PREDICTING UL 94 RATINGS for PLASTICS based on MCC MEASUREMENTS

Mahmood Tabaddor, PhD, Priyansu Panda & Thomas Fabian, PhD
Commercial & Industrial, UL LLC

INTRODUCTION

As manufacturers are always under pressure to be more innovative and faster to market with new product offerings, relying on complicated testing of physical samples can be a competitive disadvantage. However, for any manufacturer, from material-scale to system-level, there are many physical tests that must be conducted to help determine performance. With the growing use of Machine Learning (ML), there are now possibilities that physical tests can be replaced with computer algorithms which can provide high accuracy in predicting performance and so would be an economic boon to manufacturers. Alternatively, if the outcome of a resource-intensive physical test can be predicted based on the results of a simpler physical test, then ML can help learn that relationship and provide tremendous value to manufacturers. Yet the challenge is to understand how to build and validate these ML-based models.

For one set of manufacturers, plastic manufacturers, trying to predict the performance of new materials is challenging. The UL 94[1] test for flammability is one such test. The outcome of the testing is a flammability rating, one of four that are possible. Sample preparation time, amount of materials and length of testing time all motivate the search for a simpler test that could be predictive of the outcome of the UL 94 test. One candidate test is the Micro-scale Combustion Calorimetry (MCC)[2]. MCC is a thermal analysis technique that applies controlled heating to a sample to high temperatures driving combustion. The MCC test does not require molded flame bars typical of the UL 94 test; samples on the order of milligrams are sufficient. Since both tests, UL 94 and the MCC, push the materials into a state of combustion, there might be a technical basis for a link between the outcome of the MCC test and the UL 94 for a material, the first requirement for using ML-based approach.

Brief Literature Survey

Several studies have been carried out in the last few years looking at a possible relationship between MCC and UL 94. In one study[3], the authors visualized a few MCC measurements and UL 94 ratings to uncover a potential relationship. Figure 1, from this study, shows a plot of two MCC variables with the data colored by the resulting UL 94 rating. In this sample set, only two out of the four possible ratings from UL 94 are plotted. Nevertheless, the data does suggest that one or more of the MCC measurements may help the outcome of the UL 94 test. The authors then proceed to develop a flammability index based on three of the MCC parameters but with no stated link to the UL 94 rating.

In other studies[4,5,6], the authors have looked at the relationship between MCC measurements and the associated UL 94 ratings. Figure 2 shows one example from a study[4] where a single parameter, the heat release rate, measured by the MCC test is shown versus UL 94 rating. Again, the data analysis is suggestive of a potential relationship between UL 94 rating category and the MCC test. For this specific example, once again a threshold appears that may help separate one category, V0, from all the other categories. Yet there is overlap and the parameter cannot help distinguish between all four of the UL 94 categories.

Figure 1 (from reference 3)

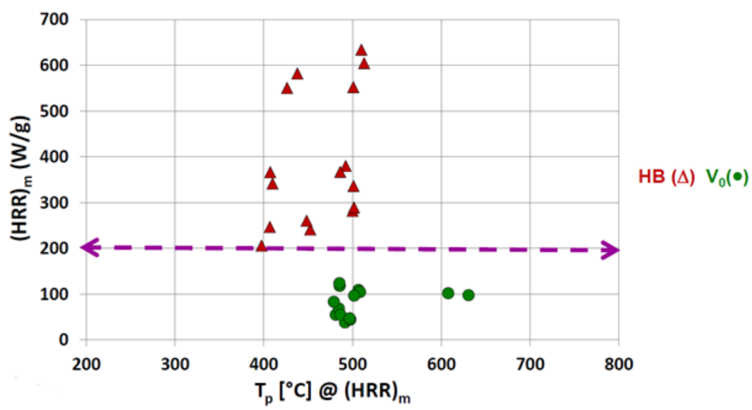
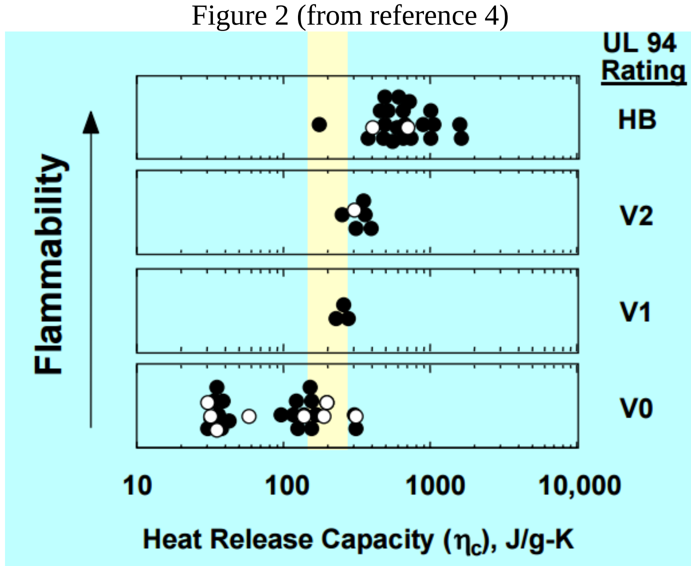


