Augmentation of an Artificial Neural Network (ANN) Model with Expert Knowledge of Critical Combustion Features for Optimizing a Compression Ignition Engine Using Multiple Injections

2017-01-0701
Published 03/28/2017

Aaron M. Bertram and Song-Charng Kong
Iowa State University

CITATION: Bertram, A. and Kong, S., "Augmentation of an Artificial Neural Network (ANN) Model with Expert Knowledge of Critical Combustion Features for Optimizing a Compression Ignition Engine using Multiple Injections," SAE Technical Paper 2017-01-0701, 2017, doi:10.4271/2017-01-0701.

Abstract

The objective of this work was to identify methods of reliably predicting optimum operating conditions in an experimental compression ignition engine using multiple injections. Abstract modeling offered an efficient way to predict large volumes data, when compared with simulation, although the initial cost of constructing such models can be large. This work aims to reduce that initial cost by adding knowledge about the favorable network structures and training rules which are discovered. The data were gathered from a high pressure common rail direct injection turbocharged compression ignition engine utilizing a high EGR configuration. The range of design parameters were relatively large; 100 MPa - 240 MPa for fuel pressure, up to 62% EGR using a modified, long-route, low pressure EGR system, while the pilot timing, main timing, and pilot ratio were free within the safe operating window for the engine. The limits included restricting the dwell between injections, the upper and lower limits of main timing, and the pilot ratio which was limited to 50%. The outcomes of the research are expected to provide important insight for accelerating and augmenting engine testing by offering an alternative to exhaustive phenomenological modeling. Quantification of the models ability to represent the experimental data is done using cross validation and various metrics are used to prevent overtraining while also maintaining high computational performance. A brief overview of the relative performance of the methods is also provided.

Introduction

Compression ignition engines have undergone major technological revisions over the last decade due to various regulatory and economic pressures. While the engines in the marketplace right now rely on the same fundamental mechanisms that have been in place for over a hundred years the technology they employ to deliver fuel and control the combustion conditions is somewhat revolutionary. Technological advancements in manufacturing, electronics, and materials have all provided significant contributions to the modern combustion engines' specifications. On the other hand, design engineers need new design tools to keep up with this changing landscape of engine design.

The engine design process made up of many phases and while the general trend of design paradigms indicates that products will be designed as fully as possible before being built, engine manufacturers, like many others, will struggle with implementing this design method completely. Combustion is an inherently difficult process to model and predict. Still, models are useful tools in all phases of design, though they need to be refined and tuned to each application.

While substantial parts of the design phase occur in a virtualized environment, engine testing is done on real engines after manufacturing to provide the emissions and fuel consumption data which are required by several key stakeholders. Engine testing is the critical phase that connects the prior design process with the potential customer experience. Beyond emission certifications and fuel economy results engine testing provides design engineers with feedback about the success of the design and the predictions made about the design performance. On the other hand, engine testing is costly, and little can be done to accelerate the testing process due to the time required to reach steady state in a typical engines. Improving this testing time substantially would require radical approaches in material selection to reduce the thermal capacitance and heat losses without substantial purpose other than to accelerate testing, which would be impractical. Still, heat loss mitigation such as thermal coatings may reduce the time required to reach thermal equilibrium and may improve transient performance of the engine, a worthy design objective. Certainly, thermal coatings are popular in research although not yet applied liberally in commercial production, and in the meantime other methods to accelerate testing are beneficial.

