
**A Critical Analysis of Power Curves Prediction in Small Scale Compressed Air Energy
Storage**

By

Timothy Linus THOMAS
Department of Mechanical Engineering
University of Port Harcourt
Rivers State

&

Nnanna Ukwa NNATE
Department of Mechanical/Production Engineering
Nnamdi Azikiwe University
Anambra State

ABSTRACT

The key to optimizing the system is to know the operating point of the system at the time of loading. This operating point can also be expressed as the power curve. The conventional method of identifying the operating points is through mathematical modelling. To model a good identification curve, some component's characteristics need to be known in addition to proper observations, especially if the component variable is unknown. Typically, this type of scenario usually lead to a long identification process. This exploratory research presents the way to find out the power curve of a system without going through the mathematical modeling process, but by using the polynomial regression technique. This regression technique involves the use of empirical data of the power curve to form parameter on ss-caes prototype. The method is based on five approach model. In this model, the variation of loading sampling data used is aimed at finding the best sampling of prediction. The data is analyzed in the form of statistical parameters and graphs to show the evaluation process of the technique. From the results of the regression, it can be concluded that the power curve of ss-caes can be identified with a high correlation value of 0.997 (99,745% accuracy) and the best way to take samples of data to be used in this technique is presented in the paper.

KEYWORDS: Empirical data; observation; prediction; power curve; regression; small scale-CAES.

Introduction

It is obvious that technological developments in energy storage are growing due to the high energy use derived from renewable energy sources that can be combined to support peak loads on the power grid (Lv et al. 2016; Ji et al. 2017), one of which is the use of

compressed air energy storage (CAES). This technology is one of the promising technologies because it is characterized by high reliability and low environmental impact (Luo et al. 2014; Chen, et al. 2018). The main disadvantage of big systems of this technologies is the reliance on geological formations required for energy storage such as A- CAES, D-CAES, and I-CAES (Guo, et al. 2018; Widjonarko et al. 2019). Because of that, the Small Scale-CAES (SS-CAES) system is an alternative solution with a more compact storage tank, high system portability and higher adaptability even with distributed or standalone energy production (Kokaew, et al. 2014; Proczka, et al. 2013).

On topics related to SS-CAES, to be able to perform system performance optimization, the operating point characteristic of the system must be known, and the performance can be optimized (Majid et al. 2014; Luo et al. 2008). The operating point is known as power curves. The power curves are the relationship between several parameters contained in the SS-CAES to the power produced (Ashikur et al. 2011). Some of these parameters are air-motor speed (ω), air pressure (p_{Air}), and electrical energy generated by the generator (P) on the loading. By knowing these characteristics, it will be easier to operate the system, such as knowing how much pressure is required and the speed to be achieved to achieve specific power or to achieve the maximum power transfer (Sinha, 2013). To be able to know the characteristics of the SS-CAES system, it can be done by entering the variable value of speed and pressure on the mathematical equations from the system that has been created. So the value of specific components of the generator or air motors, such as armature inductance of the generator, armature resistance of the generator, and others must be known before. After that, the values of parameter power curve characteristic of SS-CAES (pressure and speed) entered into the mathematical equation that has been formed from the system, then the operating point of the system can be known (Chang, et al. 2013). By reviewing existing research to know the power curve characteristic of the system, they are still using a mathematical model. This technique still will be used by the researcher because of the importance of the role to know the characteristic of the curve to be optimized of a system. However, the problem with existing techniques is that some parameter values must be known first. So that if there is an unknown component variable value, it needs to be observed, either by measuring or dismantling the component such as to know the air-motor component (Ostertagová, 2012) and then the mathematical model of the system can be solved. Therefore in this paper will be discussed how to be able to know the characteristic curve of SS-CAES by using another model to solve that problem.