Figure 2 (from reference 4)

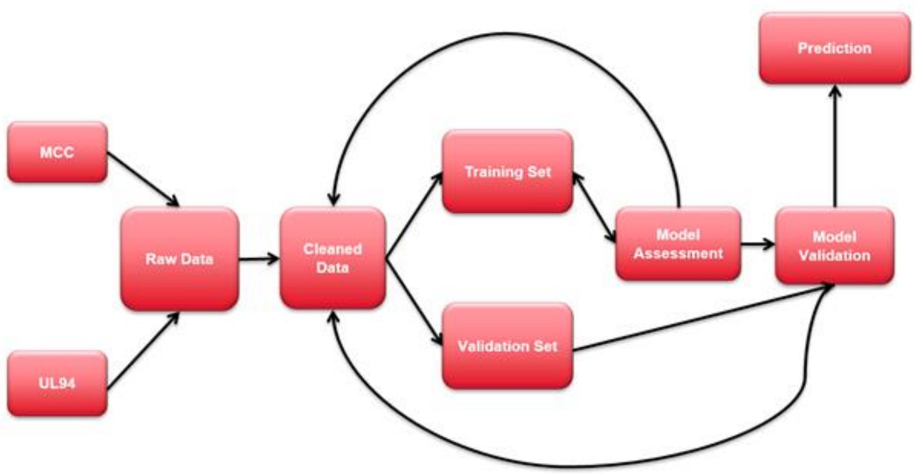


ML-based Modeling Approach

There is a long history in using ML-based algorithms, such as neural networks[7], to predict the properties of polymers. However, these sorts of tools were only available to academics with the time to handle and troubleshoot these complicated algorithms and immune to the pressures of practical product development. Over the last few years, with the boost in cheaper and powerful computational machines and advances in ML algorithms, ML-based modelling is now a more accessible tool for practical product development. Yet, there is a need for a careful and expert oriented process to ensure the modelling process does not generate misleading and incorrect models. Unlike the clichéd garbage in, garbage out, the key concern is that even with good inputs, the modelling process can be carried out incorrectly as the ML algorithms are vastly more complicated than the traditional linear analysis that has been the work horse of the regression in statistical analysis.

Figure 3 shows a schematic of the key steps in an ML-based modelling building process. Without going into details there are several key steps. One is exploratory data analysis. In this step, the data is reviewed, cleaned and visualized to ensure that an understanding of the quality and potential gaps are identified early. Some visualizations look for correlations amongst the list of candidate input variables. Any correlations can help simplify the list of potential input parameters and lead to a more efficient modelling process. Also there is quite a bit of learning that can happen in this stage even before a model is even built. Once the data is assessed and streamlined, then the next key step is to divide the data into two sets (for more advanced analysis, three sets). The purpose of one set, the training set, is to help the ML algorithm to learn all it can about potential patterns within the data. This is where model parameter tuning and selection occur. Then once a model(s) is selected, the accuracy of the model is determined by trying to predict the outcomes given in the validation set. This separation process is one of the keys to the integrity of the model building process. The division of the total data set into a training and validation sets must meet certain requirements not detailed here.

Figure 3



For this project, since the intent is to predict a UL 94 rating, classification ML algorithms will be chosen. The target, or the outcome, that is being predicted is simply the following ratings: V0, V1, V2 and Fail. The inputs are all the measurements from the MCC test and also the thickness of the sample during the UL 94 test. Not all MCC parameters are expected to be predictive of the UL 94 rating but this is part of the learning process. The ML process will help filter those inputs that are not strong predictors of the outcome of interest.

