Robust Surrogate Modeling of Lithium Ion Batteries

Rajeswari Chandrasekaran[1, 2], Jonathan Denis[1], Michael Skinner[1], Thomas F.Fuller[1, 2]

(1) Center for Innovative Fuel Cell & Battery Technologies, Georgia Institute of Technology, Atlanta, GA 30332, (2)School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332.

Introduction

Significant research is being done in the field of lithium ion batteries for hybrid electric vehicles and plug-in hybrid vehicles by different automakers. Various physics based models exist in literature that describe in detail the initial performance and also the cycle life or storage issues. But the knowledge from physics based models with respect to both performance and durability should be incorporated in the vehicle level system model earlier in the design phase. The emphasis on this approach is increasing as the paradigm is shifting to design for affordability. Incorporating physics based models in system model is complicated due to different languages in which they are coded and also because of the slower computation speed of the physics based models as numerical methods are involved. Hence a systematic framework is necessary to approach this problem. Robust design through surrogate model methodology is useful in this regard to develop simpler models without losing much of the fidelity and they are easy to integrate in vehicle model. Response surface methodology is a key enabler for surrogate models. This methodology allows one to explore the design space for multiple responses of interest –capacity, power, specific energy, cycle life, etc. and determine the optimal settings by means of a desirability function approach. Overall desirability can be obtained from the individual desirability functions of the responses. This means a trade-off between multi- objectives is obtained. The robust design provided by the surrogate model methodology is also less sensitive to uncertainties or noise factors. Both physics based models or experimental data can be used to arrive at response models. Experiments avert the need for assumptions associated with physics based models, whereas the latter is needed for cases where historical or experimental data are unavailable.

In the current research, the surrogate model approach is introduced for analyzing performance of lithium ion batteries. This can be extended in the future to include cycle and calendar life as well to study degradation of batteries at system level.

Response Surface Methodology (RSM)

Even though the system models that are available currently have battery models as one of the components, most of the battery models use look up tables for OCP or SOC values. The disadvantage of look up tables is that they have limited scope in terms of data available, e.g., OCP data may not be available at different temperatures. Another disadvantage is that the interpolation procedure used can lead to unknown errors. Also, determination of SOC of a battery onboard a vehicle is difficult and is a source of uncertainty. Simple battery models in current system models may be sensitive to such uncertainties. Hence, there is need for a surrogate model based on RSM which facilitates a robust design space solution insensitive to uncertainties. Discussion on RSM can be found elsewhere (1). The simple response surface equation (RSE) is a quadratic equation based on Taylor series approximation.

Application of RSM for battery surrogate model

Lithium ion cell from Valence technology with rated capacity of 1.1 Ah (C rate) is chosen for the experiments. The controllable factors chosen were temperature of operation and rate of discharge. It should be noted as the experiment proceeds, there will be change in temperature of the battery, especially at high rates of discharge and there will definitely be a temperature gradient within the battery. In this way, the temperature of the battery is a noise factor (or uncontrollable factor). But for the time being, constant temperature of operation of the battery is considered. Once a good RSE is developed, one can add random distribution to the output to reflect these noise factors during analysis. For each of the independent variable, three levels are chosen. A central composite design of experiments is created and total of 9 runs were conducted using an Arbin battery cycler (BT-2000) and HD-508 environmental chamber from Associated Environmental Systems. The design of experiments is shown in Table 1. The CCD design gives the same number of runs as a full factorial in this case since there are only two independent variables. All rates of discharge are denoted in reference to C rate.

Temperature (^oC)	Rate of Discharge (C rates)
0	0.5
0	1.75
0	3
25	0.5
25	1.75
25	3
50	0.5
50	1.75
50	3

Table1. CCD Design of Experiments

Results and Discussions

The discharge curves for the above 9 cases for Valence cell are given in Figure 1.