Engine calibration, as the final stage of design, provides the data to feed back into the iterative design loop. Ideally, engine calibration would follow a rigorous process testing each potential combination of factors to a resolution that would produce a complete and smooth

response surface of the engine output. That approach represents engine calibration as it was done 20 years ago, when the design factors in the calibration phase were far fewer than at present. Today, the feasibility of complete engine testing has been eroded by the increasing number of dynamic tuning factors available to calibration engineers. Engineers must either divide the problem of optimizing the engine calibration down into smaller pieces or they must augment their data with external models. To provide a confident global result, testing smaller ranges of the design space is not very useful. Other tools include heuristic optimization methods such as the genetic algorithm and particle swarm algorithm, which have both gained some reputation for being successful ways of finding near-optimal regions of operation within the design space. Still, such heuristics can fail to locate global optimums and they do still require much more data as design spaces increase, and furthermore, they can become trapped in local optimums. Data driven models can fill the gap between engine testing and optimization routines by providing a more rapid and economical results to the optimizer. Once the optimum location is determined, further engine testing in the reduced space can produce more confident results than modeling alone, while locating the optimum much faster than a full factorial test of the engine can. The role of data modeling in this process will grow as new technologies develop and engine mechanical properties remain rather constant.

Models used for this purpose are generally classified as either abstract or physical, meaning, phenomenological. Physical modeling is appealing, especially to general scientists, because once the model is complete and tuned, it can provide a vast amount of data about the process itself. Physical models aim to provide a complete knowledge of the entire process as it happens. They are generally very sensitive to certain input data, require a very deep and rigorous attention to the system boundaries and engineering assumptions made, and usually have a high fixed cost per iteration, once developed. On the other hand, abstract modeling approaches aim to connect the inputs and outputs in the most direct way possible. Abstract modeling is advantageous to physical models in that the cost to develop the initial model is the primary cost and the iterative cost of the model is very low. Also, being a data driven approach, unknown factors can be accounted for before they are understood completely. Abstract models fall short of providing the kind of insight about the process that phenomenological models provide, but with the advantage of being more rapid to develop and more flexible once built. These modeling techniques, though competitive, provide different results which cater to different goals in their use.

Abstract modeling has enjoyed a long history of use dating back to C. Gauss and Legendre in the first decade of the 1800s, when Gauss' method was used to predict the emergence of the planet Ceres without solving the non-linear equations of motion developed by Kepler. Generally speaking, it is the process of creating a model and optimizing the parameters of that model to fit some data with the least error. The Least Squares method is a specific application of the generally abstract models. The abstraction characteristic comes from the fundamental assumption that there are unknown factors at work in the model and the parameters which reduce the error most will not necessarily elucidate those factors.

Most recently abstract modeling has appreciated increased attention in the fields of Artificial Intelligence and Machine Learning. These methods can be classified as Regression, Regularization, Instance-based, Trees, Bayesian, Clustering, Association Learning, Artificial Neural Networks (ANN), Deep Learning, Dimensionality Reduction (DR), or Ensemble methods. The methods most prudent and realistic for the problem of optimizing engine calibrations are Regression, Regularization (e.g. Bayesian), ANN, DR, and Ensemble methods, as these methods deal strictly with continuous variables while the other methods listed are primarily used for logistical problems, which may still arise in situations such as cyclical variation analysis.[1] Taken as individuals, and in no specific context, these methods are often referred to as "black boxes" and, while that term makes sense, it isn't technically precise. These methods can utilize a priori information and the methods can be used in conjunction with known models that use known constructions which lead to "gray box" method. These methods can also be used to extract rule sets to garner information about the system and processes, though such work is beyond the scope of this application.

Several studies show that ANNs have become popular for predicting pressure features [2] and emissions characteristics [2, 3, 4]. ANNs have been used to predict emissions [2, 3, 4, 5, 6], ID [3], Combustion Duration (CD) , EGT, Brake Thermal Efficiency (BTE) [2, 3, 4, 5], Peak Pressure (PP) [2], Peak Pressure Rise Rate (PPRR) [2], wall heat transfer, cylinder temperature [2]. ANNs have been used to predict the effect of fuel blend [6], injection schedule [3], fuel pressure [5–6], EGR rate, valve timing, compression ratio, cetane number, boost pressure, and more. Some studies sought to resolve the cylinder pressure and temperature in time [2] while others sought only to identify the critical features of the pressure trace.