Finally, there is another challenge specific to predicting the UL 94 rating. The rating assignment is based on what we call a “bad apple” approach. A single bad reading can lead to the overall assigned lowered rating from the sample of 5 that are measured. So basically, if only one sample performs sufficiently poor it will override the performance of the other samples and determine the final rating. This is not something that can be expected to be predicted accurately unless there is a reason to believe that the MCC would more sensitive to that individual sample than all the others. We would expect that the model can predict the overall behaviour of all the samples of materials (such as an averaged value), not that of an individual sample the deviates from the others.

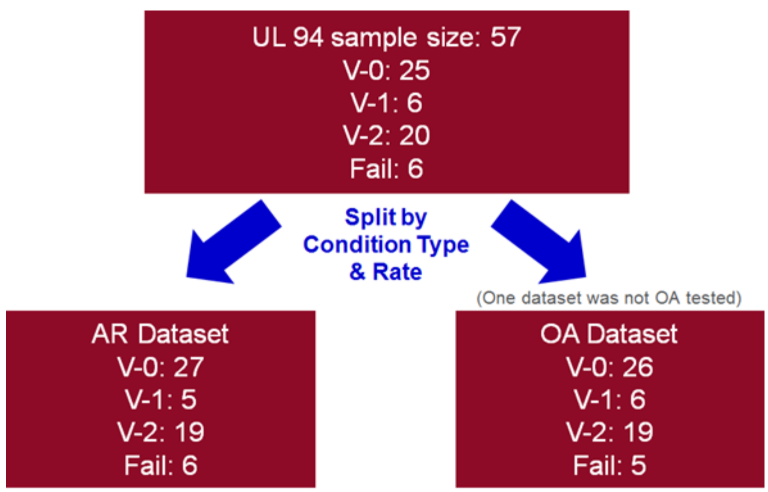
To address this concern, two different methods of assigning UL 94 ratings will be examined and the effect on the accuracy of the model will be shown.

THE DATASET

Data was gathered on 39 independent samples. This is generally a small test size for ML based learning but will likely represent the reality of many similar projects. Still, it is a good start and will help identify the potential in ML modeling and determine if there is value in investing in building the dataset size. The 57 samples were manufactured using 25 unique grades of polymers, including both thermoset and thermoplastics, 13 different generic resin types and different thicknesses. MCC data was generated suing Method B with 5 replicates. Measurements included PHRR(W/g), T@PHRR(°C), Onset HRR (10) Temp(°C), Onset HRR(25)Temp(°C), HOC(kJ/g), Residue(%), and Heat Release Capacity(J/g-K).

For the UL 94 test, multiple thicknesses were used with final ratings covering all four categories and 5 replicates for each sample conditioning type. For the UL 94 rating test, 5 samples were tested as received (AR) with no prior conditioning. 5 other samples were conditioned by oven aging (OA) prior to running the UL 94 test. No films were tested. Hence the dataset was divided into two sets: one for AR and one for OA (Figure 4). Each receives its own UL 94 rating. Hence, a different model would need to be built for each one. The samples for the MCC test do not undergo any conditioning and so a model between MCC and AR UL 94 data would be expected to perform better. The UL 94 rating shown follows the procedure described in the standard.

Figure 4



Exploratory Data Analysis

Exploratory data analysis (EDA) is an approach to visualize datasets especially those with many input parameters that define a hyperspace. The objectives of EDA are to suggest hypotheses about the causes of the observed data, support in modelling of the data, and guidelines for future data collection. One such visualization tool is a scatterplot (Figure 5) where all the input variables are plotted against each other in pairings. This helps show the myriad of possible relationships between the parameters.

From this scatterplot, certain patterns are observed and quantified for their strength. For the MCC attributes, there is a strong positive linear correlation between PHRR and HRC. There is a similarly strong positive correlation between OnsetHRR25Temp and OnsetHRR10Temp. Finally, there is a somewhat strong negative linear correlation between HOC and Residue. Looking at the first pairing, PHRR and HRC, the two attributes are perfectly correlated. This is due to the fact that one parameter is derived from the other, under constant heating rates. In building the model, it is not necessary to keep both attributes and the list of inputs to the model is reduced by one.

