Michał KEKEZ, Leszek RADZISZEWSKI
KIELCE UNIVERSITY OF TECHNOLOGY,
Aleja 1000-lecia Państwa Polskiego 7, 25-314 Kielce

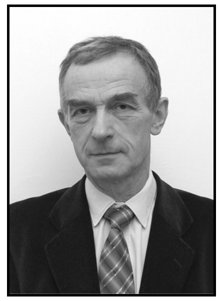
Fuel recognition in compression ignition engine in the real time

Dr. Eng. Michał KEKEZ

Lecturer in the Department of Mechanics, Faculty of Mechatronics and Machine Building, Kielce University of Technology, Poland. Research interests: artificial intelligence, internal combustion engines.

e-mail: m.kekez@tu.kielce.pl

Dr. habil. Eng. Leszek RADZISZEWSKI



Professor in the Department of Mechanics, Faculty of Mechatronics and Machine Building, Kielce University of Technology, Poland. Research interests: acoustics of internal combustion engines.

e-mail: Lradzisz@tu.kielce.pl

Abstract

Contemporary engines allow controlling the fuel injection process, which should be adjusted to a given fuel. On test bench the engine was fuelled by diesel oil, RME or its blends with diesel oil. In this paper selected artificial intelligence methods are used to build classifiers which recognize type of fuel using cylinder pressure curves recorded and averaged for 20, 30, 40, or 50 consecutive engine working cycles. The accuracy of these methods is compared. There is presented the estimation of the minimum number of consecutive engine cycles during which the pressure curves are recorded, required for recognition of type of fuel by a classifier.

Keywords: diesel engines, cylinder pressure, type of fuel, artificial intelligence.

Rozpoznawanie spalanego paliwa w silniku o zapłonie samoczynnym w czasie rzeczywistym

Streszczenie

Silniki o zapłonie samoczynnym mogą być zasilane różnymi paliwami. Współczesne silniki pozwalają na sterowanie procesem zasilania, który powinien być przystosowany do danego paliwa. Na stanowisku badawczym, silnik był zasilany olejem napędowym, paliwem RME (estry metylowe kwasów tłuszczowych oleju rzepakowego) oraz mieszankami tych paliw. W artykule zastosowano wybrane metody sztucznej inteligencji w celu zbudowania klasyfikatora, który rozpoznaje typ paliwa na podstawie przebiegów ciśnienia w cylindrze zarejestrowanych i uśrednionych dla 20, 30, 40 oraz 50 kolejnych cykli roboczych silnika. Przedstawiono porównanie dokładności zastosowanych metod (sztuczne sieci neuronowe oraz drzewa decyzyjne CART i CHAID, zaimplementowane w pakiecie Statistica Data Mining, a także drzewo decyzyjne See5). Przedstawiono oszacowanie minimalnej liczby kolejnych cykli pracy silnika, podczas których rejestrowane są przebiegi ciśnienia, niezbędnej do rozpoznania typu paliwa przez klasyfikator. Zaproponowano również implementację klasyfikatora na mikrokontrolerze, pozwalającą na rozpoznawanie typu paliwa w czasie rzeczywistym.

Słowa kluczowe: silniki o zapłonie samoczynnym, ciśnienie w cylindrze, rodzaj paliwa, sztuczna inteligencja.

1. Introduction

The work of a diesel engine fuelled by various fuels can be modeled by the methods based on the first law of thermodynamics [4], or the computational fluid dynamics methods [1]. The artificial intelligence methods are also used for modeling selected aspects of engine work [2]. Fuel type can be recognized (based on cylinder pressure courses) with use of some of these methods, as proposed by the authors.

2. Fuel recognition

Contemporary engines allow controlling the fuel injection process, which should be adjusted to a given fuel. Automatic recognition of the fuel type will allow switching between control programs for particular fuels.

