University of Utah
UNDERGRADUATE RESEARCH JOURNAL

USING REGRESSION MODELING FOR ACCURATE PREDICTION OF POWER GENERATION IN WEARABLE ENERGY HARVESTERS

Julia Dominesey, Dr. Shad Roundy
Department of Mechanical Engineering

I. Introduction to Wearable Energy Harvesters and Project Goal

The regression models developed predict the power generated from previously designed wrist-worn energy harvesters.

These devices convert the kinetic energy imposed on the device by the user into electrical energy which allows the device to power itself. In this experiment, four different devices were tested: Generation 2 with a spring, Generation 2 without a spring, Generation 3 with a spring and Generation 3 without a spring.

For these designs to be feasible, the amount of power generated needs to meet power requirements. Our goal is to predict power output given excitation imposed by the user. Two different modeling techniques were employed to attempt to accurately predict the amount of power generated: a physics based Ordinary Differential Equation (ODE) model and an empirical, statistically based multi-linear regression model.

The multi-linear regression models are the focus of this research; its accuracy is compared to the accuracy of the ODE model for its validation of use.

II. Physics Based Modelling Approach Overview

The physics based model uses physical parameters specific to the device type such as rotational inertia, number and strength of magnets, number of coils, etc. to create equations that are unique to each device and produce a power estimate as an outcome. The ODE model uses instantaneous values for acceleration, and rotation rate as well as model parameters to simulate the average power dissipated in the load resistor during the steady state time interval. The ODE model uses the entire rotation rate and acceleration traces of each sample to form its simulation.

III. Statistics Based Modelling Approach

a. Method for creating Regression Models

A multiple linear regression equation is created from experimental data to determine if and what significance of relationship an output has with its inputs. Linear regression equations have the same form as a line equation($y = ax+b$, where y is the output and x is the input), except they have multiple independent variables.

When creating the equations for the wearable energy harvesters, the goal was to minimize the number of input variables without sacrificing the accuracy of the equation itself. In order to do this, the MATLAB function “stepwiselm()” and “fitlm” was used. These functions automatically calculate the significance of each variable. The stepwiselm function differs from the fitlm function in that stepwiselm uses forward selection to add or take away a term and calculates a corresponding p-value for the combination of variables. If the calculated p value is less than the threshold p value then the operation stays, ultimately creating a regression equation with the least amount of necessary terms. The fitlm function must be altered manually. The functions’ input is a matrix X (the root mean square acceleration and rotation values in all three axes from a steady state time trace) and a vector y (the measured power output).

Once this equation is created, multiple root mean square acceleration and gyration values can be plugged into the equation to create a predicted power value, these equations are listed in Table 1.

b. Drawbacks to the Regression Model

The linear regression equations constructed are limited in that they are determined based on the means of the independent variables (acceleration and rotation rate in all three axes) and the dependent variable (the amount of power produced). Therefore, the models are not intended for extremely low levels or high levels of excitations (i.e. those outside of the set of data from which the model was generated). These models are not based on a physical understanding of the devices, but rather empirical data.

c. Benefits of the Regression Model

As opposed to the ODE model, the regression models are simple and general. This means that these equations are faster and easier to interpret. ODE models also require a full IMU trace as an input to create its necessary equation, but the regression equations do not. The regression equations only requires averaged data on a few parameters (acceleration excitation in 2 axes for example) which are available for a much larger set of the general population. are easy to communicate and use in multiple different applications in a way that the ODE equations cannot.

IV. Results

a. Methods for Analyzing Predictive Accuracy of each model

After the physics based and statistics based models were created, the accuracy and therefore the validity of each model needed to be compared. To do this, an error analysis was conducted.

The error for each subject was calculated by taking the absolute value of the difference of the actual power generated and subtracting the predicted power generated, then dividing by the actual power. From here, an average error was calculated per design (Generation 2 with a Spring, Generation 2 without a Spring, Generation 3 with a Spring, Generation 3 without a Spring). Simply comparing which error is lower indicates which model is more accurate; however, it does not provide information on whether or not the difference in accuracy is significant.