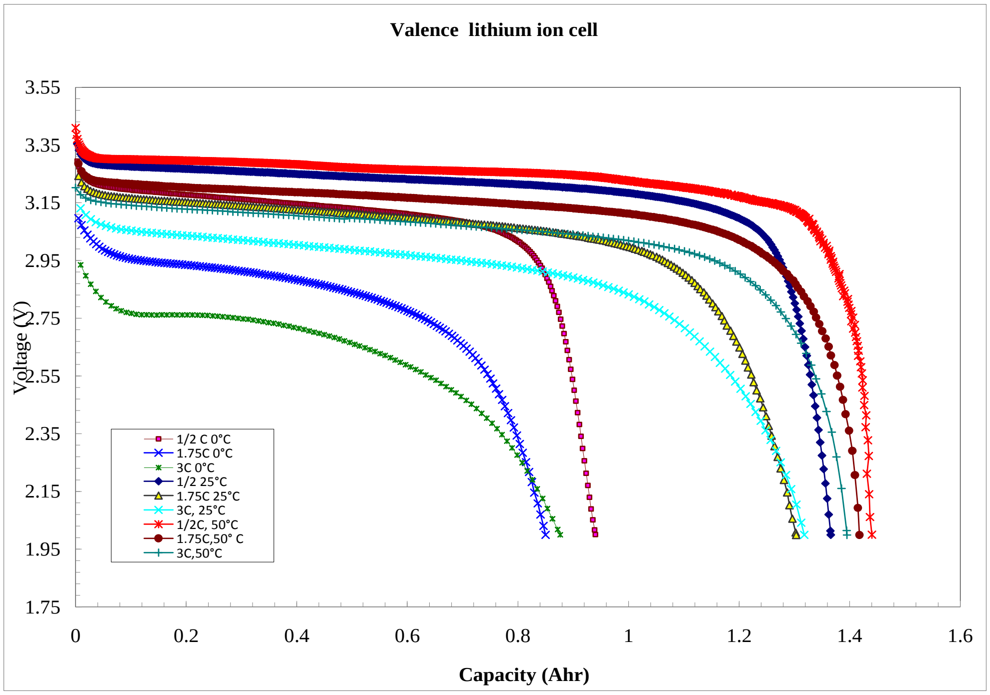

Figure 1 Discharge curve of Valence 1.1 Ah lithium ion cell

Response Surfaces were created for the capacity obtained in the above runs as a function of temperature and rates of discharge. The results are now analyzed using statistical goodness of fit tests.

Actual vs. Predicted Plot

Figure 2 gives the actual vs. predicted plot for the capacity (Ah). The R^2 value is 0.99. All data points lie close to the Perfect Fit line (diagonal). But they don't all seem to be distributed evenly and this might occur when a single variable is the dominant factor. The mean of the response is also slightly shifted upwards because of the uneven distribution. More cases could be run to check if this observation still exists.

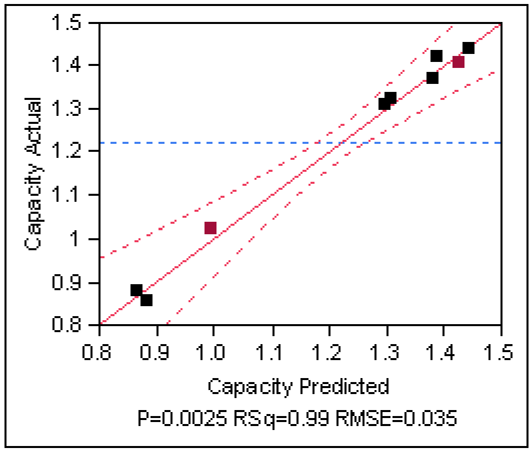

Figure 2 Actual vs. Predicted plot of capacities (Ah) of Valence lithium ion cell

Residual vs. Predicted Plot

Figure 3 gives the residual vs. predicted plot. There is no distinguishable pattern observed in the residual by predicted plot. A good random distribution of the error would imply that it was fine to have discarded the higher order terms and interactions from the Taylor series expansion for the assumed model. But slight data clumping towards right corner is seen. This could again imply that single variable is driving the response in the regions of interest. The total span of error is 10% of the minimum of the predicted capacity values. Usually it is desirable to have this as low as 5% as a rule of thumb for validity of the second order meta model.

