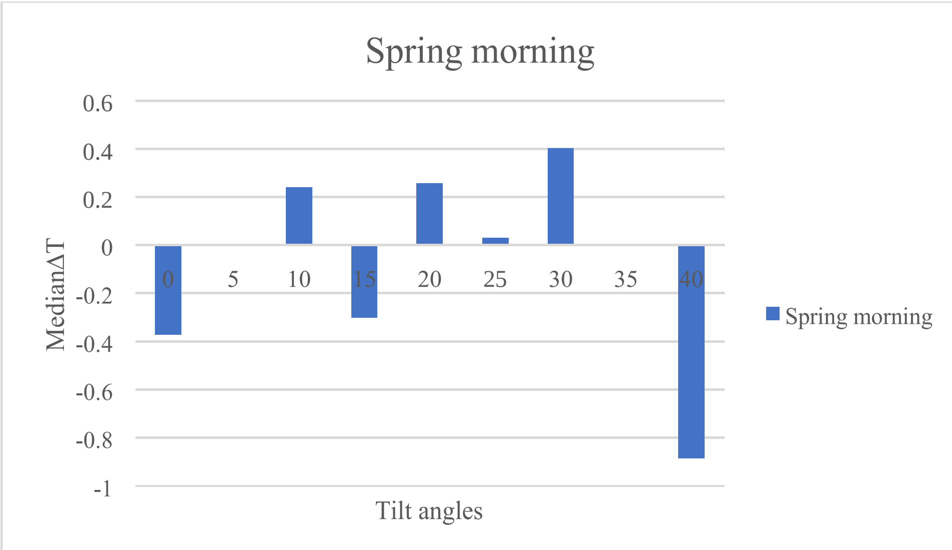
Time series plot of SLF for 5° of tilt angle



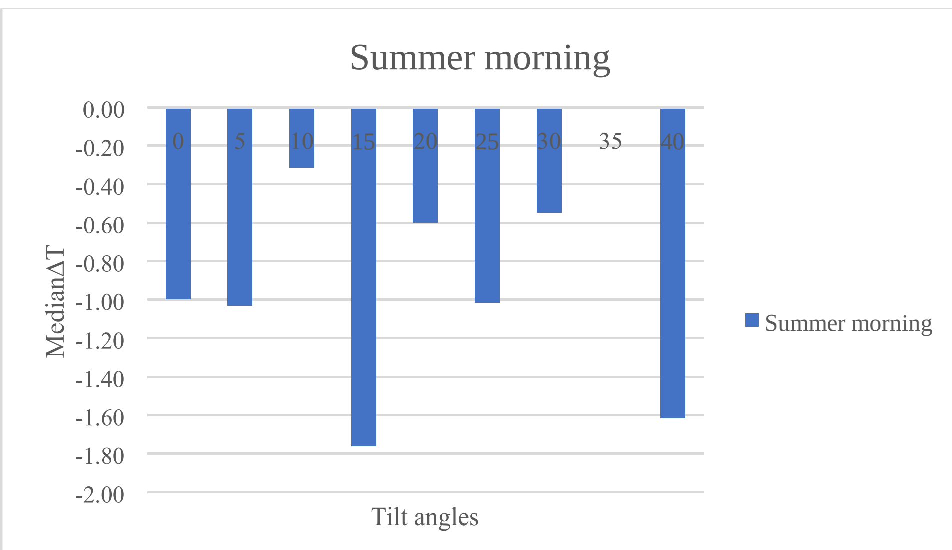
Time series plot of SLF for 10° of tilt angle