In this study, the researchers will explore the use of polynomial regression models (Rawlings, et al. 1998) to be able to predict the prototype power curve by using the sampling data from the empirical data output of SS-CAES. This research was conducted by making an SS-CAES prototype. Then the data from the SS-CAES prototype will be taken at specific loading to obtain the power curve forming parameters. By using the regression technique, the power curve prediction will be based on five approaches in which variation of load criteria of data sampling to determine the value of SS-CAES power generation parameters (air-motor speed (ω), air pressure (p_{Air}), and power generated by the generator (P) on the loading). Then after getting the five sampling of data, it will be processed by using polynomial regression to predict the power curve of the prototype. The results of this method or this regression technique will

be compared with the overall data power curve of the system taken directly on the prototype to determine the accuracy and error values of the methods used. Also, the result is displayed in the form of statistical performance (SSE and RMSE), correlation value and in the form of graph fitting to know the difference in the trendline observation data with prediction data. Therefore the result in this study, two objectives will be obtained. First about the accuracy of the use of this regression to be able to predict the power curve of the SS- CAES prototype. The second explores the best data sampling method in the use of this technique so that when this technique is used, other researchers do not need to get all the data from the prototype. Researchers only need to use some data with the sampling method found in this study so that other researchers will get the best accuracy from this technique.

Methods and Materials

Small Scale Compressed Air Energy Storage

For applying this regression technique, empirical data from the real SS-CAES is needed. Therefore, the SS-CAES prototype was created to obtain the empirical data. The SS-CAES prototype made in this study has four main components [Fumo and Biswas, 2015]. These components include air tanks, air valves, air motors, and DC generators.

The SS-CAES prototype is also installed with several sensors. The sensors include air pressure sensor (MPX 5010), speed sensor (hall effect sensor), current sensor (ACS712) and voltage sensor (using the voltage divider by resistor). All sensors have been calibrated with measuring instruments in general, so that the data generated by the sensor is accurate data. The purpose of installing this sensor is to make data retrieval easier for researchers when the system is running. Each parameter forming the power curve (air motor speed (ω), air pressure (p_{Air}), and power generated by the generator (P)) can be easily retrieved by the researcher by getting the data from the sensor and then store it into the data logger.

Data Observation

As stated by the researcher in the previous section, the SS-CAES prototype was created in this study only use four main components. In the SS-CAES prototype that had been built, the heating components commonly used on large-scale SS-CAES (De Lieto, et al. 2014) are not used. It because the prototype that had been constructed just generated the maximum power 60W, so the heating components not used and changed to air valve for regulating the air pressure.

In the retrieval data process, the researcher must get the data of air motor speed (ω), air pressure (p_{Air}), and power generated by the generator (P) to form the power curve, so the prototype must be running. However when the system is running, the system conditions must be changed to get the several response systems of three parameters of the power curve. The condition of the system will be changed by installing a specific load on the system, then pumping the air pressure up slowly. For the specific load, ten loads will be applied.

On Tables 1 and 2 are some of the data taken consisting of 4 data received, such as voltage (volt), current (ampere), power (watt) (the result of multiplication between voltage and current), the speed of the air motor rotation (RPM) and the pressure through the prototype

(bar). Tables 1 and 2 are data from loads 9,4 and 12 Ω. For the data of the other, that is eight loads at loads 38.7, 56, 100, 313, 560, 691, 1000 and 2200 Ω, are performed in the same manner and stored in data logger for analysis in subsequent processes.

Table 1. Observation data load on 9.4 Ω.

V	I	P	RPM	pAir
3.9	0.415	1.6185	414	0.5
7.92	0.842	6.66864	1004	1
12.15	1.29	15.6735	1548	1.5
15.61	1.66	25.9126	2049	2
17.47	1.859	32.47673	2434	2.5
20.12	2.14	43.0568	2851	3
21.56	2.29	49.3724	3092	3.5

Table 2. Observation data load on 12 Ω.

V	I	P	RPM	pAir
3.9	0.415	1.6185	414	0.5
7.92	0.842	6.66864	1004	1
12.15	1.29	15.6735	1548	1.5
15.61	1.66	25.9126	2049	2
17.47	1.859	32.47673	2434	2.5
20.12	2.14	43.0568	2851	3
21.56	2.29	49.3724	3092	3.5

After getting all the observed data, the next step is to make the predictive power curve. The predictive power curve from the data sample will be divided by the five sampling approach methods. That five sampling approach methods are: the first approach is to take four loads of observations at the highest loads (loads 1 to 4), the second approach by making the four lowest load data (loads 7 to 10). The third approach by taking two of the highest load data and two middle loads (loads 1, 2, 5 and 6), the fourth approach takes two of the load data (loads 1, 2, 9 and 10), and the last is the fifth approach by taking two middle load data and two data with the lowest load (loads 5,6,9 and 10). The purpose of using five sampling models is to find the best sampling point for predicting SS-CAES power curves.