This study examined the use of post-hoc analyzed cylinder pressure data to augment the model structure. The cylinder pressure data was not considered as an input to the combustion system, nor were the pressure features desired as outputs. Instead, this corollary data was used to enhance the overall model reliability in predicting the outputs from the inputs since the pressure data is informative and almost always collected despite not always being used directly. In this study, statistical tools are used to first predict the pressure features from the input data. After the initial modeling of the pressure trace features the input data are collected with the pressure predictions and then all together are used to predict the final output. A demonstrative graphic of this process using just three inputs, one predicted factor, and one output, is shown in Figure 1. Using the key features of the combustion trace in addition to the input parameters enhances the prediction of the output data, provided the pressure prediction models have high certainty. As a necessary step towards this model structure the pressure features which can be modeled most accurately must be identified.

Many pressure trace and heat release features have been found to be useful in understanding the mechanisms at work during the combustion process. These features were extracted from the total pressure trace, the combustion pressure decomposition, and the heat release rate calculation. Important features from literature included absolute magnitudes, magnitude of derivatives, and the crank angle timing of such features, as previously identified. Singular event related features include the Start of Combustion (SOC) while other event related features are related to a duration, such as the angle of 50% combustion completion (MFB50). These features are summarized in Table 1.

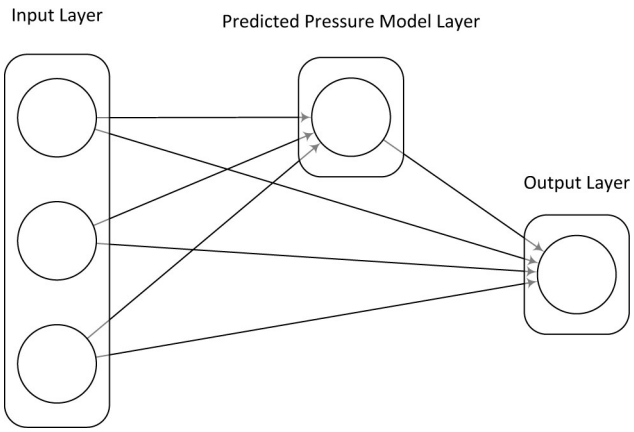

Figure 1. Simplified illustration of the method showing the two-layered approach to first modeling an intermediate piece of data and then using the prediction along with the input layer to predict the output. Behind each connective arrow is a hidden neural network layer. This study used five inputs and many intermediate predictions,

Table 1. Collection of relevant pressure trace features from literature, accounting for all commonly used combustion trace features.

Raw Pressure Max	Pressure Derivative Max	Timing of Both
Ave. Pressure Max	Average Derivative Max	Timing of Both
Combed Pressure Max	Combed Derivative Max	Timing of Both
Filtered Pressure Max	Filtered Derivative Max	Timing of Both
Harmonic Pressure Max	Harmonic Derivative Max	Timing of Both
Combustion Press. Max	Combustion Deriv. Max	Timing of Both
Heat Release Rate Max	HRR Derivative Max	Timing of Both
Start of Combustion		Timing only
Mass Fraction Burned	MFB 5,10,50,90,95,98,99	Timing only

The scope of this study of the input variables was limited to include common tuning parameters for a dual injection scheme with variable fuel pressure. A modified exhaust gas recirculation (EGR) system allowed a wide range of EGR rates by utilizing long route and external pumping assistance. The emissions measured included hydrocarbons (HC), carbon monoxide (CO), carbon dioxide (CO2), nitrogen oxides (NOx), along with particulate matter (PM) and fuel consumption (SFC). The torque and speed were kept constant along with the inlet conditions for air and fuel. The data were gathered using a heuristic algorithm to obtain near-optimal results while representing the global optimization approach. From a statistical standpoint, this data gathering method has a high potential to create leverage and distribution problems as the data are not spread evenly, and they may be outside the region of interest. Data which had high leverage were removed from the original data set to reduce the impact of leverage and clumping without reducing the quality of the final models. It was prudent to do so, as well, because the points which have very high emissions and are far away from the optimum are not of interest to those optimizing the calibrations.

