A deterministic model for wear of piston ring and liner and a machine learning-based model for engine oil emissions

by

Chongjie Gu

S.M., Mechanical Engineering

Massachusetts Institute of Technology, 2017

B.Sc., Mechanical Engineering
Purdue University, 2015
Shanghai Jiao Tong University, 2015

Submitted to the Department of Mechanical Engineering in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy in Mechanical Engineering

at the
Massachusetts Institute of Technology

September 2021

Signature of Author:

_____

Department of Mechanical Engineering
August, 2021

Certified by:

_____

Dr. Tian Tian
Principal Research Engineer, Department of Mechanical Engineering
Thesis Supervisor

Accepted by:

_____

Nicolas Hadjiconstantinou
Professor, Department of Mechanical Engineering
Chairman, Committee for Graduate Students

A deterministic model for wear of piston ring and liner and a machine learning-based model for engine oil emissions

by
Chongjie Gu
Submitted to the Department of Mechanical Engineering on August 6, 2021
in Partial Fulfillment of the Requirements of the Degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

Nowadays, more constraints are required for design of internal combustion engines, to meet the energy saving and the emissions standards in the new era. Engine emissions and engine durability are two of the most important factors in the development of IC engines.

Engine particulate emissions are strongly correlated with the lubricant oil consumption. On the other hand, the carbon soot particles mixed in the lubricant from the combustion are the major source for long term wear of the piston, piston ring, and cylinder liner. Costly engine tests are required to develop the new system to meet emission and durability requirements. More advanced data analytics and models connecting critical design and operating parameters to performance will help shorten the development lead time for more efficient and cleaner engines.

This thesis work aims to model the engine wear during break-in and steady-state stages, capture oil emission correlations with engine operating parameters, and provide engine design guidance. This work is the first time to build deterministic physics-based wear models to perform systematic level engine wear simulations, including the effect of the liner topography. The wear simulation results are compared to experimental outcomes for both engine stages. It is also the first try to model the oil emission based on machine learning and connect the data-driven results with different engine ring-pack designs. The results suggest a good consistency of the machine learning analyzation and the underlying oil emission physics. The entire defined data-driven procedures show a promising future to accelerate engine development cycle, reduce engine testing cost, and help understand oil transport mechanisms and design influences.

Thesis Supervisor: Dr. Tian Tian

Title: Principle Research Engineer, Department of Mechanical Engineering, MIT

Acknowledgements

In the past years at MIT, I received support and contributions to my thesis work from many people around me. Without their help, it would be much more difficult for me to overcome the barriers in my study life at MIT. This work has been sponsored by Daimler, Argonne National Laboratory, and the consortium on lubrication in internal combustion engines. The consortium members were Mahle, MTU, Peugeot, Renault, Shell, Toyota, Volkswagen, Volvo Trucks, and Weichai Power.

First of all, I would like to thank my advisor, Dr Tian Tian, for his consistent support and guidance throughout my research at MIT. He always encourages me to broaden my perspective and try different methods to solve problems. I highly appreciate his knowledge, rich experience in academic areas and the logical thinking habit, from which I can benefit for my whole life. Besides his guidance in the academic area, he is also a valuable friend in my life. I have learned a lot from his experience and insights of life.

I would also like to thank the other members in my thesis committee, Professor Anette Hosoi, Professor Jung-Hoon Chun and Dr. Enrique Lizarraga-Garcia. They have given me a lot of advice from various perspectives to help me make my work more comprehensive.

I am gratitude to the members of the Sloan Automotive Laboratory for their kind help and friendship. They are: Yang Liu, Tianshi Fang, Sebastian Ahling, Zhen Meng, Qin Zhang, Wang Zhang, Jerome Sacherer, Zhe Chen, Xinlin Zhong and Koji Kikuhara. I would also like to thank former students in this group, especially Renze Wang, Haijie Chen, Steve Przesmitzki and Dallwoo Kim. They provide me both theoretical models and experimental results. Special thanks to Janet Maslow for the daily assistance in the lab.

Finally, I would like to thank my friends and family members for their love and support throughout my life.

Chongjie Gu, Cambridge, MA, August 2021

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Background and motivations

Nowadays, more constraints are required for design of internal combustion engines, to meet the energy saving and the carbon emission standards in the new era. Engine emission and engine durability are two of the most important factors in the development of IC engines.

Engine emission is strongly correlated with engine efficiency and lubricant oil usage. Around 10% of the total fuel energy is dissipated to heat due to mechanical friction, among which 20% is caused by the contact between the cylinder liner and the piston rings [1]. The behavior of the friction between the piston rings and the liner, in general can be divided to two stages: the break-in period and the steady state. As illustrated in Figure 1.1, the friction level between piston rings and the liner reduces during the engine break-in period, followed by a stable friction magnitude for more or less the rest of the engine life. This can also be reflected in the engine Stribeck curve change during the break-in. As illustrated in Figure 1.2, the Stribeck gradually shifts towards the horizontal axis during the break-in, indicating a reduction of the friction level in the mixed lubrication regime. The friction change in the break-in period is mainly caused by the liner surface topography evolution, which is a direct result of engine break-in wear [2].
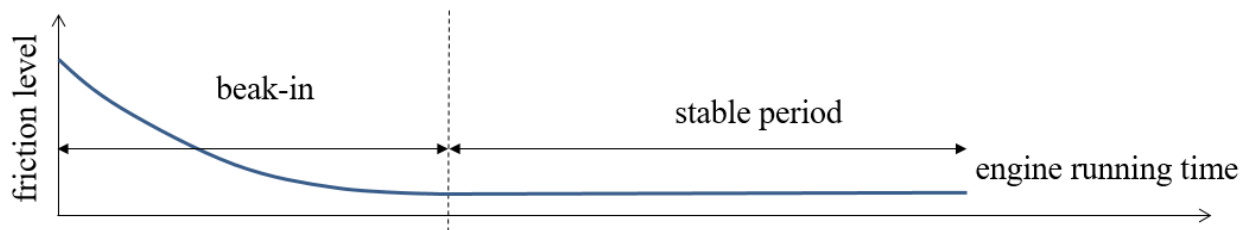
Figure 1.1. the friction change during engine running
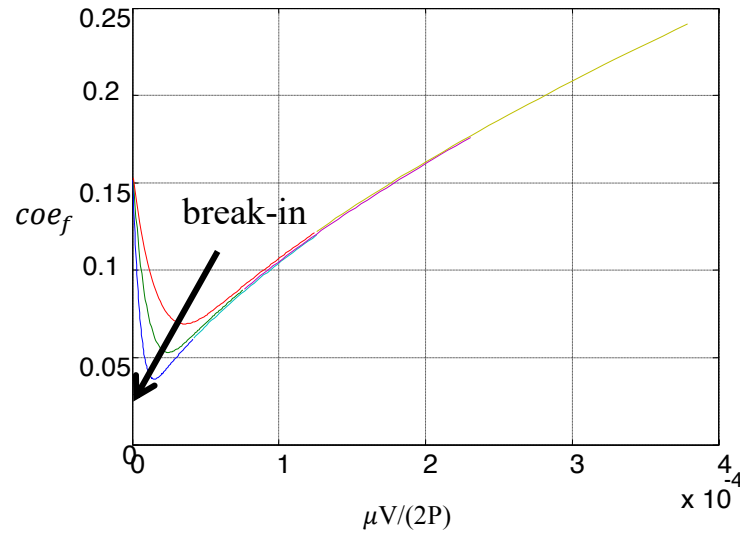
Figure 1.2. the engine Stribeck curve change during the break-in. coe_f is the friction coefficient. μ is the lubricant viscosity. V is the sliding speed. P is the normal load [2]

In addition to the engine efficiency, the transport and emission of lubricant oil is critical to engine emission, especially carbon particle emission. Experiments show that the particle emission has a positive correlation with oil emission. As plotted in Figure 1.3, in real engine tests, it is observed that the high peaks of the oil emission time series have significant overlapping with the particle number sequence. This indicates an important particle emission source: oil emission. When lubricant entering the combustion chamber through the piston ring-pack system, carbon particles will be generated and become part of exhaust gases. Therefore, from the aspects of engine efficiency and carbon particle generation, it is important to understand the process of friction evolution during the engine break-in and the oil emission, leaking through the ring-pack system.
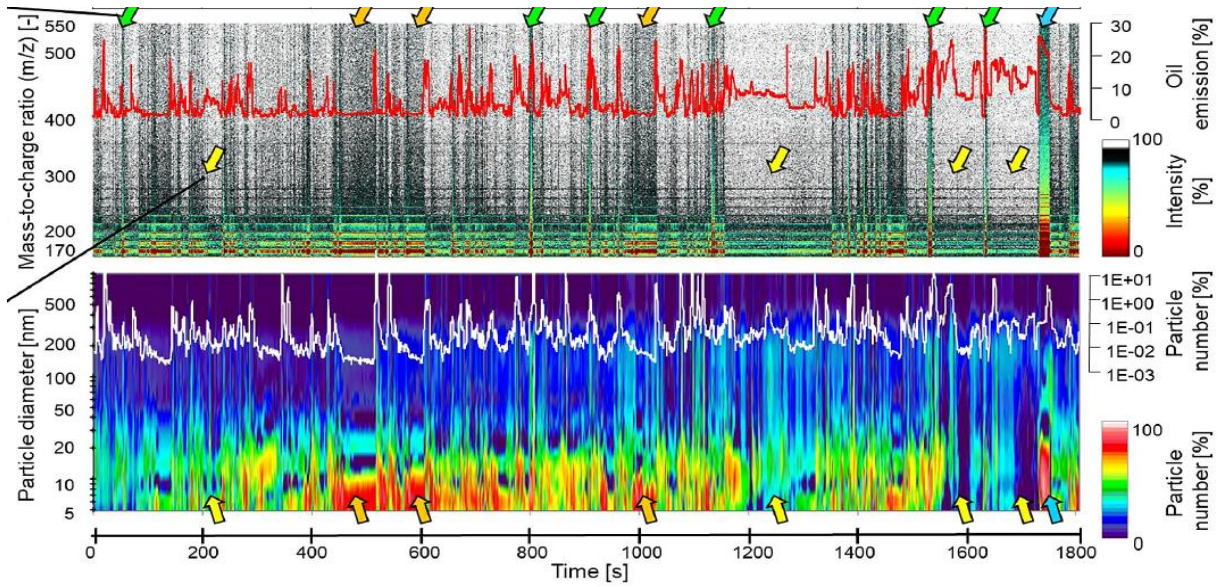
Figure 1.3. oil emission and particle emission [3]
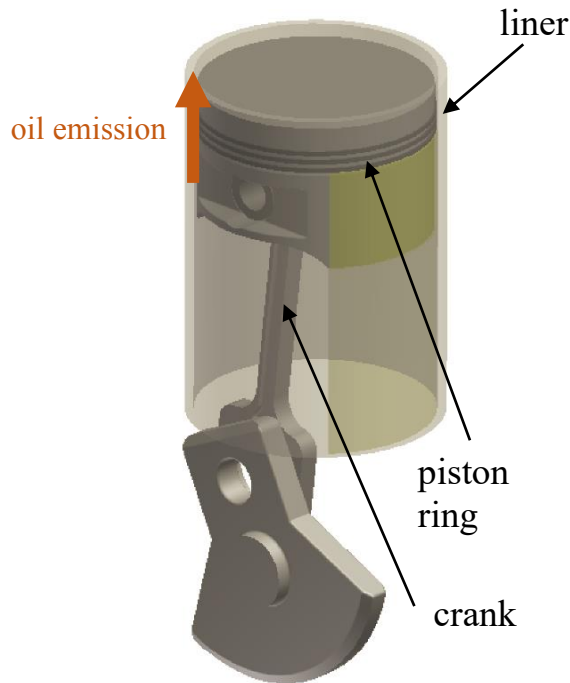
Figure 1.4. oil emission through the ring-pack system

The second design constraint is the engine durability. The durability of internal combustion engines depends on the life times of different engine components, among which piston rings and the cylinder liner are critical to engine failure. These two components are in contact and under

cyclic loadings, making them easy to experience severe wear. For example, in the following figure, a liner has experienced severe wear processing, with numerous abrasive wear scars in the piston sliding direction. Such a liner, after severe wear, loses its tightness with the piston ring, resulting an engine failure. On the other hand, the ring surface is even under a more vulnerable condition because it always contacts with the liner throughout the engine running process, although it has a higher harness than the liner. As a result, providing understandings and models of the steady-state engine wear process is essential to developing sustainable engines.

Figure 1.5. a cylinder liner with tribological failure

While sufficient lubricant supply helps reducing wear and friction, inadequate control of lubricant transport results in unacceptable oil emissions. This thesis work focuses on these two competing processes to provide more advanced understanding and models to develop more efficient and cleaner engines.

1.2 The general understanding of the physics for engine wear and engine oil emission

Wear of the piston rings and the cylinder liner can be divided into two stages based on the friction behavior change, break-in and steady-state. The dominant wear mechanisms in these two stages of engine life are different, and are discussed and modeled separately for the two stages.

Wear is a complex process, with many different mechanisms coupled together. In this work, only mechanical wear mechanisms are considered, because these mechanisms contribute the most for

surface topography changes. These mechanisms, including asperity flattening, fatigue wear, third-body and two body abrasive wear, cause surface topography change in the scale of micrometers or sub-micrometers. At this scale of wear, the surface contact and lubrication conditions will be influenced. Other mechanisms, such as molecular level physical or chemical wear processes, are not modeled in this work.
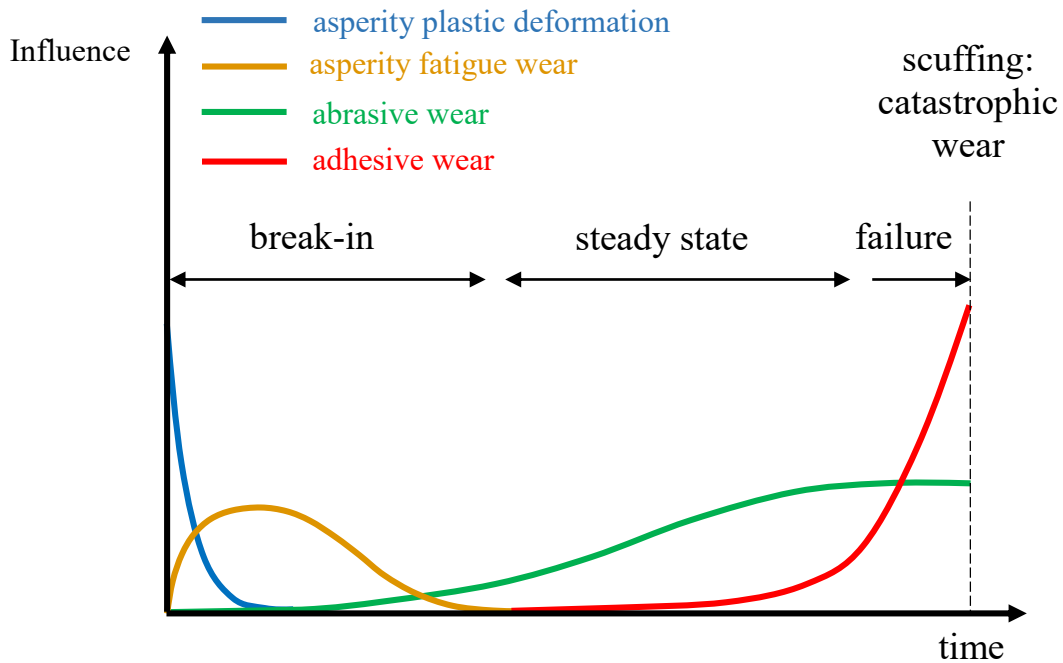
Figure 1.6. the dominant mechanical wear mechanisms at different stages of engine running

Based on the current understanding of engine wear, different mechanical mechanisms dominate in different stages of engine running because of the topography changes of the ring surface and the liner surface. After an engine starts running, surface asperities immediately experience a plastic flattening process. This process happens very fast, with only several hundreds of engine cycles, the original asperities on the liner surface are flattened [4]. After that, as asperities on the liner surface are under cyclic loadings, asperity fatigue wear starts to become the dominant wear mechanism. Surface cracks and subsurface cracks initiate and propagate under the external loadings, eventually leading to fractures of the liner asperities. During this process, the ring surface has little wear because the hardness of the ring coating is remarkably higher than the liner. As a result, the ring surface approximately remains smooth during the engine break-in period. The asperities with high peaks on the liner surface are gradually removed and the average

size of the contacting asperities are getting larger. Together with more and more hydrodynamic support, the fatigue wear rate diminishes at the end of break-in stage for at least majority of the liner along the stroke.

After the break-in period, as the third-body abrasive wear becomes significant to cause wear on both the ring surface and the liner surface around the TDC and BDC areas where boundary lubrication between rings and liner still exist. Because the ring surface is gradually roughened with third-body abrasion, two-body abrasive wear also initiates and damage the liner surface. Combining two-body and third-body effects, abrasive wear is the dominant wear mechanism in the engine steady state, as many abrasive wear scars are observed during experiments of this stage [5,6,7]. Figure 1.7 and Figure 1.8 illustrate the different dominant wear mechanisms during the break-in and the steady-state engine periods.
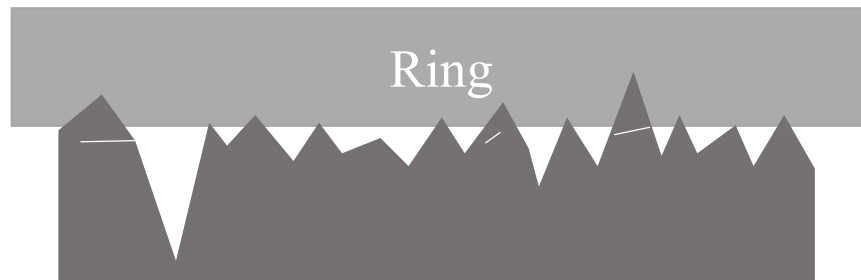
(a)
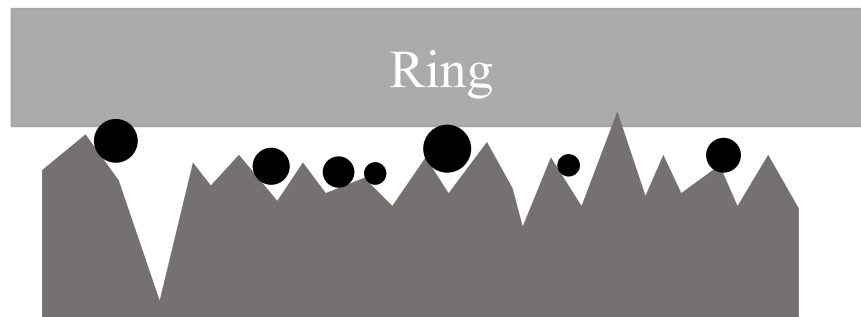
(b)

Figure 1.7. the different dominant wear mechanisms at different engine running stages (a) asperity fatigue wear (b) abrasive wear

<div align="center">(a)   (b)</div>
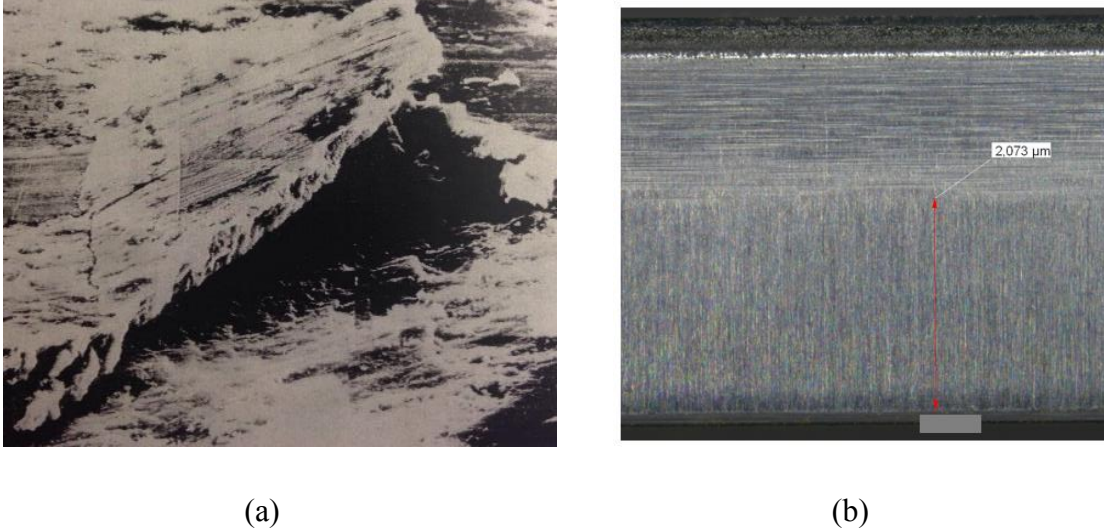
Figure 1.8. the observed different mechanisms of wear (a) fatigue wear [8] (b) abrasive wear [7]

In this work, the engine break-in wear and steady-state wear are modeled separately with asperity fatigue wear and abrasive wear, respectively. In reality, the dominant wear mechanism transition does not have a clear turning point for break-in and steady-state running. However, for the model simplicity, the two different wear mechanisms are modeled and simulated in a separate way.

Engine oil consumption is known as one of the most difficult problems in engine development. Most of time, many design iterations are needed to find optimal solutions. Among other, large number of involved design and operating parameters, lack of sufficient understandings and predictions of oil transport, complexity of transient oil consumption behaviors, and difficulties in measuring real time oil consumption are the burdens for solving oil consumption problems. One particular area that has brought much attention especially recently, is the oil emission and its impact on particle number (PN) emission during real driving emission cycles. The engine OEMs are obliged to ensure the engines to meet emission standard in real driving conditions. As the instantaneous oil consumption under varying speed and load is not uniquely defined by steady-state oil consumption at the present speed and load, a simple steady-state oil consumption measurement is not able give sufficient information to predict arbitrary driving cycle. As such, on the physics side, understanding the oil accumulation history and its relationship with the oil consumption rate, and on the experimental side, ability to design a test cycle whose results can

be used to predict any driving cycles, bear practical interest to understanding and predicting the transient oil consumption as well as minimizing the engine test time.

1.3 Existing work and the thesis objectives

To meet the current challenges in developing internal combustion engines, this work aims to generate new knowledge and models on the wear of piston rings and liner as well as on the oil emissions. As mentioned in the previous section, engine oil emission is strongly correlated with engine particle emission. To design low-emission engines, the oil emission characteristics need to be connected to engine design parameters as well as engine running cycles. However, because of the complexity of oil transport mechanisms contributing to the oil emission and the involvement of vast amount of relevant design and operating parameters, predicting oil emission reliably is still out of reach. Currently, evaluating the oil emission is still largely relying on engine tests, which is expensive in both time and cost [9,10,11]. Moreover, the oil emission residence time and extra oil consumption during the transient operation have been one of the major concerns for the engine lubricating-oil transport and consumption with previous research works dedicated to this topic [12,13]. However, these works focus on step transient with constant speed and load before and after the step transient. In contrast, the real engines experience various speed and load changes.

to the first objective of this work is to employ Machine Learning (ML) to extract important parameters and relations dictating the transient oil emission from the oil emission measurements under Real Driving Emission (RDE) cycles and to predict oil emission of other driving cycles based on the training results. The central questions to be addressed include

- If the current oil emission does not solely depend on the engine speed and load of the present time, how far back of the history of the operating conditions, hereafter called residence time, plays a determining effects on the present oil emission?
- Purely based on oil emission measurement, what other parameters are also important to determining the present OE?

- What are the implications to the oil transport mechanisms from the parameters and relations gained from the ML?
- What is the shortest running cycle whose OE measurement can be used to predict the OE of arbitrary driving cycles without further tests.

Previous wear models, particularly fatigue wear and abrasive wear models, can be roughly divided into physics-based models and mathematical models. Physics-based models usually can be used to simulate the material loss or surface height change at the contacting spots, without including the effect of surface roughness [14,15,16,17]. Previous physics-based fatigue wear and abrasive wear models are based on a single contacting point. Statistical models, on the other hand, focus more on the change of important surface roughness parameters during the wear process [18,19,20,21,22]. Some statistical models can also generate a worn surface topography with some artificially generated scratching lines, but the physics processes are missing [2,23]. This work is the first time to build deterministic physics-based wear models to perform systematic level engine wear simulations, including the effect of the liner topography. With the built model, it is able to predict the wear coefficients in engine break-in and steady-state stages and estimate engine life time under given conditions. Design guidance can be provided to build more sustainable and more efficient engines.

The objectives of the second part of the thesis work are to develop deterministic models for the break-in of the liner roughness and the long-term wear of the liner roughness and rings. The work focuses on the mechanical processes only. A fatigue delamination wear model is employed for the break-in period while an abrasion model is employed for the long-term wear of the liner roughness and the rings. Third body particles are introduced for the abrasion model in order to bring wear on the coated ring whose material in general is harder than the liner. Finally, the modeling results will be compared with the measurements on friction and ring wear.

1.4 Scope of the thesis

Machine learning based oil emission modeling is discussed in Chapter 2. The data processing, model selection and fitting procedures are defined first. Important system properties are also

studied and estimated, such as the oil emission residence time, engine acceleration effect and quasi-static sequences. Point prediction, interval prediction and total oil emission prediction are also discussed with different application purposes.

The break-in wear model is physics-based and comprised of different sub-models. In Chapter 3, the mechanisms during the engine break-in are modeled and integrated. Wear simulation results are also compared to the existing empirical wear law. Some friction experimental measurements during the break-in are also presented and compared in Chapter 3.

Abrasive wear modeling of engine steady state is discussed in Chapter 4, starting from the basis of all the sub-models of third-body transport and abrasion. The two important influencing factors are investigated: the particle size distribution and the ring coating material. These are critical design components for engine sustainability.

Chapter 5 summaries the thesis work and provides the potential future work for improvements.

Chapter 2. Learning Oil Transport Mechanisms from Oil Emission Measurement via Machine Learning

2.1 General introduction

As discussed in Chapter 1, engine particle emission is positively correlated with engine oil emission. Designing low oil emission engines is the key to reduce emissions in the automotive industry. However, development of new engines has two nontrivial challenges. First, the internal mechanisms of engine emission are complex, with many different physical and chemical processes coupled together. As a result, it is difficult to build models of engine oil emission and obtain insights of some important system properties. Second, the oil emission testing process of new engines takes long time, at least one to two weeks, leading to a slow development cycle and high design and testing cost. In this chapter, we propose a set of data-driven procedures based on experimental oil emission measurement to improve the design and testing process of engine development.

Due to the complexity of oil transport and emission process, data-driven methods is more efficient and feasible to conduct modeling work in a fast paced engine development environment. The entire set of data-driven procedures is summarized in Figure 2.1, as a form of a flowchart. As the input of the data-driven procedure, oil emission measurements provide simultaneously-recorded engine speed, load and oil emission sequences. After data processing and feature engineering, the most suitable model is selected to capture the correlations between oil emission and engine running conditions. With a chosen machine learning model, some important system properties can be calculated and evaluated, such as the residence time and input data uncertainty. The uncertainty level of the input data can be used to provide feedbacks on engine experimental setup, such as data collection methods, sensor setup or the amount of data required for oil emission analysis. Another important system property is the residence time, which will be applied in data augmentation and total oil emission prediction. Uncertainty and the residence time are also useful in designing engine testing cycles to minimize the testing duration and cost.
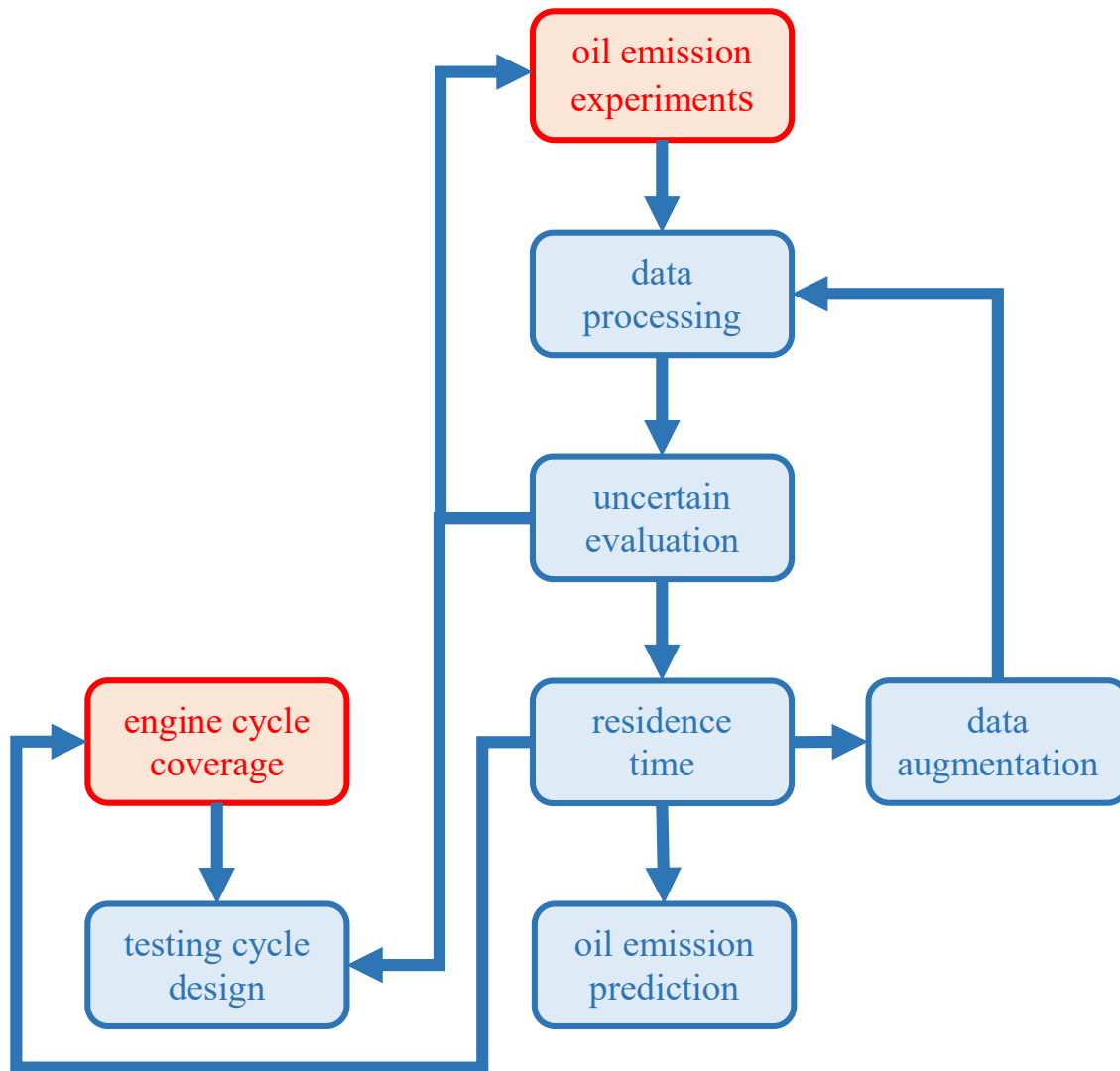
Figure 2.1 the overall scope of the data-driven procedures defined in this project for new engine development

The structure of this chapter is also based on the oil emission dataset analysis procedures. First, the basic data structure and the piston ring-pack designs used in this project are introduced. Data processing and feature engineering are discussed next with a chosen machine learning model for the oil emission regression task. With the chosen model, important system properties are investigated, followed by predictions and design of engine testing cycle.

2.2 Oil emission datasets and engine testing conditions

Among all engine design elements, ring-pack design is the critical part to control oil emission. Figure 2.2 contains two pictures of a real running engine piston ring pack using two-dimensional Laser Induced Fluorescence [24]. The horizontal dark bars are the piston rings while the oil exists in the bright areas, with more oil accumulating in brighter locations. there accumulates large amount of oil in all regions of the ring pack during the closed-throttle condition, particularly below the second ringAfter the throttle is opened, it [24] appeared that most of the oil in the third land and oil control ring was quickly released downwards while some of the oil leaked upwards passing the top ring to the crown land. This leaked oil will never return to the engine lubrication system and eventualy become engine emission. Therefore, engine oil emission is remarkably influenced by the ringpack design and the accumulation process of the lubricant.