Figure 5

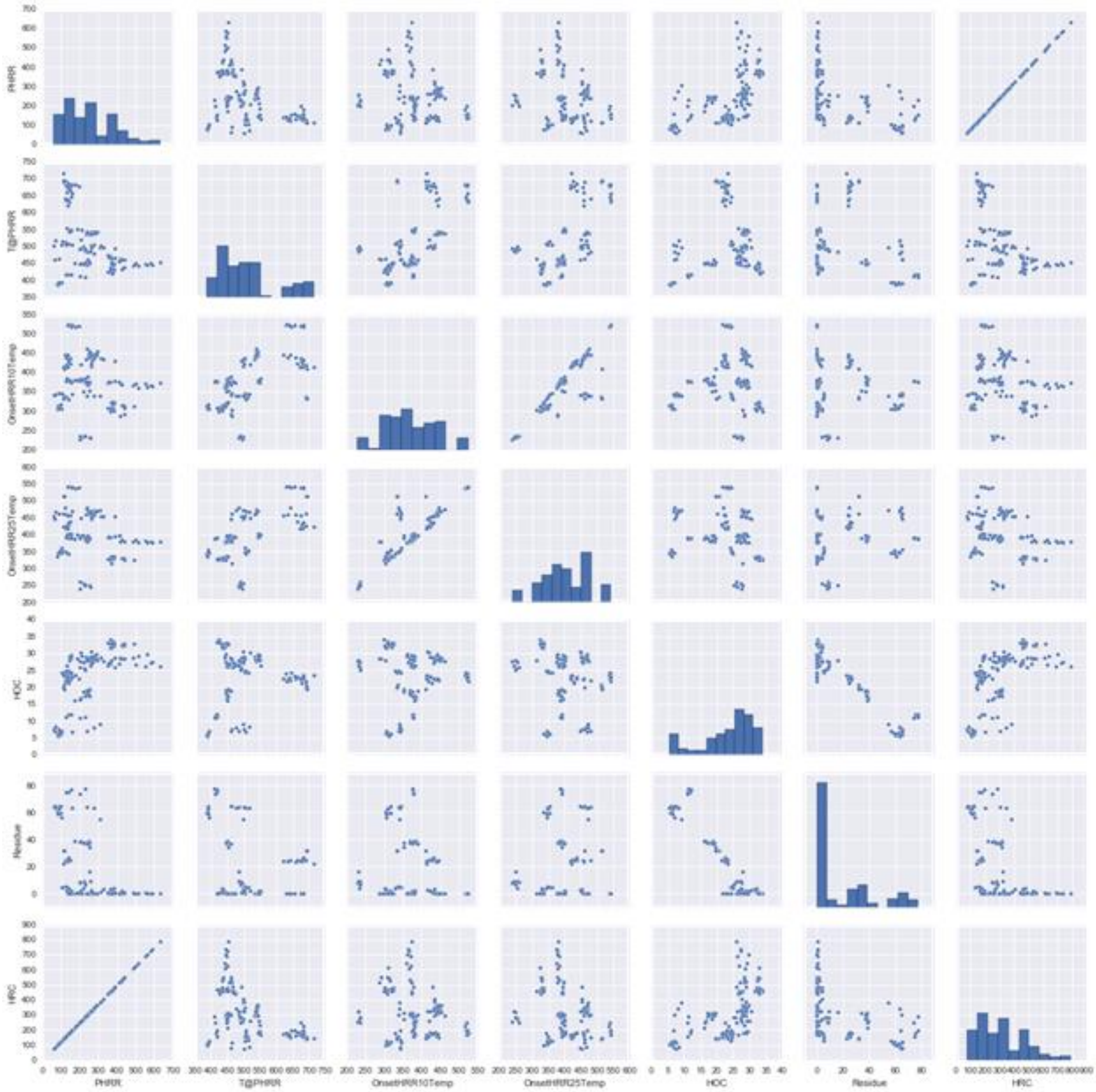
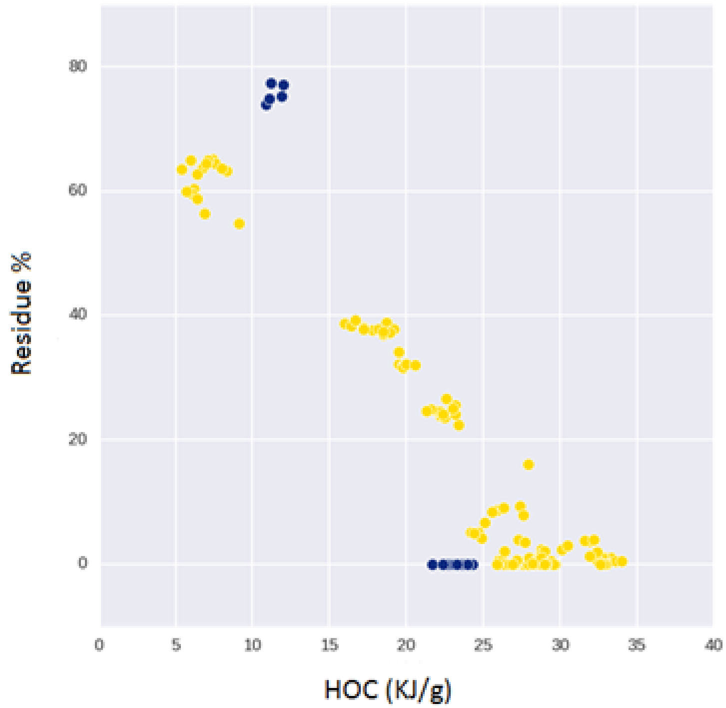