In experiments on a test bench five different fuels were used: diesel fuel, RME (methyl esters of rapeseed oil), and the blends of diesel fuel with RME, called B10, B20, and B30, which contain 10, 20, and 30 percent of RME, respectively [3]. The experiments were carried out at the maximum load and for several rotational speeds. In the all experiments the data was collected for 50 consecutive engine cycles.

Our aim was to construct a classifier which recognizes the fuel type by using values of the cylinder pressure and the injection pipe pressure.

The classification (recognition) of the fuel type by the See5 Release 2.06 software [5] required preparation of the dataset which was used for creation of a decision tree (classifier).

The dataset consisted of records containing 5 attributes (or variables): fuel type (y), maximum cylinder pressure (x_1), minimum cylinder pressure (x_2), maximum injection pipe pressure (x_3) and minimum injection pipe pressure (x_4). The maximum cylinder pressure values vary for the all tested fuels. The biggest differences occur at the rotational speed of 1200 rpm (revolutions per minute), and therefore the data for this speed was used for preparation of the training dataset.

The experimental data was averaged for each possible 40 consecutive engine cycles taken from 50 consecutive cycles recorded for one fuel. This method produced 11 records for one fuel in a dataset.

The See5 built quite a simple classifier (decision tree shown in Fig. 1 that can be transformed into 5 rules), which recognizes the fuel type with 100% accuracy.

```
x1 <= 8.480765:
:...x2 <= 1.539154: B10 (11)
:   x2 > 1.539154: RME (11)
x1 > 8.480765:
:...x2 <= 1.707941: diesel (11)
    x2 > 1.707941:
    :...x3 <= 20.757: B20 (11)
        x3 > 20.757: B30 (11)
```

Fig. 1. Decision tree obtained by See5
Rys. 1. Drzewo decyzyjne zbudowane przez See5

The classifier has the following form:

y=diesel if x_1>8.480765 and x_2<=1.707941
y=RME if x_1<=8.480765 and x_2>1.539154
y=B10 if x_1<=8.480765 and x_2<=1.539154
y=B20 if x_1>8.480765 and x_2>1.707941 and x_3<=20.757
y=B30 if x_1>8.480765 and x_2>1.707941 and x_3>20.757

When the training dataset was simplified, and contained only the fuel type and maximum cylinder pressure values, i.e. only one input variable instead of four, the classifier obtained by See5 method was also simpler, and still had 100% accuracy:

y=B10 if x_1 <= 8.344099
y=RME if 8.344099 < x_1 <= 8.480765
y=B30 if 8.480765 < x_1 <= 8.608824
y=B20 if 8.608824 < x_1 <= 8.651529
y=diesel if x_1 > 8.651529

The number of cycles necessary to recognize the fuel plays important role in practical applications, because it is linearly proportional to the time required for gathering the data.

When the experimental data was averaged for each possible 30 consecutive engine cycles taken from 50 consecutive cycles recorded for one fuel, this produced 21 records for one fuel in the dataset. As in the previous example, there was one input attribute in the dataset.

The classifier obtained by the See5 method (Fig. 2) was similar to the previous one, and still had 100% accuracy; only the borders of decision regions slightly changed:

y=B10 if x_1 <= 8.347212
y=RME if 8.347212 < x_1 <= 8.491647
y=B30 if 8.491647 < x_1 <= 8.617451
y=B20 if 8.617451 < x_1 <= 8.664569
y=diesel if x_1 > 8.664569

```
x1 <= 8.491647:
:...x1 <= 8.347212: B10 (21)
:   x1 > 8.347212: RME (21)
x1 > 8.491647:
:...x1 <= 8.617451: B30 (21)
    x1 > 8.617451:
    :...x1 <= 8.664569: B20 (21)
        x1 > 8.664569: diesel (21)
```

Fig. 2. Decision tree obtained by See5, when 30 consecutive engine cycles were used
Rys. 2. Drzewo decyzyjne zbudowane przez See5, na podstawie danych dla 30 kolejnych cykli pracy silnika

When the experimental data was averaged for each possible 20 consecutive engine cycles taken from 50 consecutive cycles recorded for one fuel, it produced 31 records for one fuel in a dataset.