(a)   (b)

RPM 3502   Intake 54
Blowby 1.7   Laser CA 75
CCase 1010

(c)

RPM 3498   Intake 56
Blowby 0.5   Laser CA 75
CCase 1009

(d)

RPM 3499   Intake 945
Blowby 3.2   Laser CA 75
CCase 1031
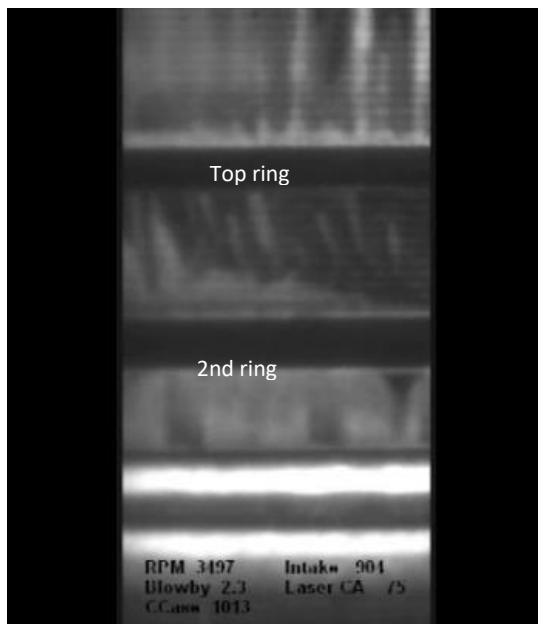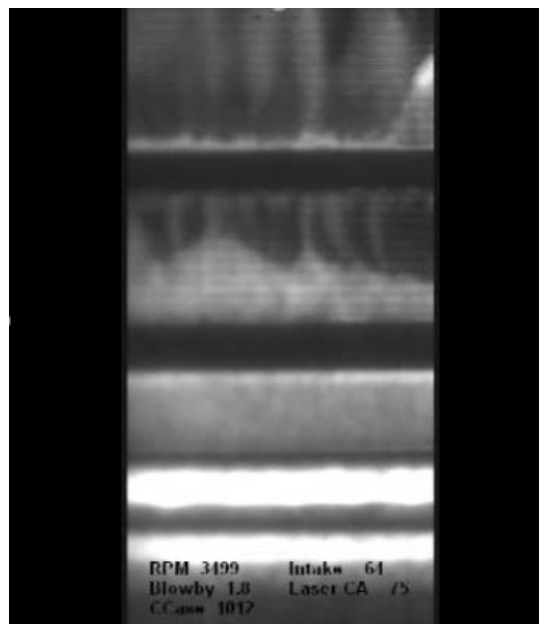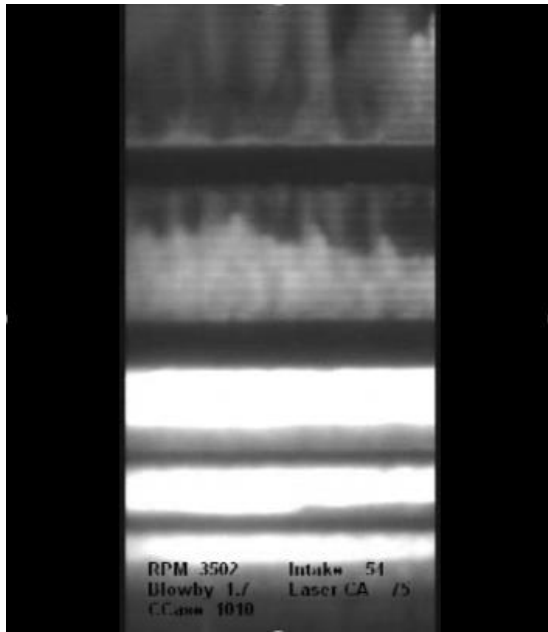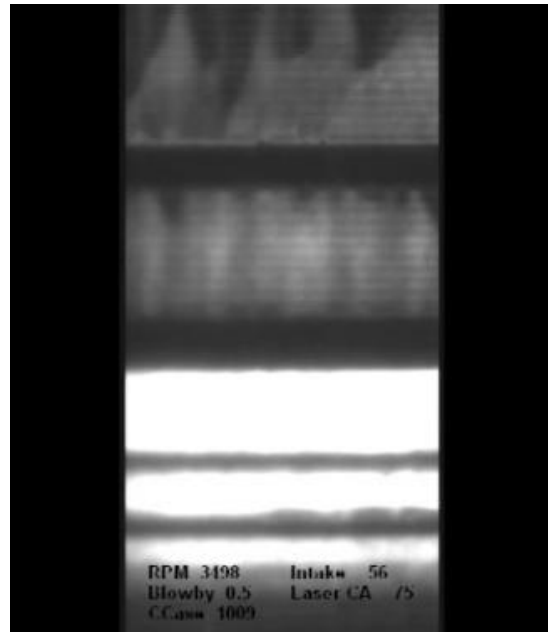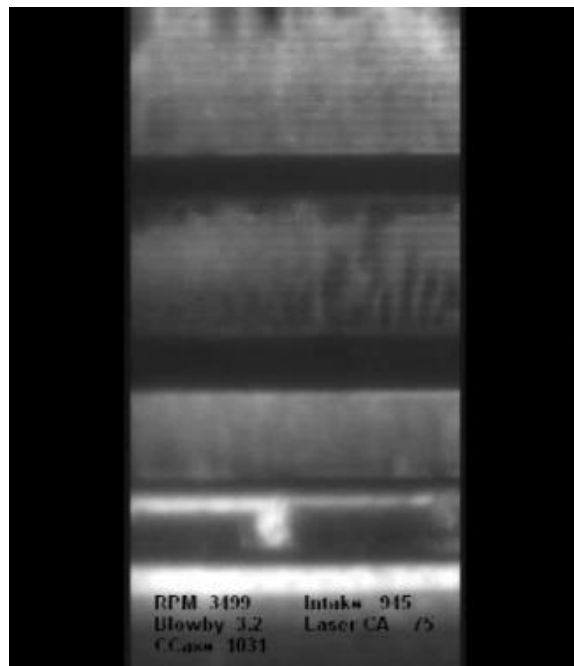
(e)

Figure 2.2 oil emission mechanisms near the ringpack system. (a)-(d) the oil accumulation process during closing throttle (e) opening throttle, oil released from the ring-pack system to form oil emission [24]

In this work, the oil emission measurements by Mahle [11] of a turbo-charged Si engine with three different oil control rings are used to illustrate how the defined data-driven procedures are used to perform system analysis, predictions and engine testing designs. The first two oil control rings have the same design, namely, three-piece oil control ring, but different tensions. The first has a standard tension and the second has 1/3 of the standard tension. The thrid oil control ring is Uflex. A rail gap is designed to make it easier for assembly, but it also allows oil leakage through the oil control ring. The only differernce between design 1 and design 2 is the tension within the expander. Design 1 contains high tension in the expander compared to design 2 so that the bore and the ring fitting is tighter. Design 3 does not uses three-piece oil control ring. Instead, there is only one piece in the oil control ring for design 3. This makes it more flexible to fit into the bore with changing shapes, but the notches on the ring provide additional paths for oil leakage.
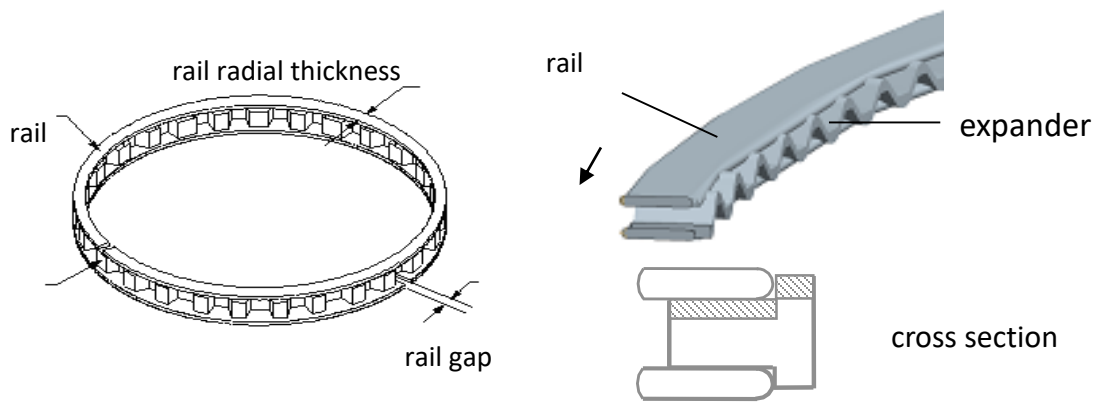
Figure 2.3 the oil control of ring-pack design 1 and design 2
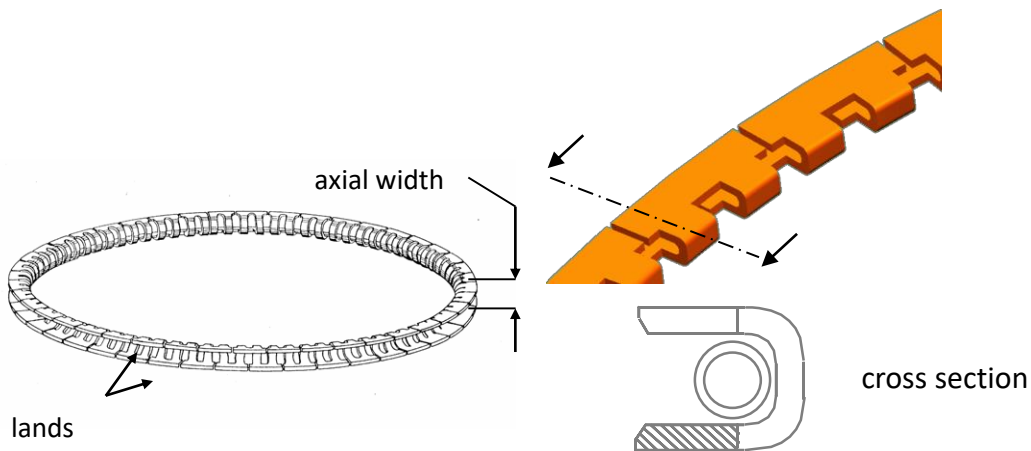
Figure 2.4 the oil control of ring-pack design 3

The collected data from oil emission experiments of different ringpack designs are presented as time sequences. The sensors record the engine status during the testing process, including engine speed, engine load and oil emission. These variables are all recorded instantaneously with a fixed time interval, 0.55 second. The load and speed sequences are predefined from a Real Driving Emission (RDE) cycle while the oil emission series is the corresponding system response. Therefore, in order to eliminate the random effect of a single engine test, the experiment is conducted 6 times with the same predefined speed and load. The averaged result is used as the final experimental oil emission measurement datasets.
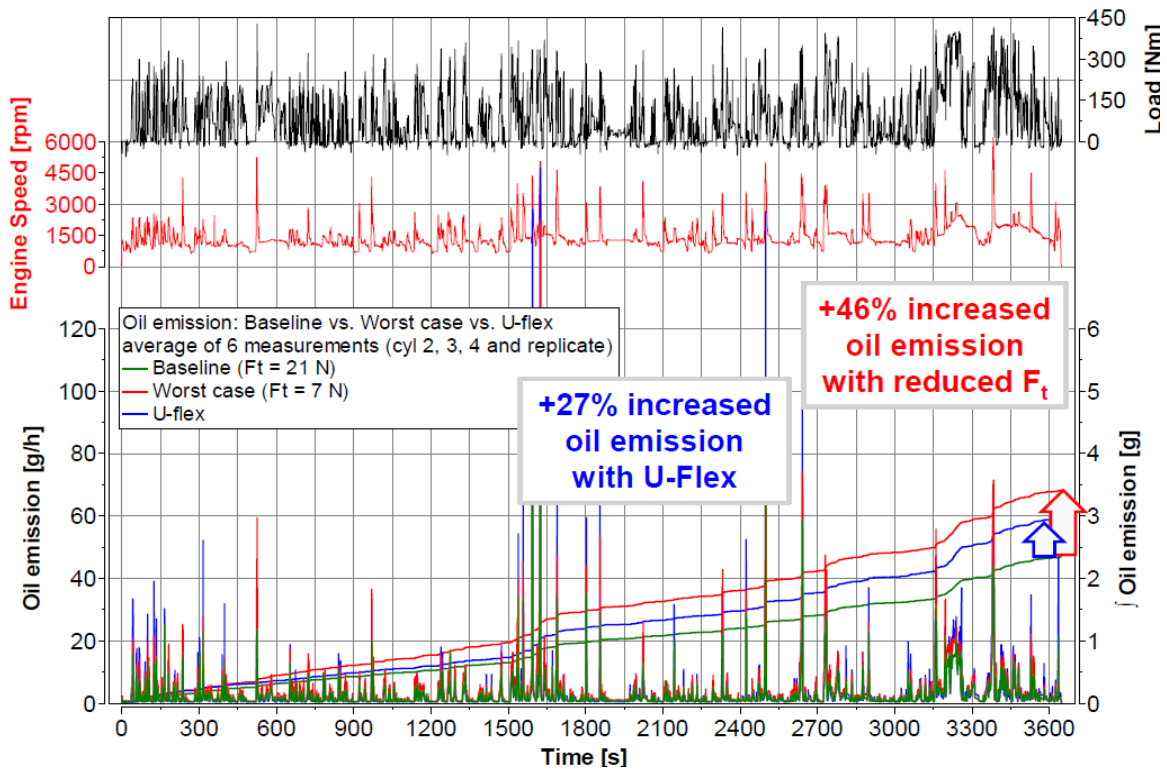
Figure 2.5 oil emission datasets for the three different ringpack designs.

baseline: ring-pack design 1; worst case: ring-pack design 2; u-flex: ring-pack design 3

Figure 2.5 shows the datasets of oil emissions measured in real engine experiments. The ring pack design names, 'baseline', 'worst case' and 'u-flex', are the original names provided used by

33

Mahle [11]. For simplicity, we use ring-pack design 1 to 3 to refer to the designs. Oil emission sequences are also plotted with different colors in Figure 2.5, for different ring-pack designs. Compared to design 2, ring-pack design 1 has the lower accumulated oil emission amount, because of its higher tension. Design 1 also achieves lower oil emission than design 3.

2.3 Data processing, feature engineering and model selection

The measurement of the oil emission test series was presented as a set of time sequences, including engine speed, engine load and real time oil emission. Several steps should be made with the purpose of capturing internal mathematical correlations within the oil emission datasets. Data processing, feature engineering, and choosing appropriate models are the essential three steps for data-driven modeling for engine oil emission. It is inefficient or sometimes even unfeasible to feed raw measurement data into machine learning models. Based on the physics properties of the problem and measurement data structures, necessary data separation and rescaling are needed for data processing. Feature engineering is another preparation step for algorithm fitting, which turns a physics problem into a mathematical one through defining independent variables and the corresponding dependent variable. After data processing and feature engineering, a set of models should be chosen to fit the data and catch the correlation. In this step, several candidate models are compared under a carefully selected evaluation metric before the best is selected.

2.3.1 Data processing

For a dynamic system like engine oil emission, a commonly used practical analysis is separating the response into steady-state and transient parts [9,13,25]. The total oil emission at any moment is the summation of steady-state oil emission and transient oil emission:

$$OE_{total} = OE_{ss} + OE_{tr}$$

OE_{ss}: steady-state oil emission

OE_{tr}: transient oil emission

More specifically, OE_{ss} is only determined by the engine speed and engine load at the corresponding moment, while OE_{tr} is influenced by previous load and speed history as well. Due to the existence of system memory, the number of previous influencing steps is also a limited number, T. Therefore, we have:

$$O_t^{ss} = f_{ss}(S_t, \ L_t)$$

$$O_t^{tr} = f_{tr}(S_t, \ S_{t-1}, S_{t-2}\ldots \ S_{t-T}, \ L_t, \ L_{t-1}, \ L_{t-2}\ldots \ L_{t-T}, \ O_{t-1}, \ O_{t-2}\ldots \ O_{t-T})$$

For the steady state component, 16 independent responses are recorded with each engine ring pack design. A simple third order polynomial fitting is adopted to fit the steady-state oil emission correlation:

$$O_t^{ss} = f_{ss}(S_t, \ L_t) = A_{30}L^3 + A_{21}L^2S + A_{12}LS^2 + A_{03}S^3 + A_{20}L^2 + A_{11}LS + A_{02}S^2 + A_{10}L + A_{01}S + A_{00}$$

Where L and S are the normalized speed and load at each moment. The above expression is a third-order polynomial fitting, but in application the final choice of polynomial order is decided by the fitting error on the testing set, not limited to third-order.

The reason to fit the correlation after normalizing the speed and load data of the 16 points is to eliminate the effect of different magnitude of the two features. Min-max scaling is used as the normalization method:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Furthermore, with the same purpose, the entire load and speed sequences are processed with min-max scaling where the minimum and maximum values are defined by the extremums in the steady-state dataset.

As an example, here we use the ring-pack design 1.

Engine speed [rpm]	Engine load [N]	Oil emission [g/h]
700	-0.1	0.23
1550	-23.1	1.96
1550	10	0.56
1550	122	1.39

1550	243.1	3.52
1550	374.9	7.82
3000	-27.6	2.93
3000	10	0.97
3000	131	2.93
3000	262	8.58
3000	390	18.91
4600	-35.4	6.48
4600	10	1.31
4600	128	6.08
4600	256.1	13.56
4600	390.9	39.3

Table 2.1 the steady-state engine status of ring-pack design 1

Through a third-order polynomial fitting, the correlation of baseline steady state can be mathematically represented by the following curve. The highest point corresponds to the most extreme steady state testing condition where the engine speed and load are 4600 rpm and 390.9 N, respectively.
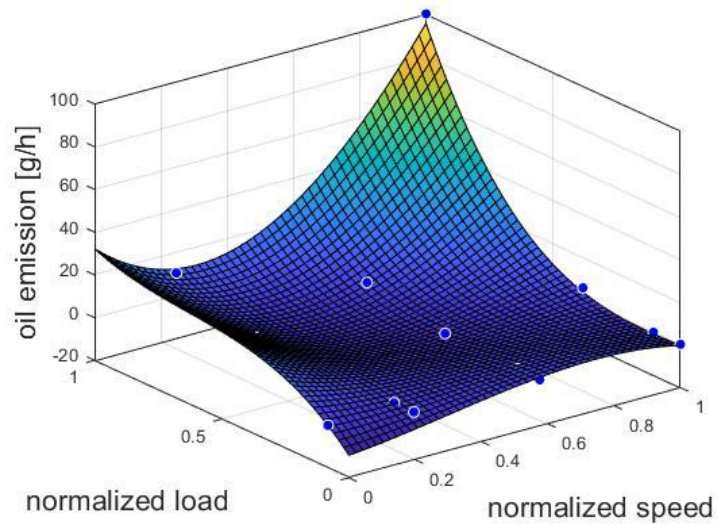
Figure 2.6 steady-state correlation of ring-pack design 1
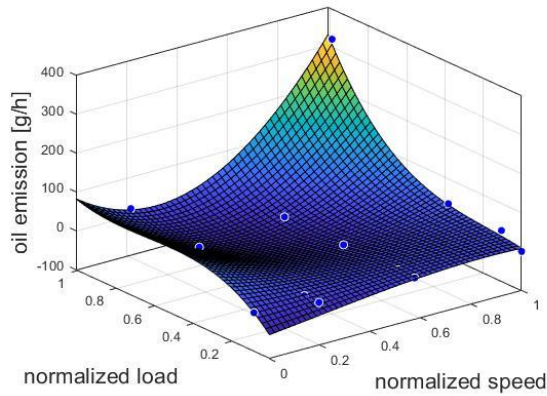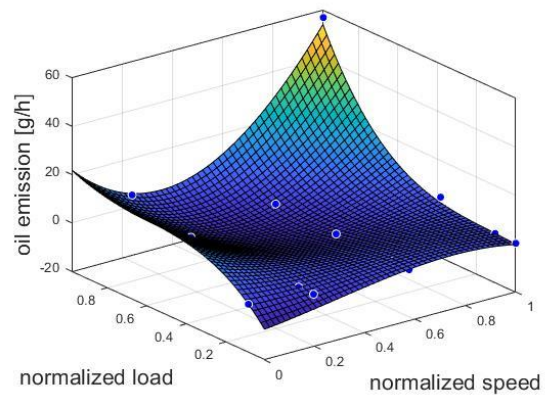
| (a) design 2 | (b) design 3 |

Figure 2.7 steady-state correlation of design 2 and design 3

2.3.2 Feature Engineering

The oil emission dataset is a set of time series. With the purpose of predicting in the oil emission series, supervised learning is applied to uncover the correlation of transient oil emission component. The essential part of feature engineering in supervised learning is defining the independent variables and dependent variables, among which we would like to find a mathematical mapping. In the task of predicting a single step of oil emission, the dependent variable is the oil emission at each moment while the independent variables are the speed, load and oil emission at previous steps and the speed and load values at the current moment.

Although the system memory duration is not determined yet (will be investigated in the next section), it is clear that the oil emission at each moment is only correlated with the status of the recent history. In other words, the system has a residence time. From the physics of the oil transport process, it is natural to draw such a conclusion. Oil emission is largely dependent on the oil accumulation in the ring pack (above the oil control ring). Moreover, oil accumulation is not an instantaneous process, which requires a certain number of engine cycles to achieve equilibrium. The subscript T in the equation above represents the number of previous steps used as input features. The optimized value of T is determined by the system residence time, which will be discussed next. In Section 2.5, a number of system properties will be investigated under

37

the scope of machine learning, including residence time, acceleration effect, quasi-static oil emission. With different regression purposes, the corresponding feature set is also different. But, the scope of using previous speed, load and oil emission steps remains the same.

Normalization is applied to all the input features in order to avoid the potential effect of magnitude inconsistent across speed, load and oil emission feature groups. Although some machine learning models are insensitive to feature normalization, it is still applied as a standard procedure. For example, using tree-based models, such as gradient boosting and random forest, does not require feature normalization as an indivisible procedure, but utilizing neural networks needs normalization to eliminate feature magnitude effect. In this project, two normalization methods are adopted: min-max scaling and standard normalization.

Min-max scaling: $X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$

Standard normalization: $X_{norm} = \frac{X - X_{mean}}{\sigma}$

X and X_{norm} are the original value and the normalized value, respectively. $X_{min}, X_{max}, X_{mean}$ are the minimum, maximum and average values of the feature set. Speed, load and oil emission are rescaled separately. The normalized values, as the outcome of normalization, cover different ranges based on different normalization approaches. Min-max scaling leads to the normalized value in the range of [0, 1], while standard normalization makes the result in both negative and positive regions. The final choice of normalization methods depends on the final performance on the testing dataset.

2.3.3 Model selection

For the oil emission dataset, several models can be used to fit the correlation and investigate underlying system properties. With the purpose of selecting the most efficient model, the original dataset is divided into a training set and a testing set while the final performance comparison relies on the testing set. As the dataset is a time sequence set, the split of training and testing sets is along the time axis: the first 80% of the total measurement sequence is used as the training set with the rest 20% as the testing set. The prerequisite assumption to make such a partition is that

the engine status keeps stable across the time domain of oil emission series so that the correlation remains the same in the training set and the testing set. This assumption is valid once the engine experiment goes beyond the beginning break-in period. The separation of the time series data is illustrated in the following figure.
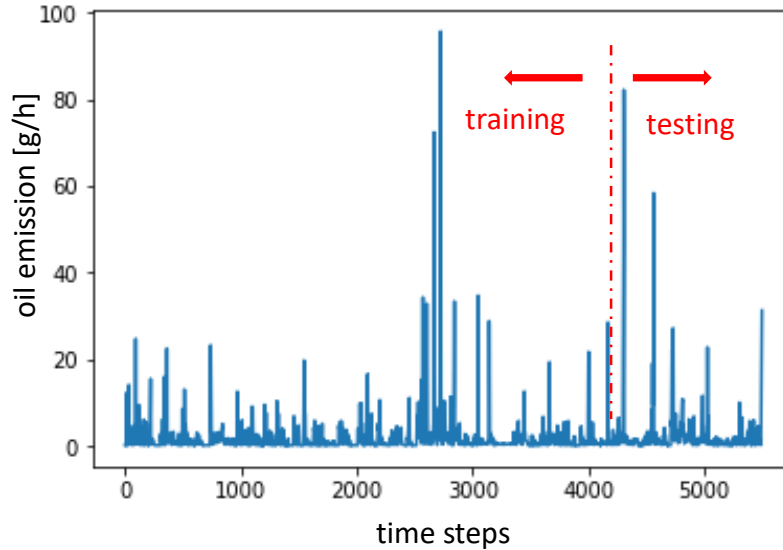
Figure 2.8 the split of the training and the testing datasets

For sequential datasets, many models can perform the role of correlation fitting. Among those models, several widely-used models are considered in this project and compared to each other to select the best model [25]. Auto regressive integrated moving average (ARIMA), recurrent neural network (RNN), gradient boosting and random forest are the most commonly utilized model for time series datasets [26,27,28].

The first model, ARIMA, is a traditional statistical time series model, very efficient in capturing linear serial correlations. The mathematical expression of ARIMA is:

$$Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

Where Y_t to Y_{t-p} are the time series values and ε_t to ε_{t-q} are the error terms with Gaussian distribution. The model tries to find the best fitting parameters for α_1 to α_p and θ_1 to θ_q. ARIMA is computationally efficient because it is a generalized linear model. Therefore, it is difficult to fit the transient oil emission data which contains a significant portion of nonlinear correlation.

The second model is a special type of neural network: recurrent neural network (RNN). RNN can find both linear and nonlinear serial correlations through regression of successive terms in a series. The following figure shows the working mechanisms of RNN. It uses a single neuron in an iterative way to perform the memory function so that the model a deal with time series data. Contrary to ARIMA, RNN can detect nonlinear correlations in time series as well. However, the RNN model is a kind of deep neural networks, requiring a large set of training data so that the model variance can be reduced to an accepted range. In this project, the oil emission sequence only contains thousands of measurement points, which is not sufficient for training an RNN. The testing accuracy also confirms the limitation of application of RNN on the oil emission problem. Little training set size will lead to a high variance level along with a risk of overfitting.
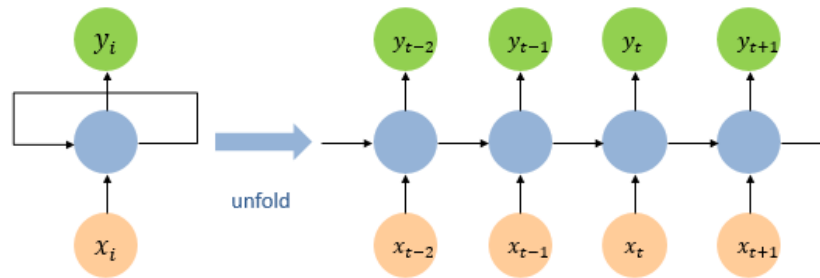
Figure 2.9 the structure of RNN

The third model is gradient boosting. Gradient boosting is a kind of resemble method, which performs classification or regression on the results of many basis learners. The final prediction of the entire model depends the way to integrate those basis learners: majority voting for classification or weighted average for regression. Because gradient boosting algorithm builds each basis learner based on the residual error of all previous leaners, it can significantly improve the fitting via bias reduction. Gradient boosting is pragmatically useful for tabular datasets. Meanwhile, the oil emission system has a residence time, which makes it possible to convert the time series data into a tabular set. Therefore, gradient boosting is suitable for this oil emission regression problem.

The fourth model is random forest. Similar to gradient boosting, random forest is a kind of resemble methods as well. The symbolic difference is that random forest builds each basis learner independently while gradient boosting builds each learner based on previous residuals.

As a result, the averaging regression result on all the basis learners in random forest can keep the entire model variance at a low level. Therefore, which model can better perform the task of oil emission regression depends on the fitting would benefit more from bias reduction or variance reduction. Both gradient boosting and random forest are good at fitting tabular data which can be converted from oil emission time series.

Based on the regression on the testing set, gradient boosting performs the best with the lowest fitting error. The metric used to evaluation model fitting is mean absolute percent error (MAPE):

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{x_i - \hat{x_i}}{x_i} \right|$$

Where x_i and $\hat{x_i}$ are the real value and the predicted value for each testing sample point. The main reason to use this metric is the existence of many high peaks of oil emission response. For instance, if the traditional metric squared loss or absolute loss is utilized, the overall loss function will be consisted mainly from fitting errors of those emission peaks, leading to the condition that oil emission measurements are unevenly evaluated. The following table contains the testing error of the four models.

Testing error (MAPE)	ARIMA	RNN	Gradient boosting	Random forest
Ringpack design 1	0.327	0.312	0.188	0.221
Ringpack design 2	0.257	0.199	0.197	0.285
Ringpack design 3	0.409	0.391	0.307	0.641

Table 2.2 the fitting results of different statistical time series models

From the regression results on the testing set, gradient boosting is the best candidate regression model achieving the lowest testing error. Therefore, gradient boosting is chosen as the model for further analysis of the oil emission problem.

2.4 Data fitting

In Section 1.2, the final chosen machine learning model for the oil emission regression task is gradient boosting. In this chapter, the details of model fitting will be presented, along with the best model hyperparameter combination for each ring-pack design.

2.4.1 Model fitting

Gradient boosting, as a resemble method, requires defining the type of basis learners as the first step. A commonly used basis learner is classification and regression tree (CART), a non-parametric decision tree learning technique being able to perform either classification or regression. CART could build variance tree structures in order to minimize an objective function. Compared to each basis learner, gradient boosting can reduce bias by fitting model residuals in every iteration. Therefore, each CART can grow relatively larger and deeper without too much concern of overfitting. More specifically, tree-based gradient boosting is the best model for the project.

To uniquely determine a tree-based gradient boosting model, several hyperparameters should be fixed. The major hyperparamters are listed in the table below.

notation	
min_samples_split	the minimum number of samples (or observations) which are required in a node to be considered for splitting
min_samples_leaf	the minimum samples (or observations) required in a terminal node
max_depth	the maximum depth of each tree
learning_rate	a weighting factor for the corrections by new trees when added to the model.

n_estimators	the number of basis learners
subsample	the fraction of observations to be selected for each tree

Table 2.3 the hyperparameters of tree-based gradient boosting

The best parameter combination is selected through 10-fold cross validation. Among many different combinations, the hyperparameter set with the minimum is the searching target. However, a complete search of all hyperparameter combinations is time consuming and computationally expensive. As such, we employed two techniques to reduce the computation cost: grid search and pairwise parameter tuning.

Grid search consists of two steps: coarse search and fine search. A set of hyperparameter combinations is tested, followed by a fine search of the hyperparameter space near the best combination. This two-stage search method is possible to miss the optimized combination, but the final choice, being able to beat most of the parameter combinations, can be used as an approximation of the optimizated combination. Pairwise parameter tuning makes use of the internal connections among the tuning hyperparameters to reduce the amount of calculation. In this model for instance, learning_rate and n_estimators is a pair of related hyperparameter. Lower values of learning_rate would require larger number of trees to model all the relations because the weight of each tree is lower in the final result. In the parameter tuning process, lower learning_rate corresponds to higher value of n_estimators, making the search space smaller compared to unconstrained search.

With the hyperparameter tuning procedure defined, the optimized hyperparameter combinations are obtained for all the three ring-pack designs. As mentioned above, although those hyperparamter combinations are not guaranteed to be the global optimality, they are close to the most optimized solution. The optimal hyperparameter combinations are presented in the following table. It is important to point out that the correlation fitted with the hyperparameters defined in the table is only the transient component. Along with the steady-state correlation, the entire oil emission sequence can be predicted.

	Design1	Design2	Design3
min_samples_split	1	3	2
min_samples_leaf	2	4	1
max_depth	4	6	5
learning_rate	0.2	0.1	0.15
n_estimators	280	400	300
subsample	1	0.95	1

Table 2.4 the optimized hyperparameters of tree-based gradient boosting

2.4.2 Comparison of different data processing methods

In the interest of comparing the performance of different data processing methods and showing the robustness of separating steady-state and transient component, we provide the exact prediction results of oil emission on the testing series. The predicted oil emission is illustrated both as sequences and testing accuracies.
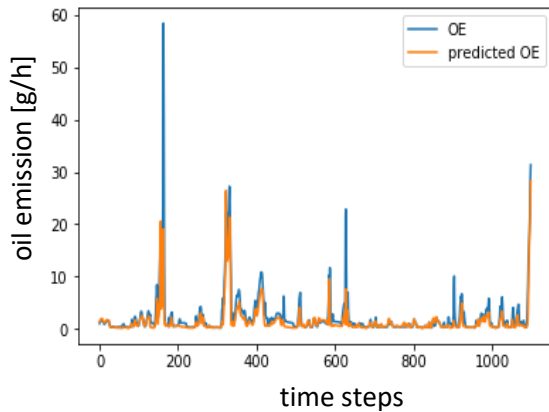
(a) Ring-pack design 1   (b) Ring-pack design 2
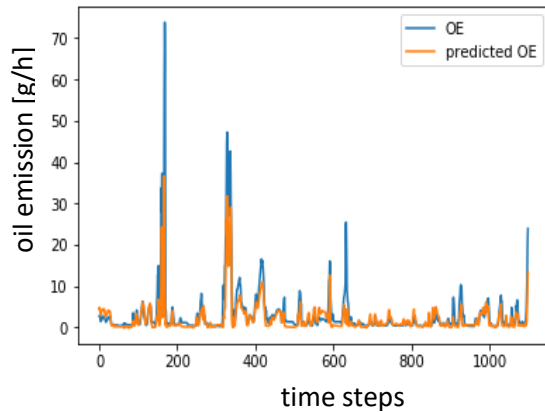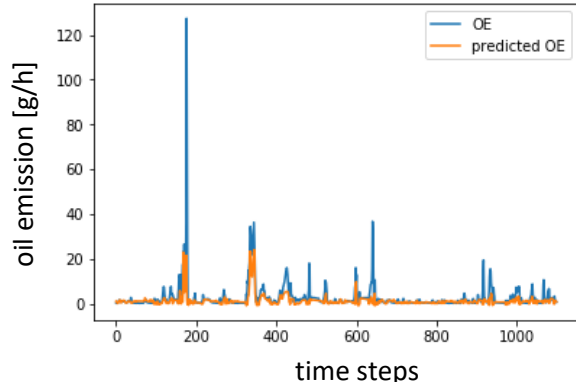
(c) Ring-pack design 3

Figure 2.10   oil emission prediction, steady-state only
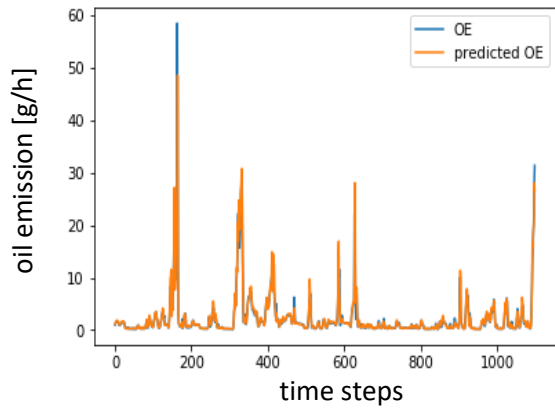
(a) Ring-pack design 1
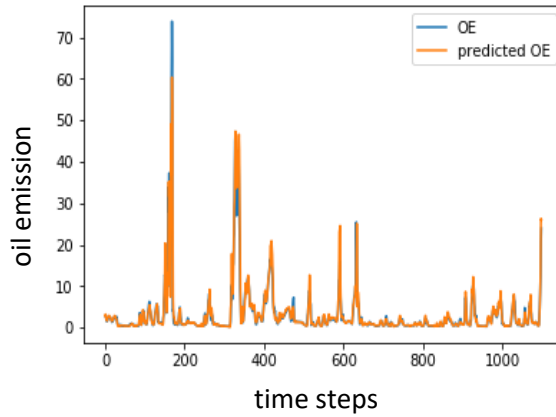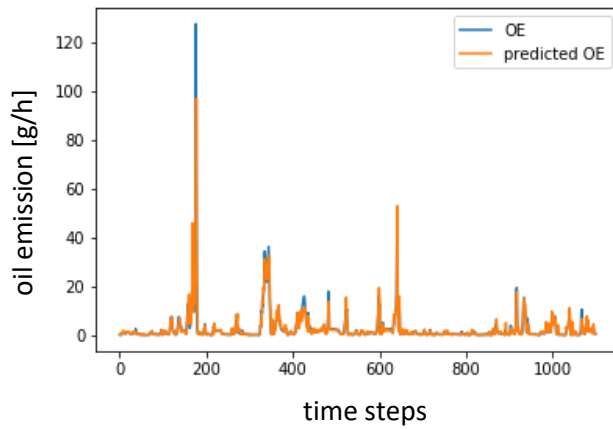
(b) Ring-pack design 2

(c) Ring-pack design 3

Figure 2.11   oil emission prediction, separating steady-state and transient

In Figure 2.10, the three plots show the predicted oil emission series on top of the real measured oil emission based on steady-state part only. The most difficult part for steady-state prediction is the condition that oil emission has an abrupt increase due to sudden change of load or speed. These peak positions contain high level of transient component. With the help of machine learning based transient correlation, we can dramatically improve the oil emission prediction, as illustrated in Figure 2.11. The improvement is mainly from oil emission peaks as transient part is properly added into consideration. Quantitatively, the prediction accuracies of different data processing methods are listed in the table.

Data processing	Testing error
No separation for steady state and transient, fitted with machine learning	Design 1: 0.432
	Design 2: 0.471
	Design 3: 0.797
Steady state only	Design 1: 0.358
	Design 2: 0.536
	Design 3: 0.871
Separating steady state and transient components	Design 1: 0.188
	Design 2: 0.197
	Design 3: 0.307

Table 2.5 testing errors of different data processing methods

2.5 System investigation

Up to now, the data processing and model fitting procedures are well defined to properly model the oil emission sequence. With the determined feature engineering and fitting model, essential system properties can be investigated and connected to oil transport physics and system design. The studied content via machine learning includes: residence time estimation, engine acceleration effect on oil emissions, effect of previous oil emission steps and quasi-static oil emission sequence generation.