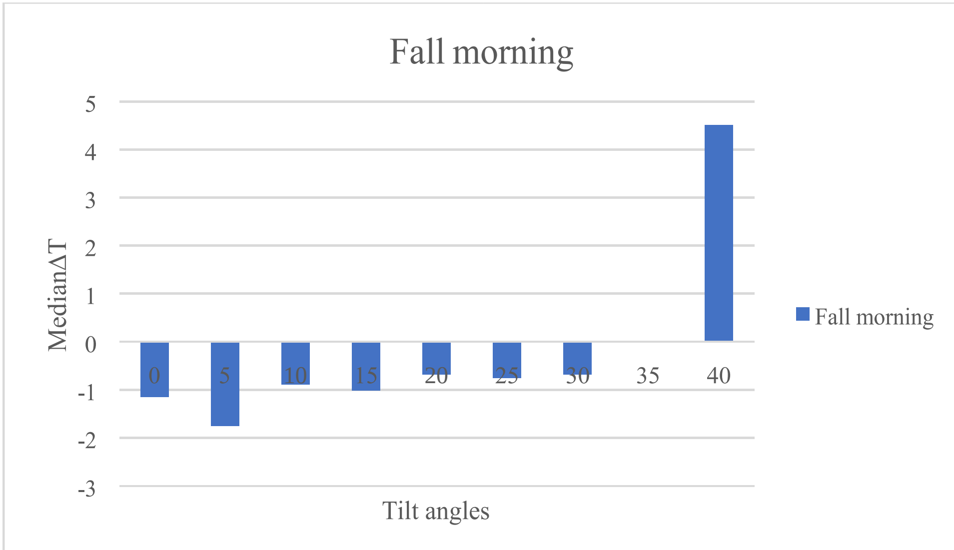
Median ΔT versus tilt angles for morning irradiance in winter



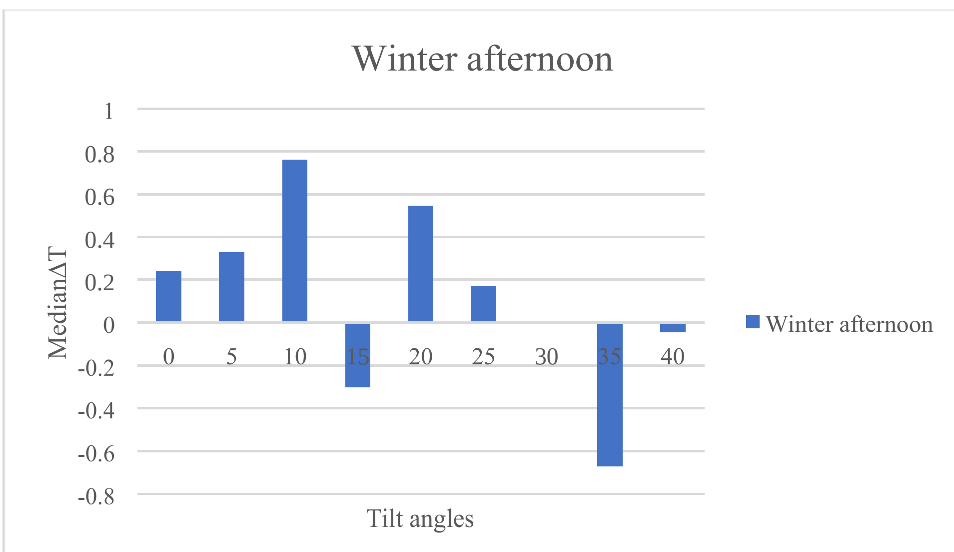
Median ΔT versus tilt angles for morning irradiance in spring