To answer whether or not this difference is negligible, a two-sided T-test for the difference of two means is conducted. The test calculates whether or not the difference between the two means of error is statistically significant. If the difference is significant, than one technique is likely to be a better predictor than the other, but if the difference is insignificant, than either model could be used with, theoretically, the same level of accuracy. The results of this test are shown in the following section and divided by design.

b. Regression Equations

The following equations were developed from a linear regression function in MATLAB that used data from the experiment discussed earlier. Each equation was derived from the data of subjects at three different walking speeds and two different locations; the three different speeds are 2.5mph, 3.5mph and 5.5mph and locations are the left and right wrist. MATLAB creates a multiple linear regression equation with the root mean square acceleration and rotation rate values during these excitations in those locations and then evaluates the accuracy of the equation as it adds and subtracts variables, resulting in a simplified equation with the most significant variables present. The data, and equations, are divided by design: Generation 2 with a Spring, Generation 2 without a Spring, Generation 3 with a Spring and Generation 3 without a Spring. The equations, as well as the coefficient of determination are listed in the table.

Design	Equation	R^2 Adjusted
Generation 2, Sprung	y = -185.67+30.83Axrms+19.99Ayrms-1.17Gyrms	0.954
Generation 2, Unsprung	y = -17.51+27.172Axrms-0.892Gyrms	0.967
Generation 3, Sprung	y = -45.28 + 26.96Axrms	0.932
Generation 3, Unsprung	y = -81.27 + 32.3Axrms	0.942

Table 1: Regression Equations

c. Table of Results of Error Analysis

The table below, Table 2, is organized by design and compares the respective ODE and Regression errors. The percent errors are reported with the p-value to compare the actual and statistical significance of each result.

Category	ODE Error %	REG Error %	t-test result	p-value
Generation 2, Sprung	38.2683	39.9976*	No Difference in Means	0.8806
Generation 2, Unsprung	89.2782	455.59	Difference in Means: ODE Better	0.027

Generation 3, Sprung	50.5117	70.5767	No Difference in Means	0.322
Generation 3, Unsprung	117.1516	369.76	Difference in Means: ODE Better	0.0032

Table 2: ODE and Regression predicted power errors by Design

Table 3 takes the same approach as was used to create Table 2, but divides the categories again by speed. High speed is the data where the subject was jogging at 5.5 mph, low speed is the data where a subject is walking at either 2.5mph or 3.5mph, all taken from the left or right wrist.

Category	ODE Error %	REG Error %	t-test result	p-value
Generation 2, Sprung, High	17.2955	11.1796	No Difference	0.3621
Generation 2, Sprung, Low	49.3067	132.2374	No Difference	0.1475
Generation 2, Unsprung, High	43.8739	8.4358	Difference:REG Better	0.000000937
Generation 2, Unsprung, Low	113.1752	679.1729	Difference:ODE Better	0.0186
Generation 3, Sprung, High	17.862	15.0553	No Difference	0.613
Generation 3, Sprung, Low	66.8366	98.3373	No Difference	0.2581
Generation 3, Unsprung, High	61.2906	49.6771	No Difference	0.843
Generation 3, Unsprung, Low	145.0821	529.8146	Difference:ODE Better	0.000585

Table 3: ODE and Regression predicted power errors by Design and Speed

*The Generation 2 with Spring regression error has two outliers removed from the original dataset. These outliers were removed because they skewed the error by more than 50% and their traces, Figure 7, were abnormal compared to other traces of their same type.

d. Graphs of Error Analysis

The first two figures in this section compare the predicted power output and the actual power output of all of the trials with devices with springs in them.