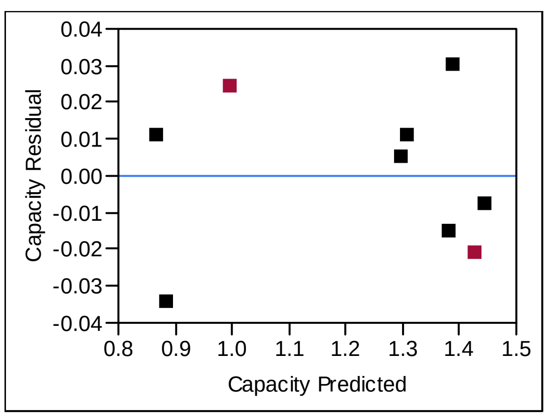

Figure 3 Residual vs. Predicted plot of capacities (Ah) of Valence lithium ion cell

Model Fit Error (MFE)

This statistic gives an idea of how well the model fits the data points in the Design of Experiments. Ideally, it is desired that the error distribution resemble a normal distribution with mean zero and standard deviation of one. Figure 4 gives the model fit error for the predicted capacities.

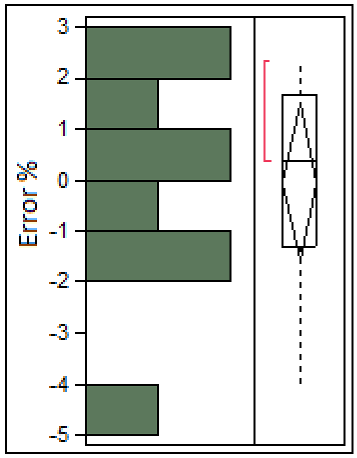

Figure 4 Model Fit Error Distribution

As can be seen in Figure 4, the upper and lower bounds of the error are less than 5%. Usually, it is desirable to minimize the bounds. The mean and standard deviation of the above distribution are 0.051 and 2.01 respectively. It is desirable to have mean approximately equal to zero (as obtained) and standard deviation less than one. In the above distribution, if the outlier is neglected, then the standard deviation reduces to 1.43 and the error bounds are also minimized to 2.3 and -1.3 % respectively. But caution has to be applied while neglecting outliers to avoid significant correlation between the independent variables. Usually, it is okay to neglect about 7-8% of the outliers. Since in this work there were only nine cases to begin with, neglecting of outliers is avoided. The correlation between the independent variables with all 9 runs is zero. It is necessary to try and fix the shortcomings in the assumed model until this stage and get the goodness of fit statistics better before proceeding further. But as a learning exercise, more in this methodology is presented here. A point to be noted above is that there were fewer runs and hence normal distribution of the error within this small data set might be slightly difficult.

Model Representation Error (MRE)

This statistic shows how well the assumed model predicts the actual response for the design settings not used in the creation of model in the entire design range of interest. For this, random settings of the independent variable are chosen and experiments performed. A subset of a Latin hypercube sample was used to arrive at some of the cases below. Usually it is desired to have at least 20% of the original number of runs to have a sufficient sampling of the space to check the MRE. The RSE developed above is used to predict the responses and arrive at MRE by comparing it to the actual capacities.

Temperature (^{o}C)	Rate of Discharge (C rates)
6	1.07
11	2.03
18	3
22	0.74
50	0.5
0	1
0	2
25	1
25	2

Table 2 Random data for Model Representation Error Distribution

Figure 5 shows the MRE distribution. It can be seen that the lower error bound has increased as well as the mean (-2.08) and standard deviation (3.63). Even though it is understandable to see an MRE with poorer statistics than MFE, a good RSE will lead to a better MRE. It should also be noted that if the data at zero degrees are removed from the above table, the mean and standard deviation of the MRE distribution as well as the lower bound significantly reduce. This implies that the predictive capability of the assumed model reduces at $0^{o}C$.

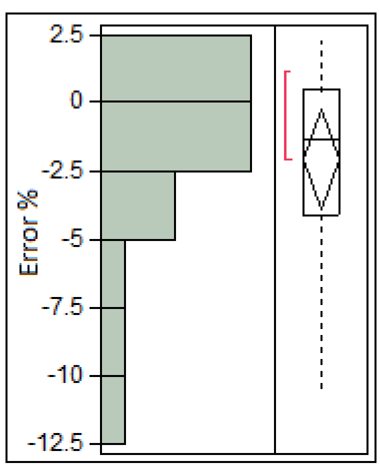

Figure5 Model Representation Error Distribution

Latin hypercube (LH) sampling is basically a space filling design and complements the central composite design. In order to improve the assumed model, a space filling design could be used alongside CCD to create the RSEs. As a first attempt, the extra cases in Table 2 were used alongside Table 1 to create the RSEs and the goodness of fit statistics was observed. The statistics didn't improve much. So higher order terms could be employed or a proper LH design could be employed to arrive at a rigorous RSE.