The overall methodology used to investigate system properties is balancing the tradeoff between variance and bias. Statistical models cannot perfectly describe the true nonlinear correlations of many engineering processes, with the residual error being classified into two kinds: variance and bias. The error due to variance is taken as the variability of a model prediction for a given data point. This is usually caused by high stochasticity of the training set or overcomplicated model, leading to a capture of unnecessary correlations. On the contrary, bias is taken as the difference between the model prediction and the true value of the system which we are targeting to predict. Bias is often caused by oversimplified model or the missing of important influencing factors. The relation between bias and variance can be illustrated in a scatter plot.

In Figure 2.12, the origin point represents the real target value while the solid points are the predicted values by the model. In order to consider the fact that any prediction is only a characteristic value of an output distribution, the prediction here is represented by a set of different points. The ideal situation is the case with low bias and low variance, but the tradeoff between bias and variance always exists. In Figure 2.13, the magnitudes of bias and variance have opposite trends with changing model complexity. When the model becomes more complex, bias is decreasing because of the enhancement of model expressiveness. However, variance increases in this process because of overfitting. Correspondingly, the overall error can achieve a minimum value with a good balance between the pair of tradeoffs. Model complexity depends on many things, such as the model type, the number of model parameters used for fitting and the choice of features.
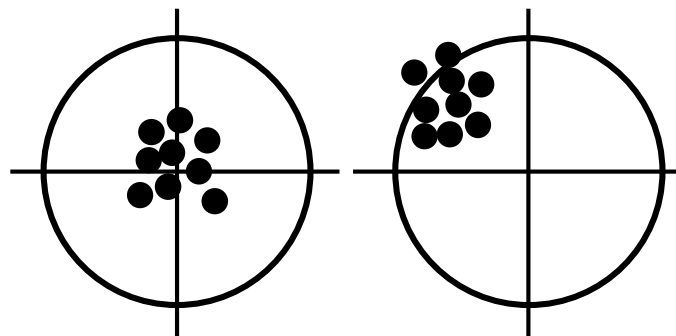
(a) low bias, low variance   (b) high bias, low variance
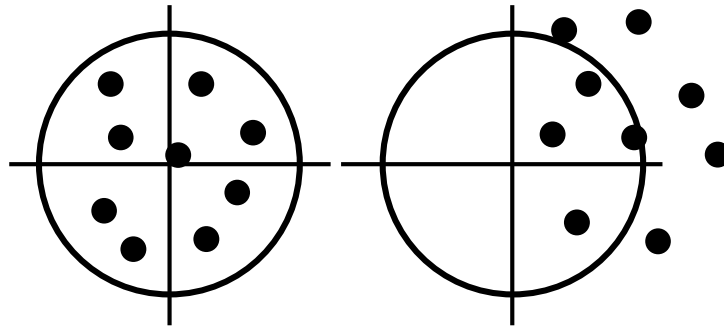
(c) low bias, high variance   (d) high bias, high variance

Figure 2.12 the effect of bias and variance on predictions
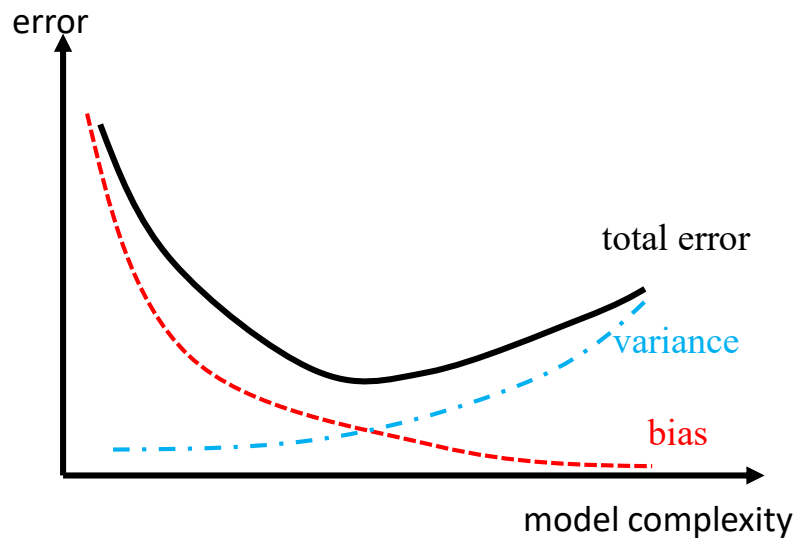
Figure 2.13 the bias-variance tradeoff

In this section, the only influencing factor of the model complexity is limited to the chosen features. This is conducted through a fixed model and optimized hyperparameter tuning. The model utilized in this section is tree-based gradient boosting. Although it is possible that switching to another time series model could achieve better fitting results in some conditions, we still use gradient boosting to eliminate the effect of different models. Hyperparameters are selected based on cross validation results to make sure that the regression error reduces to minimum value once the model is defined. Therefore, the only influencing factor left is the choice of features.

2.5.1 Residence time

The oil emission at each moment is a function of the oil emission, speed and load of some period before that moment:

$$O_t^{ss} = f_{ss}(S_t, L_t)$$

$$O_t^{tr} = f_{tr}(S_t, S_{t-1}, S_{t-2} \ldots S_{t-T}, L_t, L_{t-1}, L_{t-2} \ldots L_{t-T}, O_{t-1}, O_{t-2} \ldots O_{t-T})$$

Here, T represents the number of previous steps included in feature engineering to correlate oil emission at each moment. According to the theories of bias-variance tradeoff and model complexity, if the model type and feature processing procedure are fixed, the only thing that can influence the model complexity is the choice of features. With various values of T to fit the model, the optimal number of T can be determined to reach the best balance between bias and variance.

The prediction MAPE as a function of T is plotted in Figure 2.14. All the three ringpack designs show the same trend that the fitting error decreases, followed by an increment as T becomes larger. This agrees with the tradeoff between bias and variance. When T is small, bias contributes the most to the total error because the model is oversimplified without some important steps considered. When T becomes large enough, bias reduces to a low level but variance gets significant. The model is overcomplicated since irrelevant steps are included, leading to model overfitting. The optimized value of T, with the best variance-bias balance, indicates the most proper choice of features. This value is defined as the residence time of each ring-pack design. It is important to mention that this statistical estimation is an overall fitting average residence time, different from the real system residence time. From the physics point of view, the residence time of the entire oil emission system is not a constant, but a varying value based on the instantaneous system status. However, this optimal value of T can be used as a characteristic value of the system residence time.
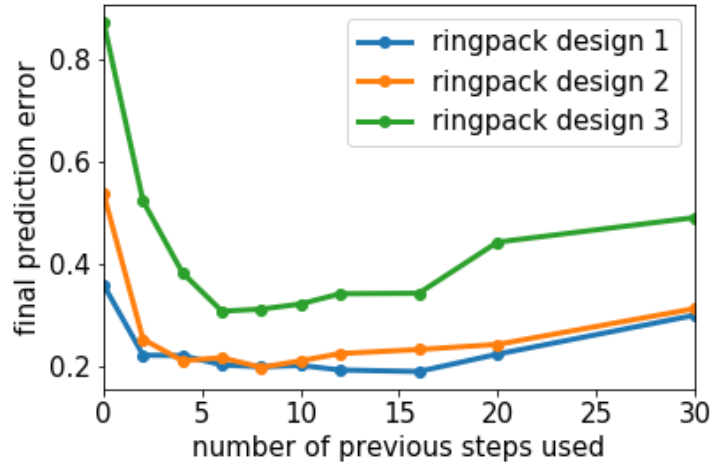
Figure 2.14 the prediction error with different number of previous steps

	Ringpack design 1	Ringpack design 2	Ringpack design 3
Residence time [# steps]	12-16	8-10	5-8
Residence time [seconds]	6.6-8.8	4.4-5.5	2.7-4.4

Table 2.6 estimated residence time

In Table 2.6, the residence time of each ring-pack design is listed. It is also converted to the real time duration in seconds, with 0.55 second for each time step interval. Considering the uncertainty of the model fitting, an interval near the optimization T is chosen as the residence time range. Ring-pack design 1 has the longest residence time, 6.6-8.8 seconds. This result from machine learning explains the ring-pack design logic. Ring-pack design 1 and design 2 have the same three-piece oil control ring setup, with different ring tensions. Comparing Designs 1 and 2, Design 1 provides lower oil supply rate to the ring pack region above the oil control ring due to its higher tension. Thus, it may be reasonable to expect that the present oil accumulation in the ring pack with Design 1 takes longer time to reach than Design 2 with higher oil supply rate. For Design 3, there are about 50 small gaps on the oil control ring and they may provide direct paths

between the region below the OCR and above. Thus the findings from the model that Design 3 has the shortest residence time also make sense.

The existence of residence time can be reflected by calculating the influence of each feature as well. In Figure 2.15, the influence of oil emission, speed and load of previous steps (features) are quantitatively calculated and plotted, with higher weight indicating larger contribution in correlation fitting. The horizontal axis represents how many steps those features before the targeted oil emission moment. The system memory exists due to the decay of feature influence. These results show that the features of the previous a few time steps are most important to the outcome of the oil emission and then the influence of the past features decreases dramatically before becoming irrelevantly small.
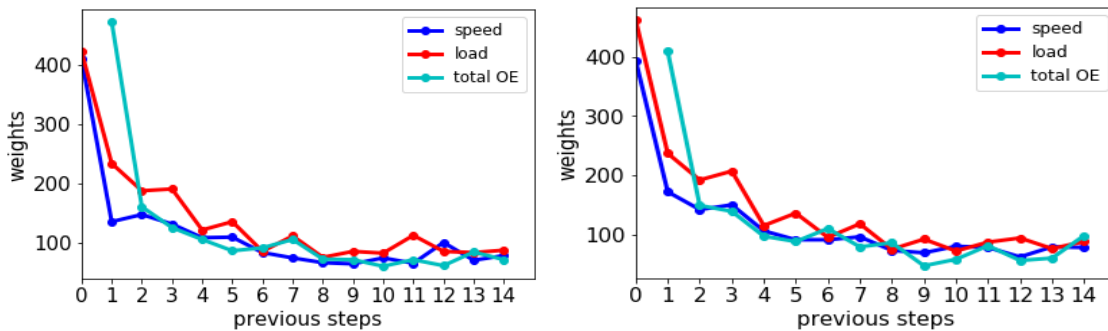
(a) Ringpack design 1   (b) Ringpack design 2
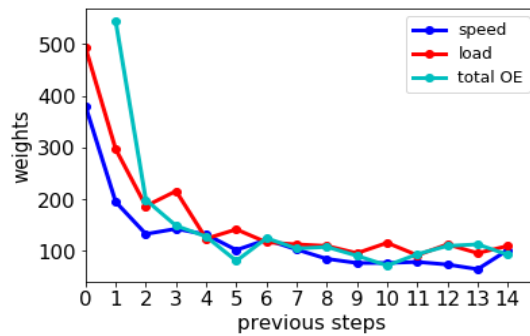
(c) Ringpack design 3

Figure 2.15 The contribution of each feature for oil emission prediction

The calculation of the feature weights relies on the internal mechanisms of tree-based gradient boosting, more specifically, the model of CART. As explained in Figure 2.16, the split of each

tree node includes two steps: selecting a feature (F) and defining a best threshold value f_0. The choice of F and f_0 is usually done by brute force searching or random searching. After defining the two steps, data samples can be divided into the two sub-nodes of the parent node. During this process, the regression error is reduced by some amount, $k(F)$. Based on the built gradient boosting model, the total contribution of each feature can be calculated as:

$$weight(F) = \sum_i w_i \sum_j k_i^j(F)$$

Here index i denotes different trees and index j denotes different node splits. w_i is the weight of each tree in the calculation of final regression results.
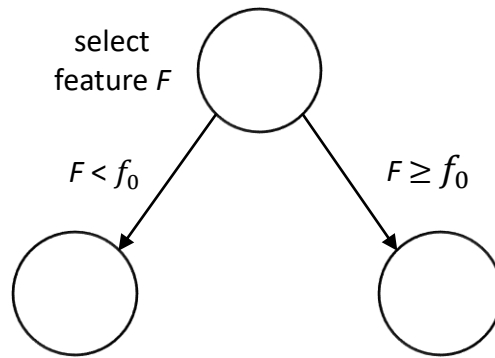
Figure 2.16 The mechanisms of CART

2.5.2 Engine accelerations

Ground on the physics of oil transport in the ring-pack system, engine acceleration should be an important input feature because inertia force of oil is determined by acceleration. The calculation of engine acceleration is fairly straightforward: the first-order differenced speed sequence is the acceleration series.

$$A_t = S_t - S_{t-1}$$

A_t is the engine acceleration. The feature choice in previous sections includes the speed terms, but acceleration history is not added into the model. Theoretically, if the machine learning model is perfectly intelligent, it should be able to fully discover the correlation between oil emission and engine acceleration because the speed information is already provided. However, in practice,

the effect of engine acceleration can only be partially detected via machine learning if the acceleration terms are not explicitly added into the feature set.
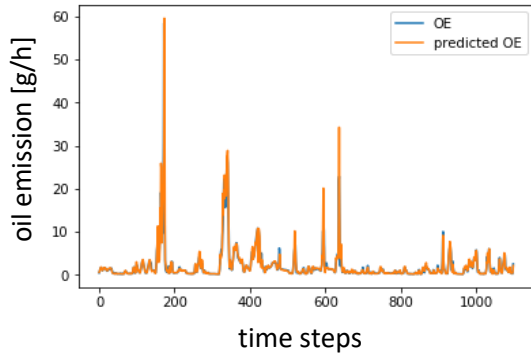
(a) Ring-pack design 1　　　　　　　　　　　(b) Ring-pack design 2
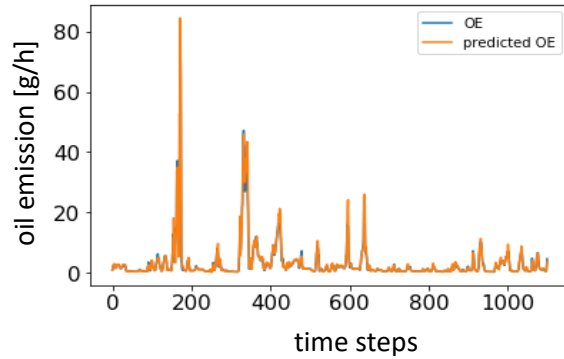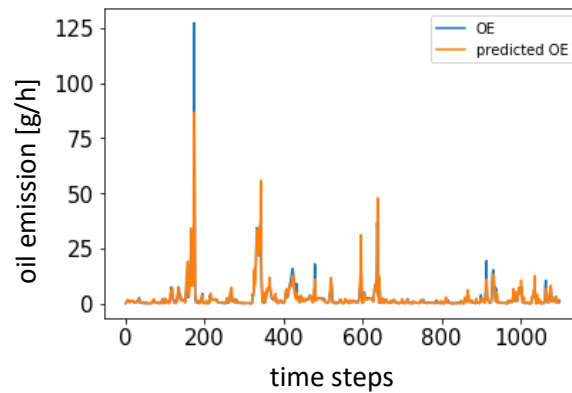
(c) Ring-pack design 3

Figure 2.17 oil emission prediction, with acceleration added into the feature set

Figure 2.17 contains plots of predicted oil emission sequences, adding previous acceleration terms. Compared to the regression results without explicitly including accelerations, the prediction of oil emission peaks is improved, which is usually the difficult part to forecast. For example, the first significant oil emission peak of ring-pack design 1 testing series is very close to the predicted value.
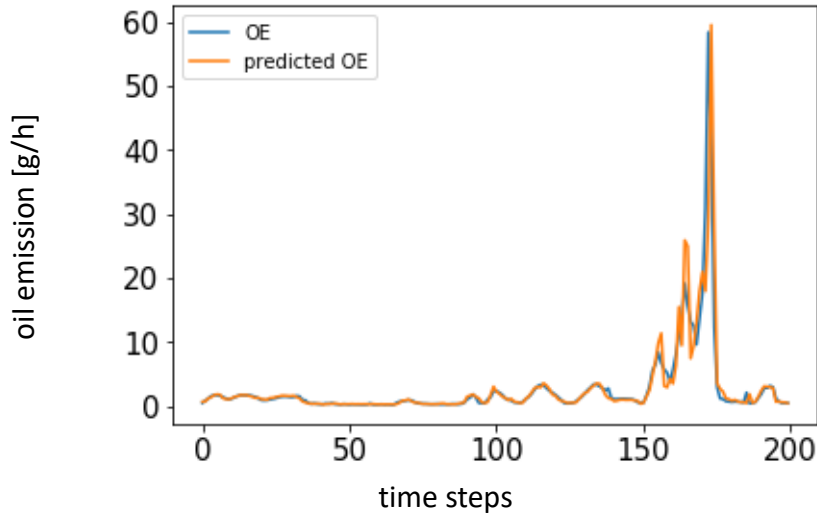
Figure 2.18 prediction around oil emission peaks, ring-pack design 1

The quantitative testing errors are presented in Table 2.7. The results show that adding acceleration as input features can improve the predictions, achieving lower errors. The amount of error reduction through adding acceleration is small, because the original fitting already includes the speed sequence, with partial acceleration correlation captured.

	Without acceleration	With acceleration
Ring-pack design 1	0.188	0.171
Ring-pack design 2	0.197	0.190
Ring-pack design 3	0.307	0.302

Table 2.7 prediction errors of the three designs, with or without acceleration included

2.5.3 Correlation of previous oil emission steps

When we are defining the residence time in the previous sections, it refers to the status of the entire system with a period of time. Thus, the feature set contains speed, load and oil emission status within the residence time period. In this section, an important property is investigated: the Markov property of oil emission. The Markov property means that evolution of one variable in the future depends only on the present state and does not depend on past history. In order to

54

investigate this property of oil emission, the feature set only includes the last oil emission step before each moment:

$$O_t^{tr} = f_{tr}(S_t, \ S_{t-1}, S_{t-2}... \ S_{t-T}, \ L_t, \ L_{t-1}, \ L_{t-2}... \ L_{t-T}, \ O_{t-1})$$

The speed and load terms are selected same as before, with T also denotes the number of previous steps included in feature engineering.
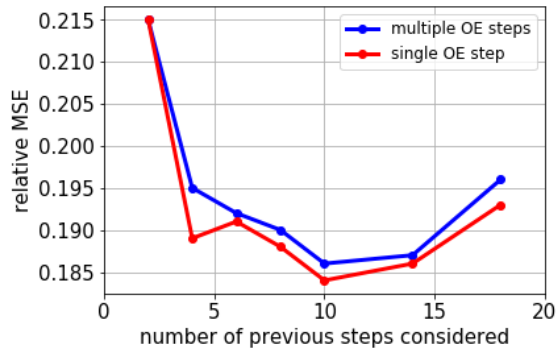
(a) Ringpack design 1
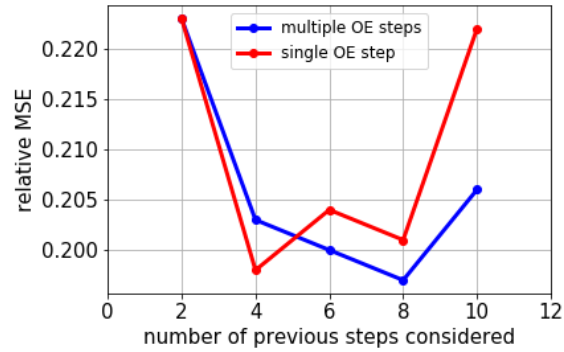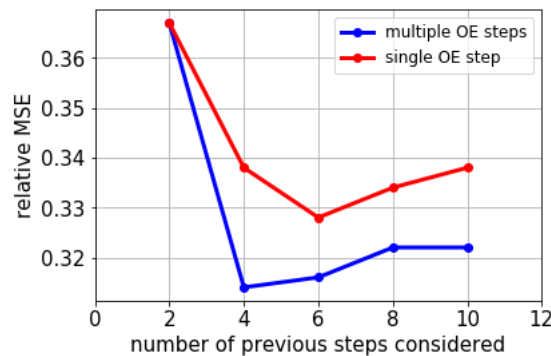
(b) Ringpack design 2

(c) Ringpack design 3

Figure 2.19 testing errors of different feature choices: multiple oil emission steps and a single oil emission step before every moment

The testing results with a single oil emission step are illustrated in Figure 2.19. In these plots, the horizontal axis is the number of previous speed and load steps considered, the variable T. Comparing to the results with multiple previous oil emission steps considered, the regression has

very close prediction error if only one oil emission step is kept. Although the regression errors have some difference, the magnitude is small enough to draw the conclusion that the current oil emission values depend mostly on the previous oil emission step. Oil emission has incomplete Markov property. The dependency of the previous speed and load is because they contribute to the oil accumulation below the top ring, which is the oil source passing the top ring to the crown land to add more oil to the crown land. The oil emission is coming from oil escaping the crown land (above the top ring), while the last time step reflects the oil accumulation on the crown land.

2.5.4 Quasi-steady oil emission sequence generation

The original measured oil emission sequence consists of both steady-state and transient components. With the found characteristic residence time, it is possible to generate a quasi-static sequence using the steady-state correlation. There are mainly two motivations to generate artificial oil emission series. First, the generated sequence can be utilized to verify the existence of system residence time and whether the found residence time is in a reasonable range. Second, this generating process can be applied as part of data augmentation, which is useful when the data volume for training is inadequate. Through adding reliably generated data, machine learning can better find the underlying correlation and make improved predictions. The quasi-static oil emission generation process is explained visually in Figure 2.20 and is divided into three steps:

1. The original load and speed sequences are elongated τ times, where τ is the characteristic residence time in terms of measurement steps. This means in the original sequence, the two adjacent steps have a time lag Δt. Now it becomes $\tau \Delta t$.

2. Linear interpolation is used to fill the gaps. Because the machine learning fitting requires a fixed time interval, it is necessary to perform interpolation between any two adjacent points in the elongated sequence. In order to maintain the same time interval Δt, $\tau - 1$ artificial points are interpolated within each gap.

3. The elongated sequence is quasi-steady, we use steady-state correlation to obtain the corresponding OE sequence. The load and speed sequences are elongated through linear

interpolation while the corresponding OE is calculated with the fitted polynomials for the SS OE.
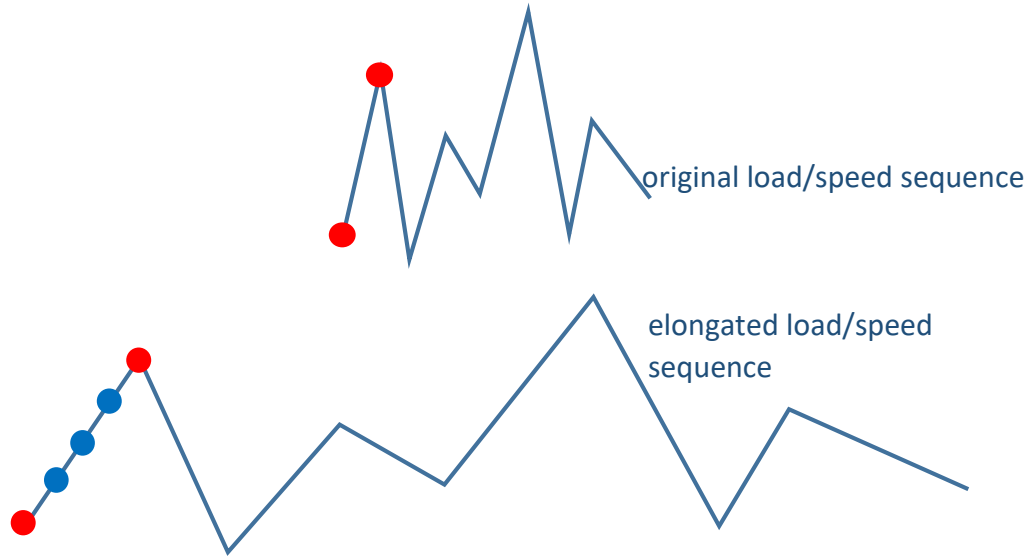
Figure 2.20 the quasi-static oil emission generation process

After obtaining the generated sequence, the training data set can include the generated data as well. With the same data processing and feature engineering procedures, the entire training set volume is expanded τ times because the length of the generated sequence is τ times longer than the original one. Table 2.8 shows that the prediction error reduces by adding to new quasi-static sequence into the training set. This improvement confirms that the generated data can enhance the training set and do contribution to transient correlation capturing. This is a good indication that the found residence time is in the correct range through getting the best bias-variance balance.

	Without quasi-static sequence	Adding quasi-static sequence
Ringpack design 1	0.188	0.182
Ringpack design 2	0.197	0.187
Ringpack design 3	0.307	0.289

Table 2.8 prediction errors of the three designs, with or without generated sequences included

It is critical to point out that the error reduction is limited to a small amount through adding the new sequence, although the entire training set becomes larger. This is because the new sequence is quasi-static covering only a portion of the transient correlation. By distracting the steady-state component, the transient magnitude left is zero. Therefore, the enlarged portion of training data of transient set always has the same training target value:

$$O_t^{tr} = 0 = f_{tr}(S_t, \ S_{t-1}, S_{t-2}... \ S_{t-T}, \ L_t, \ L_{t-1}, \ L_{t-2}... \ L_{t-T}, \ O_{t-1})$$

This limited coverage of transient oil emission leads to insignificant error reduction.

2.5.5 Connecting data-driven results to oil transport and emission physics

Based on the system analysis in Section 2.5.1-2.5.4, the data-driven system investigation results can be utilized to build the connection between engine operating conditions and the oil transport and emission process. This provides essential understandings about the oil transport mechanisms, which can be applied to designing emission reduced engines.
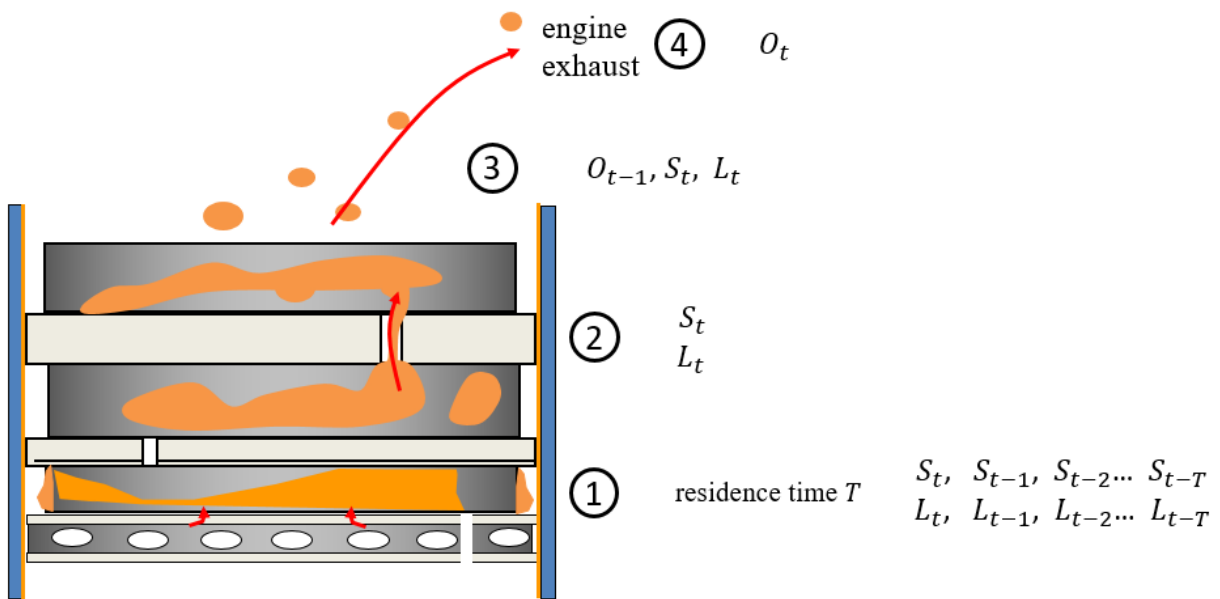
Figure 2.21 different stages of the oil transport process through the ring-pack system

As illustrated in Figure 2.21, the oil transport and emission process through the ring-pack system can be divided into four stages. The first stage is the oil leaking process through the oil control ring, leading to oil accumulation above the oil control ring. For each moment, this process is correlated with the engine operating conditions within the residence time. In addition, the residence time is determined by the design of the oil control ring. For instance, design 1 has the longest residence time because it uses the three-piece oil control ring maintaining high tension within the expander. As a result, oil accumulation above the oil control ring is very slow, therefore, the corresponding system residence time is the longest. On the other hand, the oil control ring of design 3 has many notches on it, providing direct oil flow paths. The oil accumulation is fast, resulting a shorter residence time. The second stage of oil transport is the process of oil leakage through the top two rings This process is correlated with the current load the and the current speed of the engine. The third stage is the oil accumulation process on the piston crownland (the area above the first ring) and the transport to the engine exhaust. As presented in Section 2.5.3, the current oil emission step is only correlated with the last step of oil emission, meaning the measured oil emission in the exhaust is directly coming from this stage, determined by the last oil emission step and the current engine operating condition. The last stage is the oil emission in the engine exhaust, which is measured and recorded in the oil emission sequence.

2.6 Prediction

The core application of machine learning is performing predictions of complex systems. The prediction can be applied in three aspects: point prediction for online monitoring system, interval prediction for uncertainty evaluation and total oil emission estimation. Interval prediction can also be considered as part of future data processing since it can provide feedback to the data collection procedure or experimental setup.

2.6.1 Point prediction

Point prediction is the most commonly used prediction type. From a probability point of view, any prediction is a probability distribution because prediction always contains some level of uncertainty. Point prediction of a model gives only a single value as prediction output. That value can be the mean, median or other statistics of the output variable distribution. The type of the predicted statistic depends on the objective function setup for training. The most widely used objective function in regression problems is the squared loss function, through minimizing which can provide the mean of the prediction distribution:

$$\min_{\hat{x}} \sum (\hat{x} - x_i)^2 \Leftrightarrow \hat{x} = \int xf(x)dx$$

Where $\hat{x}$ is the predicted statistic, x_i are the samples observed and $f(x)$ is the underlying probability distribution of the predicted variable.

For application in the oil emission system, point prediction can be utilized in online monitoring system. Based on the current engine status and the previous history within one residence time period, the algorithm can predict the future oil emission value. Once the expected oil emission value exceeds a critical value, the system will give out a warning of abnormally high oil emission coming next. If it is necessary, the engine testing should be suspended with inspection possibly required.

2.6.2 Interval prediction

Mentioned in the previous section, any prediction contains an underlying probability distribution, which is very useful for system design, such as risk analysis and tolerance design. However, obtaining the precisely description of an estimated output distribution is sometimes difficult or even intractable. Instead of trying to obtaining a complete distribution over an entire domain, it is often practically more useful to perform an interval prediction. Interval prediction can provide a range of the predicted variable at some certain level of confidence. On systematic level, one important application of interval prediction is uncertainty quantification.

Before using obtained data to train a machine learning model, it is always important to check the quality of the data. If the data is not qualified or contains high level of stochasticity, the training

model may have high variance because noise could also be included as some part of captured correlation. Therefore, quantifying the uncertainty level is a necessary step before performing any prediction.

Within a dataset, uncertainty can be classified to two kinds: aleatoric uncertainty (or statistical uncertainty) and epistemic uncertainty (or systematic uncertainty). Aleatoric uncertainty is the uncertainty in the input dataset, which can also be interpreted as the stochasticity of the dataset itself. This kind of uncertainty is caused by the data generation or data collection process, such as unstable experimental setup, measurement noise or simply due to improper sampling frequency of a time series. Epistemic uncertainty, on the contrary, is the uncertainty caused by the model used to fit the data. These two kinds of uncertainty are summarized in the table below.

	Uncertainty sources
Statistical uncertainty	• Measurement noise • Unstable experimental setup • Improper sampling frequency • Not enough data …
Systematic uncertainty	• Error of fitting • Stochastic procedures in the algorithm

Table 2.9 the two types of uncertainty

In this oil emission project, we only focus on the first kind of uncertainty, statistical uncertainty. This is the uncertainty reflecting the data quality. In order to eliminate the effect of systematic uncertainty, the randomness of tree-based gradient boosting is reduced to zero by setting the hyperparameter subsample to be 1. This means the built of each tree is optimized through all features, rather than a random set of features (for example, 80% of all features). Therefore, in this section, the tree-based gradient is purely deterministic.

Quantile regression is applied to predict probability intervals. In Section 1.4.1, it is discussed that the type of the predicted statistic depends on the objective function setup for training. The objective function used for quantile regression is called quantile loss function:

$$L(\hat{x} - x_i) = \begin{cases} \alpha(\hat{x} - x_i) & if\ \hat{x} \geq x_i \\ -(1 - \alpha)(\hat{x} - x_i) & if\ \hat{x} < x_i \end{cases}$$

Where $\hat{x}$ is the predicted statistic, x_i are the samples observed and α is the quantile value, with $0 < \alpha < 1$.
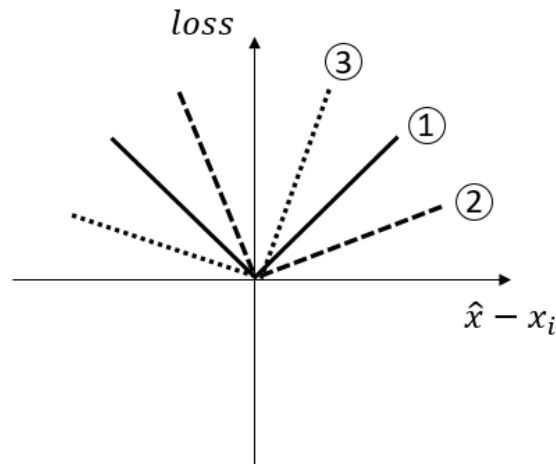
Figure 2.22 quantile loss functions with different α values

Where $\hat{x}$ is the predicted statistic, x_i are the samples observed and α is the quantile value, with $0 < \alpha < 1$. As plotted in Figure 2.22, the quantile loss function can be geometrically considered as a rotated absolute loss function, with different rotation angles corresponding to different quantiles. For instance, loss function 1 is the standard absolute loss function, through minimizing which output median can be obtained. Loss function 2 and 3 are the corresponding loss function of 25% quantile and 75 % quantile, respectively. One problem of this loss function is that it is not derivable at the origin, making it difficult for the optimization algorithm to minimize the objective function. As an alternative, Huber loss function, a very simple piecewise function, is adopted here. Huber loss uses squared loss when $|\hat{x} - x_i|$ is small and keeps the same format as the original quantile loss function when $|\hat{x} - x_i|$ is larger than a threshold value.

Figure 2.23 to Figure 2.28 are the interval prediction results of the three ring-pack design oil emissions, with two different quantile intervals, 90% and 70%. There are three oil emission sequence sections chosen for each case for better comparison. The 90% quantile interval band can cover almost every real oil emission measurement point, except for some rare conditions, such as the two peak oil emissions in ring-pack design 2. The 70% quantile band is apparently narrower, leaving more real measurements uncovered by the interval. The predicted quantile intervals of ring-pack design 3 are significantly wider than the other two ring-pack designs. This

indicates that the prediction of ring-pack design 3 oil emission contains the highest level of uncertainty. As discussed before, the only uncertainty source in this analysis comes from the data itself.