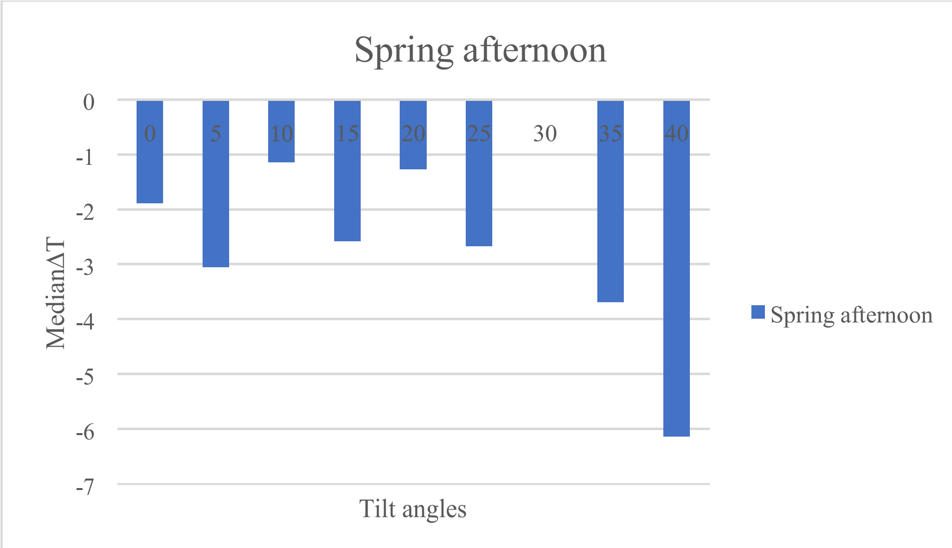
Median ΔT versus tilt angles for morning irradiance in summer



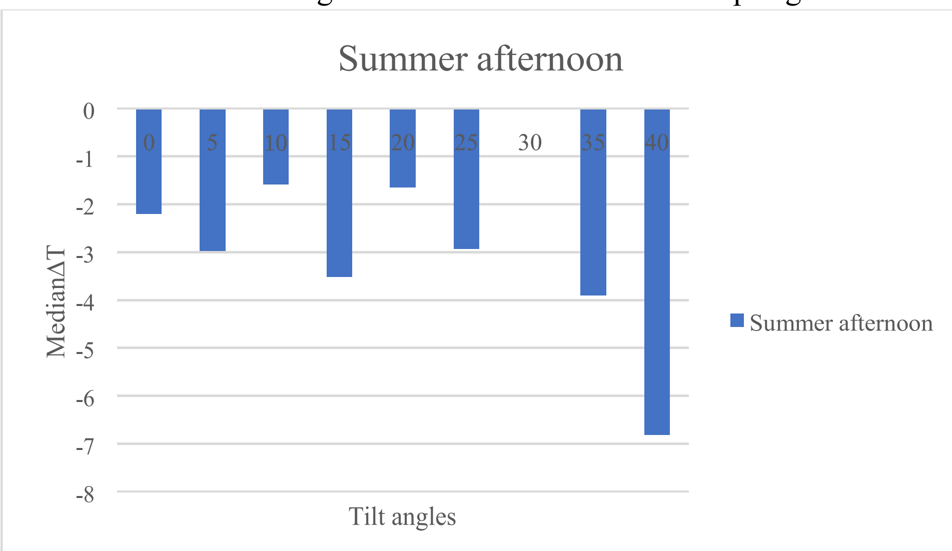
Median ΔT versus tilt angles for morning irradiance in fall



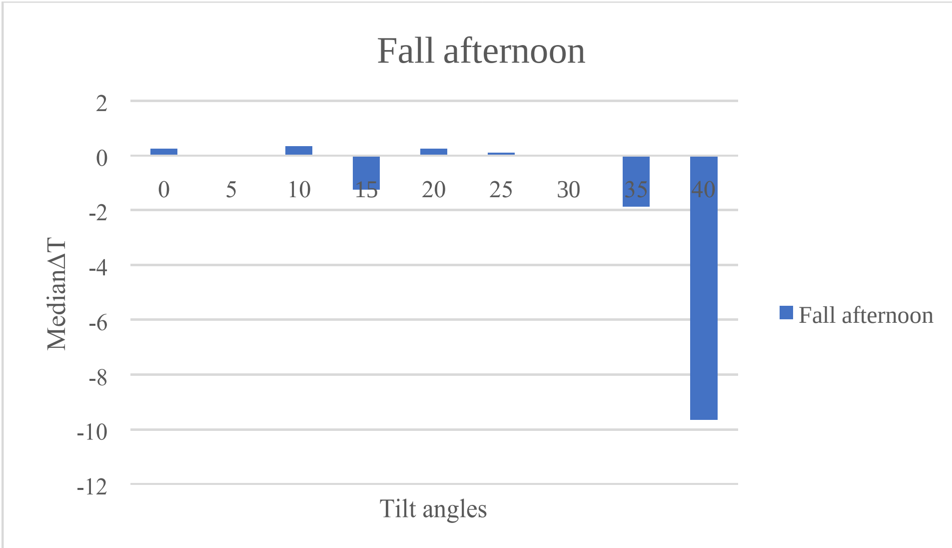
Median ΔT versus tilt angles for afternoon irradiance in winter