Response Surface Equation (RSE)

Once a good RSE is obtained, it represents the meta model of the metrics considered as a function of the variables. Figure 6 shows the response surface of capacity as a function of rate of discharge and temperature.

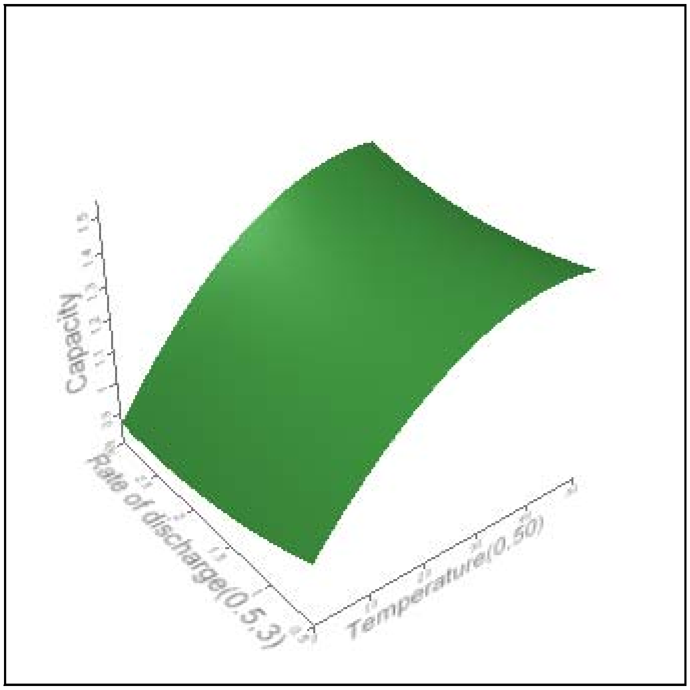

Figure 6 Response Surface of capacity of 1.1 Ah Valence lithium ion cell

Analysis of influence of the independent variables

Figure 7 shows the Pareto plot. This plot gives the relative influence of the variables and the interaction between variables on the response, i.e. capacity. It can be seen that temperature has a significant effect on the capacity.

Term	t Ratio
Temperature(0,50)	17.63075
Temperature*Temperature	-6.51896
Rate of discharge(0.5,3)	-2.61147
Rate of discharge*Rate of discharge	1.89793
Temperature*Rate of discharge	1.61650

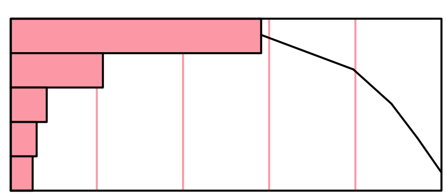

Figure 7 Pareto plot

The influence of the individual variables on the response can be viewed either in Figure 8 in a prediction profiler or in Figure 9 in the Scatter plot matrix. This gives the change in capacity of the cell with respect to temperature or rate of discharge alone while the other variable remains fixed at a value within the range of interest. One can then obtain simulated responses using the RSEs created for, say 5000 runs, by using a random uniform distribution on temperature as well as rate of discharge within the design range. A random noise can also be added on the capacity values. In this case random noise with standard deviation of 0.0349 is added to the responses. The resulting capacity distribution is shown on the right most plot in Figure 8. The corresponding scatter plot matrix is given in Figure 9.