To quantify the uncertainty level of the three datasets, average quantile interval size over the entire testing sequence is calculated for each design. Larger average interval size means the higher stochasticity involved in the dataset. Figure 2.29 is the calculated average quantile interval size of the three ring-pack designs, with varying quantiles. These intervals are all centered with the output median. It is easy to observe that the oil emission data uncertainty of ring-pack design 3 is remarkably higher than the other two oil emission sets. This agrees with the wide quantile bands plotted in Figure 2.27 and 2.28. It is also interesting to interpret this high level of uncertainty of the physics and experimental perspective. The oil control ring in design 3 has many small gaps, providing paths for oil to leak from the oil reservoir. On the measurement side, although the leaked oil contains both liquid and vapor, the oil emission measurement used in this experiment can only detect the vapor part. Thus, one possible scenario is that the lube oil leaking out of the top ring with Design 3 may have more varieties in oil droplets sizes, which contribute to the oil vapor differently. Although there is no clear reason for this hypothesis, it directs the right attention when the oil transport is further investigated based on the findings here.
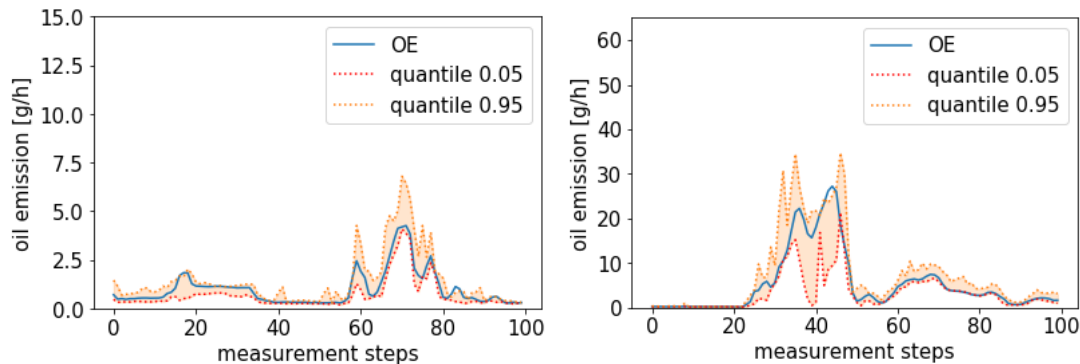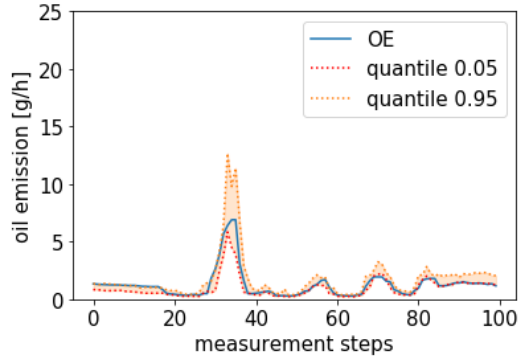
Figure 2.23 90% interval prediction of ringpack design 1, different oil emission sections
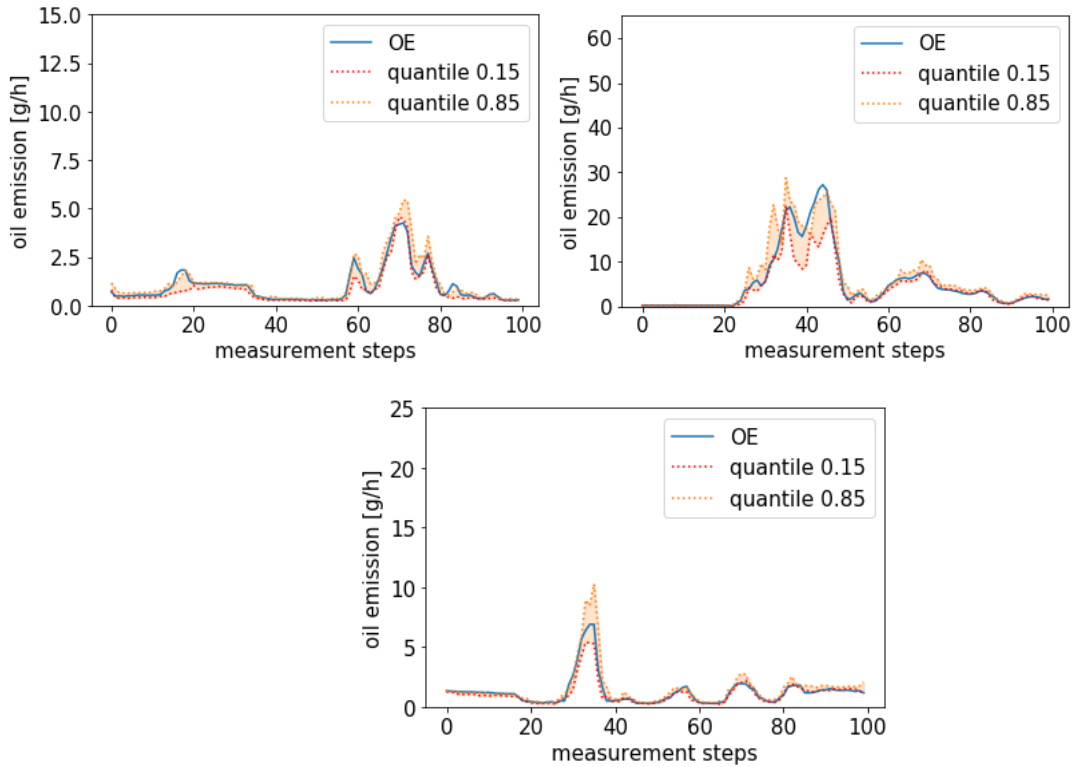
Figure 2.24 70% interval prediction of ringpack design 1, different oil emission sections
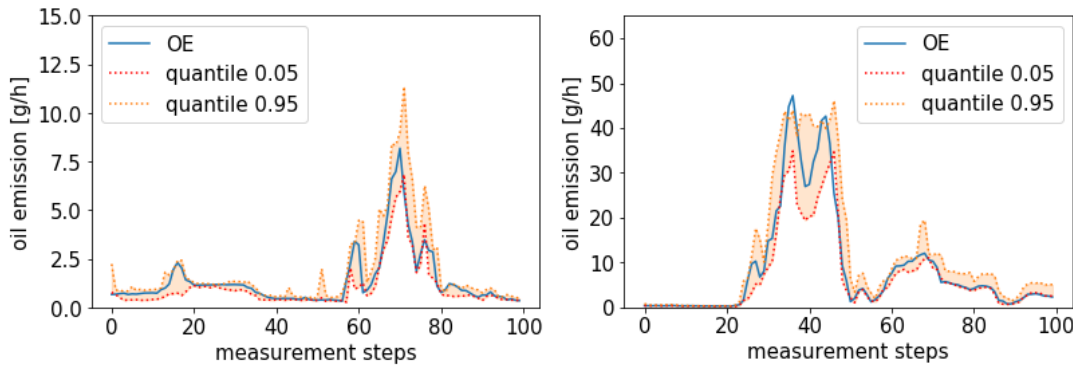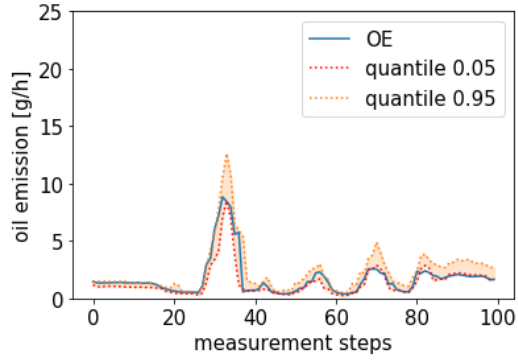
Figure 2.25 90% interval prediction of ringpack design 2, different oil emission sections
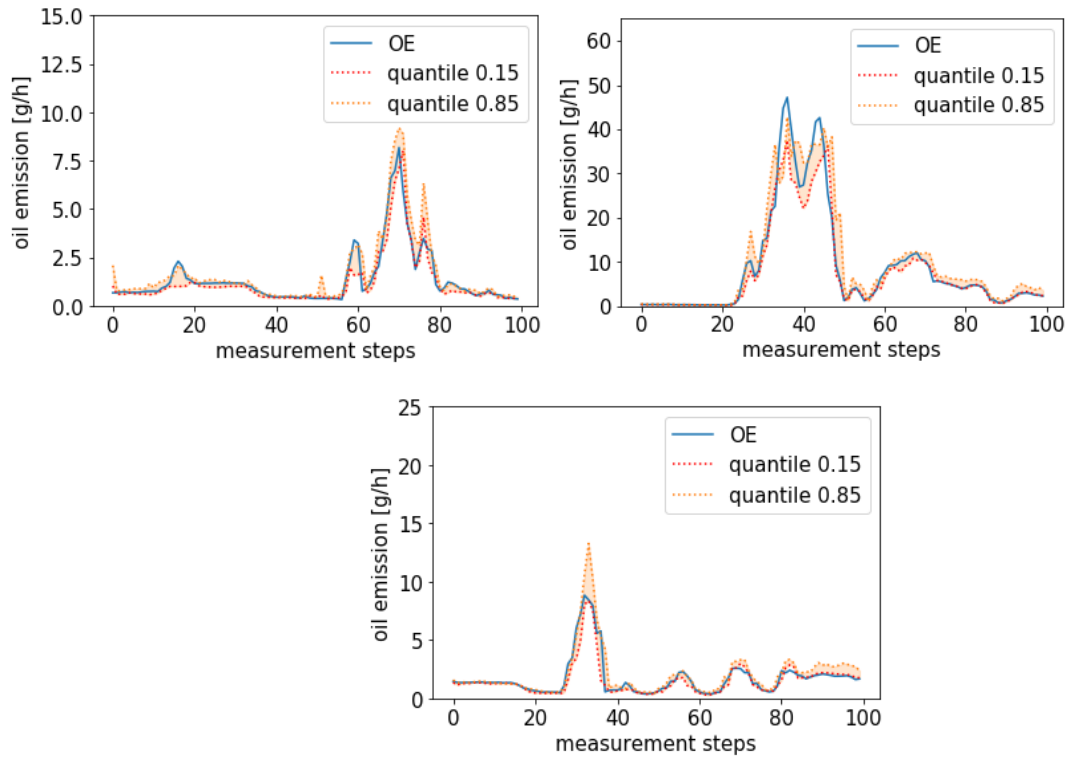
Figure 2.26 70% interval prediction of ring-pack design 2, different oil emission sections
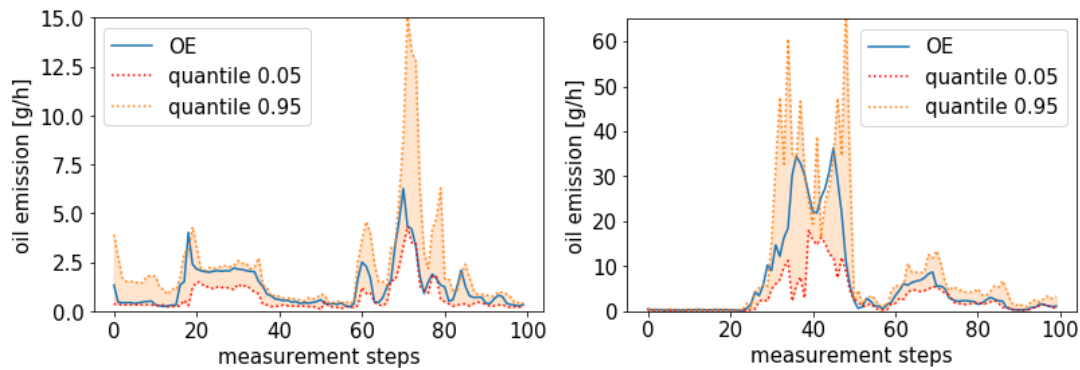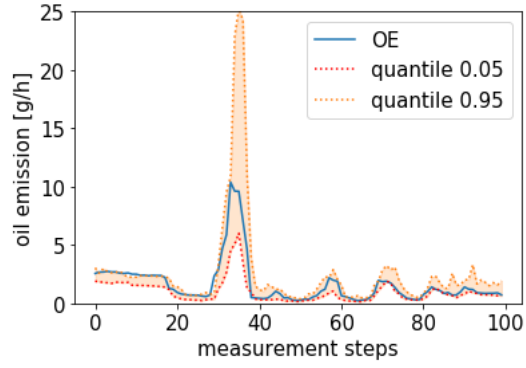
Figure 2.27 90% interval prediction of ringpack design 3, different oil emission sections
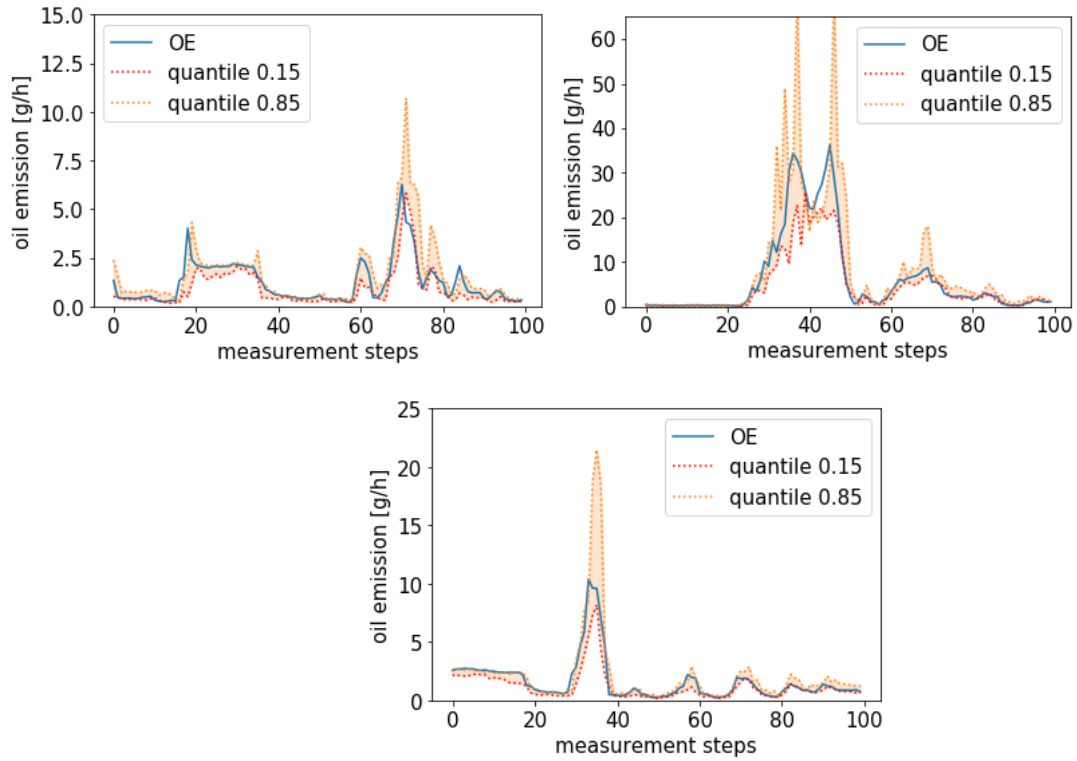
Figure 2.28 70% interval prediction of ringpack design 3, different oil emission sections
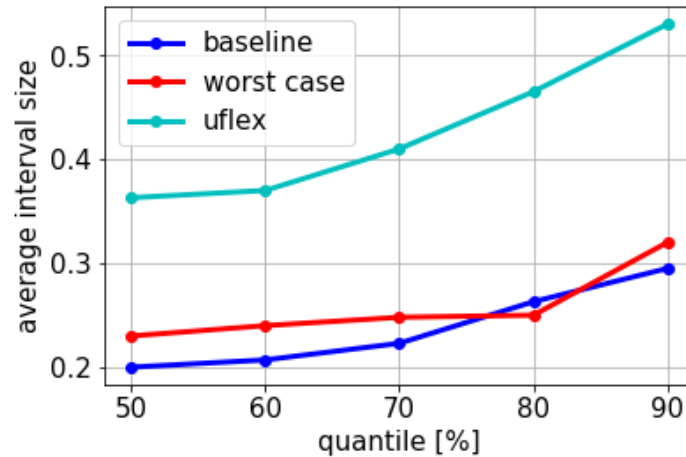
Figure 2.29 average quantile interval sizes

In Figure 2.29, the minimum quantile interval plotted is 50%. However, the trend of the interval size is disturbed by quantile cross when the quantile is low. Quantile cross means the predicted value of the higher quantile is even smaller than the value of lower higher quantile, which is mathematically wrong. The problem of quantile cross becomes significant when we decrease the quantile range. Based on the current estimation of the oil emission datasets, about 1/4 of the predicted intervals have the problem of quantile cross when the quantile range is reduced to 10%.

The machine learning based solution for quantile loss is called composite quantile regression, which predicts multiple quantiles simultaneously. In the objective function of composite quantile regression, one additional term is assigned to penalize quantile cross for any two adjacent predicted quantiles.

However, composite quantile regression can be implemented with neural networks but not tree based regression. In neural networks, it is easy to set the number of neurons in the last layer equal to the number of quantiles required to predict at the same time. Meanwhile, in the objective function of the entire network, penalty terms can be added to the quantile cross. It is difficult for a tree-based algorithm to predict multiple targets simultaneously. The CART does not have a network architecture, with the algorithm always trying to find the best path from the root to a leaf to predict just a single value, as illustrated in Figure 2.30.
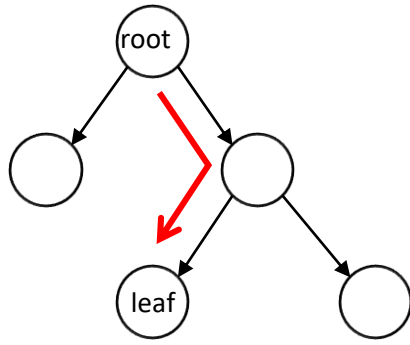
(a) prediction based on a single path   (b) prediction based on the final layer
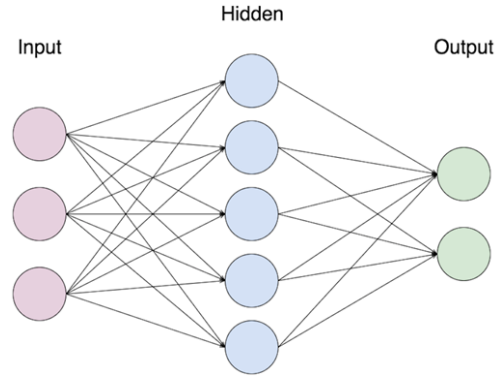
Figure 2.30 different prediction mechanisms of a regression tree and a neural network

2.6.3 Total oil emission prediction

The point prediction and the interval prediction discussed previously are all online predictions in the set of time series problems, containing oil emission history before the predicted moment. However, online prediction cannot meet the demand of total oil emission prediction task, which is an offline prediction problem. The total oil emission prediction task is providing an oil emission estimation based on the given load and speed sequences without conducting an engine experiment. This could save a lot of time and cost from reducing the number of experiments.

It is discussed in the system investigation section that oil emission has an approximate Markov property, depending on the oil emission status of the previous moment. However, performing an offline prediction means that there is no information about oil emission provided. One possible solution is using rolling prediction. With rolling prediction, the oil emission value at each moment is calculated based on the speed, load and oil emission before that moment while the oil emission history is from prediction as well. Through iterating the entire sequence and adding the obtained oil emission sequence together, the total oil emission amount can be estimated.

Rolling prediction of oil emission has a severe limitation, error accumulation. As each oil emission step is calculated based on the previous estimated oil emissions, prediction error will accumulate from the beginning of the sequence to the end. As a result, after some number of oil emission prediction steps, the calculated value will have nontrivial error.

In fact, with the purpose of obtaining the total emission over a given engine series, it is not necessary to calculate the entire corresponding oil emission sequence. The machine learning based mathematical mapping can be optimized through dividing the sequence into chunks with a fixed length. Instead of trying to get the oil emission at every moment, considering the total oil emission amount of these divided chunks is a more reasonable approach to avoid error accumulation.

Summarized in Figure 2.31, the procedure consists of two steps: sequence sectioning and training variables setup. The entire oil emission time series dataset is divided into chunks with length τ, the residence time in terms of measurement steps. The speed, load and oil emission sequences are all partitioned with the same length of chunks. The selected regression target, the dependent variable, is the total oil emission amount of every chunk. Because the objective is obtaining thetotal amount of oil emission, the calculation of the exact value at each time step is not needed. The features, the independent variables, are the speed and load steps of the current chunk and the previous chunk. The speed and load information of the previous chunk is also included because they could have influence on the oil emission at the current chunk. Since the chunk length is τ, no further speed or load steps have correlation with the training target.
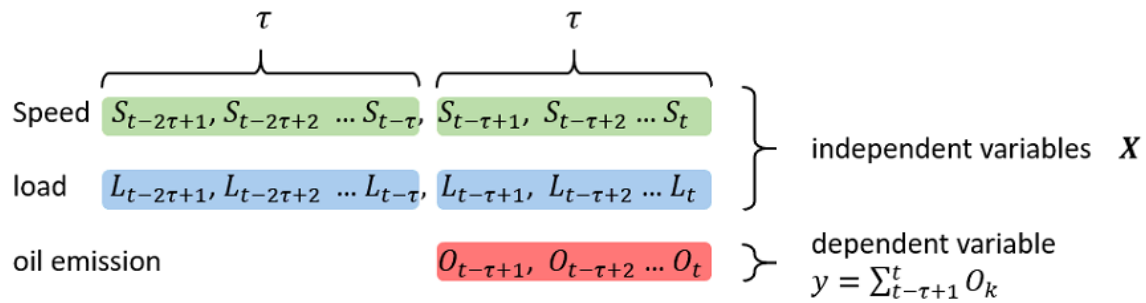
Figure 2.31 the chosen independent variables and the dependent variable in the total oil emission prediction task. The supervised learning model is built on these variables

The targeted correlation of supervised learning in this section can be expressed as:

$$O_{tot} = f_{tr}(S_{t-2\tau+1},\ S_{t-2\tau+2}, \ldots S_t,\ L_{t-2\tau+1},\ L_{t-2\tau+2}, \ldots L_t)$$

For prediction, the first step is also chopping the sequence into segments with length τ with the given speed and load sequences. Upon that, the trained machine learning model is used to predict the OE over each segment. The final total OE is the summation of prediction results over all chopped segments, as presented in Figure 2.32. One additional improvement to eliminate the effect of chopping position, all τ possible chopping positions will be tried and an averaged result will be given as the final prediction.

$$OE_1 \quad + \quad OE_2 \quad + \quad OE_3 \quad + \quad \cdots \quad + \quad OE_k \ = total\ OE$$
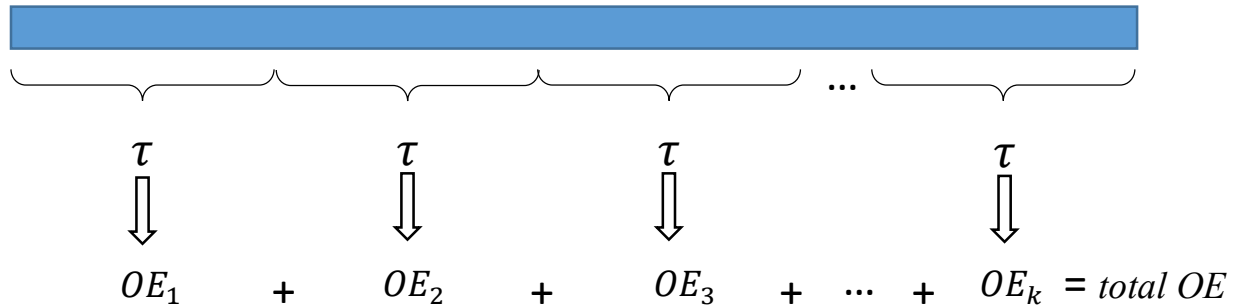
Figure 2.32 prediction of total oil emission over a given engine status sequence

MAPE	Ringpack design 1	Ringpack design 2	Ringpack design 3
Sequence of length τ	0.15	0.21	0.29
Sequence of length 500	0.05	0.09	0.09
Sequence of length 1000	0.05	0.10	0.08

Table 2.10 total oil emission prediction results with different engine sequence lengths

The total oil emission prediction errors are listed in Table 2.10. When the given engine operating sequence is long enough, the prediction error goes below 10% for all the three ringpack designs. Since the total OE is obtained through the summation of all chopped segments, the absolute error is smaller over a long sequence than the unit length τ. This is due to the reduction of variance when the predicted sequence is longer.

2.7 Design engine testing Cycles

Besides prediction, another essential application of this project is to provide engine testing guidance to save time and cost for manufacturers. Before the engine product is launched to the market, thorough engine testing should be conducted so that the oil emission condition of any customer driving cycle is within expectation. Because conducting a real engine testing is expensive, it is important to find a short but efficient engine testing sequence to cover all possible driving conditions. Different from the previous supervised learning setup, this testing cycle design requires an unsupervised model to verify whether or not a given testing cycle is representative of a specific engine. This task needs to provide design feedback without obtaining the corresponding oil emission series, leaving only load and speed sequences as inputs. In this project, one-class support vector machine (OCSVM) is utilized as the unsupervised learning model.

2.7.1 One-class support vector machine and novelty detection

Support vector machine (SVM) is a traditional machine learning model, more commonly used for binary classification problems. It splits the labeled samples into two groups in the nonlinear kernel space separated by a hyperplane with maximum margins from the samples. OCSVM, as a variant of SVM, is an unsupervised learning model, usually applied to novelty detection. Same as SVM, the first step of OCSVM is mapping the samples into a new kernel space. Instead of finding a hyperplane, OCSVM is search a hypersphere with minimum radius to contain all the samples in the kernel space:

$$\min_{R,c} R^2$$

Subject to: $\|x_i - c\|^2 \leq R^2$ for all $i = 1, ... n$

Where R is the radius of the hypersphere, c is the center of the hypersphere. x_i is the training sample point. The hypersphere in the kernel space corresponds an enclosed surface in the original feature space, which is difficult to obtain without mapping to the kernel space, as illustrated in Figure 2.33.
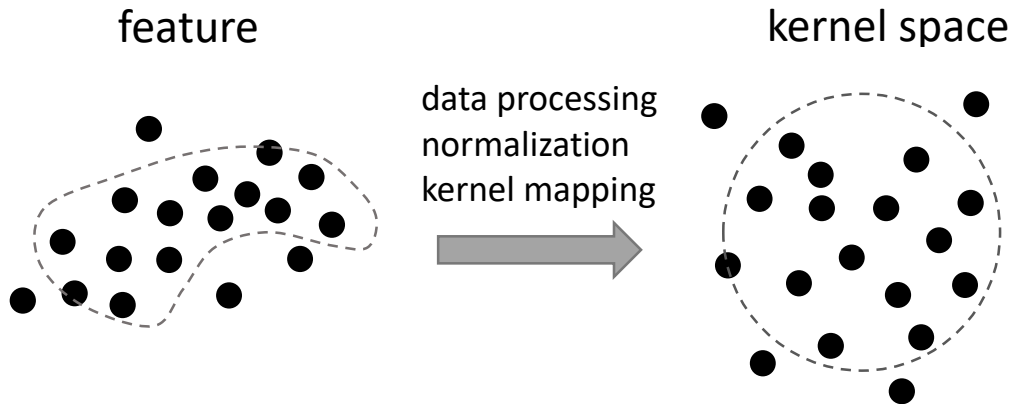
Figure 2.33 OCSVM for training

If the uncertainty of the dataset is considered, not all of the training points are necessary to be included with the hypersphere. Some of the points should be allowed to locate outside of the hypersphere to counteract the data uncertainty. In this situation, the distance from the points outside of the hypersphere to the decision boundary can be added to the objective function as penalty terms:

$$\underset{R,c}{min} \; R^2 + \beta \sum_{i=1}^{n} \xi_i$$

Subject to: $\|x_i - c\|^2 \leq R^2 + \xi_i$ for all $i = 1, \dots n$

$$\xi_i \geq 0$$

The parameter β decides the magnitude of the penalty on the points locating outside of the hypersphere. This parameter should be correlated with the uncertainty level of the oil emission dataset. If the uncertainty level is high, β will be set to a relatively lower value to reduce the penalties for the points outside of the hypersphere.

Once the optimized hypersphere is found, we can check an arbitrary given sample point, whether or not it is predictable (within the hypersphere). In addition, the distance from a point to the hypersphere surface is defined positive for interior points and negative for points outside. Figure 2.34 shows how the testing points are compared with the hypersphere boundary.
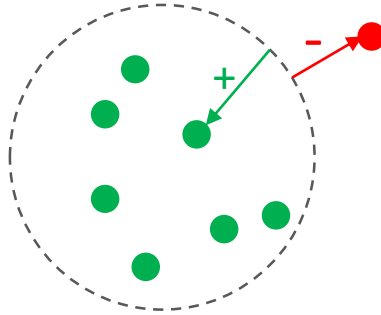
Figure 2.34 OCSVM for testing

2.7.2 Application to engine testing design

Before getting to the procedure of engine testing cycle design, we first use OCSVM to verify whether the testing sequence is covered by the training sequence in Section 1.4. In this unsupervised learning task, the only focused variables are the input features (independent variables).
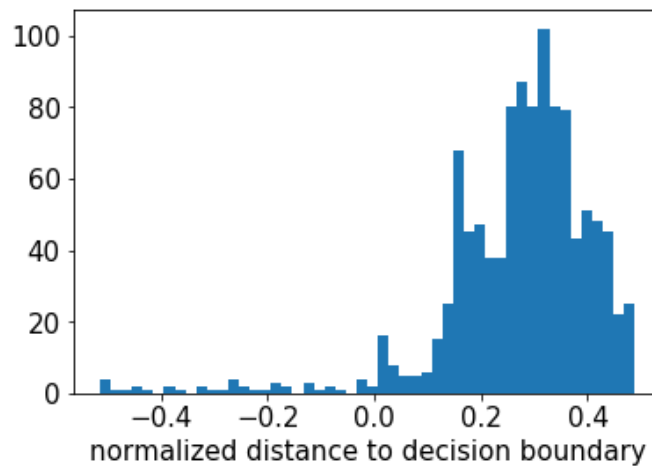
Figure 2.35 the distances from points to the decision boundary

By mapping the original feature space into a Gaussian kernel space, the hypersphere is found to enclose the training samples. As an example, the dataset of ringpack design 1 is used. Because the evaluated uncertainty of design 1 experimental data is low, all training sample points are enclosed within the decision boundary of OCSVM in the training process. For the testing sequence, the normalized distance of every testing point to the hypersphere is calculated and

73

plotted in the histogram of Figure 2.35. Most of the points have positive distances, suggesting that the majority of the testing sequence falls into the same region as the training set in the feature space. The rest 3% of the testing samples, having negative distances, are outside of the decision boundary. The prediction of those points may not be accurate enough as the training set does not cover the feature space region of them.

It is interesting to plot the obtained distances in Figure 2.35 in the time domain as a whole sequence. Based on Figure 2.36, the locations of the distances are not random, with a continuous sequential pattern. There are some periods with significantly lower distances to boundary, suggesting the low confidence of the fitted model to predict these time sections. Compared to Figure 2.11, these regions are the peaks of oil emission in the testing series, leaving observable difference between the predictions and the real measurements. The results of these two sanctions now agree about the conditions difficult to predict based on the training set.
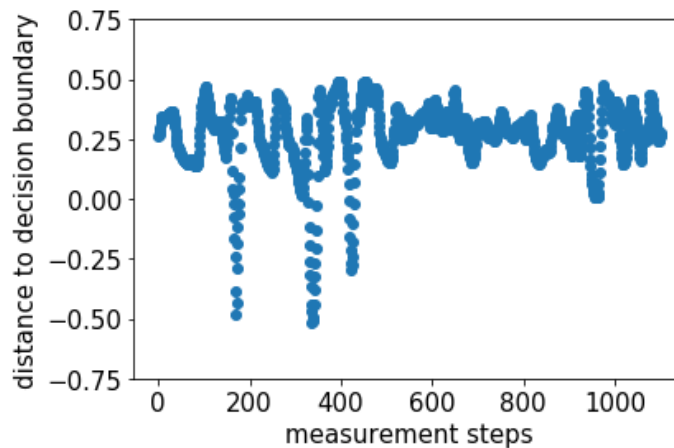
Figure 2.36 the distances from points to the decision boundary, as a time series

OCSVM is also very useful for engine testing sequence design. With a given engine, the speed range and the load range are determined, so that in the real customer driving conditions the speed and load values will not exceed these ranges. Under the given speed and load ranges, using design of experiments principles, such as Taguchi method, can generate the full set of target engine sequences that the testing cycles should cover. The number of all target engine sequences is finite with a defined residence time. The target engine operating sequences should be presented as a tabular set of data, with the number of features twice as the residence time steps τ.

74

In this design procedure, only speed and load are the input features of engine experiments design while the oil emission is just the corresponding response.

	L_1	L_2	...	L_τ	S_1	S_2	...	S_τ
1								
2								
...								
N								

Table 2.11 the format of the target engine status, generated by design of experiments

The generated target engine operating series is used as the testing set of OCSVM. The training set can be chosen from existing customer driving cycles or engine experiments. OCSVM is utilized to verify whether the target engine status sample locate within the space covered by the engine experiments in the feature space. Among all possible testing cycles, the shortest one satisfying the statistical coverage criteria of OCSVM can be chosen as the ideal testing cycle to push experimental cost to the minimum amount.

2.8 Conclusion

In the project of engine oil emission, machine learning and other statistical models are applied to find correlations, analyze system properties, perform predictions and design for engine testing. A complete set of data-driven procedures are defined in this chapter, including data processing, featuring engineering, model section, system parameter estimation, prediction and experiments design. These procedures are coupled together through important system properties, such as uncertainty and system residence time. On the application level, both supervised learning and unsupervised learning models are applied with the purpose of predictions and engine testing designs, respectively. As the examples used in this project, the oil emission testing datasets of the three different ring-pack designs are analyzed, fitted and connected to the physics of the oil transport mechanisms. The results suggest a good consistency of the machine learning analyzation and the underlying oil emission physics. The entire defined data-driven procedures

show a promising future to accelerate engine development cycle, reduce engine testing cost, and help understand oil transport mechanisms and design influences.

Chapter 3. Deterministic modeling of liner wear during the break-in period and its impact on piston ring lubrication

3.1 General introduction

To design engines with low greenhouse gas emissions, it is important to reduce friction in the contact system of piston rings and cylinder liners. Besides friction loss, the contact of a piston ring and liner also leads to wear of both surfaces of the ring and the liner, which can cause negative effect in friction reduction. The reason is directly related to the surface plateau-honed structures of cylinder liners. A critical manufacturing step in the manufacturing process of liners is called the plateau-honing process. The process includes two stages: rough honing, creating the valley part with grooves and fine honing leading to much smaller plateau roughness compared with the valleys [29,30]. As some features of the plateau roughness are quickly worn off during break-in period, engine lubrication condition and friction will change significantly. Therefore, it is crucial to understand and simulate the break-in wear of the liner surface with the target of understanding engine lubrication change and generating better liner roughness design.

The two-body asperity fatigue wear model presented here is physics-based and comprised of different sub-models. First, the surface plastic flattening was considered as the mechanism responsible for the changes of the surface during the initial contact and it was modeled with an asperity contact model and a simple flattening process. Then the asperity fatigue wear was considered as the dominant process for the change of the surface during the rest of break-in period. As the successive mechanism, asperity fatigue wear is studied, starting from the theory of delamination, which was first proposed by Suh [8]. Two sub-models are discussed to complete the simulation of fatigue wear: crack initiation and crack propagation. Because fatigue wear happens at the asperity level and near the surface, surface crack and subsurface cracks are modeled separately due to different initiation positions and driving forces for subsequent crack growth.