Multiple Regressions

After data is collected and separated according to each sampling method, furthermore, to form and extract the equations from the power curve prediction on SS-CAES will be used multiple regression. Multiple regression is an application of regression technique in which there is more than one variable. This multiple regression includes a technique called polynomial regression, so by using the polynomial regression, the dependent variable can be derived in its independent variable. (Darlington, et al. 2015)

The basic equation of multiple regression of the dependent variable on a set of independent variables k (predictor) on the variables X_1, X_2, \dots, X_k can be described as follows

$$Y_j = i_1 + b_1X_{1j} + b_2X_{2j} + \dots + b_kX_{kj} + e_j \quad (1)$$

or more complex can be represented as an equation:

$$Y_j = \sum b_i X_{kj} + e_j \quad (2)$$

where Y_j is the predicted value of the dependent variable, b_i is the regression coefficient for the predictor variable X_i , X_{ij} is the case measurement j on the predictor variable i , k is the slope of the regression surface against the variable X_j and e_i is the random error component for the i -th case. In order to know the accuracy of this technique, there are several statistical parameters that can be used as a reference in this regression process.

Goodness of fit on polynomial regression

To know the success of the regression process that has been done, some statistical parameters can be used as a reference in this regression process as known as "goodness of fit". The first statistical parameter is the SSE (sum of squares due error). This statistic measures the total deviation of the response value from the corresponding response value. This is also called quadratic squared residue. If the value of SSE approaches the value 0 then the smaller the random value generated by the equation curve that is formed.

$$SSE = \sum (y_i - \hat{y})^2 \quad (3)$$

where the y is the i -th value, n is the n -th data, and \hat{y} is the overall mean value.

The second is the RMSE (root mean square error) which serves as a standard error gauge fit and standard error regression. Or it can be said that RMSE is a standard estimate of the deviation of random components in data. If the RMSE value is closer to the value 0 then the better the equation to predict a model.

$$RMSE = \sqrt{\frac{SSE}{n}}$$

The third is R-Square which is the matching result between the equation curve formed with the variation of data taken from the observation data. If the value of R-Square closer to the value of 1 then the better the resulting equation (Darlington and Hayes, 2016).

By using the three parameters of the goodness of fit, the regression process will be evaluated to get the best result. Because R-Square is the basic equation in forming other parameters in the goodness of fit, so in this study, researchers prefer to focus on the use of R-Square parameters to determine the best results from the prediction curve that is formed. But other parameters will also be presented to be a consideration in determining the final decision.

Regression Process

At the regression process, researchers will get three results. The first result is a mathematical formula that will represent a curve formed by sampling data. The second result is a power curve in 3D that represents the prediction of the power curve, and the third result is the statistical performance of the predictive curve.

The next result is to observe the results of the power curve prediction using statistical performance analysis. In the analysis of performance statistics, R-Square, SSE and RMSE data analysis are displayed. However, it should be noted that the results of this statistical performance are the performance of the predictive power curve on the sampling data approach (five data sampling method). Therefore this statistical data cannot represent the level of accuracy of the polynomial regression method in predicting the whole of the SS-CAES power curve. The analyze of the statistic performance can be seen in Table 3.

Table 3. Statistical performance of the power curve prediction.