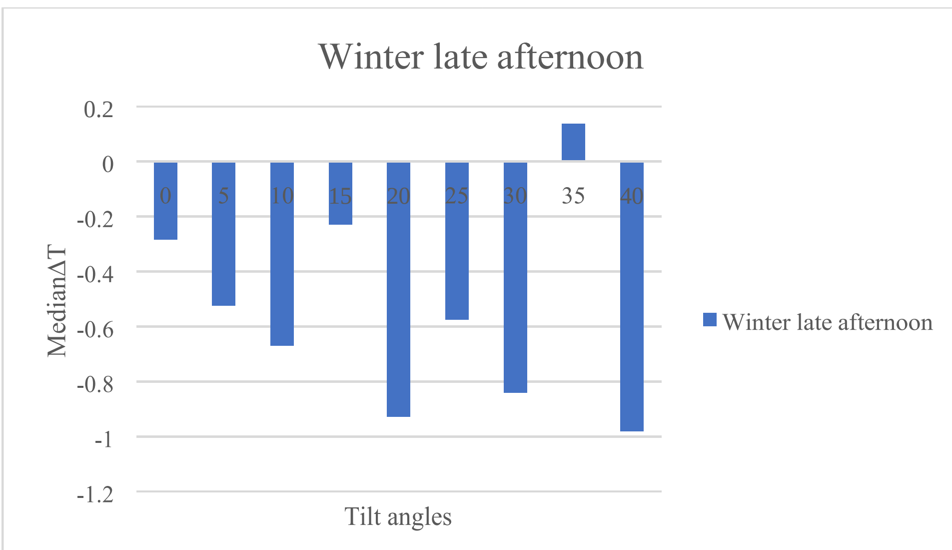
Median ΔT versus tilt angles for afternoon irradiance in spring



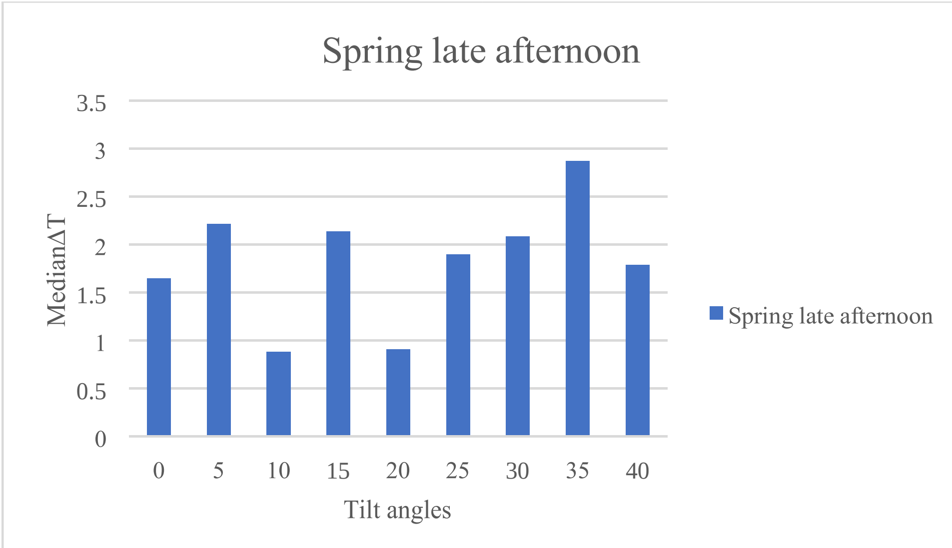
Median ΔT versus tilt angles for afternoon irradiance in summer