Methods

Engine Testing

The test apparatus consisted of a John Deere 4045T 4-cylinder turbocharged engine, with detailed specifications as in Table 2, coupled to a GE Dynamometer. The engine was equipped with a specialized EGR system allowing for temperature and flow control. The exhaust gas was sampled using a Horiba MEXA 7100 gas analyzer and an AVL smoke meter. The engine was equipped with a Kistler piezoelectric pressure transducer, fitted into the glow plug hole, and charge amplifier. Data was captured with a National Instruments data acquisition system and control of the engine was done using a proprietary tool to modify the ECU settings to deliver fuel and control fuel pressure.

Table 2. Specifications for the John Deere 4045T engine used in the study.

Displaced volume	4499 cc
Cylinders	4
Stroke	127 mm
Bore	106 mm
Connecting Rod	203 mm
Compression ratio	17:1
Number of Valves	4

Data Context

The data set was generated using a hybrid (particle swarm and genetic algorithm) heuristic optimization routine as found in [7]. The resulting data set was the result of a goal-driven process rather than a designed experimental study, which was more representative of an industrial calibration process. Data points with high leverage and regions with tightly clustered data were trimmed manually reducing the data set from 163 data points to 159 data points. As the data was not distributed evenly and the study was not “designed” traditional methods of Response Surface Modeling (RSM) are not valid. In this context, abstract modeling provided a better solution to those more straightforward methods previously established in RSM.

The data set contained combustion of 3 characteristic regimes; ignition before TDC with a single HRR event, ignition after TDC with a single HRR event, and a regime with two HRR events. This challenge is particularly difficult to account for in analysis when comparing one regime with another. Premixed combustion and diffusion combustion happen at vastly different rates and thus, the model should be able to detect the type of combustion, though the relevance of the key parameters may be dramatically different for each type.

Data Analysis

The pressure data was gathered for 500 engine cycles and averaged by the data acquisition system. Pressure and system state data were recorded and analyzed offline. Pressure data was filtered using established methods of rolling average (digital comb) and a 5th order Butterworth filter, in forward and reverse, to eliminate angle domain shift. Heat release was calculated using the method outlined by Heywood and summarized in Ferguson and Kirkpatrick [8–9], by assuming a low wall temperature, 600 K justified by the conditions of high EGR and late injections, as used in previous LTC studies on this particular engine.

Cylinder pressure was decomposed using the polytropic model with an index of 1.35, which has been traditionally used on this apparatus and under these conditions. Non-traditional applications were also explored to compare the effectiveness in terms of producing accurate

and repeatable predictors. The angle (timing) of maximum pressure derivative was taken from the unfiltered, filtered, and decomposed pressure data. The angle and value of heat release rate maximum were also considered along with the SOC. The combustion progress pinpoints, commonly referred to as MFB10, MFB50, MFB90, and MFB95, were also investigated, though the expectation was that since they are calculated from the heat release integration they may not provide more useful data. The use of the EOC is investigated but due to the uncertainty involved in determining the EOC it was not expected to provide useful input to the neural network.

ANN Structure

In this work artificial neural networks were used to create abstract models of the combustion process. While other model methods may be effective for modeling in engine testing, in general, the ANN was chosen to accommodate the features of the data collection method, as this analysis is meant to augment the calibration procedure. The methods used in this analysis accommodate for small sample size, high order input space, unknown correlation between inputs, and unequally distributed sample data. More traditional methods such as principal component analysis (PCA) are sensitive to these types of data while online calibration using only heuristics cannot provide the type of data required for those statistical methods. Furthermore, methods such as support vector machine regression lack the ability to ignore extraneous factors in the final output model. A potential shortcoming of simplified methods of data reduction is the potential to ignore a substantially different mode in a group of data, for example, PCCI regimes with traditional regimes, which would mean that the assumption of common variance may not be applicable in such data sets.