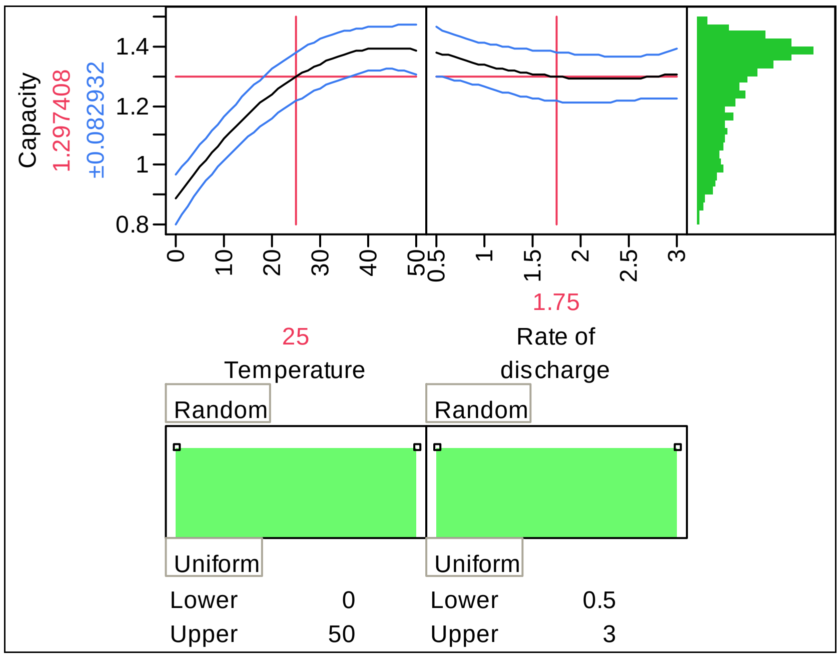

Figure 8 Prediction Profiler

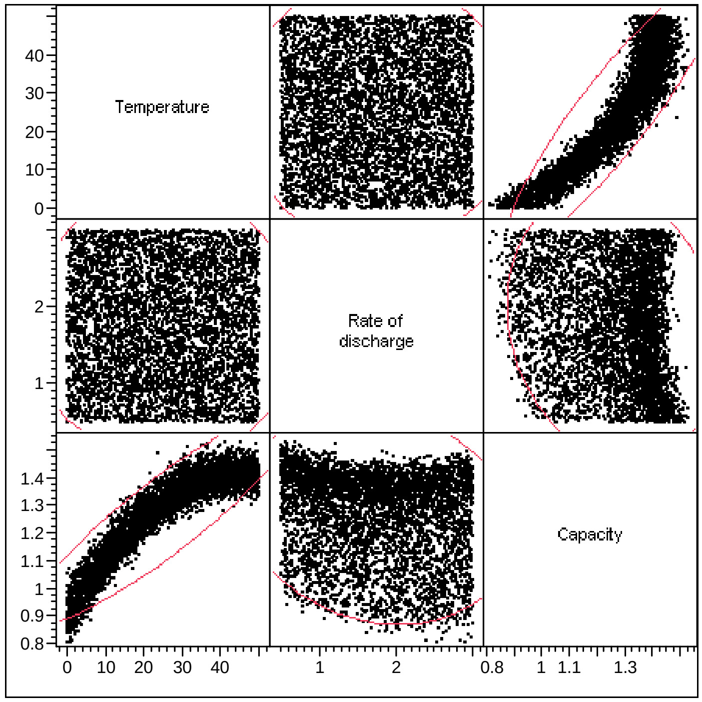

Figure 9 Scatter Plot matrix

Desirability functions can also be used to arrive at optimal settings of operation to achieve the objective.

Suggestions for future work

The RSEs when used outside the design range, e.g., at -25^{0}C or at rates lower than 0.5C (e.g. 0.1C) or greater than 3C (e.g. 4C), give greater error. Hence in order to obtain good RSEs, one might have to expand the design range. At the same time, if the design space is expanded too much, the central composite design of experiments alone will not be sufficient to obtain good RSEs. It has to be coupled with one of the space filling designs. Several other performance metrics like power, specific energy can also be studied and multiple RSEs can be created. This methodology can also be extended to understand the influence of variables like electrode thickness, porosity, etc. on the capacity by using physics based models to obtain the responses that can be used to develop response surface equation. Finally, the RSM can be extended to study degradation in batteries by having number of cycles or storage time as one of the independent variables of interest. The interaction of temperature with cycle life and calendar life on capacity will be of significance also.

Acknowledgements

Georgia Tech Research Institute is gratefully acknowledged for financial support.

Reference

(1) Rajeswari Chandrasekaran, Wu Bi and Thomas F. Fuller, *J.Power Sources, 182(2), Aug 2008*, *546-557.*