In this case, the obtained classifier had lower accuracy (98.7%), but the form of the decision tree, presented in Fig. 3, was almost the same as previously. In Fig. 3 the numbers in parentheses stand for the number of records correctly and incorrectly classified by a given rule. The latter are written after the slash (/).

```
x1 <= 8.505765:
:...x1 <= 8.350143: B10 (31)
:   x1 > 8.350143: RME (31)
x1 > 8.505765:
:...x1 > 8.670647: diesel (31)
    x1 <= 8.670647:
    :...x1 <= 8.62: B30 (33/2)
        x1 > 8.62: B20 (29)
```

Fig. 3. Decision tree obtained by See5, when 20 consecutive engine cycles were used
Rys. 3. Drzewo decyzyjne zbudowane przez See5, na podstawie danych dla 20 kolejnych cykli pracy silnika

When the dataset was created with four input attributes, but using data for 20 consecutive cycles, the See5 built a classifier, which recognized the fuel type with 100% accuracy (Fig. 4). The classifier is similar to that shown in Fig. 1, because it uses the same input variables (the minimum and maximum cylinder pressure and also the maximum injection pipe pressure).

```
x1 <= 8.505765:
:...x2 <= 1.540253: B10 (31)
:   x2 > 1.540253: RME (31)
x1 > 8.505765:
:...x1 > 8.670647: diesel (31)
    x1 <= 8.670647:
    :...x3 <= 20.954: B20 (31)
        x3 > 20.954: B30 (31)
```

Fig. 4. Decision tree obtained by See5 for dataset with 4 input attributes, when 20 consecutive engine cycles were used
Rys. 4. Drzewo decyzyjne zbudowane przez See5, na podstawie danych zawierających 4 atrybuty wejściowe, dla 20 kolejnych cykli pracy silnika

The data averaged for only 10 consecutive cycles allowed creating the decision tree with 6 leaves (Fig. 5) which has 99.5% accuracy. The physical sense of the usage of the x_4 variable (the minimum pressure in an injection pipe) is doubtful.

```
x1 <= 8.511647:
:...x2 <= 1.549044: B10 (41)
:   x2 > 1.549044: RME (41)
x1 > 8.511647:
:...x2 <= 1.726765: diesel (41)
    x2 > 1.726765:
    :...x3 <= 20.904: B20 (34)
        x3 > 20.904:
        :...x4 <= -0.534: B20 (8/1)
            x4 > -0.534: B30 (40)
```

Fig. 5. Decision tree obtained by See5 for dataset with 4 input attributes, when 10 consecutive engine cycles were used
Rys. 5. Drzewo decyzyjne zbudowane przez See5, na podstawie danych zawierających 4 atrybuty wejściowe, dla 10 kolejnych cykli pracy silnika

For the dataset with 2 input attributes, the maximum cylinder pressure (x_1), and the minimum cylinder pressure (x_2), the obtained classifier (Fig. 6) had 7 leaves and 96.6% accuracy.

```
x1 <= 8.511647:
:...x2 <= 1.549044: B10 (41)
:   x2 > 1.549044: RME (41)
x1 > 8.511647:
:...x2 <= 1.726765: diesel (41)
    x2 > 1.726765:
    :...x1 > 8.640588: B20 (24)
        x1 <= 8.640588:
        :...x2 > 1.768882: B30 (20)
            x2 <= 1.768882:
            :...x1 <= 8.605294: B30 (16/2)
                x1 > 8.605294: B20 (22/7)
```

Fig. 6. Decision tree obtained by See5 for dataset with 2 input attributes, when 10 consecutive engine cycles were used
Rys. 6. Drzewo decyzyjne zbudowane przez See5, na podstawie danych zawierających 2 atrybuty wejściowe, dla 10 kolejnych cykli pracy silnika

The type of fuel can be recognized with a satisfactory accuracy by using only 10 consecutive engine cycles, but the usage of 20 consecutive cycles suffices to recognize the type of fuel with 100% accuracy. The form of the obtained classifier is so simple that the classifier can be implemented even in a slow microcontroller. The transparency of the classifier is very high. The decision regions in See5 classifiers are rectangular (i.e. decision borders cannot be oblique to the axes), and this fact may limit the accuracy.