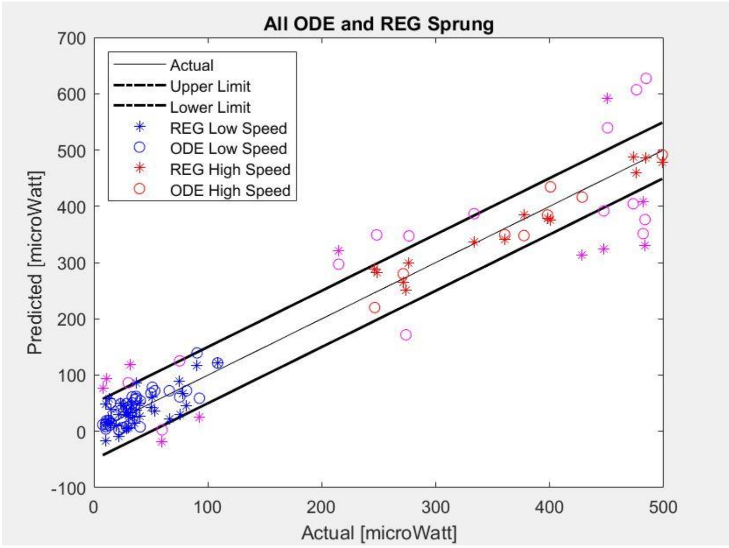

Figure 1: Sprung Device, Predicted vs. Actual Power Output

The legend in the left-hand corner differentiates the symbols that appear on the graph. The values that are colored magenta are outside of the upper limit or lower limit. This boundary was determined arbitrarily but represents a +/- 50 microwatt range. The same conventions are used for Figure 2, below.

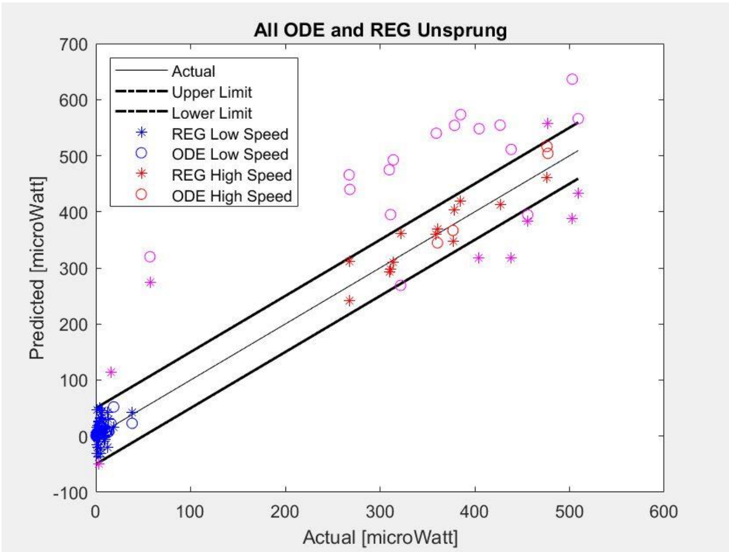

Figure 2: Unsprung Device, Predicted vs. Actual Power Output

Figures 3 through 7 plot the absolute percent error of the ODE and regression models.