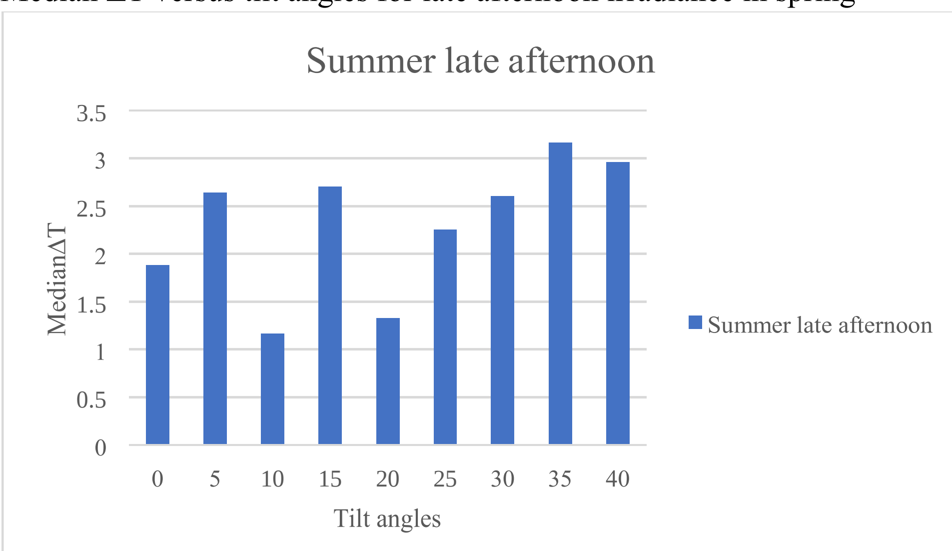
Median ΔT versus tilt angles for afternoon irradiance in fall



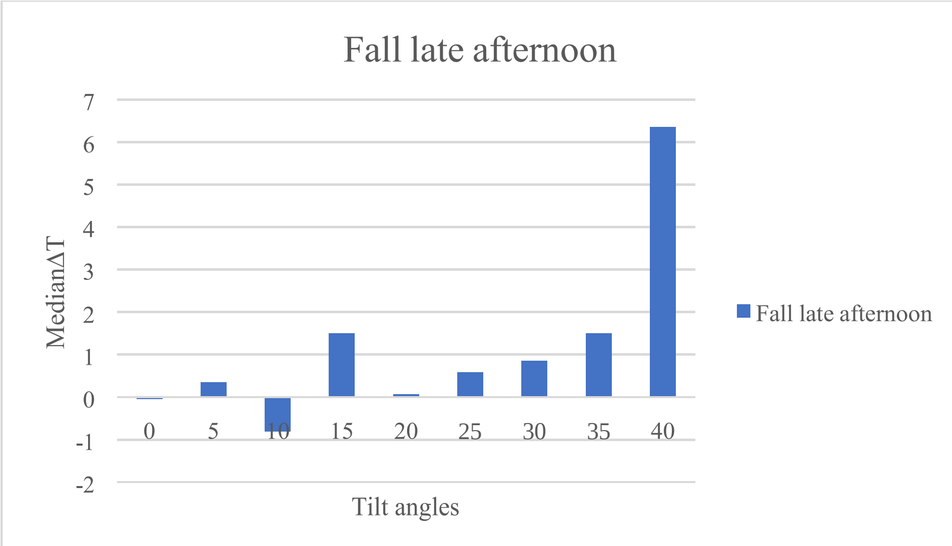
Median ΔT versus tilt angles for late afternoon irradiance in winter