Approach Model	R-Square	RMSE	SSE
Model 1	0.9987	0.7268	7.395
Model 2	0.8858	0.7769	10.86
Model 3	0.9973	1.085	16.49
Model 4	0.9982	1.03	13.79
Model 5	0.9945	0.9596	17.5

In the analysis of statistical performance, it can be seen that in all three sampling methods it does have a reasonably good value. It can be analyzed in R-squared analysis, that are in model 1 with the value of R-Square 0.9987, model 4 with a value of 0.9973 and model 5 with a value of 0.9945. However, in the analysis of the RMSE and SSE variables, the three sampling methods did not show a good thing. As for the RMSE value in model 1, the value is lower than that of model 5. SSE values also have a similar situation, namely the SSE value in model 3 is higher than the SSE value in model 2. In theory, the best sampling method should indicate good value in both the R-Square, RMSE, and SSE parameters. However, the

regression results that have been obtained show the opposite. This phenomenon is reasonable because the data used in the regression process on statistical parameters is sampling data, not the overall data of observation. Therefore the results in the statistical parameters have not been able to show the accuracy of this method.

As stated in the previous discussion, the two results of the last process regression have not been able to represent the accuracy of the polynomial regression technique in predicting the SS-CAES power curve. Therefore, to determine the level of accuracy of this technique is to extract mathematical equations from the prediction curve on the five methods of approach. In the extracted mathematical equation the value of the variable forming the equation will be entered to determine predictive power. These parameters are parameters of air pressure and air motor speed. For the mathematical equation can be seen in Equation (7) to Equation (11). On that equations are consisting of three variables, that the variable x representing the air pressure, y representing the speed of the air motor and f (x,y) representing power.

Equation of Prediction Curve 1:

$$f(x,y) = 28.87 + 1.63x + 15.47y - 8.334x^2 + 5.496xy - 2.678y^2 + 0.6549x^3 - 3.215x^2y - 0.3593xy^2 + 0.8707y^3 - 0.1285x^3y + 0.6649x^2y^2 + 0.5406xy^3 - 0.1293y^4 \quad (7)$$

Equation of Prediction Curve 2:

$$f(x,y) = 7.777 + 8.027x - 9.332y - 3.915x^2 + 6.99xy - 9.751y^2 + 2.65x^3 - 0.01228x^2y - 2.42xy^2 - 0.9329y^3 \quad (8)$$

Equation of Prediction Curve 3:

$$f(x,y) = 26.76 - 4.383x + 16.59y - 9.311x^2 + 0.3154xy - 0.3038y^2 - 0.6662x^3 - 2.332x^2y - 0.422xy^2 + 0.8875y^3 + 2.461x^3y - 4.224x^2y^2 + 3.535xy^3 - 0.6435y^4 \quad (9)$$

Equation of Prediction Curve 4:

$$f(x,y) = 25.79 - 9.616x + 15.97y + 12.8x^2 + 2.227xy - 3.084y^2 - 3.332x^2y - 1.347xy^2 + 4.834y^3 + 0.8784x^2y^2 - 0.08372xy^3 + 0.8799y^4 - 0.7256x^2y^3 + 1.343xy^4 - 1.751y^5 \quad (10)$$

Equation of Prediction Curve 5:

$$f(x,y) = 17.23 - 17.45x + 17.94y - 5.126x^2 - 1.601xy - 1.419y^2 - 2.444x^2y + 1.744xy^2 + 0.4526y^3 \quad (11)$$

The extracted equations from the regression process (Equations 7 to 11) will use to represent the regression predictions. The value of regression predictions can be known by entering the value of power curves parameter according to the parameters that describe each variable of power curves (f for power, x for air pressure and y for air motor speed). Then the value of the results of these equations will be as the predictive data value (power data), and then

the data prediction will be compared with real observation data on the prototype to determine the level of accuracy. Comparison of prediction data and real data (observation data) will be discussed in the next chapter.

Results and Discussion

In this study, researchers have the work to predictions of the power curve characteristics of SS-CAES. To be able to make predictive curves from SS-CAES prototypes, polynomial regression is used with five sampling approach methods. The purpose of using these five sampling methods is to explore the best sampling to be used in forming the power curve. Then this technique used with the five methods with different sampling method in each model to determine the best way to predict the power curve.