The See5 with enabled "boosting" option builds a committee consisting of maximum of 99 classifiers. This allowed achieving 100% accuracy even by using the data built from only 15 consecutive cycles, and with one input variable (the maximum cylinder pressure). The number of rules in each classifier (in a committee) varied from 5 to 11. The accuracy of a single classifier varied from 51.7% to 98,9%, but the fuel recognition made by the committee of classifiers was carried out with 100% accuracy. Further analysis showed that this committee of classifiers was rather "overfitted" to the training data.

Next, the C&RT and CHAID algorithms, implemented in Statistica software, were used. The experimental data was averaged for each possible 20 consecutive engine cycles taken from 50 consecutive cycles recorded for one fuel, and with only one input variable, the maximum cylinder pressure, present in a dataset.

The classifier (decision tree) built by C&RT (Fig. 7) had the same accuracy as the classifier built by the See5 (Fig. 3), namely 98.7%, which meant that only 2 records were misclassified. The tree has 5 leaves, and therefore it can be transformed into 5 rules.

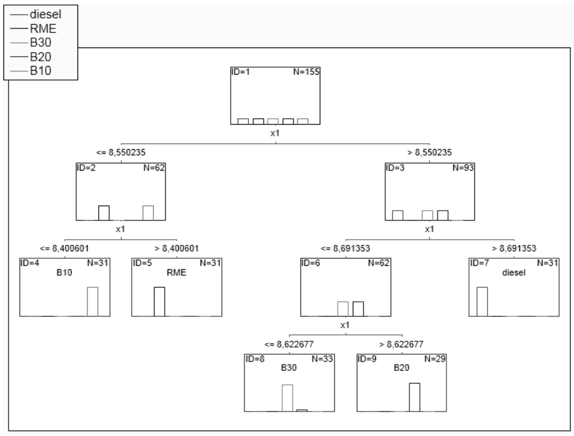


Fig. 7. Decision tree obtained by C&RT, when 20 consecutive engine cycles were used
Rys. 7. Drzewo decyzyjne zbudowane przez C&RT, na podstawie danych dla 20 kolejnych cykli pracy silnika

The CHAID method (with default options) created the decision tree presented in Fig. 8. The accuracy is worse, because 10 records are misclassified (93.5% accuracy).

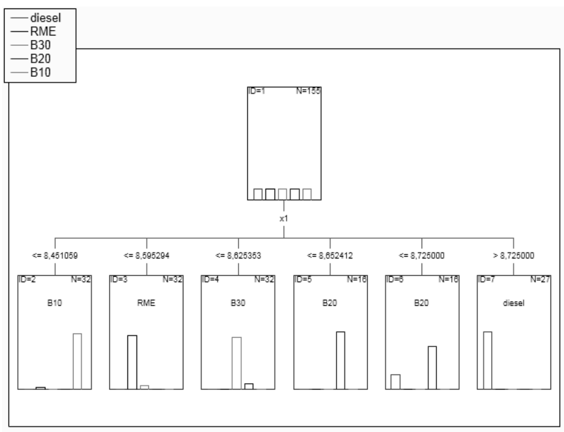


Fig. 8. Decision tree obtained by CHAID, when 20 consecutive engine cycles were used
Rys. 8. Drzewo decyzyjne zbudowane przez CHAID, na podstawie danych dla 20 kolejnych cykli pracy silnika

The Statistica Data Mining software package offers also boosted decision trees (a committee of classifiers). Using the same dataset as for C&RT and CHAID, the algorithm determined that the optimal number of boosted trees was 68. The "committee" of these classifiers has 99.4% accuracy (only one record from dataset is misclassified). The process of selection the optimal number of trees is shown in Fig. 9.