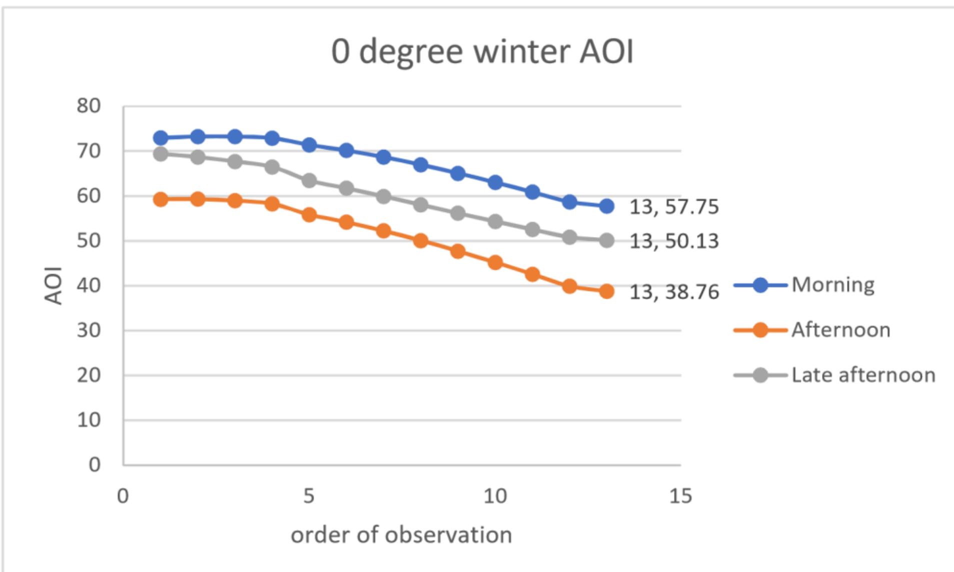
Median ΔT versus tilt angles for late afternoon irradiance in spring



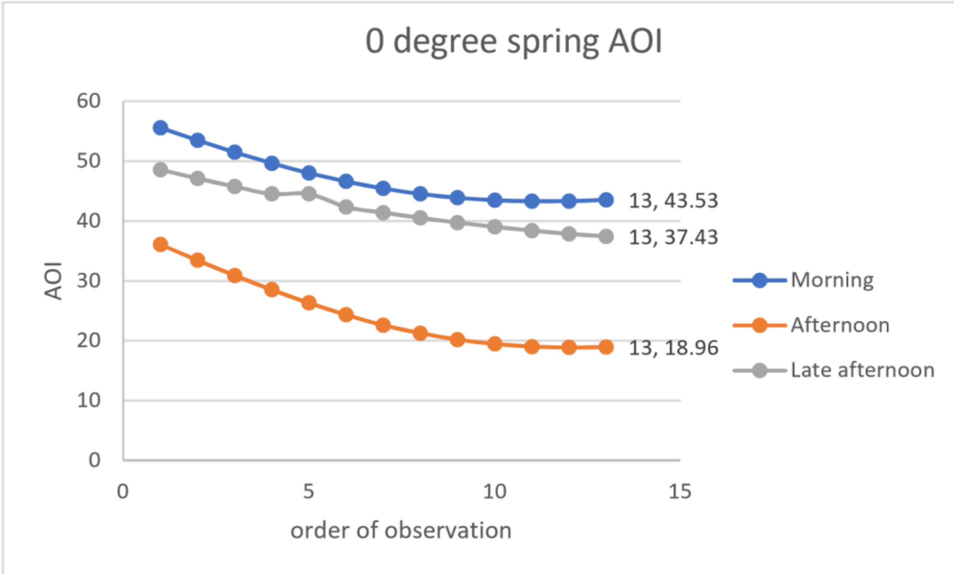
Median ΔT versus tilt angles for late afternoon irradiance in summer