Many previous physics-based models were developed to simulate the fatigue wear process. Lundberg-Palmgren model is a representative work of early fatigue wear [14]. Lundberg and Palmgren proposed the first mathematical model for predicting contact fatigue life. The later studies by Tallian added more physics-based understandings and the proposed models are based on a crack propagation and elastic/plastic properties of materials [15,16]. Kudish and Burris improved the previous models by adding the effect of normal stress, size of the contact, and friction coefficient [17]. Previous works were also conducted for modeling of engine break-in wear [18,19,20]. However, without engaging detailed mechanisms of break-in wear, these models mainly focus on how to model the wear processes in a mathematical way. This thesis work for the first time applies the fatigue wear theories to liner roughness break-in wear. Furthermore, the functional change in ring lubrication and friction caused by the wear investigated here will be demonstrated and the comparison with the experimental measurements on friction will be presented.

Asperity Fatigue Wear Model	Engine Cycle Friction Model

Asperity Fatigue Wear Model

- Fatigue models and parameters
- Fracture time for each asperity
- New liner surface topography

Engine Cycle Friction Model

Different correlations in engine cycles:
Hydrodynamic pressure correlation
Hydrodynamic friction correlation
Contact pressure correlation

- Oil viscosity, density
- Lubrication theory
- Friction change during break-in

Figure 3.1. The four parts of the break-in liner wear project

The work is consisted of four parts, as presented in Figure 3.1. As the model input, liner surface topography is measured and represented as a roughness matrix, with each entry as the height of the corresponding location. The asperity contact model is used to determine the clearance between the two surfaces based on the contact condition of each asperity. For the asperities under cyclic load, both surface cracks and subsurface cracks initiate, propagate and cause asperity fracture when they meet each other. As the last part, friction change is also calculated and compared

3.2 Elements of break-in asperity fatigue wear

In order to perform physics-based deterministic modeling of break-in liner wear, the two basic processes are modeled: asperity contact and asperity fatigue. The governing equations and approximations are presented in this section.

3.2.1 Asperity contact

Due to the design convention that ring surface is significantly smoother compared with the liner surface, the deformation of asperities on the liner surface first experiences a flattening process, with a fast change for the plateau roughness. This plastic flattening process is significant

especially during the first several cycles of engine running. Although this process is short in duration, the change of surface topology significantly affects the behavior of the asperity contact. Therefore, it is essential to consider this mechanism, and integrate it into the entire two-body break-in wear simulation.

This asperity contact model is computationally efficient under several assumptions. First, the contact occurs between a smooth surface and a rough surface, with the smooth surface considered as a rigid body. In the ring liner contact system, ring is prominently smoother and harder compared with liner. Second, the interactions among asperities are neglected, so the deformation of each asperity only depends on the contact situation between the ring and its own. This assumption is sufficient for a sparse contact pattern, with the real contact area only consists of a small portion of the nominal contact area. The final assumption is that all asperities have spherical summits, which contact with the ring.

Based on the work of Chang [31], the contact model for each single spherical asperity can be mathematically expressed as:

$$A_c = \pi R h; \quad P_a = \frac{4E}{3\pi}(\frac{h}{R})^{1/2}$$

$$A_c = \pi R h [1 - 2\left(\frac{h-h_1}{h_2-h_1}\right)^3 + 3\left(\frac{h-h_1}{h_2-h_1}\right)^2]; \quad P_a = H - H(1-k)\frac{\ln h_2 - \ln h}{\ln h_2 - \ln h_1}$$

$$A_c = 2\pi R h; \quad P_a = H$$

In equations above, A_c is the real contact area, P_a is the mean contact pressure. R is the radius curvature of the asperity tips, H is the hardness of the material, and k is the mean contact pressure factor, which represents that if the mean normal pressure P_a exceeds kH, initial yielding occurs. For a liner made of cast iron, k is usually taken around 0.4. The determining parameter in the model is the interference depth h, which is the asperity tip height reduction after contact. Figure 3.2 shows the geometric properties and asperity deformation of the contact model.
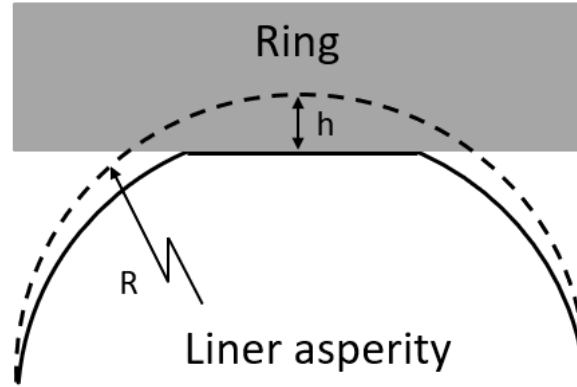
Figure 3.2 asperity contact model

The contact pattern for each asperity always falls into one of the three contact categories: elastic, elastoplastic contact and fully plastic. The three categories are identified by two critical interference depth: $h_1 = (\frac{3\pi kH}{4E})^2 R$ and $h_2 = (\frac{3\pi H}{2E})^2 R$. When $h \leq h_1$, the contact is purely elastic, which corresponds to the condition that the normal pressure everywhere on the contact surface is smaller than material hardness. Elastoplastic condition is a transition between elastic contact and fully plastic contact, with the center part of the contact region reaches plasticity. As the interference depth further exceeds h_2, the contact becomes fully plastic. Equations above correspond to the analytical expressions of real contact area and mean contact pressure for the three conditions, respectively.

Based on the asperity contact model discussed above, the average clearance between the ring and the rough liner surface under a given pressure can be solved iteratively. The summation of contact forces over all individual asperities equals the total load of solid-to-solid contact.

For asperities contacting elastically with the ring, they will recover their original shape right after the ring leaves. However, when the deformation contains some portion of plasticity, permanent flattening effect would occur.

The detailed asperity plastic deformation is not trivial. Because of different shapes, contact pressures, contacting angles, all asperities on the liner surface could have various plastic deformation patterns. However, under the assumption that asperities all have spherical summits, calculating the deformation for asperities is still feasible. An efficient wear model should adopt a

relatively universal method to perform asperity deformation analysis over the entire simulating surface. As proposed by Sören and Söderberg, asperity plastic deformation can be modeled as plastic flattening with an elastic spring back [32]. If the contact deformation involves plasticity, the asperity will never get back to its original shape but has a small spring back. The amount of elastic spring back at the center of the contact asperity is given by KL Johnson [33]:

$$u_{el} = 2(1 - v^2)\frac{H}{E}\sqrt{\frac{A_c}{\pi}}$$

where v is the Poisson ratio, H is the material hardness, E is the elastic modulus, and A_c is the real contact area. Once the contact is beyond the Hertzain range, identified by the penetration depth $h > h_1$, this elastic spring back is used to simulate the asperity height after surface flattening. Because of the existence of spring back effect on each asperity contact plastically, asperities are not completely flattened.

3.2.2 Asperity fatigue

Among all mechanical wear mechanisms, surface plastic deformation, fatigue wear and abrasive wear are the most essential ones leading to the surface topology change. Delamination theory considers the fatigue as the core mechanism in the wear process. In the engine break-in period, this is the dominant wear mechanism as the ring surface is still smooth, without abrasion of third-body particles.

In the delamination theory, wear starts with asperity deformation and fracture due to surface contact under certain external load. This process starts immediately after contact begins, gradually switching plastic contact into elastic contact. Meanwhile, the delamination process also begins at the surface and subsurface of the contact material through the process of crack initiation. It is critical to point out that the delamination process includes not only crack propagation and fracture of surface material, but also crack initiation.
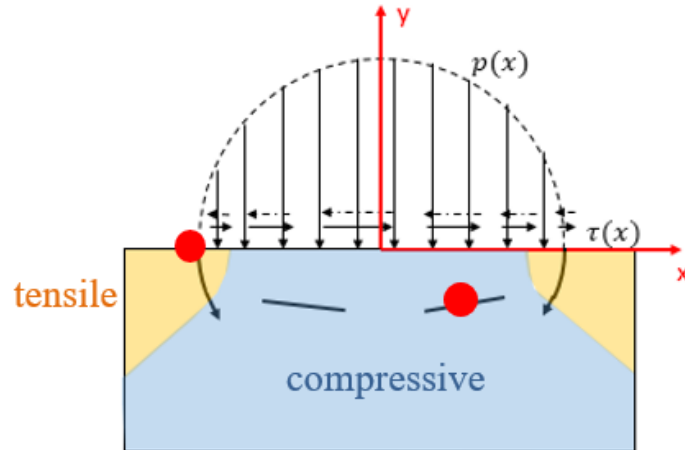
Figure 3.3 Crack initiation. Surface cracks initiate from the edge of the contact region where the surface tensile stress reaches maximum. Subsurface cracks initiate from maximum shear stress region.

As illustrated in Figure 3.3, according to the principal normal stress, the subsurface under each contact can be divided into tensile region and compressive region. Both surface cracks and subsurface cracks are initiated based on small material defects although the forces leading to crack initiation are different. Surface crack initiation relies on surface tensile stress as tensile component on the surface can create a trend for cracks to form and grow (compressive stress will not lead to crack initiation as it always helps close the existing defects). Therefore, surface cracks always initiate at the edge of contacting region at which surface tensile stress is maximum. The damage force for subsurface crack initiation is shear stress. This is because subsurface cracks initiate in compressive region where no tensile component would help create cracks. The only source for crack initiation is the shear component. As a result, the initiation position for subsurface cracks is the maximum shear stress position.

	Surface crack	Subsurface crack
Damage force	surface tensile $\sigma_{xx}\|_{y=0}$	shear τ_{xy}
Initiation position	edge of the contact region	location with maximum τ_{xy}
Propagation mode	mode I, mode II combined	mode II only
Driving force	tensile circumferential stress: $\sigma_{\theta\theta} = \frac{1}{\sqrt{2\pi r}}\cos(\frac{\theta}{2})(K_1 \cos\left(\frac{\theta}{2}\right)^2 - \frac{3}{2}K_2\sin(\theta))$ K_1, K_2: stress intensity factors for mode I and mode II	shear stress at crack tip: $\tau = \frac{1}{2\sqrt{2\pi r}}[K_1{}^2 \sin^2\theta + 2K_1K_2\sin 2\theta + K_2{}^2(4 - 3\sin^2\theta)]^{0.5}$
Propagation direction	perpendicular to $\sigma_{\theta\theta\,max}$	along τ_{max}

Table 3.1. Crack initiation and propagation mechanisms. As surface cracks and subsurface cracks locate in different regions under contact asperities, the mechanisms for initiation and propagation are different.

After the initiation of surface and subsurface cracks, propagation of cracks would ultimately lead to fracture of contact asperities. Similar to initiation of cracks, cracks in tensile region and compressive region have different propagation modes and driving forces. For a surface crack, both fracture mode1 and fracture mode2 contribute to propagation as tensile stress and shear stress exist simultaneously in the tensile region. The driving force for surface crack propagation is the tensile circumferential stress $\sigma_{\theta\theta}$ near the crack tip, with the corresponding propagation

direction perpendicular to the maximum $\sigma_{\theta\theta}$ direction. For a subsurface crack located in the compressive region, the driving force is the shear stress at the crack tip while the crack propagates to the direction of maximum shear stress τ_{max}. The crack initiation and propagation mechanisms are summarized in Table 3.1.
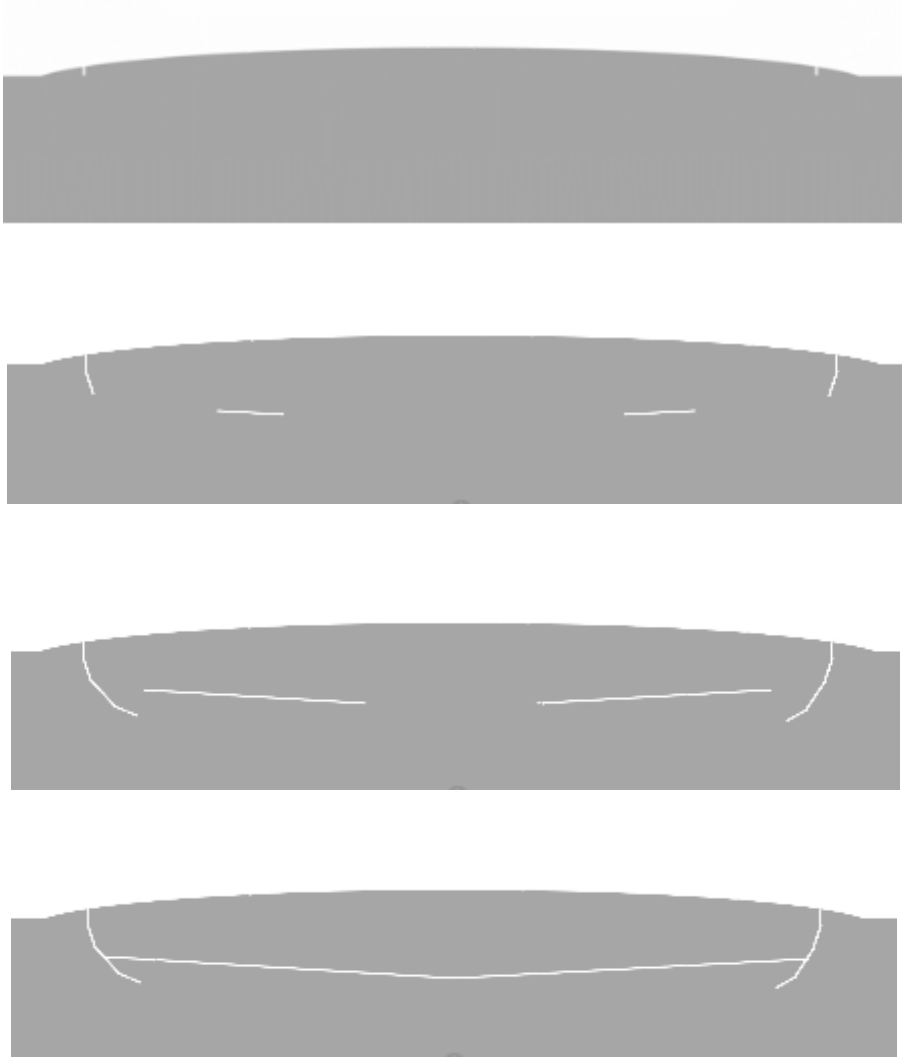
Figure 3.4 initiation and propagation of surface cracks and subsurface cracks of one asperity under repeated load

The crack propagation is calculated by the Paris-Erdogan law [34] under the assumption of small yielding around cracks. Kudish and Burris presented an efficient method to calculate the stress

intensity factors K_1 and K_2 under each contact asperity [17]. An asymptotic solution is given by the authors to solve K_1 and K_2 in an explicit way.

3.2.3 Algorithm of the deterministic modeling

Based on sub-models introduced in the previous sections, a final asperity fatigue wear model of liner surface in break-in period is developed. This wear model includes the following essential steps: asperity contact, surface plastic deformation, crack initiation, crack propagation and asperity fracture. These steps are applied for every sliding cycle iteratively. The flowchart in Figure 3 shows the algorithm of this fatigue wear model, which integrates all these critical mechanisms in the fatigue wear process.

In Figure 3.5, the asperity fatigue wear algorithm needs two inputs: rough liner surface S, represented by a matrix and an external nominal pressure P. First, the algorithm finds the clearance h between the liner base altitude and the flat ring surface, according to the asperity contact model with the given liner roughness measurement and pressure. All asperities in contact A(n) are numerically defined once the clearance is determined, including each asperity's radius of curvature R(n), average normal pressure P(n) and real contact radius r(n). With all the information of each asperity, the algorithm performs the surface plastic flattening calculation, getting updated surface S and a new clearance h. As the first step into the fatigue part, each asperity is identified whether it is an asperity coming to contact for the first time. If it is a new contacting asperity, surface cracks C_s and subsurface cracks C_{sub} are assigned to the corresponding initiation positions. Otherwise, if an asperity is already in contact, the lengths of cracks are updated based on the Paris-Erdogan law. As the following step, the algorithm checks for each asperity whether the surface cracks meet with the subsurface cracks. If these cracks meet with each other, the asperity is fractured, resulting in a new surface topography. This new surface is then used as the input surface of the next cycle. After a specific number of sliding cycles, this fatigue wear algorithm provides us a simulated surfaces roughness after wear.
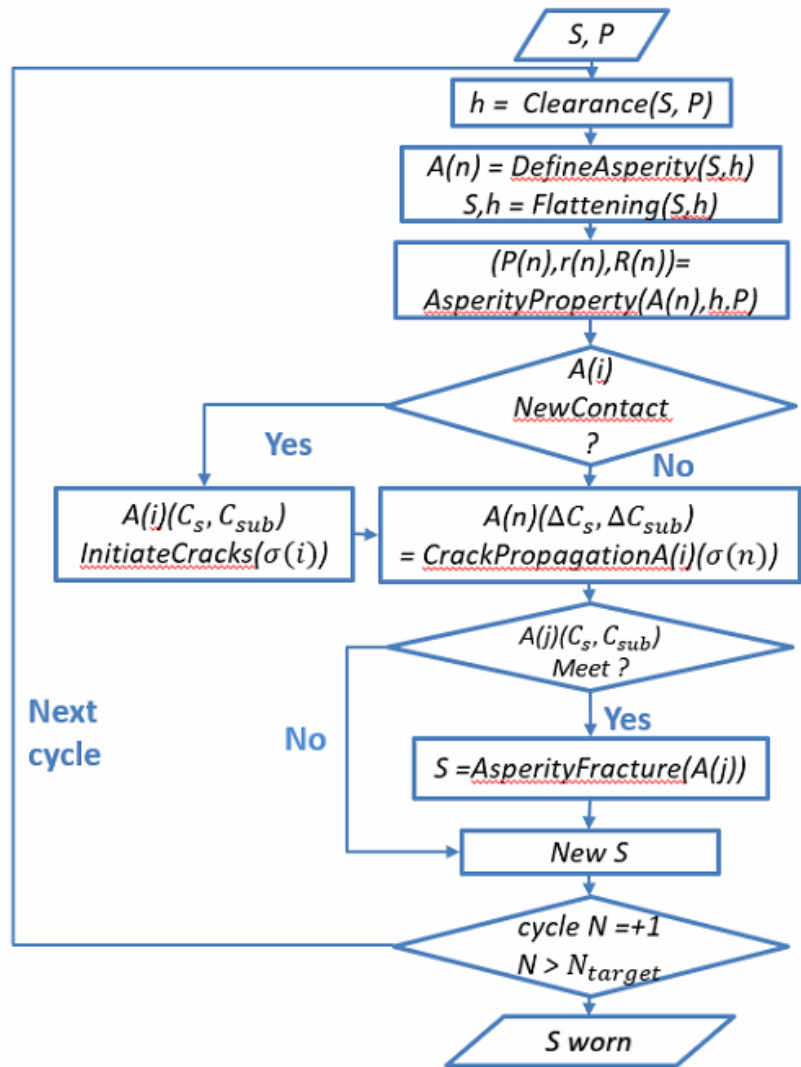
Figure 3.5. algorithm of the asperity fatigue wear model

S	liner surface
P	nominal load
h	clearance between two surfaces
$A(n)$	set of all asperities in contact
$P(n), r(n), R(n)$	set of asperity properties for all contact asperities, average pressure, contact radius and asperity radius of curvature
$A(i)$	set of asperities coming into contact for the first time in this cycle
$A(j)$	set of asperities which are already in contact with the ring
$A(f)$	set of fracture asperities
$\sigma(n)$	set of stress fields
C_s, C_{sub}	surface crack and subsurface crack
$\Delta C_s, \Delta C_{sub}$	propagated lengths for surface crack and subsurface crack
N	cycle number

Table 3.2 variable notations in the algorithm of the asperity fatigue wear

3.3 Fatigue wear results and the friction change during break-in

Based on the sub-models and algorithms in the previous section, liner surfaces with different roughness honing processes are used for simulation of liner wear in the break-in period. Results reveal that the wear rate of each surface is high at the beginning, followed by a constant wear rate. The constant wear rate obtained through simulation agrees with the empirical Archard's law of wear. The material criteria of Archard's law during the break-in period of also discussed. With plotting complete Stribeck curves of different liner surfaces, the friction change in the ring-liner system during the break-in is calculated and compared.

3.3.1 Tested surfaces

For purpose of investigating the effect of liner surface roughness on break-in wear, simulations are conducted based on the five different liner surface finishes: GG07, GG09, GG21, GG28 and GG30. The classification of different liner surface finishes is summarized in Figure 3.6 [35,36]. These liner finishes are classified according to surface roughness parameters: plateau roughness R_{pk}, core roughness R_k and valley roughness R_{vk}. Rougher ones are located relatively on the right of the figure as they have larger total roughness. Therefore, GG07 has the largest roughness among all the five finishes. Furthermore, if several finishes locate colinearly, such as GG21, GG28 and GG30, they are manufactured by the same tools for both rough and smooth honing while fine honing process happen at different heights. The three surfaces have the same plateau roughness and the same valley structure with various plateau ratios.
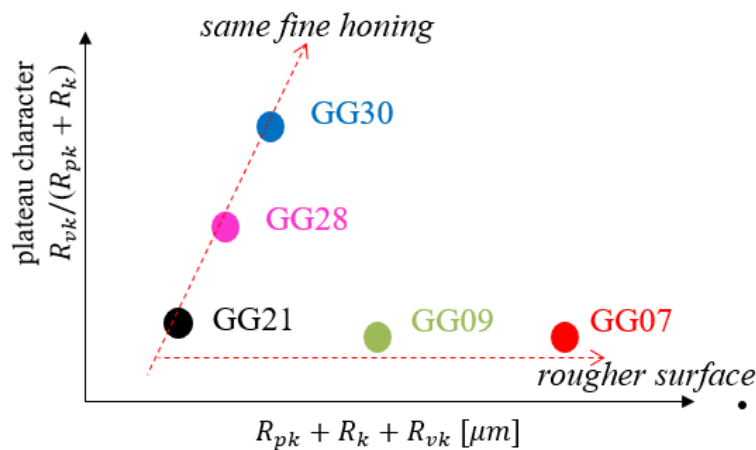
Figure 3.6 different liner surface finishes

3.3.2 Results of liner wear during break-in period

In this section, liner wear break-in simulation results are presented, based on the asperity fatigue wear model. Liner surfaces with different finishes are tested to investigate the impact of

roughness, denoted as GG07, GG09, GG21, GG28 and GG30, shown in Figure 4 with the Daimler liner finish classification [36]. The surfaces used here as the inputs of the wear model are generated numerically based on the real optical measurements [37]. For each surface, external nominal pressure varies from 1 MPa to 4 MPa in order to study the effect of pressure on fatigue wear as well.

According to the results of liner wear simulation, wear rate of the liner is relatively large at the beginning, but gradually reduces to a steady value for all different nominal stresses. Moreover, steady-state wear rate is proportional to normal pressure, as indicated in Figure 3.7. The vertical axis is the steady-state wear rate, which is evaluated by the average surface height reduction over each thousand running cycles. For wear of metallic materials, Archard's law is a widely used correlation which quantitatively describes sliding wear and external parameters: the volume of wear debris for metal sliding is proportional to the product of normal load and sliding distance, but inversely proportional to the material hardness [38,39].

$$V = k\frac{PL}{3H}$$

In the equation above, V is the volume of wear debris, and P, L, H are external normal pressure, sliding distance, material hardness, respectively. k is the wear coefficient, which is a constant for a certain material with a specific surface finish. As one important mechanism in metal wear, asperity fatigue wear obeys Archard's law. Therefore, it is reasonable and essential to calculate the wear coefficient of surfaces with different finishes, which is conventionally used as a parameter to evaluate wear resistance of a surface.
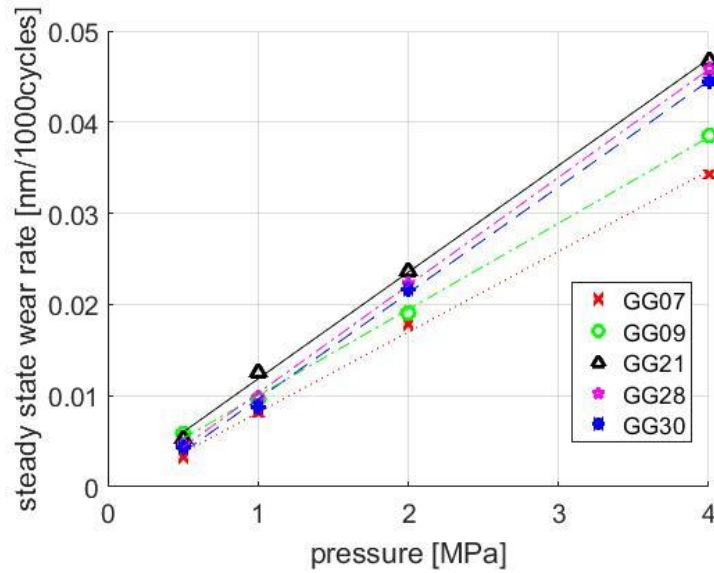
Figure 3.7 steady-state wear rate is proportional.

Surface finish	GG07	GG09	GG21	GG28	GG30
Wear coefficient ($\times 10^{-8}$)	1.14	1.22	1.51	1.53	1.50

Table 3.3 steady-state wear coefficients of different liners

Corresponding to the constant slope in Figure 3.7, the wear coefficient based on The Archard's equation is presented in Table 3.3. It is essential to point out that the wear coefficients of GG21, GG28 and GG30 liner finishes have approximately the same value, indicating steady-state wear rate of these surfaces are very close. The reason may be found from their manufacturing procedures. As illustrated in section 2.5, the only difference among GG21, GG28 and GG30 finishes is the height of fine finish. Although the plateau ratios are different for these three surfaces due to fine finish, plateau regions have the same roughness level. Additionally, the size and shape of contact asperities are statistically the same for the three finishes. Therefore, it is reasonable that their wear coefficients are so close. Furthermore, the results here imply that

steady-state asperity fatigue wear rate largely depends on the roughness level of the plateau region, resulted from fine honing.

3.3.3 More investigations of Archard's law of break-in wear

The steady-state break-in wear rate obtained through the wear simulation above is proportional to the external pressure, resulting a good agreement with the Archard's law. The Archard's law of wear is a widely used empirical correlation, but the generalizability of applying this to all break-in wear problems still remains unclear. In this section, we perform an order of magnitude estimation to investigate the criteria making Archard's law a valid correlation of liner break-in wear.
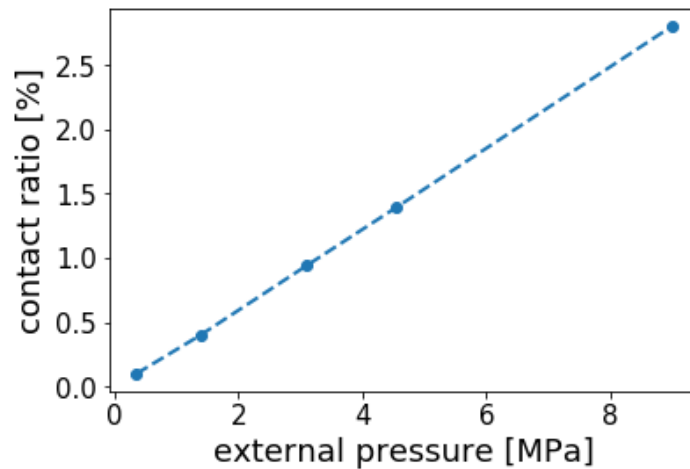
Figure 3.8 approximate linearity between the external pressure and the real surface contacting ratio, GG30

According to the results of the contact model in section 3.2.1, the real contacting area on the liner surface is approximately proportional to the external pressure exerted by the ring. As an example, Figure 3.8 shows an approximate linearity correlation of GG30.

$$P \propto A_{contact}$$

Here P is the external pressure, $A_{contact}$ is the real contacting area under the given pressure P. The wear rate in terms of volume loss can be derived as follows (a_i is the real contacting radius for one asperity).

The fractured volume:

$$V_{fracture} \propto \sum a_i{}^3$$

After plastic flattening, each contacting asperity is under elastic contact condition:

$$P_{avg} = \frac{4E}{3\pi} \left(\frac{h}{R}\right)^{1/2} \propto a_i$$

The stress intensity factor:

$$\Delta k \propto \propto a_i{}^{3/2}$$

The crack propagation speed:

$$\frac{dL}{dN} \propto C\Delta k_{equ}{}^m \propto a_i{}^{3m/2}$$

Crack length leading to fracture:

$$L \propto a_i$$

Time (number of cycles) leading to fracture of one asperity:

$$n \propto \frac{L}{dL/dN} \propto a_i{}^{1-3m/2}$$

Fatigue wear rate of the entire surface:

$$w_{surface} \propto w_{asperity} A_{contact} \propto P \sum a_i{}^{2+3m/2}$$

In the derived equations above, P is the external pressure. Through doing first-order analysis and neglecting higher order terms, the asperity fatigue wear of the entire surface $w_{surface}$ is approximately proportional to $P \sum a_i{}^{2+3m/2}$. Therefore, the criterion making Archard's law valid in asperity fatigue defends on the value of the material properties m, the power coefficient in

Paris law and the asperity size a_i. For most of the metallic materials, m is usually between 2 to 6, with some extreme cases as high as 8. This means the $\sum a_i^{2+3m/2}$ is a high order summation of different asperity sizes, dominated by the large asperities on the liner surface. For the liner plateau honing, asperities generated on the plateau are not homogenous with some asperities significantly larger than others. Therefore, the wear rate of the entire surface is approximately proportional to pressure P as the magnitude of the rest part is mainly determined by the largest asperities.

Currently, calculation of wear often relies on the empirical Archard's law, with a wear coefficient obtained from experiments as a predefined parameter. In this section, we show that the validation of such an approach for wear estimation of asperity fatigue depends on the material fatigue coefficient and the asperity size distribution.
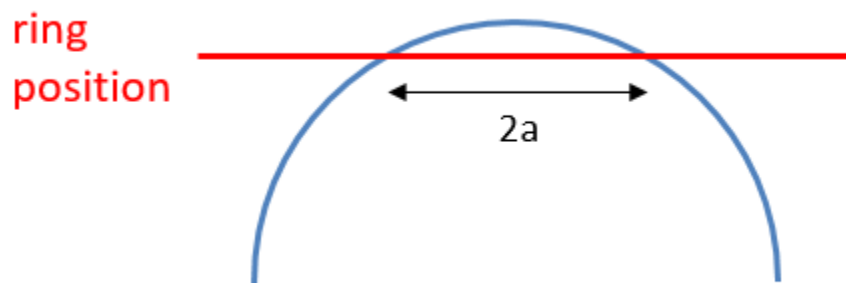
Figure 3.9 contacting condition of each asperity

3.3.4 Hydrodynamic, contact and friction correlations

In order to investigate the effect of break-in fatigue wear on engine lubrication, liner surface finishes before and after wear simulation are utilized to calculate friction. As the piston ring slides to different positions on the liner, the mean clearance also changes, leading to the change of hydrodynamic pressure. To obtain the correlations between the ring-liner clearance and hydrodynamic pressure, a deterministic model is applied. According to this model, the dependency

of the average hydrodynamic pressure on the ring-liner clearance can be expressed by the following equation [2]:

$$P_{hydro} = \frac{\mu V}{\mu_0 V_0} P_h (\frac{h}{\sigma_p})^{-K_h}$$

Here, μ_0 and V_0 are the reference viscosity and reference sliding speed, while μ and V are the corresponding real parameters. P_h and K_h are constants. h/σ_p is the ratio of average clearance and the standard deviation of plateau roughness.

Similarly, the hydrodynamic shear stress can be correlated with the ratio h/σ_p through the following analytical expression with C_i (i = 1, 2, 3) as constants [2]:

$$f_{hydro} = \frac{\mu V}{h} \left(C_1 + C_2 \exp \left(C_3 \frac{h}{\sigma_p} \right) \right)$$

The total pressure is the sum of contact pressure and hydrodynamic pressure. The contact pressure also depends on clearance, which can be determined by the contact model introduced in Section 3.2.1. So as to integrate the contact correlation into the friction model of the entire engine cycle, an analytical expression is also developed:

$$P_c = K'E'A \left(\Omega - \frac{h}{\sigma_p} \right)^Z$$

Here K', A, Ω and Z are constants for a specific surface. E' is the equivalent elastic modulus of the two contact surface, which is computed as:

$$E' = \frac{2}{\frac{1 - v_1^2}{E_1} + \frac{1 - v_2^2}{E_2}}$$

where v_1, v_2, E_1 and E_2 are the Poisson ratio and elastic modulus of the two contact materials.

The correlations of hydrodynamic pressure, hydrodynamic shear stress and contact pressure introduced before are utilized to obtain the friction curves for the entire engine cycle. As the

ring-liner clearance depends on the ring sliding speed, the friction also varies for each engine cycle, with the minimum friction value occurs at the mid stroke at low engine speed. Since the sliding speed is close to zero at the top dead center (TDC) and the bottom dead center (BDC), different engine speeds have similar friction magnitudes at the two positions. When the clearance reaches this smallest value, both contact pressure and hydrodynamic pressure do not have obvious change after wear. The simulated contact pressure and hydrodynamic pressure of GG30 surface finish are plotted in Figure 3.10 as functions of the ring-liner clearance. The wear simulation of GG30 is done by the asperity fatigue wear model, with simulated duration of five hours, engine speed 400 rpm and external nominal pressure 2MPa. Compared with the original surface, the surface after wear simulation has approximately the same hydrodynamic pressure, but the contact pressure is reduced because the asperities are flattened and gradually fractured. The hydrodynamic friction force is plotted in Figure 3.11, with fitted curves of the hydrodynamic shear force equation. Although the hydrodynamic shear stress increases after wear, the total friction between the ring and the liner usually decreases because of the loss of contact pressure.