Figure 6 shows the presence of a negative correlation between HOC and Residue. However, by color coding the data to show whether the sample was milled (blue) or bulk form (yellow), then it becomes obvious that the strong correlation holds for the bulk form of the sample. Though there is a relatively strong negative correlation, both attributes are kept for the model building process. A similar analysis was conducted for all the pairings shown to understand where gaps in the data exist and which attributes might be useful to keep in building the model.

Figure 6



MODELING

The EDA helped pare down the list of potential MCC measurements to only the PHRR, T@PHRR, HOC, Residue, OnsetHRR10Temp and OnsetHRR25Temp. In addition, the thickness from the UL 94 test is a necessary input as it is well known that sample thickness affects assigned rating. For the output, the model needs to predict one of four possible UL 94 ratings: V-0, V-1, V-2, and Fail. Hence the ML algorithms that are described as multi-class classifiers (with a cross validation approach) will be the basis for the modelling building. Specifically, the Random Forest classifier was the algorithm of choice as it performed better than a number of other classifiers. The details of this process are not presented here other than to note that using more than one model is a best practice. There is no single universal algorithm that will work best with all datasets.

Before moving ahead with the modelling building process, it is important to upfront agree on an evaluation metric, one that will help establish the accuracy level of the model and be the basis for deciding on whether a model is ready for release. For this classifier, accuracy will be evaluated on the basis of proportion of true results (both true positives and true negatives) among the total number of cases examined. Tuning of the hyper-parameters of the algorithm required that a grid search technique was used on the training set along with a k-fold cross validation.

Modeling Results

Figure 7 shows what is called a confusion matrix. It is a very simple and compact table that compares the different predictions versus the actual ratings for UL 94 for the AR samples. In this case, the overall accuracy of the model is approximately 52%. This is better than the case for the OA samples which produced a model with an accuracy of 50%. It may not seem like a big difference but for such a small dataset, it does demonstrate that there is likely a stronger relationship between AR samples from UL 94 rating to MCC. This was expected since the MCC samples are not subject to any conditioning.

Figure 7

		Predicted Rating			
		V-0	V-1	V-2	Fail
Actual Rating	V-0	384	110	95	36
	V-1	95	0	20	0
	V-2	82	10	296	57
	Fail	37	3	80	0

From the confusion matrix, it can be seen that one area that challenged the model was accurately predicting V-1 and Fail. The model performed much better for V-0 and V-2. Part of the error may be due to the test method and how a single sample results in a rating change, always lowering the rating. V-0 is considered the best rating followed by V-1, V-2 and finally Fail. The other part of the error may be due to the small number of samples that earned a V-1 or Fail rating in the larger dataset.

Another positive feature of the ML-based algorithms is that the classifier provides a relative ranking of the importance of all the input variables. Figure 8 shows the attribute rankings with thickness having the largest effect. This was expected and shows how a non-material level property influences the predictions. Next is the Residue as the most prominent MCC measurement followed by PHRR, HOC and OnsetHRR10Temp being basically equally impactful. This sort of ranking can help lead to learning as an explanation on why this ranking might be correct needs to take place.