While other modeling techniques could replace the neural network in the pressure prediction step the ANN offers the ability to improve the model iteratively as data is collected, and this feature is not replicated by the other modeling techniques, as they must define the number of necessary features strictly to eliminate correlated inputs, and yet, this is not possible before the data is gathered. Thus, the ANN provides a unique ability to ignore extraneous data while capturing highly complex relationships with data sets ranging from small to large while also using prior model data as a starting point to improve performance.

The neural networks used in this study were standard models commonly available in software packages such as the Matlab Neural Network Toolbox. All neural networks rely on the perceptron model which sums a weighted set of inputs, applies a bias, and processes an activation function to determine the neuron's output. Grouping perceptrons into layers and then a larger topology forms the actual neural network. Distinguishing factors of ANNs are the use of a fixed bias versus a complicated bias vector, the specific activation function(s) chosen, and the interconnectivity of the layers.

The second key factor in how an ANN is described is the training function. While the network properties are a feature of the ANN while the training method itself is not a property of the network, it is still often referred to when describing the ANN. More precisely, the training function is the method by which the connection weights and biases are optimized to achieve the final network solution. There is an interaction between the network structure, the data, and the training method which cannot be ignored. Generally, a network which is too small cannot fit data well enough while a much larger network can overfit or become trapped in a local optimum more easily. The analysis of the training methods available and how the network structures interact with data and training methods are beyond the scope of this study. The purpose of the networks in this study were to demonstrate how a priori knowledge and analytical techniques could impact the reliability of abstract models in predicting the outputs. The neural network tuning was performed by adjusting the network size and examining how the networks performed in cross-validation.

The ANNs used in this study were the common three-layer construction with two layers for conditioning the input and output with one hidden layer where the abstract model is contained. Specialized network structures, as with various training methods, were beyond the scope of this study. A single model for each intermediate variable and each output variable were produced. Each network to predict intermediate values was given the full set of six input variables. Many networks for each intermediate variable were trained so that the repeatability of the model structure itself could be analyzed, rather than a singular solution. This was critical as with any heuristic or stochastic method a singular result is not very informative as to the function of the overall structure. A preliminary study indicated that 40 networks were enough to get information about the repeatability of the network structure. While 40 network training sessions could provide information about the network architecture performance, 100 networks were trained for each set to address statistical quality concerns of using too few networks to characterize the topology.

Fitting

The data set of 158 test conditions was divided into training, validation, and test groups at a rate of 70%, 15%, and 15% respectively. Each network training session involved randomly dividing the validation groups, training the network, recording the fit parameters, and then repeating the process for 100 iterations. By examining the mean R^2 fit parameter an overall understanding of how well the network structure was capturing the data. At the same time, it was critical to prevent overfitting. Overfitting is characterized by high errors on the validation set with nearly interpolating results in the training set. Overfitting can be seen in the test set as well, resulting in a fit which was better for the training than the validation or test. Overfitting was prevented by tracking the data errors made by all 100 networks, and summing the square of these raw errors, as with the standard MSE. The variance of the errors was tracked, per data point, and a high variance of errors on each data point indicated overfitting. The network size which produced the lowest average variance in the errors overall data and all 100 iterations was chosen as the best network size. This did not always agree with the network performance, the overall MSE of the individual networks. Examining the MSE of each network alone would have allowed for large errors on a small set of data to be treated the same as smaller errors on the entire set, which would have allowed for overfitting, and eventually, impractically large solutions with low generalization potential.

The predictive power of the networks was averaged and reported as R^2 for each parameter. This was the average R^2 value of the multitude of networks trained for each parameter. A high correlation reported that the network was reliably predicting outputs even when retrained many times with different sets of training data. Variation in R^2, which

would have represented a failure of an individual network to tune to the data, resulted in a reduction of the mean R^2. Parameters which achieved an R^2 of 0.98 and above were considered reliable, while 0.90 to 0.97 represented some unaccounted factors, and below 0.90 was a poor fit. The ability of the networks to predict each combustion feature was tracked and compared.