This can be further revealed in Figure 3.12, which contains the height distribution curves for the liner surface before and after wear simulation. The clearance here is defined as the distance between the ring surface and the peak of the plateau. With a more flattened plateau, contact pressure reduces after wear. When the ring-liner clearance is small, the contact pressure changes little because only high asperities are fractured after 5 hours' break-in wear. Small clearance exists for low sliding conditions, especially when the piston moves to the top dead center or the bottom dead center. Range1, range2 and range3 correspond to the ring-liner clearance range when the engine speed is 100 rpm, 500 rpm and 1000 rpm, respectively.
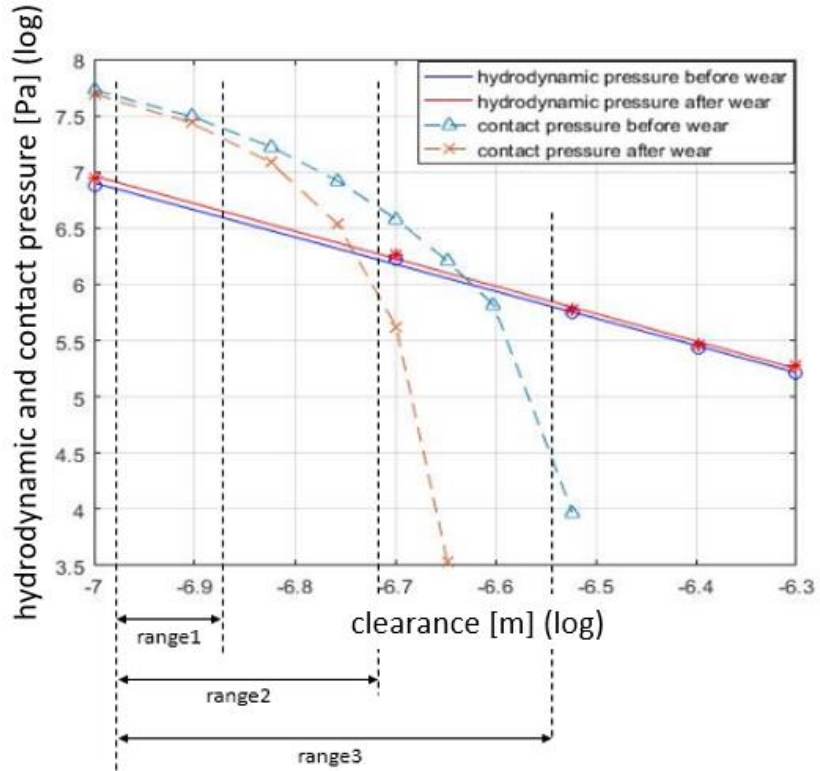
Figure 3.10 hydrodynamic pressures and contact pressures of GG30, before wear and after 5 hours running under external load 2MPa
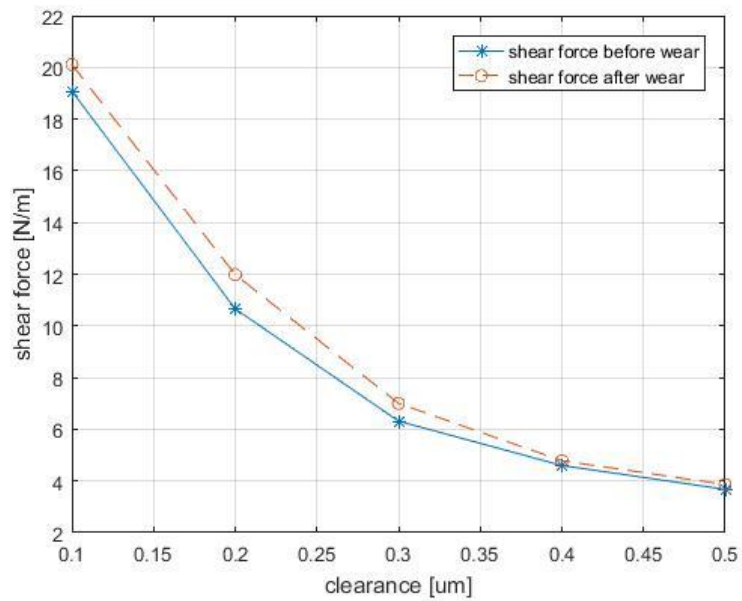
Figure 3.11 normalized shear stress of GG30, before wear and after 5 hours running under external load 2MPa
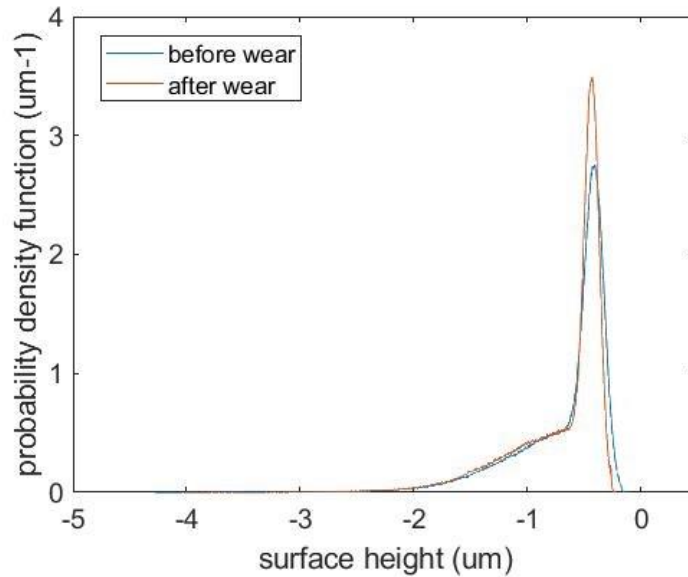
Figure 3.12 liner surface height distributions before and after wear simulation

3.3.5 Friction evolution during the break-in period

The evolution of engine lubrication is studied and compared with experimental results through the Stribeck curve. The theoretical calculation of friction is performed by both the model of hydrodynamic pressure and the model of dry contact. The wear process of different liner finishes is simulated by the asperity fatigue wear model, with friction curves plotted for surfaces before wear and after wear.

The Stribeck curve reflecting the friction coefficient as a function of sliding speed is plotted in Figure 3.13. When sliding speed is low (boundary lubrication), friction is mainly caused by solid contact. As the sliding speed increases, hydrodynamic friction gradually does more significant contribution so that it goes into mixed friction. If the speed keeps increasing, two surfaces are separated further and hydrodynamic force is the only friction source. because of the increase of shear rate, friction increases again. For GG07, because the surface is rougher, it needs larger

sliding velocity to enter hydrodynamic lubrication condition, which is beyond the range of engine speed. Unlike GG30, the contact pressure of GG07 is always playing the major role regardless of the piston position, and this situation does not change after wear. Two factors contribute to the difference with GG30. First, the GG07 surface finish is rougher than GG30, leading to a stronger contact correlation. Second, the average asperity size of GG07 is larger compared with GG30, resulting lower asperity fatigue wear rate as mentioned in the previous section. Therefore, the surface topology change of GG07 is also smaller, resulting that the friction drop of GG07 after wear is not as remarkable as GG30. The experimental measurements also support this difference between GG07 and GG30.

Based on the Stribeck curves, engine friction level at different running speeds can also be expected. For instance, when the engine speed is low, the contact pressure has a slight drop. This leads to the friction reduction after wear at the mid stroke, while contact pressure is always the dominant one, much larger than the hydrodynamic pressure. When the engine speed reaches a high level, contact pressure is much lower than hydrodynamic pressure at the mid stroke, for both conditions before and after wear. Compared with the friction measurements obtained experimentally, the trend of the relative importance between hydrodynamic pressure and contact pressure agrees well.

One important mismatch between simulation and experiment is the magnitude of friction drop of GG30. Especially for low engine speed, the friction drop in the mid stroke region is larger than the predicted result. As hydrodynamic friction is not dominant at low engine speed, this indicates the predicted contact pressure drop is smaller than experimental measurements. Therefore, the break-in wear rate is also underestimated. One source for the underestimation of fatigue wear rate comes from the assumption of small yielding for calculating crack propagation. However, it is not always necessarily that the plastic regions around cracks are small compared with crack lengths. With the extended plastic regions, cracks can propagate faster than predicted in this fatigue model, leading to higher wear rate and more dramatic friction drop. The other possible error source is the assumption that asperity fatigue wear is uniform over the entire liner surface while in reality asperity contact is different along the stroke during break-in.

(a)

(b)

101
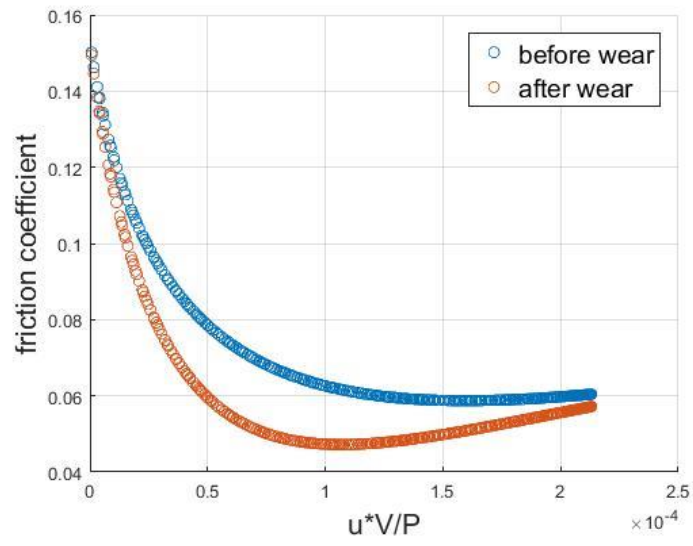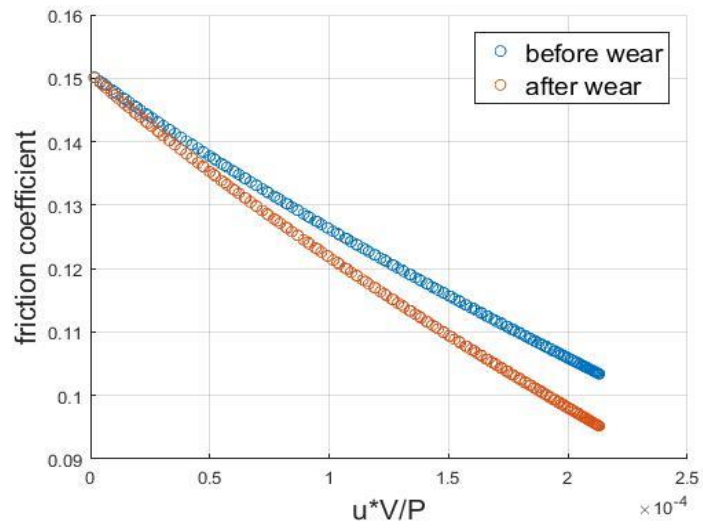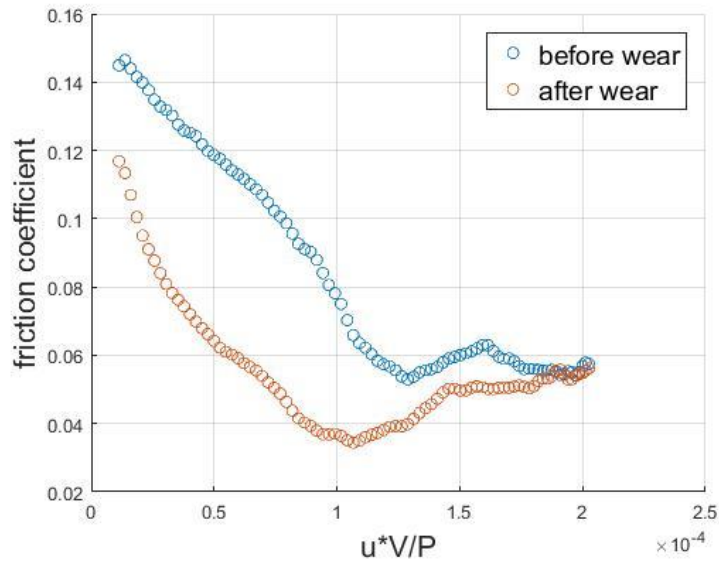
(c)
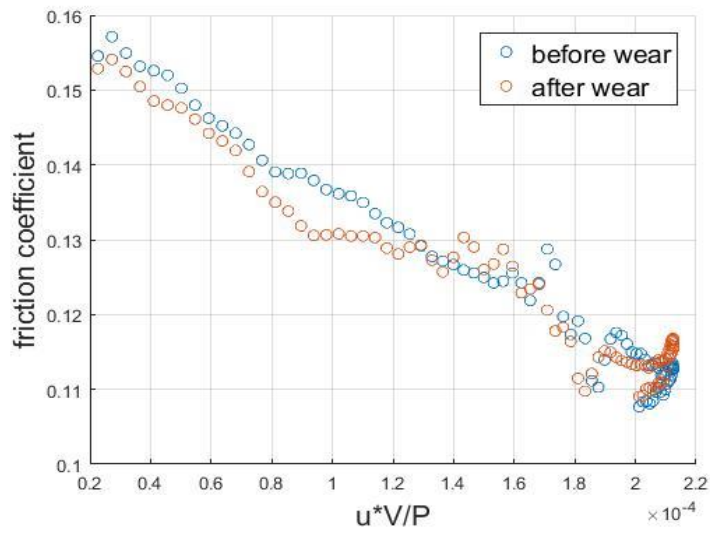
(d)

Figure 3.13 Stribeck curves for experimental measurements and simulation results. (a) simulation results, GG30 (b) simulation results, GG07 (c) experimental measurements, GG30 (d) experimental measurements, GG07

3.4 Conclusion

In this chapter, asperity fatigue liner wear model during break-in period in internal combustion engines is established. Two mechanisms leading to liner topography change has been studied: asperity plastic deformation and fatigue. Simulation results agree with Archard's wear law that for a specific liner surface finish, the steady-state wear rate is proportional to external nominal load. The criterion making the Archard's law valid during the break-in is also derived. Simulations indicate that the maximum steady-state wear rate can be obtained for a specific surface roughness level. This could be applied in the future for liner surface roughness design to minimize the break-in time.

The engine friction change due to break-in liner wear calculated by the presented wear model shows the same trend as the experimental measurements. The liner finish with smoother plateau region exhibits faster break-in than the rougher liner finish. Moreover, the evolution of the friction predicted is comparable to the measurement in magnitude, implying the potential of the model in predicting duration of the break-in period and asymptotic friction. With the assumption of small yielding, fatigue wear rate is underestimated compared with experimental measurements.

Chapter 4. Deterministic modeling of the piston rings and the liner wear during the steady state period

4.1 General introduction and background

After the engine break-in period, two significant changes occur in the system of the piston ring and the cylinder liner. The dominant wear mechanism switches from asperity fatigue to abrasive wear. Meanwhile, the friction level becomes stable, with further friction reduction stopped. The transition from break-in to steady-state wear is caused by the flattening, fracture of the original asperities on the liner surface and the third-body particles in the aged oil. In addition to third-body particles, the ring surface also creates two-body abrasive wear on the liner because the ring is gradually getting rougher. For the ring-liner sliding pair, the ring is the moving part, resulting a direct contact with the liner throughout the engine running period. Contrarily, any location on the liner only contacts with the ring twice within one engine revolution, with no contact with the ring for most of the time. As a consequence, the piston ring experiences longer duration of contact stress than anywhere on the liner and thus its wear needs to be studied together with the liner although it is remarkably harder than the liner material.

As the liner is significantly softer than the ring, it can hardly damage the ring surface by itself through direct contact. Instead, third-body particles play the major role in damaging the piston rings. The third-body particles in the engine system can be from the combustion outcome as carbon soot particles, or the wear debris. Considering the solid particles are mainly carbon soot particles and the hardness of carbon is greater than that of worn particles and the coatings on the ring, we will first focus on situation with the soot as the only third-body particles, which is the simplest situation that both liner and ring wear can be generated through abrasion process. As illustrated in Figure 4.1, the modeled three components in the steady-state engine wear system are: the ring, soot particles and the liner. With the assumption of their relative hardness, liner wear in the steady state is caused by the abrasion of both third-body particles and the roughened ring, while the damage of the ring is only caused by the abrasion of the particles. This is the overall assumption made in this project about the modeling of engine steady-state wear.
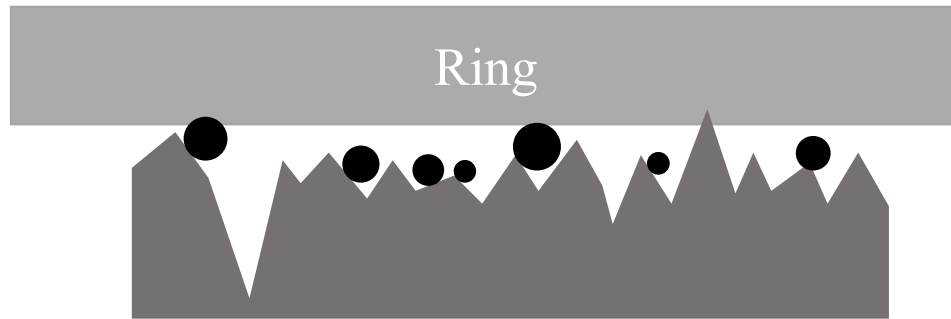
hardness comparison:
$$H_{soot} \geq H_{ring} \gg H_{liner}$$

Figure 4.1. The three components of the wear system during steady-state engine condition. The relative hardness of the three components: third-body particles are harder than the ring; the particles and the ring is significantly harder than the liner.

The nontrivial effect of soot particles is also confirmed experiments. The experiments are conducted through real engine testing, with controlling the size of the third-body particles to investigate the influence of particles on wear. A particle filter is applied to the engine system so that third-body particles with sizes larger than a certain threshold cannot enter the ring-liner system and cause further wear on the surfaces of the ring or the liner. The severity during the steady-state is evaluated from the ring side, with measuring the width of the wear track. In addition, the cross section of the ring after wear is also plotted to intuitively compare the wear volume of each test [7]. Figure 4.2 illustrates the significant influence of the third-body particles on ring wear. Without a particle filter, all soot particles will participate in the third-body wear process, resulting a wider wear track on the ring surface. If the particle filter is added, the width of the wear track becomes narrower, reducing from more than 2000 μm to less than 1000 μm. In terms of wear volume, the cross sections of the rings indicate an even larger reduction. In (c) and (d) of Figure 4.2, the wear volume, as the lost material, is greatly reduced if the particle filter is added. From this set of real engine experiments, it is important to focus the modeling work on the effect of third-body particles, which plays the key role in engine long term running condition and design for sustainability.
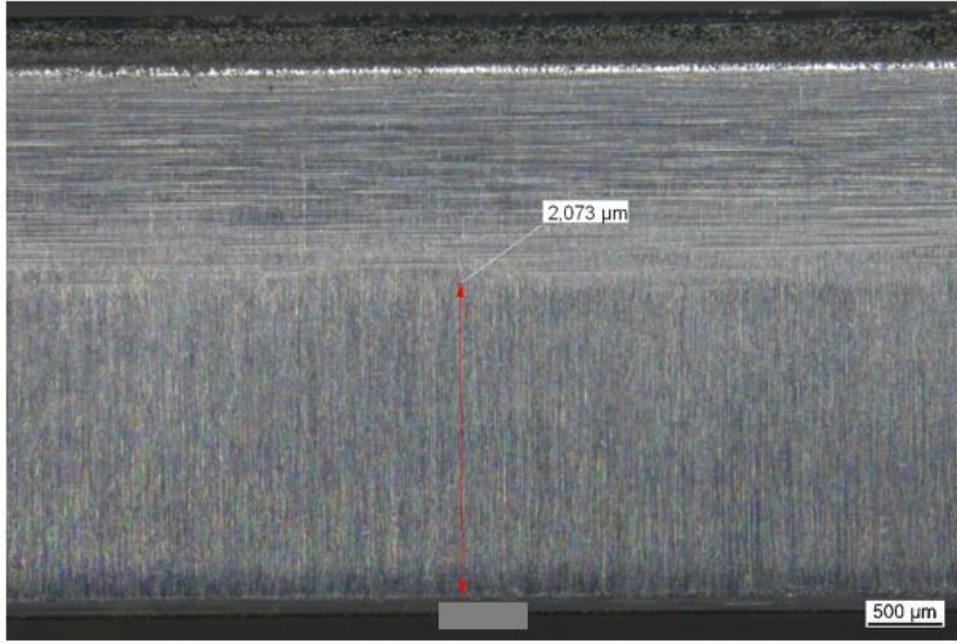
(a)
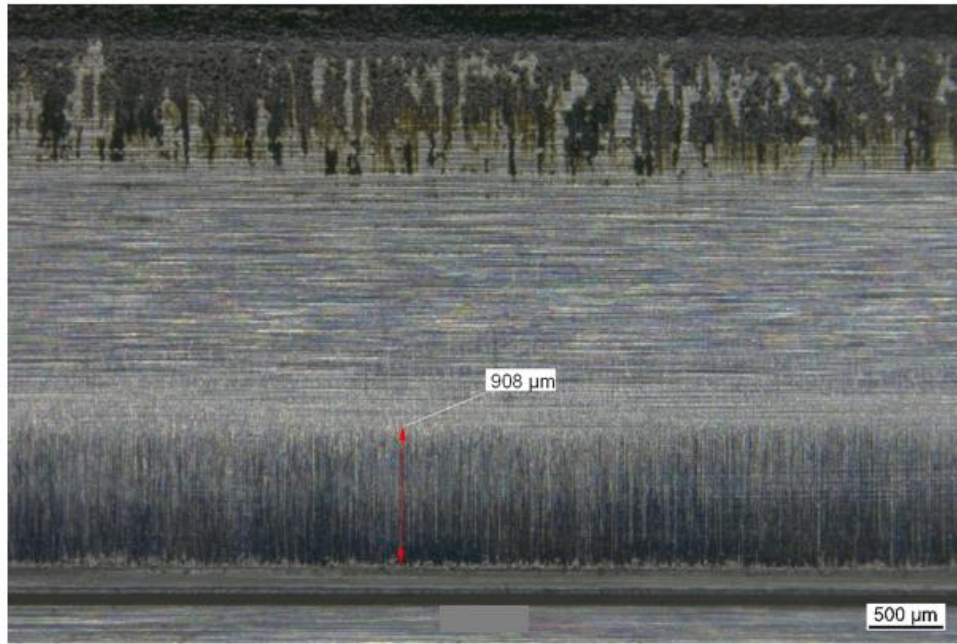
(b)
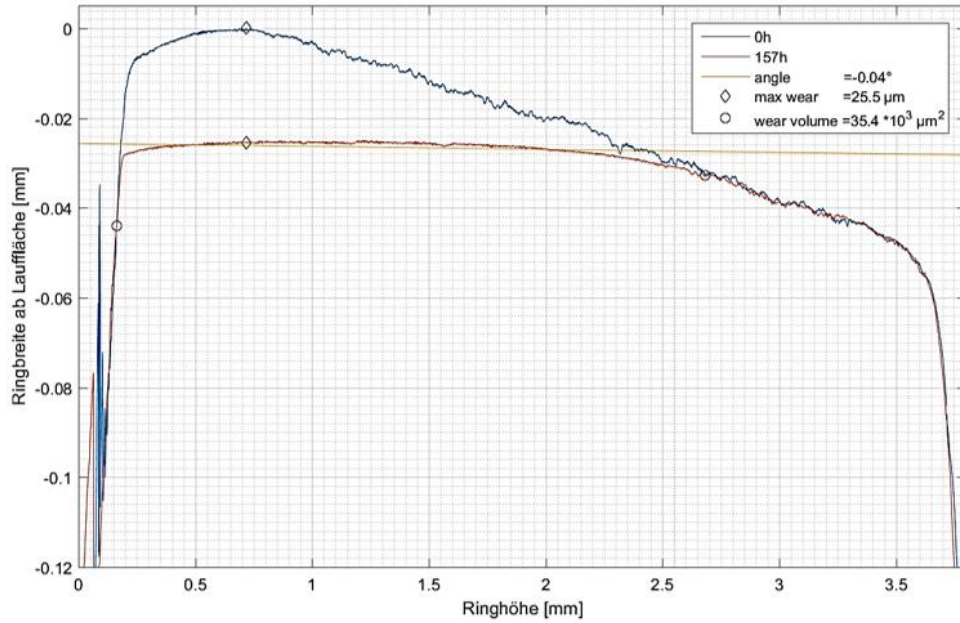
(c)
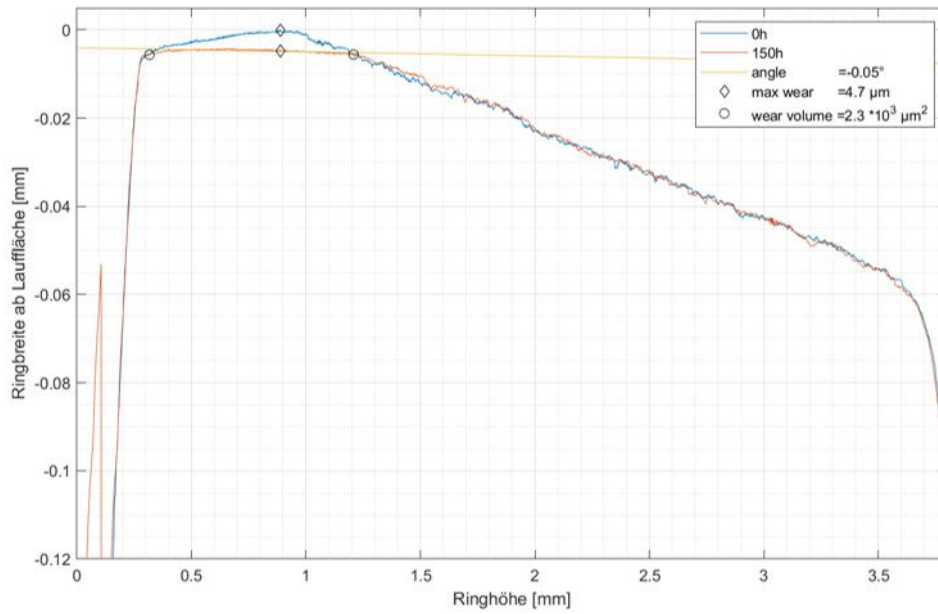
(d)

Figure 4.2. steady-state wear results with different soot particle levels. (a) the wear track on the ring without a particle filter. (b) the wear track on the ring with a particle filter. (c) the ring cross section without a particle filter. (d) the ring cross section with a particle filter [7].

On the liner side, the mechanisms become more complicated. As the softest and the most vulnerable part, the liner faces the abrasive wear from the particles and the ring. However, the wear loss on the liner surface is not uniformly distributed from the TDC to the BDC. Different areas on the liner surface experience different levels of abrasive wear as the load, clearance, lubrication condition vary at different positions. Engine experiments also show the difference of wear severity.
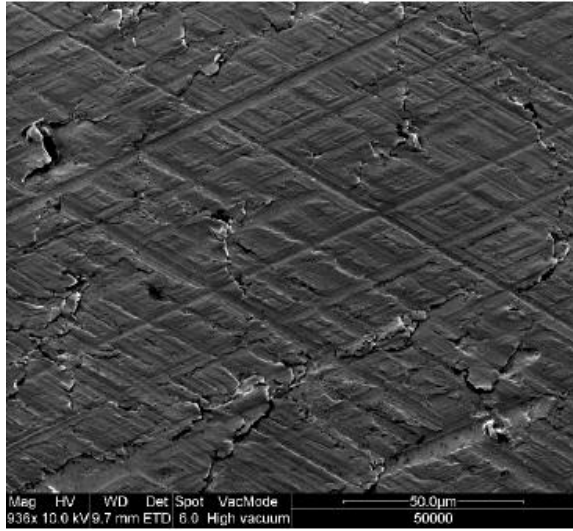
(a)
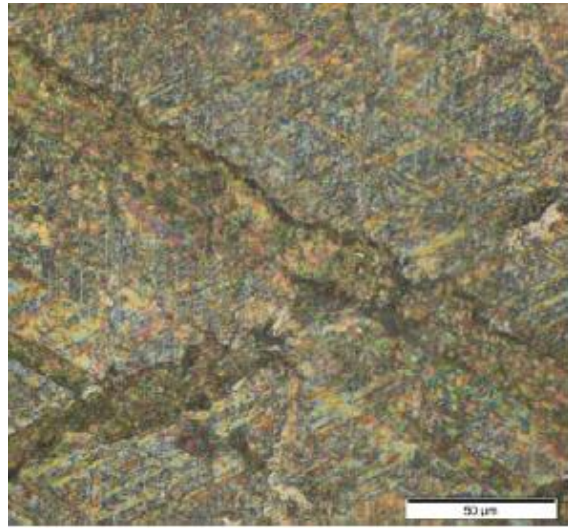
(b)
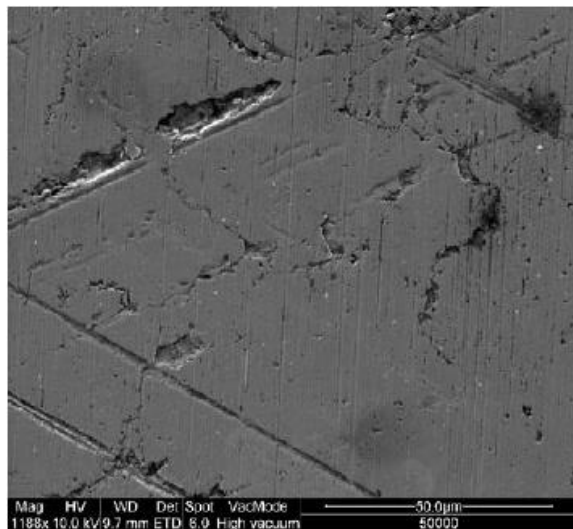
(c)
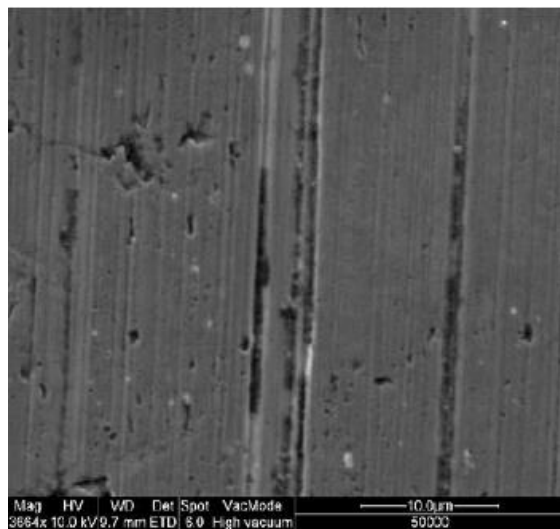
(d)

Figure 4.3. The wear conditions on the liner surface. The pictures were taken from a heavy-duty diesel engine after 80000 km driving (a) the liner surface before the test. (b) The TDC area after the wear test. (c) The BDC area after the wear test. (d) A zoomed picture of BDC [40].

When the ring slides from the TDC to the BDC, the normal load between the two surfaces is always changing. In mid stroke, the contact load is relatively low and the hydrodynamic pressure becomes significant. In such a case, wear does not happen in a fast pace because of the large clearance filled with lubricants. When the ring moves at the TDC or the BDC, the contact pressure becomes dominant, providing a potential to create severe wear. Furthermore, the conditions of the TDC and the BDC are different as well. As shown in Figure 4.3, the original liner honing structure at the TDC still exists due largely to the existence of tribo-film, meaning that the material loss due to wear is limited. However, the honing structure around the BDC position of the oil control ring is almost completely removed and there was no sign of tribo-film. The lubrication function of the original liner surface is damaged. This indicates more severe wear happens at the BDC, making it the critical area for the engine durability design. Although both the TDC and the BDC experience high contact load, a layer of tribo-film is formed in the TDC area to protect the liner from further damaging. This layer is formed based on the additives in the oil. Therefore, in this project, wear of the BDC area is considered and modeled, where most of the normal force is supported by the solid-solid contact with no tribo-film formation. The dominant wear mechanism is also reflected in the experiments. As shown in (d) of Figure 4.3, vertical scratches, scars of abrasive wear, along the sliding direction are generated. In this work, we focus on the oil control ring (OCR) and liner wear at the BDC where there is plenty of oil and the soot content in sump oil is representative to the oil here. This pair presents the purest environment for the abrasive wear with 3rd body involvement. On the practical side, wear of the oil control ring is one of the most critical issues in engine development as it directly contributes to the sustainability of the oil consumption control.

In general, the work of this chapter focuses on several aspects of the engine steady-state wear, which is the key problem of the engine durability. The size of third-body particles that do the most damage to the ring-liner system is investigated. Different processes are also studied and modeled so that more understanding of detailed mechanisms can be provided for design.

Moreover, systematic simulations of steady-state wear are conducted in a physics-based deterministic way.

4.2 Elements of steady-state wear modeling

The process of steady-state wear is complicated, involving many different physical and chemical mechanisms. These mechanisms are coupled together to make it even more difficult to model. In this project, only the mechanisms and transport processes that mostly contribute to the surface topography change are modeled. These mechanisms have the most significant influence on the abrasive wear rate and engine durability. Several essential model elements are discussed in this section: third-body particle size distribution, particle movement in lubricant, particle entrapment, particle embedment, particle transport and abrasion. The entire simulation algorithm consists of these parts to perform a systematic wear modeling of the ring-liner system.

4.2.1 Particle size distributions

As the most important abrasion source, carbon soot particles are vital for the system, whose size distribution is a decisive factor. In general, particle size distributions in different systems vary due to different particle formation mechanisms. In internal combustion engines, the sizes of third-body soot particles cover a wide range, from the order of 10 nanometers to several micrometers. Moreover, the distribution of soot particles is highly skewed because of the agglomeration effect. Some particles can grow extremely large, consisting of hundreds of primary carbon particles while the lower bound of soot particle size can never become smaller than zero. The particle size distribution in the engines can be approximately described by the log-normal distribution:

$$f(d) = \frac{1}{\sqrt{2\pi\sigma^2}}(\frac{1}{d})e^{-\frac{(lnd-\mu)^2}{2\sigma^2}}$$

In the probability density function above, d is the particle diameter. μ and σ are the two parameters originated from the standard normal distribution, but they are now the mean and the standard deviation of the log-normal distribution.

The approximate log-normal distribution of engine soot particles is also observed in engine tests. As plotted in Figure 4.4, the frequencies of different particle sizes form a bell shaped curve. Considering the horizontal axis, the particle diameter, is in logarithmic scale, the overall size distributions are close to the set of log-normal distributions, even under various of engine running conditions of load and speed.
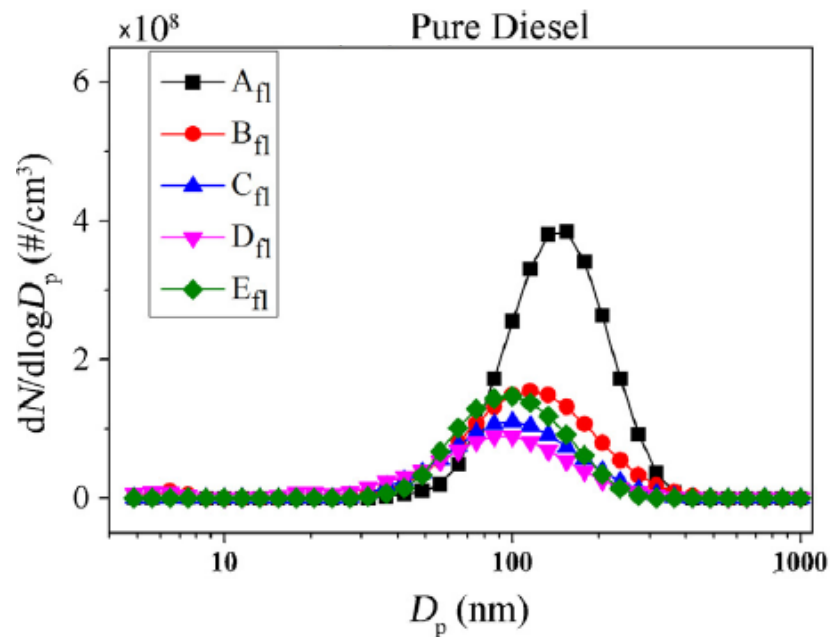
Figure 4.4. soot particle size distribution from real engine tests. Particles are collected from the engine exhaust. D_p is the particle diameter. A_{fl} to E_{fl} indicate different engine running conditions [41].