To be able to predict the power curve in the SS-CAES, several steps must be done. The first is a regression process which results. The purpose of this regression process is to obtain mathematical equations from the five sampling methods proposed. Second is to compare predictive data with overall observation data. In this process, the mathematical equations that have been obtained in the regression process will be given some values on the equation forming variables (variable air pressure and speed). The aim is to get predictive power values for each proposed model. After that, the data will be compared with the overall observation data to determine the level of accuracy. In this process, to determine the accuracy of each model will be done in three ways. The first way is to use statistical parameters, the second is to use the R-correlation parameter, and the third is the curve graph.

For the first process, that is regression process. The data that have been obtained are attached to the previous discussion can be explained as follows. First, for the performance of the power curve prediction on the result parameters in Table 3. From the three parameters of the results in the table 4, the best results obtained on the five methods of data sampling proposed is for the first method with R-Square value of 0.9987 and with the number of sum squared error (SSE) of 7.395. Some phenomena have also been explained in the regression process section. However, the result in this process is only the best result for the suitability of the curve with the data entered (data sampling used in each model) to form the prediction curve, not the overall power prediction curve of the SS-CAES prototype.

For the second result is a fit parameter between the data used to form the power curve prediction (data sampling used in the five proposed methods) with the overall data of the observed data from the power curve on the prototype which the results are shown in Table 4.

Table 4. Validation results of the power curve prediction results.

Approach Model	SSE Data Validation	RMSE Data Validation
Model 1	211.44	1.73798
Model 2	1318.15	4.33944
Model 3	68.6293	0.990161
Model 4	66.4366	0.974214
Model 5	476.475	2.60898

In Table 4, from the two validation data SSE validation and RMSE validation, it can be concluded that the best sum squared of error (SSE) value is found in the fourth method (the sampling method using the two lowest load data and the two highest load data) with SSE value of 66.4366 and the RMSE value 0.974214. Whereas for the other value that closes with the best result is the using data sampling method 3 with SSE value of 68.6293 and the RMSE value 0.990161. However, the first sampling method where the previous analysis has a superior value compared to the best of the two sampling methods (the third and fourth sampling method) has a low result. The values obtained from method 1 are SSE 211.44 and RMSE 1.73798. This value of method 1 is far from the best value in the comparative analysis with the whole data (method 4). Then it can be concluded that the fourth sampling method is the best sampling method compared to the other five methods in this analysis.

The next is analysis the correlation of the power prediction results (from 5 sampling methods) with power observation data on the SS-CAES prototype.

If it looks at the relationship between the correlation with the error value (SSE) found in Table 4, it will be concluded that the higher the SSE will make the correlation between the predicted results and the results of the observations becomes smaller. For example, in method 1 it has a correlation of 0.99745 with SSE 211.44, while for method 2 with SSE 1318.15 it has a lower correlation compared to method 1, which the value is 0.96397. For the best model approach that has the best correlation too for all model, there is a fourth approach with a correlation value of 0.99745 and SSE 66.4366. Whereas the method that is almost close to the best value is found in method three which is the correlation value of this method is 0.99744 and SSE 68.6293. Therefore, in the section, it can be seen that the best sampling method for predicting the SS-CAES power curve is to use method 4 with an accuracy of 0.99745 or 99.745%. However, to be more convincing, the analysis will be carried out in the next section using graph analysis.

Conclusions

In this study exploration of the use of regression has been carried out to be able to predict the power curve of the SS-CAES prototype. The research was conducted by taking data on the SS-CAES prototype. Then by using the 5 data sampling methods that have been obtained, the overall data will be predicted. The use of 5 methods of this approach is intended to find out the best sampling method in this regression technique. From the regression process on the five methods of using the data sampling, then the power curve equation is obtained which represents the relationship between air-motor speed (ω), air pressure (p_{Air}), and power generated by the generator (P). By entering the variable forming the power curve in the equation produced in the regression process, the value of predictive power will be obtained.

The results of the research and the results of the evaluation of each method show that the fourth sampling data model is the best model for data submission techniques with correlations 0.99745 (99,745% accuracy) and SSE 66.4366. This analysis is obtained by looking at the results of the SSE parameters, RMSE, correlation and fitting graph in each method. Therefore, by looking at these results, it can be concluded that the first in this study the use of regression can be used to predict the power curve in the SS-CAES. The two best techniques for exploration of this sampling method are using the fourth method, namely by using the lowest load data and the highest load.