This methodology follows general principles of cross-validation as well as bootstrapping and ensemble methods, but ensemble approaches were not used in generating the final models. The variation in performance of the networks was only used to provide an indication that the network topology was sufficient to produce similar results on different divisions of the training and validation data. In fact, ensemble methods leverage the differences between networks trained on slightly different data sub-sets while the methods used in this study were designed to indicate minimum differences between models trained across different divisions of the data set. Those variations could come from both under fitting and over fitting the minimum variation from model to model is a good indicator of appropriate network size, at least, for simple ANN.

Reference Model

The same neural network sizing and training methods were applied directly to the five input conditions, utilized as controls by the heuristic optimization (PSO) routine, and those networks were considered as the single-layer reference for comparison as to using the improved, two-layer approach. It was expected that the two-layer method can outperform the single-layer method in cases where the single-layer method was inadequate, whereas in cases where the single-layer method was very accurate, the additional uncertainty introduced by adding data and calculations as well as another model layer may degrade the ability of the model to predict the results. Each pressure feature model was trained on a different, random division of the data and standard early stopping criteria were used to prevent over-training. Though multiple models were created for each pressure feature to determine the appropriate network size only one model was used for each pressure feature prediction in final testing, and these predicted pressure features were combined with the five engine test condition inputs to generate one ANN model for the prediction of the engine output.

Results

Parameter Selection

The network size selection process produced ANNs ranging from 6 to 15 nodes in the hidden layers. The layer selection choices are summarized in Table 3 and indicate that some parameters were more difficult to model than others. Considering a single hidden layer, the increase in model parameters is proportional with the node quantity. The factors were grouped into three groups based on the mean model performance. Factors which achieved very high performance were considered for the two-layer solution. These included the angle of maximum combustion pressure (decomposed), angle of maximum HRR, angle of both their derivatives, the absolute pressure maximum of the raw, comb-filtered, and resonance-filtered pressure, as well as the MFB10, MFB50, and MFB90. These factors all achieved a very high repeatability, measured by low variance of model performance, and overall fit quality, measured by the correlation coefficient.

Table 3. Pressure trace features examined, optimal network size for prediction, and mean prediction performance (R^2).

Maximum / Value	Nodes	R^2
Raw Pressure Timing	11	0.90
Raw Pressure Deriv. Timing	15	0.92
Combed Pressure Timing	13	0.91
Combed Pressure Deriv. Timing	13	0.95
Filtered Pressure Timing	13	0.92
Filtered Pressure Deriv. Timing	11	0.93
Harmonic Pressure Timing	15	0.92
Harmonic Pressure Deriv.	13	0.95
Combustion Pressure Timing	11	0.98
Combustion Pressure Deriv. Timing	11	0.98
Heat Release Rate Timing	13	0.99
Heat Release Rate Deriv. Timing	11	0.96
Raw Pressure	9	0.99
Raw Pressure Deriv.	12	0.74
Combed Pressure	9	0.99
Combed Pressure Deriv.	11	0.78
Filtered Pressure	10	0.99
Filtered Pressure Deriv.	11	0.87
Harmonic Pressure	6	0.67
Harmonic Pressure Deriv.	9	0.84
Combustion Pressure	9	0.71
Combustion Pressure Deriv.	9	0.71
Heat Release Rate	8	0.87
Heat Release Rate Deriv.	7	0.83
Start of Combustion	11	0.96
MFB 10%	9	0.98
MFB 50%	9	0.999
MFB 90%	9	0.98
MFB 95%	6	0.96
MFB 98%	10	0.93
MFB 99%	11	0.89

Factors which had high fit quality but reliability issues included the SOC, MFB98, MFB95, as well as the angle of maximum pressure derivatives of the harmonic component, raw, comb-filtered, harmonic-filtered, and the angle of the maximum of the same harmonic component, and raw, comb-filtered, and harmonic-filtered pressures. These factors achieved a reliability high enough to be predicted by ANNs for analysis, but they did not achieve a high enough reliability to be considered for the two-layer approach.