It is important to point out the underlying mechanism leading to the log-normal soot particle size distribution. Based on the central limit theorem, the summation of a sequence of independent random variables is normally distributed. As a generalization, a log-normal distribution is formed through multiplying many independent random variables. In engines, the soot particles are agglomerates of primary carbon particles, leading to the random influencing factors multiplicative to each other. However, the size distribution of primary carbon particles is not

112

necessarily log-normal distribution. Instead, the size of primary carbon particles is closer to a normal distribution, as illustrated in Figure 4.5 [42].

In this project, log-normal distributions are used as the third-body particle size distributions, because almost all the particles collected in engine exhaust and aged oil are agglomerated particles. The particles are modeled as hard and spherical balls to simplify the calculation.
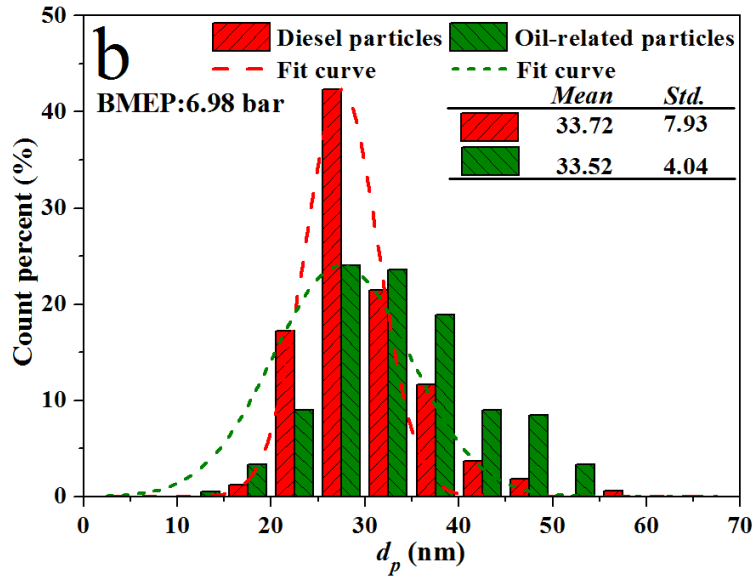
Figure 4.5. the particle size distribution of primary soot particles [42]

4.2.2 Particle movement in lubricant

If a particle is squeezed between the ring and the liner, its movement pattern will be determined by the solid to solid contact. However, before a particle gets into the clearance, it would flow with the lubricant. The Reynolds number can be estimated with a typical third-body size of 500 nm:

$$R_e = \frac{\rho U D}{\mu} \sim \frac{0.8 \times \frac{10^3 kg}{m^3} \times \frac{1m}{s} \times 5 \times 10^{-7} m}{0.01 Pa.s}$$

$$\sim 0.04 \ll 1$$

113

With a small Reynolds number, the flow can be assumed to Stokes flow for the particles. Under the Stokes flow assumption, third-body particles will simply follow the flow streamlines. Therefore, the flow condition needs to be solved in order to provide the movement of particles in the lubricant.

An existing model developed by Chen [2] is able to provide the horizontal lubricant flow field, in the sliding direction and the circumferential direction. However, the vertical direction, perpendicular to the two surfaces, is still missing. In this section, the vertical velocity profile is derived in order to obtain a complete velocity field.
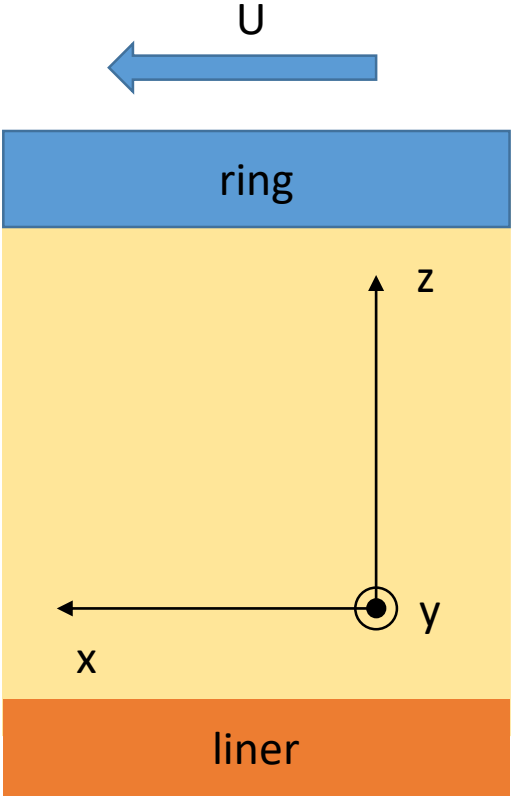
Figure 4.6. the coordinate system of the lubricant flow field between the ring and the liner

With a moving ring and pressure gradient, the flow between the ring and the liner is a mixture of Couette flow and Poiseuille flow. As indicated in Figure 4.6, the z direction is the direction perpendicular to the surfaces while the x direction is the sliding direction. The governing equations of Couette flow and Poiseuille flow are the following three:

$$u = \frac{1}{2\mu}\left(-\frac{\partial P}{\partial x}\right)z(h-z) + U\frac{z}{h}$$

$$v = \frac{1}{2\mu}\left(-\frac{\partial P}{\partial y}\right)z(h-z)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0$$

Where u, v and w denote the lubricant velocities in x, y and z directions, respectively. P is the pressure while U is the ring sliding speed. h denotes the local clearance between the two surfaces. The first two equations are the velocity equations, with the third one as the continuity equation. Based on the governing equations, the vertical velocity w should have the following format:

$$\frac{\partial w}{\partial z} = C_1(x,\ y)z + C_2(x,\ y)z^2$$

Where

$$C_1(x,\ y) = \left[\frac{1}{2\mu}\left(\frac{\partial P}{\partial x}\frac{\partial h}{\partial x} + \frac{\partial^2 P}{\partial x^2}h\right) + \frac{U}{h^2}\frac{\partial h}{\partial x}\right] + \frac{1}{2\mu}\left(\frac{\partial P}{\partial y}\frac{\partial h}{\partial y} + \frac{\partial^2 P}{\partial y^2}h\right)$$

$$C_2(x,\ y) = -\frac{1}{2\mu}\left(\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2}\right)$$

The boundary conditions are:

$$w = 0 \quad \text{at } z = 0$$

$$w = 0 \quad \text{at } z = h$$

Theoretically, solving the vertical velocity field does not require two boundary conditions. However, in this problem, the two boundary conditions are not independent. The existing model calculating the horizontal velocities already includes the mass conservation in the two directions, making the two boundary conditions above coupled through the continuity equation.

Based on the derivation above, the full velocity profile of the lubricant at any liner location can be obtained. The calculation results show that the velocity in the vertical direction is significantly smaller than the velocities in the other two perpendicular directions. For instance, Figure 4.7 contains the velocity profiles of two different locations with different local clearances. The velocity along the sliding direction is always the largest, but the velocity component w is extremely small. Therefore, it is reasonable to neglect w in the flow field. Particles are mainly flowing with the ring sliding direction. The two examples are calculated based on the following physics properties:

Ring sliding velocity $= 0.1 \mathrm{m/s}$

Lubricant dynamic viscosity $= 0.01 \mathrm{Pa. s}$

Other properties, such as the local pressure, pressure gradient and velocity v, are obtained through the existing two-dimensional flow model.
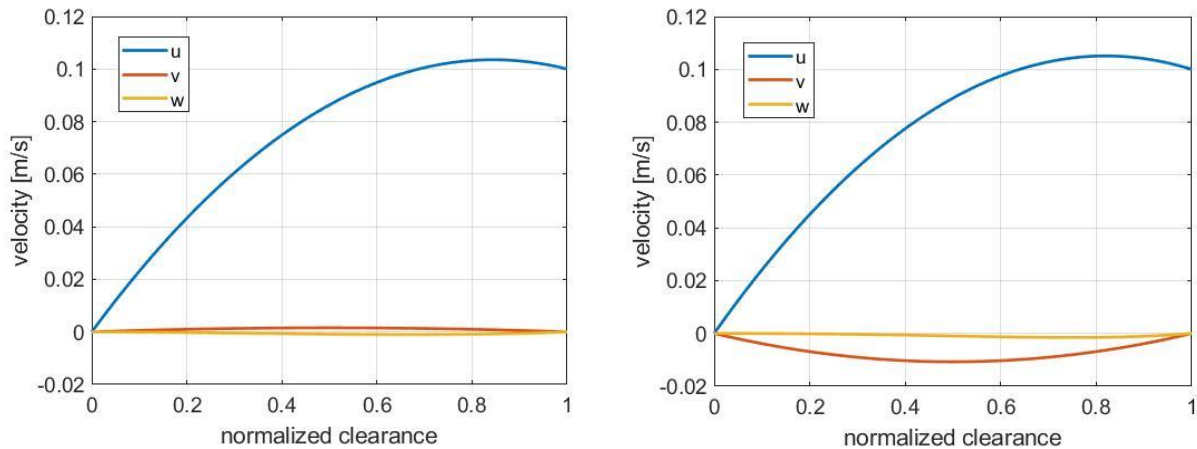
Figure 4.7. the velocity fields of the lubricant at two different locations on the liner surface

With all the analysis in this section, the particle movement in the lubricant can be simplified through two assumptions in the modeling work. First, because the velocity of the lubricant is dominated by the component in the ring sliding direction, the velocities in the vertical and circumferential directions are neglected. Second, once a third-body particle is in the nominal contact region between the liner and the ring but not stuck in the local clearances, it will follow

116

the lubricant flow in the sliding direction with the streamline at the center of the clearance. These two assumptions can make the calculation efficient without losing modeling accuracy.

4.2.3 Particle entrapment

If a particle is floating with the lubricant, it will not create any wear or surface topography change, simply because there is no direct contact. Once a particle moves to the position where the local clearance is equal to the particle diameter, it is possible for it to get trapped between the two surfaces and create damages on the ring and the liner. However, the particle is also possible to be rejected by that clearance, being pushed back to the lubricant again. Therefore, it is essential to find a criterion that a particle will further get trapped between the surfaces once it arrives the clearance entrance. As illustrated in Figure 4.8, it is discussed in this section, whether a particle would be squeezed in the clearance or return to the floating condition in the lubricant.
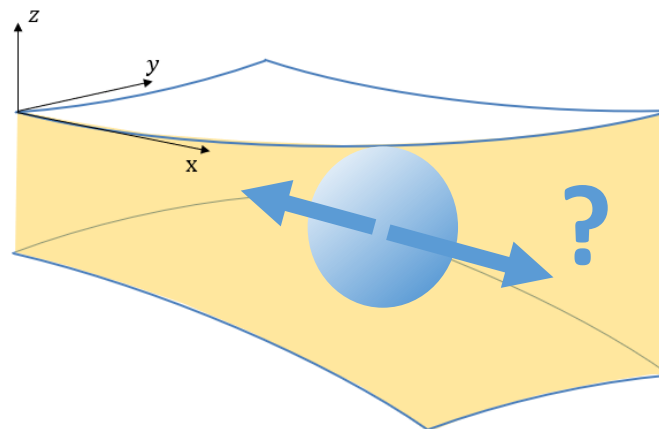
Figure 4.8. the entrapment model for a particle at the entrance of a local clearance

As proposed by Nikas [ref], the entrapment status of a particle depends on the three kinds of forces exerting on the particle: fluid forces, normal forces from the two surfaces and the corresponding friction forces. As shown in Figure 4.9, the moving ring generates a normal force N_1 and a tangential friction force T_1, which tries to drag the particle further into the clearance. As a response, the liner also generates a normal force and a friction force to counterbalance the forces exerting on the particle by the sliding ring. Meanwhile, the particle is under the influence

of the force from the lubricant as well. The of the fluid forces in three different directions can be calculated as [43]:

$$F_x^{(f)} \cong -3\pi\eta \frac{U_1}{2} d - \pi \frac{\pi + 4}{16} \left(\frac{\partial P}{\partial x}\right) d^3$$

$$F_y^{(f)} \cong -\pi \frac{\pi + 4}{16} \left(\frac{\partial P}{\partial y}\right) d^3$$

$$F_z^{(f)} \cong -3\pi\eta d \left(\frac{\partial z_1}{\partial x} \frac{U_1}{\sqrt{1 + (\partial z_1/\partial x)^2}}\right)$$

$F_x^{(f)}, F_y^{(f)}, F_z^{(f)}$ are the fluid forces in three directions. η is the dynamic viscosity of the lubricant and P is the local fluid pressure. d is the particle diameter. $\frac{\partial z_1}{\partial x}$ denotes the local gradient of the moving surface. U_1 is the sliding speed of the piston ring.
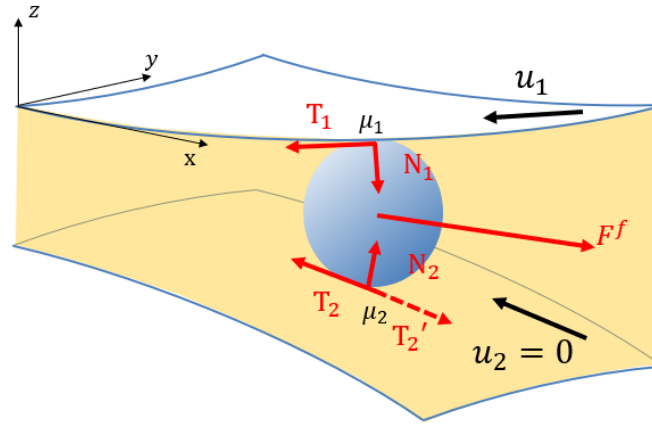
Figure 4.9. forces exerting on a particle at the clearance entrance

With the obtained fluid forces, equations of force balance can be derived. It is assumed here that the particle can only move along the sliding direction, or opposite to the ring moving direction if the particle is pushed back to the lubricant. Consequently, the force balance in the other two directions can be established [43]:

$$N_{1y} + T_{1y} + N_{2y} + T_{2y} + F_y^{(f)} = 0$$

$$N_{1z} + T_{1z} + N_{2z} + T_{2z} + F_z^{(f)} = 0$$

$$|T_1| = \mu_1 |N_1|$$

$$|T_2| = \mu_2 |N_2|$$

Where N_{1y}, T_{1y}, N_{2y} and T_{2y} are the y direction component of the corresponding forces. The same notation is also applied to the z direction. μ_1 and μ_2 are the friction coefficients of the ring and the liner surfaces. Thus, the normal forces and the tangential forces exerted by the two surfaces on the particle can be calculated through solving the set of force balancing equations. As the last step to determine whether or not the particle will further get trapped between the local clearance, the resultant force in the sliding direction is calculated:

$$F_x^{(resultant)} = N_{1x} + T_{1x} + N_{2x} + T_{2x} + F_x^{(f)}$$

The status of the particle is determined by the signs of the three calculated forces: the two normal forces N_1, N_2 and the resultant force $F_x^{(resultant)}$. If either N_1 or N_2 is negative, the particle will leave the local clearance back to the lubricant because N_1 and N_2 must be positive. If a valid force equilibrium is established, $F_x^{(resultant)}$ becomes the determining factor. A negative $F_x^{(resultant)}$ means the resultant force points into the clearance, forcing the particle getting squeezed between the ring and the liner. Thus, the particle gets trapped. Otherwise, with a positive $F_x^{(resultant)}$, the particle simply leaves the entrance of the clearance back to the lubricant.

So far, the analysis is almost adequate as the particle movement is dominated in the sliding direction. However, in order to include the possibility of a particle to escape the clearance from other directions rather than the sliding direction, a generalized version of particle entrapment is also proposed. The main logic remains the same, with one improvement. As illustrated in Figure 4.10, the tangential force T_2 is not strict to the sliding direction. Instead, it can point to any direction perpendicular to N_2. The particle will get trapped only if the particle fails to leave the clearance with all possible T_2 directions. For simplicity, the calculation can be completed with a rotating coordinate (x', y', z), in which the z axis is the rotating axis. The simulation is completed through enumerating all possible directions of T_2. In each case, T_2 and x' are always

in the same vertical plane. Within the set of force balancing equations, only the equilibrium equation of y' needs to be modified:

$$N_{1y'} + T_{1y'} + N_{2y'} + T_{2y'} + F_{y'}^{(f)} = 0$$

The balancing equation is now established in the y' direction. After obtaining the forces, the criterion of particle entrapment also has a modification. For any direction of T_2, if N_1, N_2 are all positive and the resultant force $F_{x'}^{(resultant)}$ points into the direction with a negative clearance gradient, the particle gets trapped. Otherwise, the particle will simply find one direction for T_2 and leave the clearance.
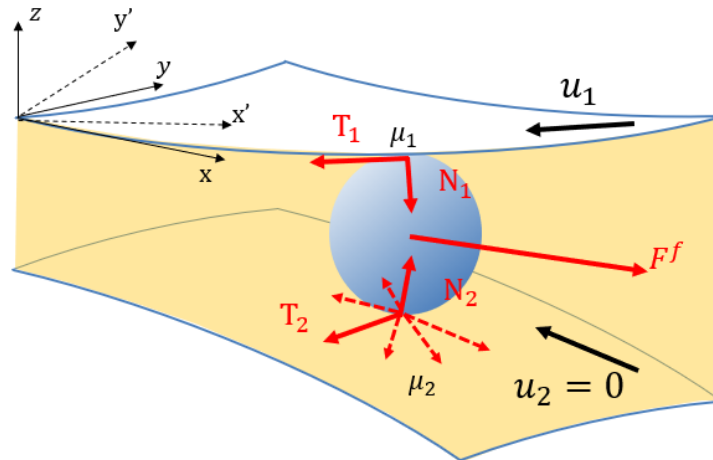
Figure 4.10. forces exerting on a particle at the clearance entrance, with all possible particle escaping directions

Simulations are performed based on the refined entrapment model to investigate important properties of the entrapment process. In the setup of the simulations, 1000 particles are randomly assigned to different locations on the liner surface, with reducing clearances in the ring sliding direction. With a fixed ring height, the diameter of each particle is set to be the same as the local clearance so that the particles are exactly at the clearance entrances. The ring surface is assumed smooth in this case.

In Figure 4.11, the influences of the ring sliding velocity and the friction coefficient are plotted. The trap ratio is defined as the proportion of trapped particles among all the simulated particles at the clearance entrances. In this simulation, the friction coefficients of the two surfaces are always equal, which is also close to the friction condition in a ring-liner system. With a fixed friction coefficient, the trap ratio is not influenced obviously by the ring sliding speed. However, the magnitude of the friction coefficient plays an important role in particle entrapment. When the sliding velocity is fixed to 1m/s, the number of trapped particles significantly increases as the friction coefficient is increasing. When the friction coefficient reaches 0.3, most of the particles get trapped into the local clearances. The results show a great influence of the friction coefficient, but the effect of the ring sliding speed is trivial.

(a) (b)
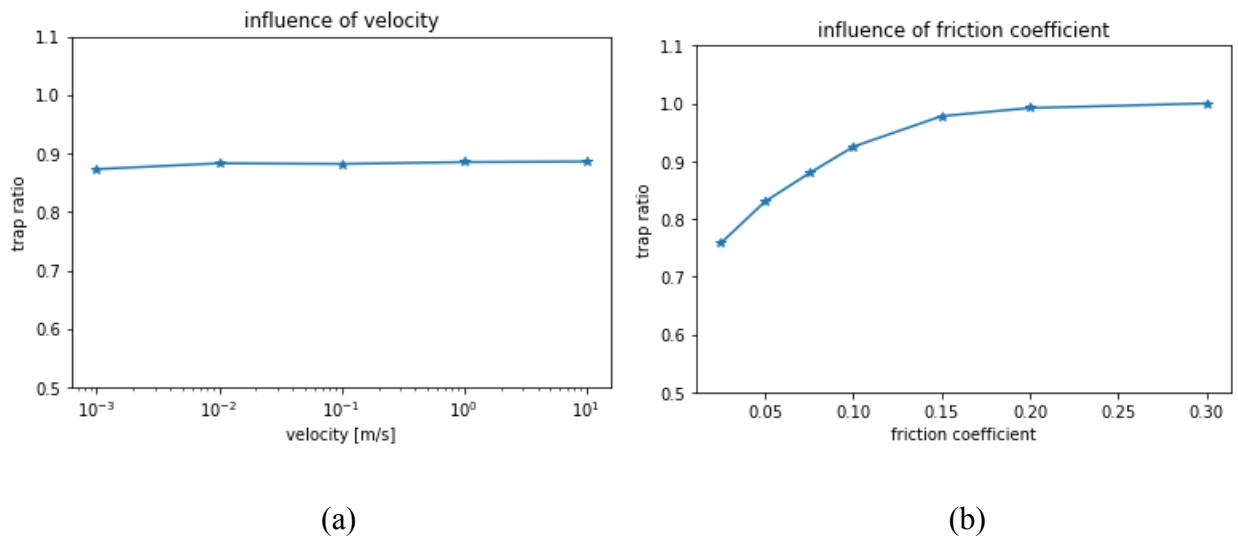
Figure 4.11. The influences of the ring sliding velocity and the friction coefficient. (a) the effect of the sliding velocity (b) the effect of the friction coefficient

It is also interesting to investigate the status of each particle and find more influencing factors. With a constant ring sliding speed 1m/s and a fixed friction coefficient value 0.1, the trapped particles and the free particles are separated into two groups. For each group, histograms are plotted with the magnitude of the local clearances as the bin axis. Plotted in (a) and (b) of Figure 4.12, nontrapped particles are distributed more in the range of large local clearance gradient compared to the trapped particles. If further decomposition of the local clearance is conducted, it is apparent to observe that the gradient in the sliding direction (x direction) plays the dominant

role in (c) and (d). The trapped particles all locate in the positions with local x direction gradient smaller than 0.2. This threshold is consistent with the friction coefficient value of 0.1 for both surfaces. If a local clearance with gradient larger than the sum of the two coefficients of the two surfaces, the entrance will become too steep for the particle being pushed further into the clearance. On the other hand, larger friction coefficient can help the particle overcome the slope of clearance entrance. It is also reflected in Figure 4.12 that the clearance gradient in the y direction is not essential. The particle status is mainly determined by the gradient in the ring sliding direction.

(a)

(b)

(c)

(d)

trapped particles — nontrapped particles

(e) (f)
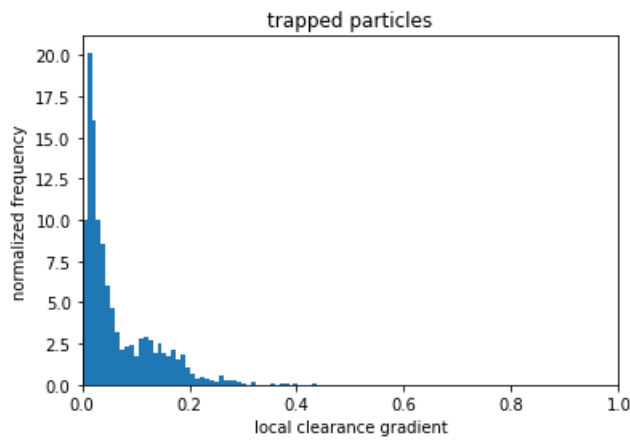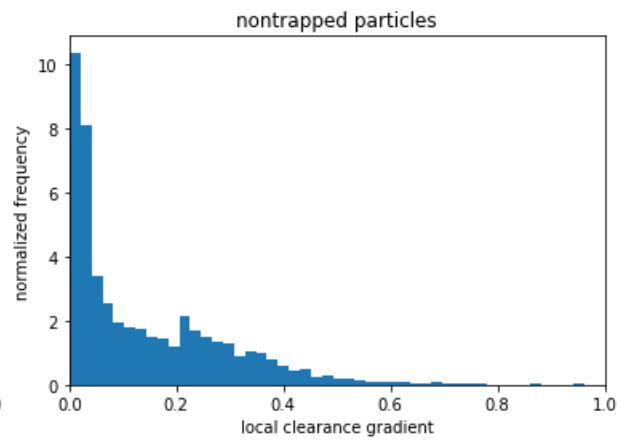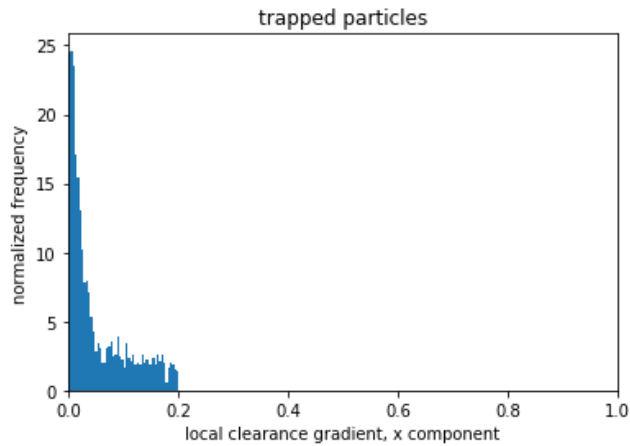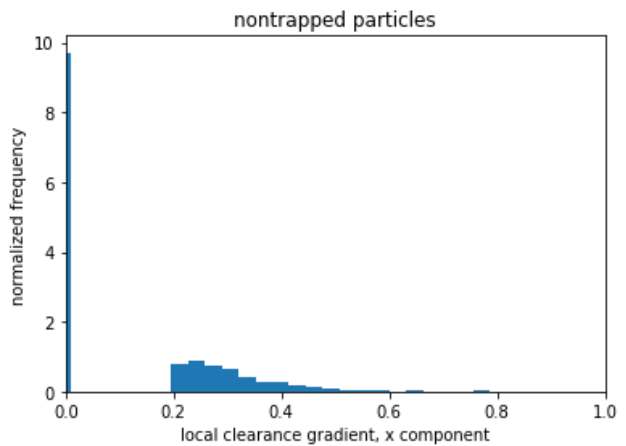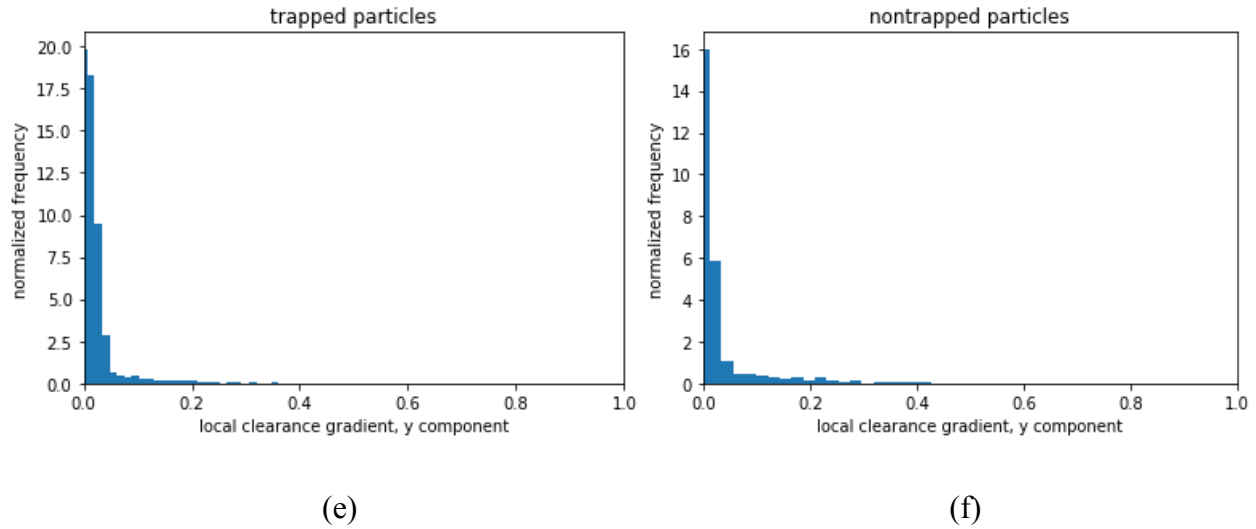
Figure 4.12. trapped and nontrapped particles locating at different clearance gradient positions. (a) the total clearance gradient distribution of trapped particles (b) the total clearance gradient distribution of nontrapped particles (c) the x direction clearance gradient distribution of trapped particles (d) the x direction clearance gradient distribution of nontrapped particles (e) the y direction clearance gradient distribution of trapped particles (f) the y direction clearance gradient distribution of nontrapped particles

In this section, the particle entrapment model is established. Based on simulations, the status of a particle at the clearance entrance largely depends on the friction level and the local geometry, especially the clearance gradient in the sliding direction.

4.2.4 Transport and embedment of trapped particles

Once a particle is getting trapped between the two surfaces, it is possible for the particle to create abrasive wear scars. Meanwhile, which surface will suffer from the third-body abrasion depends on the movement pattern of the trapped particle. For a trapped particle, three movement patterns are possibly to happen: embedding into the ring, embedding into the liner or rotating between the two surfaces. If a particle is embedded into one surface, it will slide with that surface and damage the other. If a particle is rotating between the two surfaces, it can create wear scars on both the ring and the liner.

123

The movement pattern of a trapped particle can be determined by the torques of the normal and tangential forces and the friction coefficient. As shown in Figure 4.13, with a moving ring, the normal forces from the two surfaces are not collinear, with a distance e between. Similarly, there is also a distance h between the two tangential forces. The ratio of e/h can be calculated by the following equations [44]:

$$D_1 = \frac{N}{4\pi R H_1}$$

$$D_2 = \frac{N}{4\pi R H_2}$$

$$e = \frac{4}{3\pi}(r_1 + r_2)$$

$$\frac{e}{h} = \frac{N}{2\pi^2 R^2 H_1} \frac{\varepsilon_1^{1/2} + \varepsilon_2^{1/2}}{\varepsilon_1^{3/2} + \frac{H_2}{H_1}\varepsilon_2^{3/2}}$$

$$\varepsilon_1 = 1 - (1 - \frac{N}{4\pi R^2 H_1})^2$$

$$\varepsilon_2 = 1 - (1 - \frac{N}{4\pi R^2 H_2})^2$$

Where N is the normal force, R is the particle radius. D_i, is the indentation into the surfaces (i=1, 2). r_i is the projected contact radii of the two surfaces, while H_i is the material hardness.
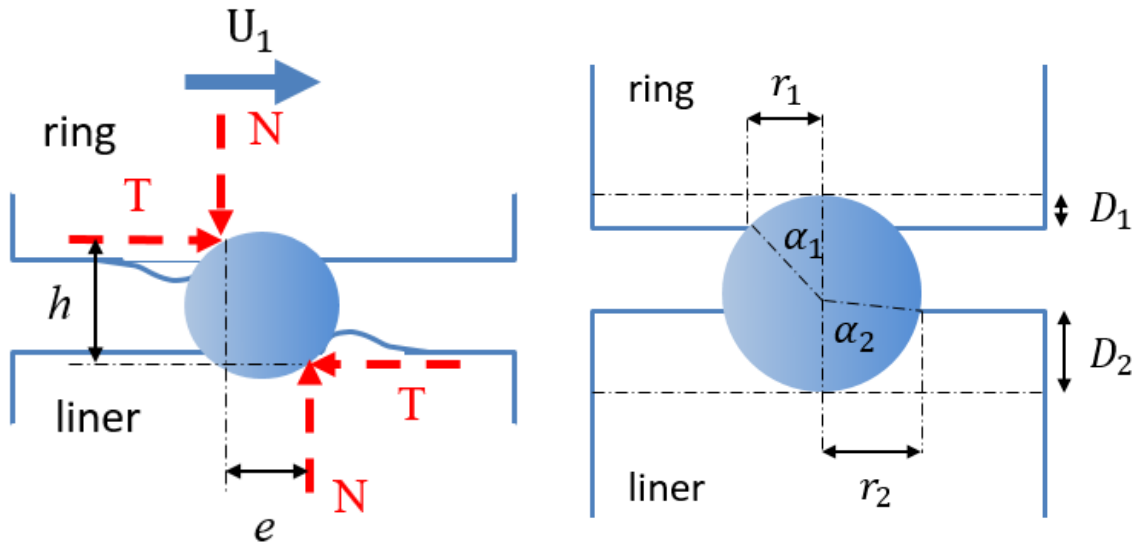
Figure 4.13. the forces and torques exerted on a trapped particle

In the calculation above, the underlying assumption is that the contact is plastic so that the contact pressure reaches the material hardness. Such an assumption is reasonable for the third-body particles in general because the elastic indentation limits of the particles are so small that almost all the trapped ones will contact plastically with the surfaces.

The determining factor of particle rolling or sliding is simply the relative magnitue between the ratio $\frac{e}{h}$ and the friction coefficient μ. If μ is larger than $\frac{e}{h}$, the particle will roll between the two surfaces. Otherwise, it will get embedded into one surface and slide against the other.
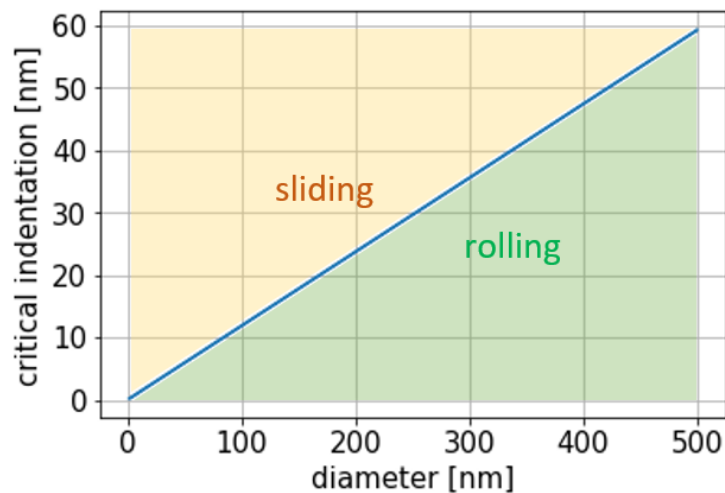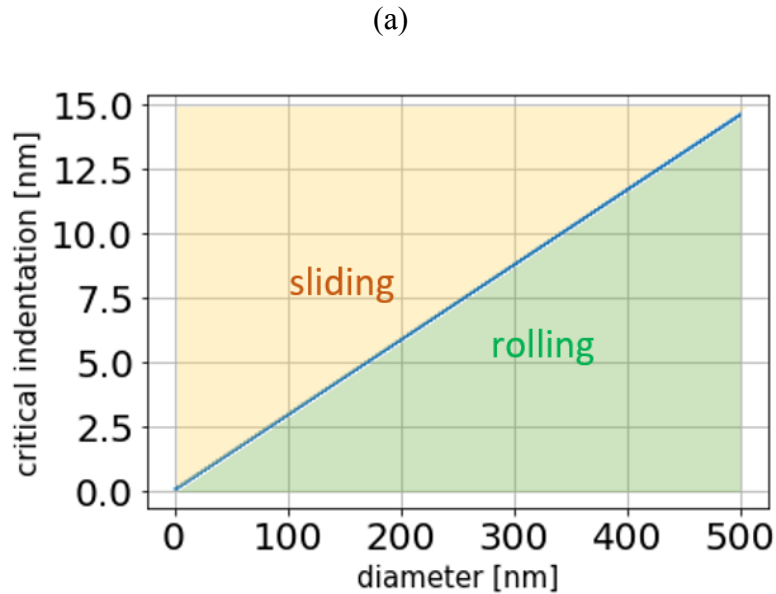
Figure 4.14. the decision boundary of particle rolling and sliding. (a) friction coefficiet 0.2 (b) friction coefficient 0.1

In Figure 4.14, the blue curves illustrate the crtical total indentation ($D_1 + D_2$) of a particle between the two transport modes, rolling and sliding with a varying particle size as the horizontal axis. If the total indentation is less than the critical value, the particle will tend to roll, corresponding to the green area below the critical total indentation curve. Otherwise, the particle will be in the sliding mode, squeezed between the two surfaces. It is also reflected in Figure 4.14 that higher friction coefficient will increase the critical total indentation value, leaving more particles rolling instead of sliding.