Figure 9 shows the confusion matrix for an alternative method for determining the UL 94 rating on the AR samples. As mentioned earlier, the normal UL 94 rating could be determined based on a single sample, the "bad apple". The theoretical foundations of predictive modeling would conclude that trying to model such a scenario will be unsuitable. So to investigate this issue, a new UL 94 rating was assigned to the sample grouping based on a "majority rules" approach. In this case, the rating that was most dominant amongst the 5 samples was selected. This did result in a change in the UL 94 ratings for some of the samples. The accuracy for the "majority rules" was 67% versus 52% for the "bad apples". This suggests that the current method in which the UL 94 rating is determined will be a limiting factor in developing a highly accurate relationship between MCC data and UL 94 ratings.

Figure 8

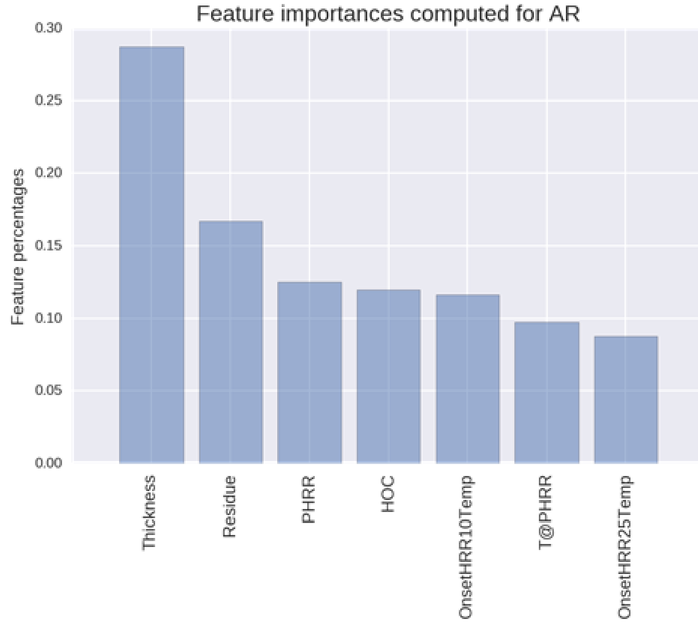


Figure 9

		Predicted Rating			
		V-0	V-1	V-2	Fail
Actual Rating	V-0	485	95	95	15
	V-1	80	0	0	0
	V-2	35	0	391	19
	Fail	22	3	65	0

SUMMARY AND RECOMMENDATIONS

Progress in ML makes it a very practical tool to help manufacturers gain insight into their product development. In this paper, the power of ML is demonstrated with the example of finding a relationship between MCC measurements and UL 94 to help predict the subsequent UL 94 rating. The methodical approach used in the model building uncovered relationships between different MCC attributes and strengthened the choice of the MCC as a surrogate for the UL 94 rating. However, gaps in the data and the quirk of the how the rating is assigned may also suggest some limits to building a very accurate model.

REFERENCES

1. **UL 94.** *Standard for Tests for Flammability of Plastic Materials for Parts in Devices and Appliances.* 2013. Edition 6.

2. **UL.** *Microscale Combustion Calorimetry (MCC) Testing for Polymeric Materials.* s.l. : Performance Materials, 2014.

3. *Micro-scale evalution of flammability for cable materials.* **Yang, HE et al.** s.l. : Proceedings for 62nd International Wire & Cable Symposium, 2013. Vol. Proceedings.

4. *Standard Test for Flammability Screening of AC Materials.* **Lyon, RE.** s.l. : International Aircraft Materials Fire Test Working Group, 2006.

5. *Correlations between Microscale Combustion Calorimetry and Conventional Flammability Tests for Flame Retardant Wire and Cable Compounds.* **Lyon, HE et. al.** s.l. : 56th International Wire & Cable Symposium, 2007.

6. *Assessing Material Consistency Using MCC.* **Fabian, Tom.** s.l. : International Aircraft Materials Fire Test Working Group Meeting, 2015.

7. *On the use of computational neural networks for the prediction of polymer properties.* **Sumpter, BG et al.** 3-4, s.l. : Journal of Thermal Analysis and Calorimetry, 2005, Vol. 46.