Use of neural networks can improve accuracy, but a disadvantage is the loss of transparency of the classifier.

The Statistica Automated Neural Networks (SANN) automatically check all possible combinations of the network architecture, activation functions, and the learning algorithm.

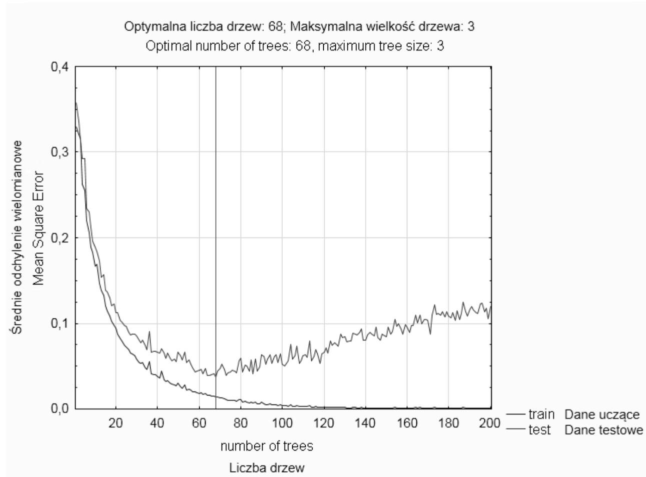


Fig. 9. Process of selection of the number of trees in the boosted trees algorithm
Rys. 9. Proces wyboru liczby drzew w algorytmie drzew wzmacnianych

The dataset with one input attribute, built from 20 consecutive cycles, allowed SANN to create an artificial neural network which misclassified only 2 cases; the accuracy was the same (99.5%) as that achieved by the See5 method.

The obtained neural network was a multilayer perceptron 1-7-5 (7 neurons in a hidden layer), with the BFGS (Broyden-Fletcher-Goldfarb-Shanno) learning algorithm, the SOS (sum-of-squares) error, and the logistic activation function.

The use of "committee" of networks should increase the accuracy. However, SANN with options "500 total networks" and "save 99 best networks" still had the same accuracy (99.5%).

3. Conclusions

Decision trees and artificial neural networks have similar accuracy in prediction of fuel type, but a decision tree with the so-called boosting is slightly better. The minimum number of consecutive engine cycles (at the rotational speed of 1200 rpm) required to predict the fuel type is about 20. This means that the fuel type can be recognized in 2 seconds, but in a small number of situations (about 0.5%) it will be difficult to distinguish between B20 and B30 biofuels. The solution to this problem is to extend the time of measurement to 3 second (30 cycles), but only if B20 or B30 were recognized after 20 engine working cycles.

Selection of attributes describing the engine work (like the maximum cylinder pressure or the maximum injection pressure) for dataset construction affects the accuracy of fuel recognition. Some other attributes (like the moment of maximum increase in cylinder pressure) should probably improve the accuracy, when placed in a dataset.

4. References

[1] Amsden A. A.: KIVA-3V: A Block-Structured KIVA Program for Engines with Vertical or Canted Valves. Los Alamos National Laboratory, LA-13313-MS, 1997.
[2] Kalogirou S. A.: Artificial intelligence for the modeling and control of combustion processes: a review. Progress in Energy and Combustion Science, Vol. 29(6), 2003, pp. 515-566.
[3] Kekez M., Radziszewski L.: Genetic-fuzzy model of diesel engine working cycle. Bulletin of the Polish Academy of Sciences, Technical Sciences Vol. 58(4), 2010, pp. 665-671.
[4] Heywood J. B.: Internal combustion engine fundamentals. 1988, McGraw-Hill, Inc.
[5] See5, http://www.rulequest.com/see5-info.html

otrzymano / received: 27.02.2013
przyjęto do druku / accepted: 01.04.2013 *artykuł recenzowany / revised paper*