

Linear Model Summaries

Model 1

Predictors	Estimate	Std. Error	t value	p-value
Intercept	990.104	108.663	9.112	<2e-16 ***
GHI	16.782	0.421	39.861	< 2e-16 ***
Temp	-59.673	3.728	-16.006	< 2e-16 ***
WS	-15.854	22.093	-0.718	0.473

Note. GHI = Global Horizontal Irradiance, Temp = Ambient Temperature (C), WS = Wind Speed (mph). *** indicates statistical significance. Model is fit to predict total output (kWh).

Residual standard error: 547.6 on 361 degrees of freedom

Multiple R-squared: 0.8394, Adjusted R-squared: 0.838

F-statistic: 628.8 on 3 and 361 DF, p-value: < 2.2e-16

Model 2

Predictors	Estimate	Std. Error	t value	p-value
Intercept	389.89375	68.21317	5.716	2.28e-08 ***
GHI	20.65018	0.52493	39.339	< 2e-16 ***
GHI:Temp	-0.30821	0.01648	-18.702	< 2e-16 ***

Note. GHI = Global Horizontal Irradiance, Temp = Ambient Temperature (C), : denotes interaction (product of two variables). *** indicates statistical significance. Model is fit to predict total output (kWh).

Residual standard error: 509.9 on 362 degrees of freedom

Multiple R-squared: 0.8603, Adjusted R-squared: 0.8596

F-statistic: 1115 on 2 and 362 DF, p-value: < 2.2e-16

Model 3

Predictors	Estimate	Std. Error	t value	p-value
Intercept	1024.1789	82.8750	12.36	<2e-16 ***
GHI	12.4247	0.4013	30.96	< 2e-16 ***

Note. GHI = Global Horizontal Irradiance. *** indicates statistical significance. Model is fit to predict total output (kWh).

Residual standard error: 714 on 363 degrees of freedom

Multiple R-squared: 0.7254, Adjusted R-squared: 0.7246

F-statistic: 958.8 on 1 and 363 DF, p-value: < 2.2e-16

Takeaways: Basing our decision off of Multiple R-squared (correlation coefficient) and lowest p-values, we can deduce that Model 2 with the highest R² of 0.8603 is the best model to use going forward. The coefficients for this model are 389.89375 for intercept, 20.65018 for GHI and -0.30821 for the interaction effect of GHI and Temp. This can be written into a linear equation of $y = 389.89375 + 20.65018(\text{GHI}) - 0.30821(\text{GHI:Temp})$. Where y=kWh, and GHI:Temp is the product of the two.

Summary Stats on Predicted vs Observed Values (Using Model 2)

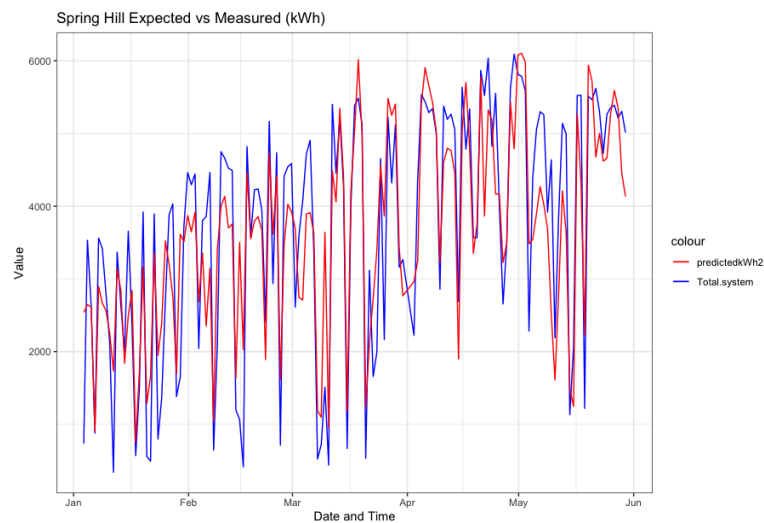
Five number summary stats:

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Energy Ratio (monthly)	0.9408	0.9544	1.0393	1.0168	1.0537	1.0956
kWh Predicted (monthly)	68545	93245	111367	104887	125447	125833
kWh Observed (monthly)	64485	96912	106291	107547	132188	137858

Model 2 predicted values summary stats:

	RMSE	R
Model 2	7010.468	0.9827195

Monthly Expected vs Measured (kWh)



Note. predictedkWh2 = Monthly sum predicted kWh. Total.system = TotalOutput observed kWh.

Takeaways: As we can see from the above plots and five number summary statistics, the predicted/expected kWh monthly values fit exceptionally well to the observed ones. The high R value indicates a strong linear relationship between predicted values and observed ones. The minimum ER of 0.9408 indicates the small average difference between observed and predicted and how well the site is performing.