Remaining factors which were not able to be reliably predicted from the input factors were the maximum harmonic pressure, the maximum combustion pressure, the maximum HRR, and the derivatives of the harmonic, raw, comb-filtered, harmonic-filtered, and combustion pressures. The derivative of HRR along with the MFB5 and MFB99 were not predicted reliably enough to be useful, nor did they have much traditional use in characterizing combustion.

Two-Layer Approach

Using the factors identified from the selection process a two-step model was created for each of the five objectives. The two-step models were compared to single-layer models using the same process of ANN structure validation and size selection. The performance results of the two-step models were compared to the single-layer models and results are provided in Table 4. The quality of fit was improved by the added factors for all objectives. Fuel consumption predictions remained difficult while emissions predictions were raised to very high predictive power. The fuel consumption was not adequately modeled by the simple ANN structure without increasing the network size. Improvements were seen in using larger, regularized networks, in general, and when a 28 node Bayesian regularized network was used an improvement in correlation from 0.90 to 0.96 was seen. In all cases, improvements were seen by adding the additional predictive factors from the first stage of modeling.

Table 4. Comparison of the single-layer model results and the enhanced two-step model results. Improved mean correlation fit parameters can be seen for all objectives. * An improved 28 node model using Bayesian Regularization was used to demonstrate a potential solution to the problem of under-performing models

Objective	One-layer model		Two-layer model	
	R^2	Nodes	R^2	Nodes
CO	0.99	9	0.99	9
THC	0.91	8	0.96	9
NOx	0.92	9	0.98	9
PM	0.86	9	0.99	9
Fuel Consumption	0.80	7	0.89	7
Fuel Consumption*	0.90	28*	0.96	28*

Discussion and Conclusion

As the results indicate an improvement in predictive power was found by adding additional factors to the input. By using only additional factors which can be modeled accurately added uncertainty was avoided. The increase in predictive power raised all models to levels which could be used in the calibration optimization process. As the margin between emission limits and actual test conditions is often quite small, this high predictive power was important for creating a relevant model. The overall cost of producing the additional models, while not trivial, was low as compared with simulation methods, taking just a few minutes each on a modern desktop PC. In particular, the smoke emissions prediction was improved dramatically, from 0.86 to 0.99, demonstrating how the additional input knowledge improved prediction reliability of the model dramatically.

These results indicate that when high-precision abstract models can incorporate extra a priori information about the experimental system the overall improvement can enhance the model from being suggestive (R^2 from 0.6-0.85) to being predictive (R^2 from 0.85 to 0.999). The value of expert knowledge can be added to abstract models without creating physical sub-models. Furthermore, data reduction using expert knowledge was validated, as the entire pressure curve was replaced with just 9 features. The process of reducing the approximately 1800 distinct pressure data measurement points (relevant to combustion) to 9 features using expert analysis meant that the model complexity and scope of the factors and observations didn't change dramatically. In other words, by adding only a small number of input factors and no new observations the modeling structure could remain largely the same. Should the new considerations have raised the number of factors above the number of observations, or raised the order of magnitude of the observations, the resultant model structure would no longer be appropriate.

The specific results indicated that the predictive power of some combustion features often discussed in literature were less than expected. One notable example was the peak pressure rise rate, and the angle of peak pressure rise rate. This characteristic feature contributes to noise and can influence longevity of components, but the ability of the ANN to predict it was low, and as such, it could not be used in the two-layer modeling approach. While the peak pressure rise rate and timing of such may be useful in general analysis, it was replaced by the extracted combustion pressure component with an improvement in overall reliability as a modeling factor. The current study is limited to modeling the parameter and does not indicate that the peak pressure rise rate, as measured data, is not relevant to predicting emissions or characterizing combustion, but rather that it cannot be modeled from inlet conditions using the simple model structure presented here and that it was not necessary to achieve reliability in the final two-layer models.

The results of this study indicate that the hybrid modeling approach of using expert analysis combined with abstract models can improve such models to a level required for use in iterative design and optimization, and potentially enough to be used in certified reliability processes. The key parameters identified by the study are relevant to compression ignition engines using multiple injection strategies with a wide variety of combustion profiles, excluding HCCI (fully premixed).