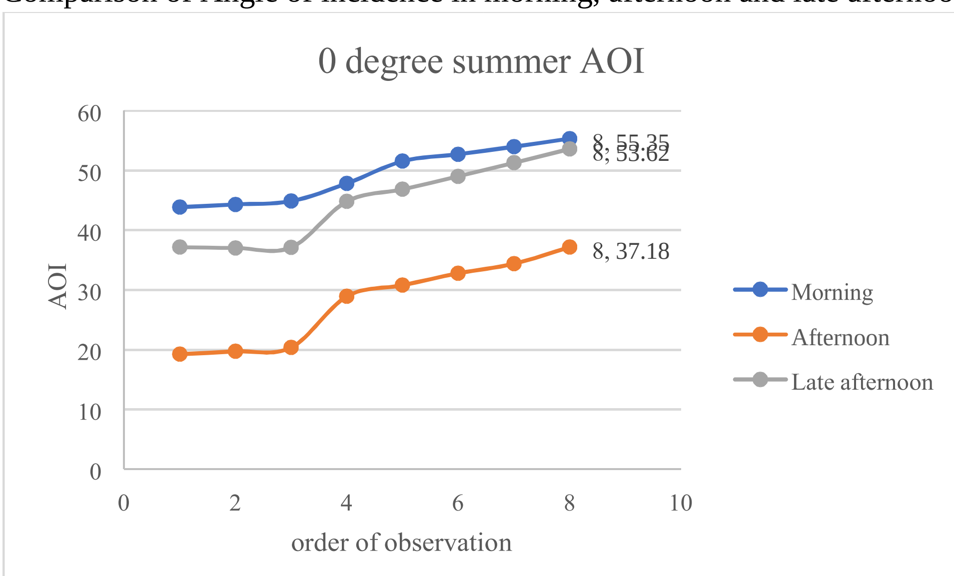
Median ΔT versus tilt angles for late afternoon irradiance in fall



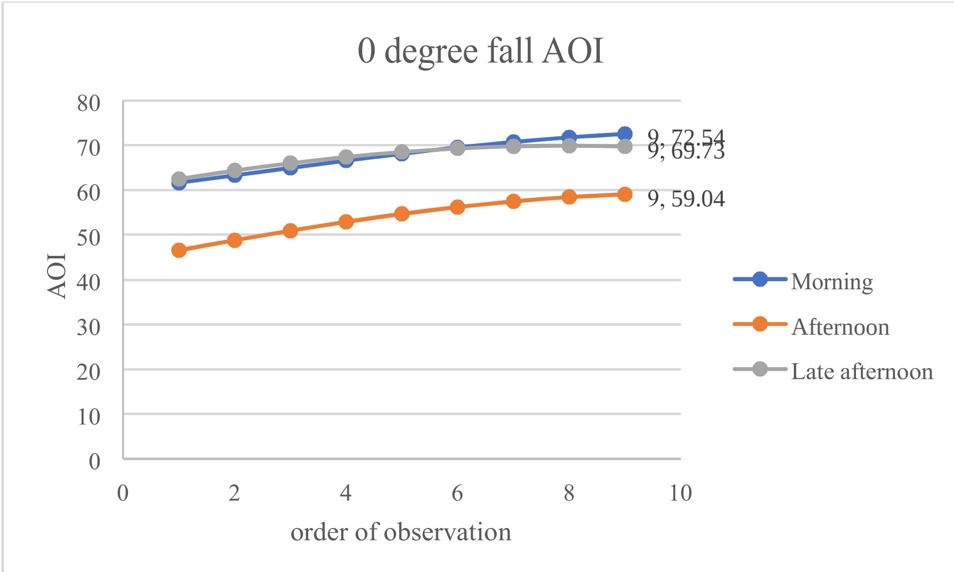
Comparison of Angle of incidence in morning, afternoon and late afternoon in winter



Comparison of Angle of incidence in morning, afternoon and late afternoon in spring



Comparison of Angle of incidence in morning, afternoon and late afternoon in summer



Comparison of Angle of incidence in morning, afternoon and late afternoon in fall