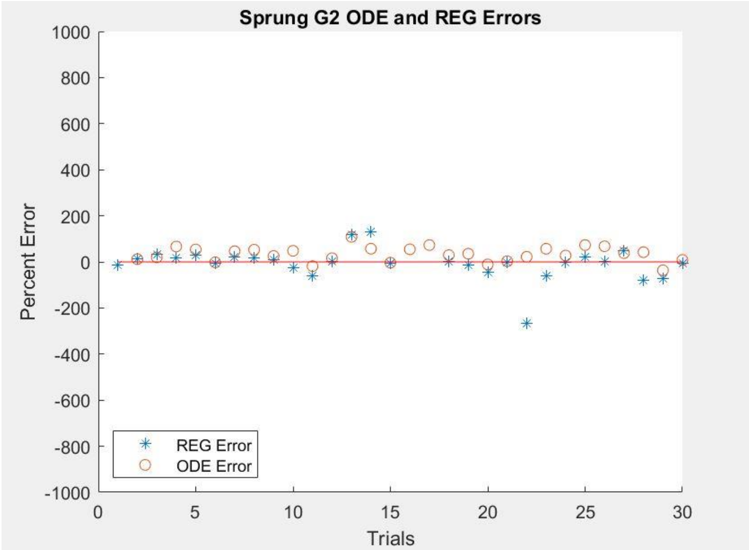

Figure 3: Generation 2, Sprung Device ODE and Regression Error Plot

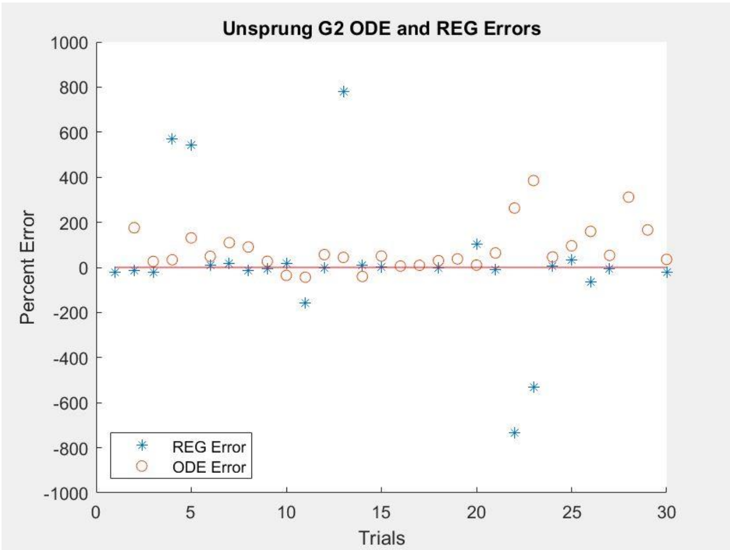

Figure 4: Generation 2, Unsprung Device ODE and Regression Error Plot

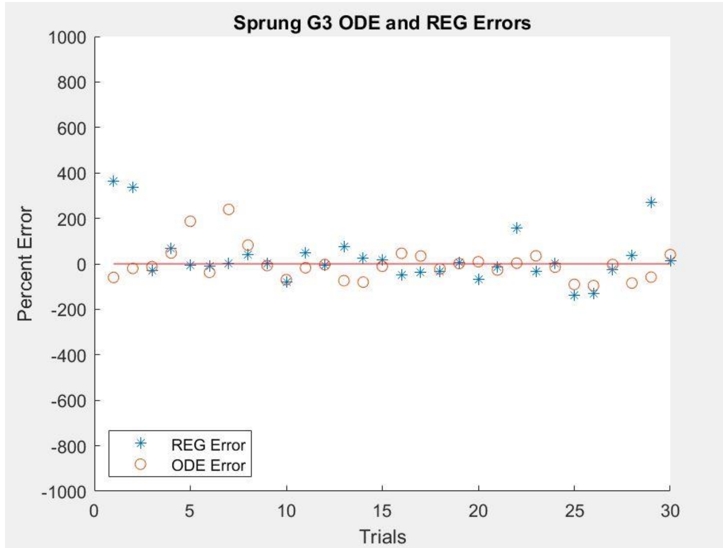

Figure 5: Generation 3, Sprung Device ODE and Regression Error Plot

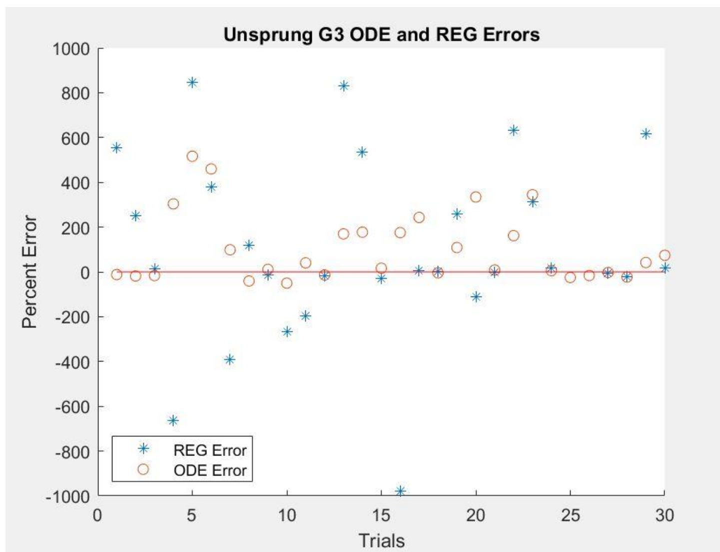

Figure 6: Generation 3, Unsprung Device ODE and Regression Error Plot

The last figure of this section concerns the * section from part c: a comparison of the outlier time traces against a standard time trace for the same location at low speed, from a separate subject.

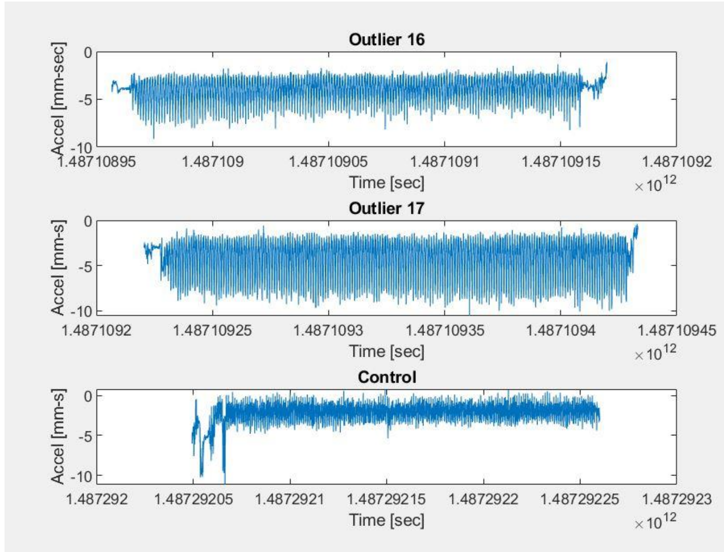

Figure 7: Outlier Acceleration time trace comparison

V. Discussion

All of the regression equations have high coefficients of determination, which indicate approximately how much of the variation in the dependent variable data (i.e. power output) is explained by the equation, yet the errors associated with these equations are not low in every case. Cross referencing Figure 1 and Figure 3 shows a reasonable explanation for this.

While Figure 1 includes both Generation 2 and Generation 3 Sprung Data, it is obvious that the lower speed predictions are more clustered, while the higher speed predictions follow a linear pattern. The coefficient of determination is high because it recognizes that the low and high speeds follow a direct, positive relationship with the accelerations. However, the error remains high because of the clustering itself which cannot be comprehensively modeled by a linear equation.

The error for the Unsprung models can be disproportionately high because the device will allow free-spin, which can be activated at any speed, and produce more power than it would if it didn't enter free spin. This free-spin condition is highly nonlinear with a threshold effect and thus difficult to model. This is why the ODE and Regression equations both are fairly inaccurate in predicting Unsprung device power output.

Both Generation 2 and Generation 3 Sprung devices have power outputs that can be reasonably predicted by the ODE or Regression models (see Table 2). Choosing which method to use will likely depend on whether or not a full position and acceleration trace is recorded. However, in the case of the Unsprung devices, the ODE models predict power output significantly better, both statistically and realistically. Unless the acceleration or rate of rotation values are consistently high in a trace (35 mm-s and above), an ODE model will predict the power output better than the Regression model.

VI. Conclusion

The goal of the regression models was to simplify predicting power outputs using data from an IMU placed on the body. The regression models for these devices can be used within reasonable accuracy once the acceleration reaches a high enough speed. At this point, the data becomes more linear and the regression does a reasonably accurate job at predicting power