One particularly important use of this modeling approach is in rapid optimization applications. The Genetic Algorithm (GA) has been used frequently with the downside of requiring many iterations, and therefore, test or simulation resources. With highly accurate, abstract models available, such as this two-layer neural network approach, algorithms such as the GA can be utilized without a substantial investment in massive computing resources. Beyond using optimization routines these types of models can produce highly dense fields of response predictions very rapidly without low computational cost and can further improve on heuristic optimization results. This type of low cost model provides a useful mitigation of the problem of local extrema trapping that occurs with global optimization methods.

Additional work should include comparison of this modeling method with other methods with the overall aim to develop accurate abstract model structures which incorporate the minimal factors necessary to achieve industry standard reliability and repeatability in product development.

References

1. Shi, Y., Ge, H.-W., Reitz, R., "Computational Optimization of Internal Combustion Engines" (London, Springer-Verlag, 2011) 125–126, doi: 10.1007/978-0-85729-619-1

2. Yusaf, T., Buttsworth, D., Saleh, K., Yousif, B., "CNG-diesel Engine Performance and Exhaust Emission Analysis with the Aid of Artificial Neural Network," *Applied Energy*87(5):1661–1669, 2010, doi: 10.1016/j.apenergy.2009.10.009

3. Kannan, G., Balasubramanian, K., Anand, R., " Artificial neural network approach to study the effect of injection pressure and timing on diesel engine performance fueled with biodiesel" *Int.J Automot. Technol*.2013 14:507. doi:10.1007/s12239-013-0055-6

4 5. He, Y., Rutland, C., "Application of artificial neural networks in engine modelling", Intl. Journal of Engine Research,5(4):281–296, 2004, doi:10.1243/146808704323224204

6. Ghobadian, B., Rahimi, H., Nikbakht, A., Najafi, G., . "Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network," *Renewable Energy*34(4):976–982, 2009, doi:10.1016/j.renene.2008.08.008

7. Bertram, A., Zhang, Q., Kong, S.-C., "A novel particle swarm and genetic algorithm hybrid method for diesel engine performance optimization," *Intl. Journal of Engine Research*17: 732–747, 2015, doi: 10.1177/1468087415611031

8. Heywood, J.B. "Internal Combustion Engine Fundamentals" (New York, NY, McGraw-Hill, 1988) ISBN: 978-0070286375

9. Ferguson, C.R., Kirkpatrick, A.T. "Internal Combustion Engines: Applied Thermosciences, Second Edition" (New York, NY, Wiley, 2001) 122, 274, ISBN: 978-0-471-35617-2

Contact Information

Aaron M. Bertram
bertram@iastate.edu

Definitions / Abbreviations

ANN - Artificial Neural Network

BTE - Brake Thermal Efficiency

CO2 - Carbon Dioxide

CO - Carbon Monoxide

DR - Dimensionality Reduction

EGR - Exhaust Gas Recirculation

EGT - Exhaust Gas Temperature

GA - Genetic Algorithm

HC - Hydrocarbon Emissions

HCCI - Homogeneous Charge Compression Ignition

HRR - Heat Release Rate

MFBXX - XX% Mass Fraction Burned (of total)

MSE - Mean Squared Error

NOx - Nitrogen Oxide Emissions

PCA - Principle Component Analysis

PCCI - Partially Premixed Charge Compression Ignition

PM - Particulate Matter

PP - Peak Pressure

PPRR - Peak Pressure Rise Rate

PSO - Particle Swarm Algorithm

RSM - Response Surface Modeling

SOC - Start of Combustion

TDC - Top Dead Center

The Engineering Meetings Board has approved this paper for publication. It has successfully completed SAE's peer review process under the supervision of the session organizer. The process requires a minimum of three (3) reviews by industry experts.

ISSN 0148-7191

http://papers.sae.org/2017-01-0701