Recommendations

1. In any exploratory research the power curve of a system need to be predicted not only modeling mathematical of the system, but with the use the polynomial regression technique as well and is highly recommended.
2. It is also recommended that the regression technique form should use the empirical data of the power curve form parameter on SS-CAES prototype.
3. Also recommended is that the method used should be based on five approach model in which is the variation of loading sampling data to be used.

REFERENCES

- Ashikur M, Khan R, Rahman MM, Kadirgama K, Maleque MA, Ishak M. (2011). Prediction of surface roughness of Ti-6Al-4V in electrical discharge machining: A regression model. *Journal Mechanical Engineering and Sciences*. 1: 16-24.
- Chang Y, Wai K, Cheng E. (2013). *Sensorless position estimation of switched reluctance motor at startup using quadratic polynomial regression*. IET Electric Power Application. 7: 618–626.
- Chen L, Chen L, Hu P, Sheng C, Xie M. (2018). *A Novel compressed air energy storage (CAES) system combined with pre-cooler and using low grade waste heat as heat source*. Energy. 13(1): 259–266.
- Darlington RB, Hayes AF. (2016). *Regression analysis and linear models: Concepts, Application and Implementation*.
- De Lieto Vollaro R, Faga F, Tallini A, Cedola L, Vallati A. (2015). *Energy and thermodynamical study of a small innovative compressed air energy storage system (micro-CAES)*. Energy Procedia. 8(2): 645–651.
- Fumo N, Biswas MAR. (2015). *Regression analysis for prediction of residential energy consumption*. Renewables and Sustainable Energy Review. 47:332–343.
- Guo H, Xu Y, Zhang Y, (2018). *Off-design performance and an optimal operation strategy for the multistage compression process in adiabatic compressed air energy storage systems*. Applied Thermal Engineering.
- Ji W, Zhou Y, Sun Y, Zhang W, An B, Wang J. (2017). *Thermodynamic analysis of a novel hybrid wind-solar-compressed air energy storage system*. Energy Conversion and Management. 14(2): 176–187.
- Kokaew V, Sharkh SM. (2014). *A hybrid method for maximum power tracking of a small scale CAES System*. International Symposium on Communication Systems, Networks & Digital Sign. 61–66.
- Luo X, Wang J, Dooner M, Clarke J, Krupke C. (2014). *Overview of current development in compressed air energy storage technology*. Energy Procedia. 6(2): 603–611.
- Luo X, Wang J, Shpanin L, Jia N, Liu G, Zinober ASI. (2008). *Development of a mathematical model for vane-type air motors with arbitrary N vanes*. International Conference of Applied and Engineering Mathematics. 1: 1–6
- Lv S, He W, Zhang A, Li G, Luo B, Liu X. (2016). *Modelling and analysis of a novel compressed air energy storage system for trigeneration based on electrical energy peak load shifting*. Energy Conversion and Management. 13(5) :394–401.
- Majid ZAA, Ruslan MH, Sopian K, Othman MY, Azmi MSM. (2014). Study on performance of 80 watt floating photovoltaic panel. *Journal Mechanical Engineering and Sciences*. 7:1150–1156.

- Ostertagová E. (2012). *Modelling using polynomial regression*. Procedia Engineering. 4(8): 500–506.
- Proczka JJ, Muralidharan K, Villela D, Simmons JH, Frantziskonis G. (2013). *Guidelines for the pressure and efficient sizing of pressure vessels for compressed air energy storage*. Energy Conversion and Management. 6(5): 597–605.
- Rawlings JO, Pantula SG, Dickey D. *Applied regression analysis: A research tool*, Second Edition. 1998. Springer US, New York.
- Sinha P. (2013). *Multivariate polynomial regression in data mining: Methodology, Problems and Solutions*. 4:962–965.
- Widjonarko, Soenoko R, Slamet Wahyudi, Siswanto E. (2019). *Comparison of intelligence control systems for voltage controlling on small scale compressed air energy storage*. Energies. 25–28.