When a particle is slding, it always gets embedded into one surface and slides against the other to create wear scars. This is because the maximum tengential force for a surface to resist the particle sliding is limited. As the two surfaces provide different maximum tengential forces, the surface with smaller critical tengential force would lose the resistence to a moving particle. As a result, the particle will be embedded into the other surface. This is also observed in experiments that a sliding particle sticks to one surface generating continous wear scars [44], as shown in Figure 4.15. The width of a third-body wear scar depends on the size and shape of the slding particle.
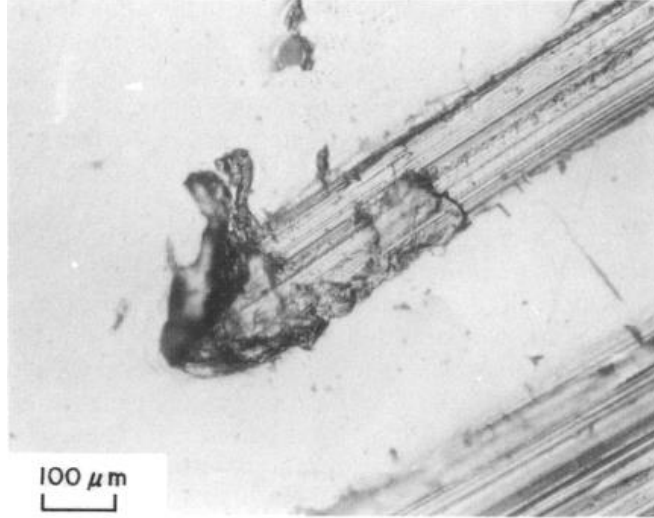
Figure 4.15. a sliding wear scar [44]

The maximum tengential force of one surface is determined by the particle size, indentation depth and the surface friction coefficient. For a spherical particle, the sliding tengential force can be estimated [45]:

$$F_f = \frac{2r^2 H}{\pi} \left[\frac{R^2}{r^2} \sin^{-1} \frac{r}{R} - \sqrt{\frac{R^2}{r^2} - 1} \right] + \mu \frac{4R^2 N}{\pi r^2} \left[1 - \sqrt{1 - \frac{r^2}{R^2}} \right]$$

Where R is the particle diameter, r is the radius of the projected contact area. μ is the surface friction coefficient and F_f is the overall tengential force of a particle to overcome the resistence and generate a path.
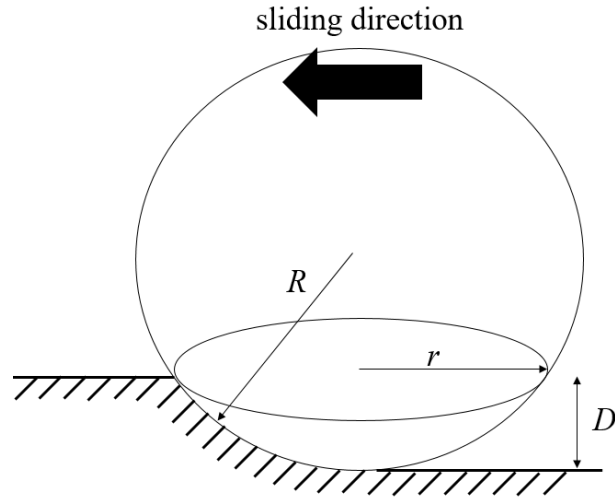
Figure 4.16. the particle sliding model

For the same surface, the critical tengential force increases with the particle indentation depth. Therefore, a particle tends to get embedded into the softer surface as the indentation is larger compared with the harder sliding counterpart. The particles have higher chance to get embedded into the soft surface, leaving the hard one vunerable in this case. However, higher hardness of the ring would help the ring to hold the particles and increase the critical tengential force. As a result, some particles would get embedded in the ring, making the liner wear with a combination of two-body and third-body abrasion.

4.2.5 The abrasion model

When a trapped particle is sliding against one surface, it generates abrasive wear scars. Three abrasion scenarios can happen in this process, depending on the particle indentation depth. As presented in Figure 4.17, the indentation is purely elastic if the indentation depth is small, resulting no permanent deformation on the surface. As the indentation depth increases, permanent plastic deformation occurs, pushing the material on the path to the two sides to form side ridges. In this scenario, plowing effect is dominant with no material volume loss. The material is merely redistributed on the surface to form new grooves with pile-ups. If the

indentation depth increases further, the plowing effect now becomes cutting, with new pile-ups forming at the two sides. The material in the cutting path is detached from the surface forever.

(a)

(b)

(c)
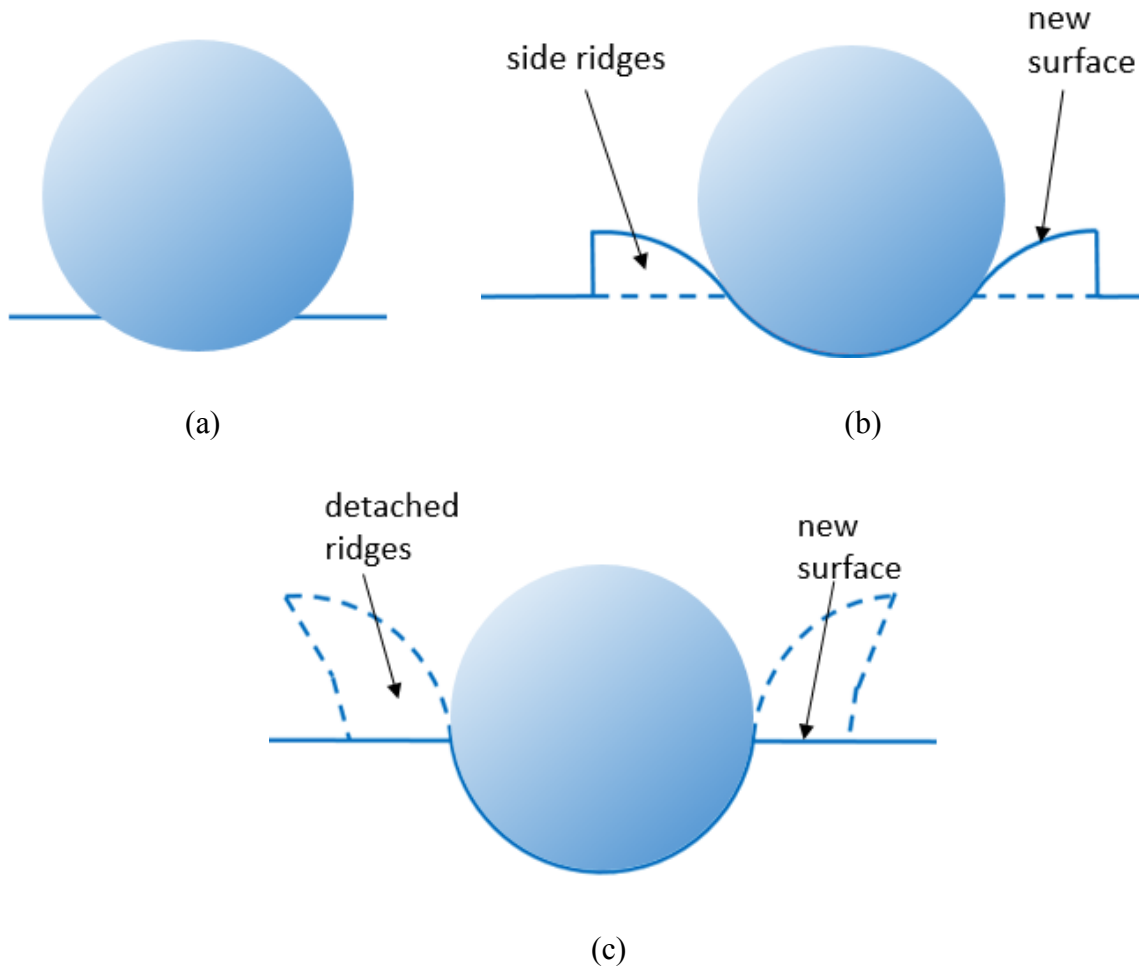
Figure 4.17. three scenarios of a sliding particle. (a) elastic indentation (b) plowing effect with no material loss (c) cutting

For spherical particles, the quantitative criteria for these three sliding scenarios are given based on the indentation geometry [46].

1. elastic indentation:

$$\frac{D}{r} < 1.78(\frac{E}{\sigma_y})^{-1}$$

2. particle plowing:

$$1.78(\frac{E}{\sigma_y})^{-1} \leq \frac{D}{r} < [1 + 0.037(\frac{E}{\sigma_y})]^{-1}$$

3. cutting:

$$\frac{D}{r} \geq [1 + 0.037(\frac{E}{\sigma_y})]^{-1}$$

Where D and r are defined in Figure 4.16. E, σ_y are the elastic modulus and yield strength of the surface.

For a sliding particle, it generates wear scars only on one surface as it is embedded into the other surface. The same abrasion model is also adopted for rolling particles. The difference is that a rolling particle creates wear scars on both surfaces with only half of the ring sliding distance. The total wear track length is the same as a sliding particle, but it is distributed on both the ring and the liner.

4.3 The modeling algorithm

With all the sub-models defined in the previous section, the systematic simulation procedure can be established. The entire algorithm is organized to provide wear modeling for both the ring and the liner during the engine steady-state condition.

The liner roughness matrix, the nominal external pressure and the soot particle concentration are provided as the model inputs. With the given particle concentration and defined particle size distribution, soot particles are randomly generated on the liner surface. For each engine cycle, the ring slides over the liner twice, in two opposite directions. For each sliding process, as the ring moves from one end to the other end on the liner, it will encounter particles on the liner surface. If a particle is at the place where the local clearance is larger than the particle diameter, the particle will flow with the lubricant. When a particle is arriving at the clearance entrance with local clearance smaller than the particle size in the ring sliding direction, the particle entrapment model is applied to determine whether it would be further trapped in the clearance. If a particle is trapped, the particle movement model is used to analyze the movement pattern, either rolling or

sliding. For a rolling particle, it travels half of the ring sliding distance, creating damages on both surfaces. For a sliding particle, the embedded surface is determined by the calculating and comparing the maximum tangential forces. After that, the counterpart surface topography is updated based on the abrasion model. As the ring surface is significantly harder than the liner surface, the two-body wear scars are also generated with each ring sliding process.

With the defined systematic modeling procedure, the wear rate and the surface topography after the steady-state wear can be obtained.

4.4 Systematic simulations

Systematic simulations are performed with the combined algorithm. In this work, it is found that two factors are decisive to steady-state abrasive wear: the particle size distribution and the ring hardness. Both short-term and long-term simulations are conducted. The short-term simulations are used to investigate the effect of the particle size distribution as the effect is obvious just within a short engine simulation period. Long-term simulations are conduced to investigate the effect of the ring properties. Long-term engine experiments are also used for comparison and model improvement.

4.4.1 The effect of particle size distribution

As discussed in Section 4.2.1, the particle size can be approximately described as log-normal distributions. In this section, two sets of simulations are conducted with different particle size distributions. Both distributions are log-normal, as shown in Figure 4.18. Particle size distribution 1 represents a set of small soot particles, without significant agglomeration effect. Most of the particles have diameters smaller than 100 nm, which is the range of primary carbon particles. Size distribution 2 corresponds to a more typical soot size range, with most of the particle diameters in the sub-micrometer range. Some large ones are in the micrometer scale.
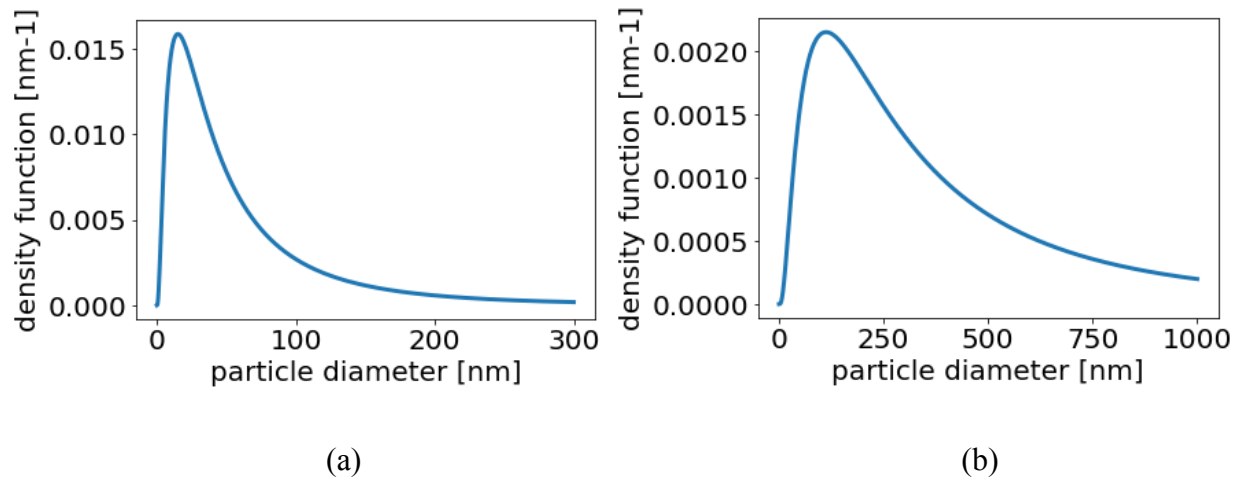
Figure 4.18. The two particle size distributions. (a) particle size distribution 1 (b) particle size distribution 2

Except for different particle size distributions, other simulation conditions keep the same for the two sets. The ring material is steel, without any coating on the surface. The elastic modulus and material hardness of the ring used in the simulation are 200 GPa and 3Gpa, respectively. The nominal contact pressure is 2MPa while the soot particle concentration is 0.1%. The material properties of cast iron are used for the liner, with elastic modulus 120GPa and hardness 915 MPa. The number of engine cycles is limited to 3000 in order to reduce the influence of two-body abrasion before the ring gets too rough due to the third-body abrasive wear.

As seen in Figure 4.19, the change of the liner surface topography is not significant. After 500 cycles, no obvious abrasive scars can be observed. Even after 3000 cycles, only a few scratches are created as a result of the abrasive wear. However, the wear rate increases apparently if the particle size distribution 2 is adopted in the simulation. With only 500 cycles, many scratches already occur on the surface, as shown in Figure 4.20.

The particle size distribution has remarkable influence on the wear rate because it determines the entrapment ratio. If a particle is small, it has lower chance to encounter with a clearance smaller than its diameter, leaving it less likely to damage the surfaces. This is confirmed with a comparison between the particle size distribution and the clearance size distribution, which also provides an insight on the particle size distribution doing the most damage.
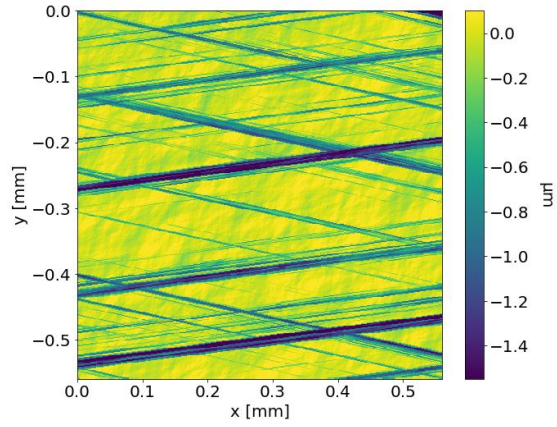
(a)
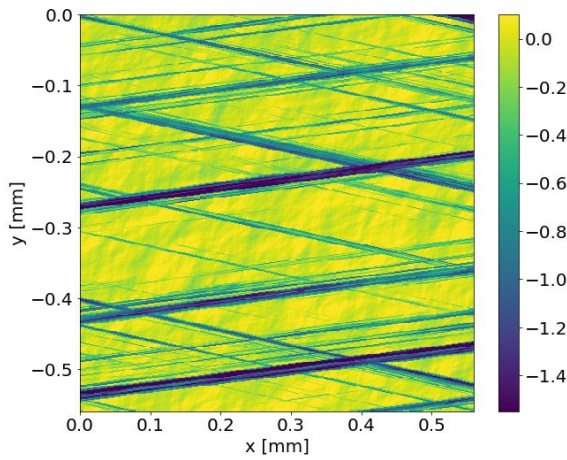
(b)
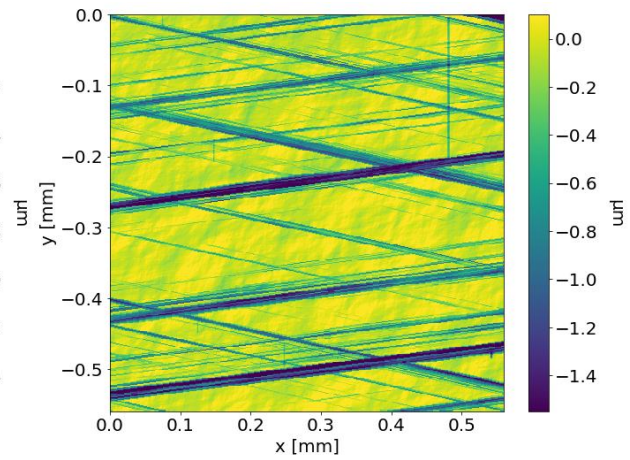
(c)

Figure 4.19. the liner topography change with particles following size distribution 1. (a) before wear (b) after 500 cycles (c) after 3000 cycles

(a)

(b)   (c)
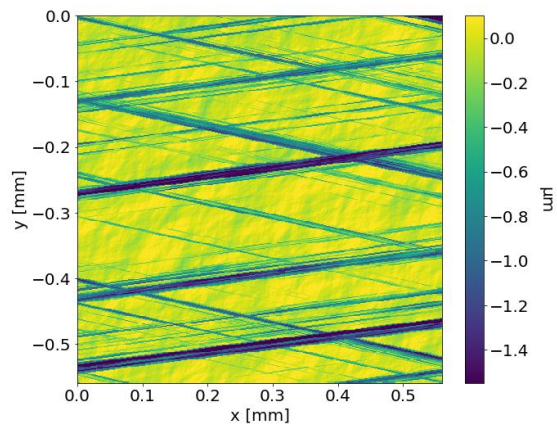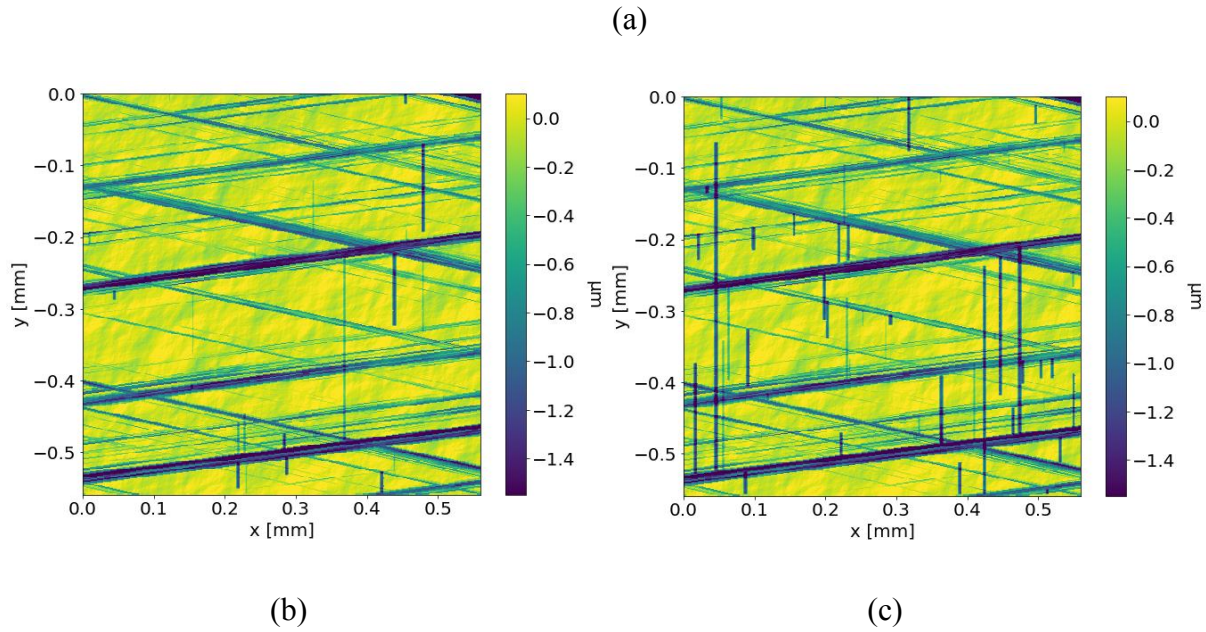
Figure 4.20. the liner topography change with particles following size distribution 2. (a) before wear (b) after 500 cycles (c) after 3000 cycles

As illustrated in Figure 4.21 and Figure 4.22, the red dashed line corresponds to the ring position on the liner surface height distribution with the given nominal pressure. On the particle size distribution plots, the blue shadow regions cover the particle size range of the frequency peaks. Most of the particles in each distribution fall into the blue shadow zone based on the particle diameters. The blue shadow regions on the particle size distributions have the identical height ranges to the red shadow regions on the surface height distributions. These plots clearly show that the third-body wear speed is strongly correlated with the relative magnitude between the clearance size and the particle size. For particle size distribution 1, the size of the majority of particles only covers a little portion of the clearance size range, leaving most particles having no chance to get trapped. Contrarily, for particle size distribution 2, most of the particle diameters locate within the majority clearance range so that particles have higher chances to meet clearances close to their sizes and get trapped.
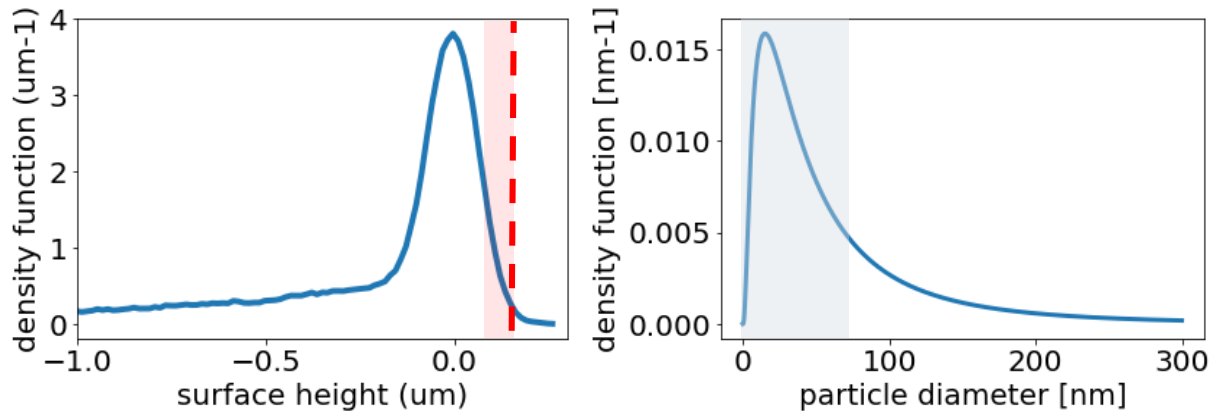
Figure 4.21. particle size distribution 1 and the corresponding clearance size range
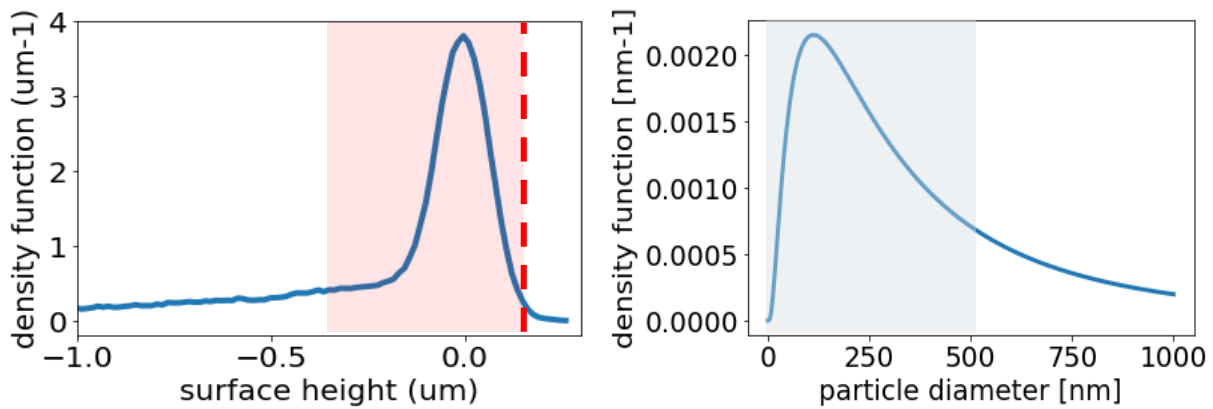
Figure 4.22. particle size distribution 2 and the corresponding clearance size range

The results of the short-term simulations show the determining effect of particle size distribution, especially the consistence with clearance distribution. If the peaks of the two distributions have a significant overlap, third-body particle abrasive wear rate will be high. Moreover, the results also indicate that the nominal load is not a dominant factor in third-body wear if the particle size distribution covers a large range. Only when the particle size distribution is concentrated in a small range, comparable to the clearance variation of different normal loads, external load would become more important. This is different from the asperity fatigue wear during the break-in period. Figure 4.23 shows the liner topography evolution when the nominal pressure is increased to 4MPa, with particle size distribution 2. The increased normal pressure does not lead to a

higher third-body wear. The liner wear volumes are 2735 um^3 and 2381 um^3 for 2MPa and 4MPa, respectively, on the simulated area after 3000 cycles.

(a)

(b)

(c)
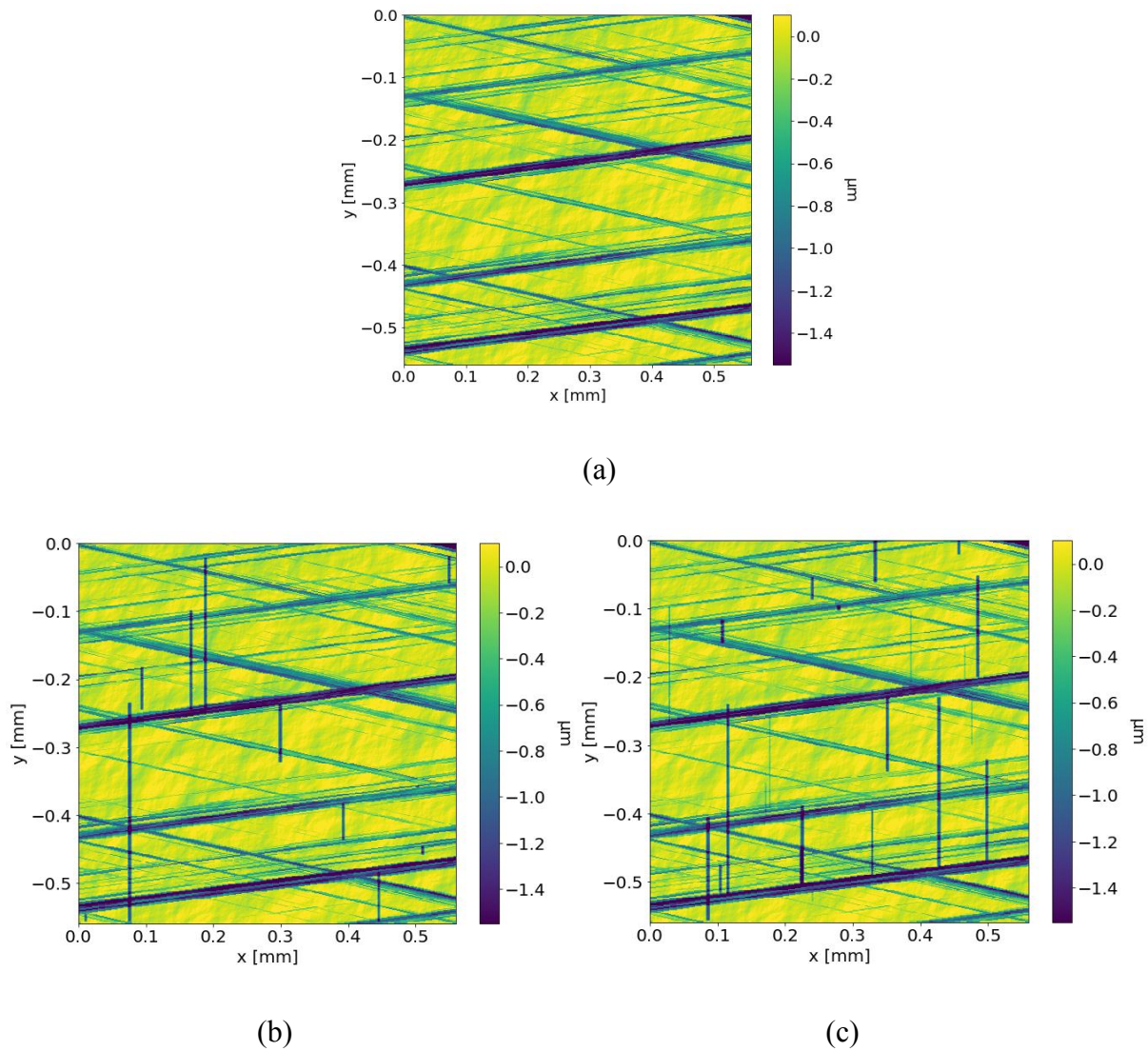
Figure 4.23. the liner topography change with particles following size distribution 2 and nominal pressure 4MPa. (a) before wear (b) after 500 cycles (c) after 3000 cycles

4.4.2 The influence of ring coatings

Another important influencing factor is the ring coating. Nowadays, the engine manufacturers try different coating materials in order to provide more sustainable engines. In this work, long-term

simulations are performed to investigate the effect of the ring surface properties, because the influence of two-body abrasion becomes significant only when the ring is roughened enough. The simulations also provide understandings about the third-body particle abrasion mechanisms during the steady state.

The simulation setup keeps the same except for the ring properties, with nominal pressure 2 MPa, soot concentration 4.5%. Particle size distribution 2 in the previous section is used in this section as well. The number of engine cycles is 500000, corresponding to about 4.2 hours of engine running. Simulation 1 and simulation 2 use two different ring coatings, PVD (physical vapor deposition) and DLC (diamond-like carbon), respectively. In the simulations, the elastic modulus and the hardness of PVD coating are 300 GPa and 10 GPa, close to the typical mechanical properties of chromium nitride coatings [47,48,49]. For the DLC coating, 330 GPa and 40 GPa are used as the material elastic modulus and hardness.

(a)   (b)
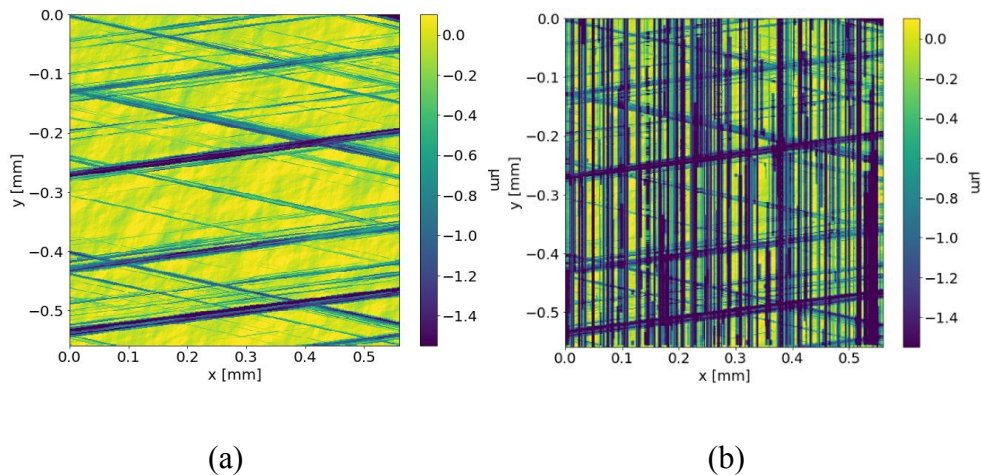
Figure 4.24. the liner surface topography of simulation 1. (a) before wear (b) after the wear simulation

In figure 4.24, the liner surface topography change can be observed. Many scratches along the sliding direction are generated, as a result of abrasion. This is a typical wear pattern created due to abrasive wear. However, the wear patterns created by two-body abrasion and third-body abrasion are still different, as illustrated in Figure 4.25. The third-body wear scars are more likely to stop in the liner valley regions because particles can fall into these valleys. Only some large particles would generate continuous wear scars, crossing the valleys on the liner. The two-

body wear pattern, on the other side, cannot be stopped by the valleys because the ring always moves from one side to the other. The wear scars of two-body abrasion are more likely to be continuous along the ring sliding direction. The ring topography after wear is also plotted in Figure 4.26, similar to the observed barcode ring wear pattern in engine experiments [7].

(a)   (b)

Figure 4.25. the two different abrasive wear patterns (a) third-body abrasion (b) two-body abrasion
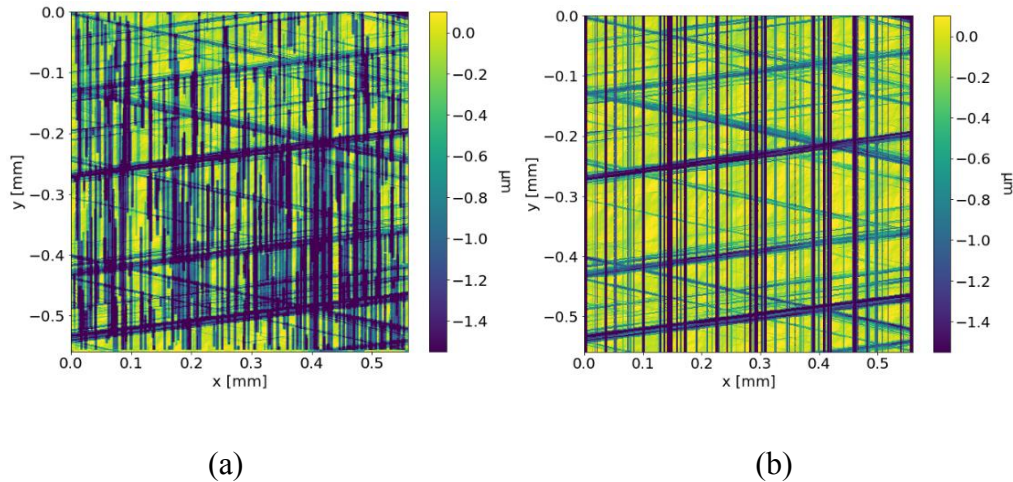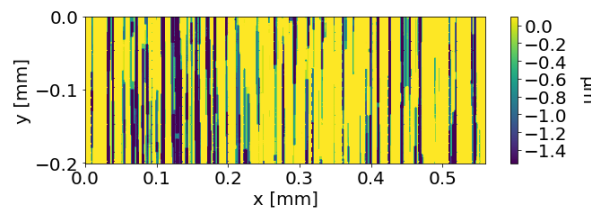

Figure 4.26. the ring surface topography of simulation 1, after wear
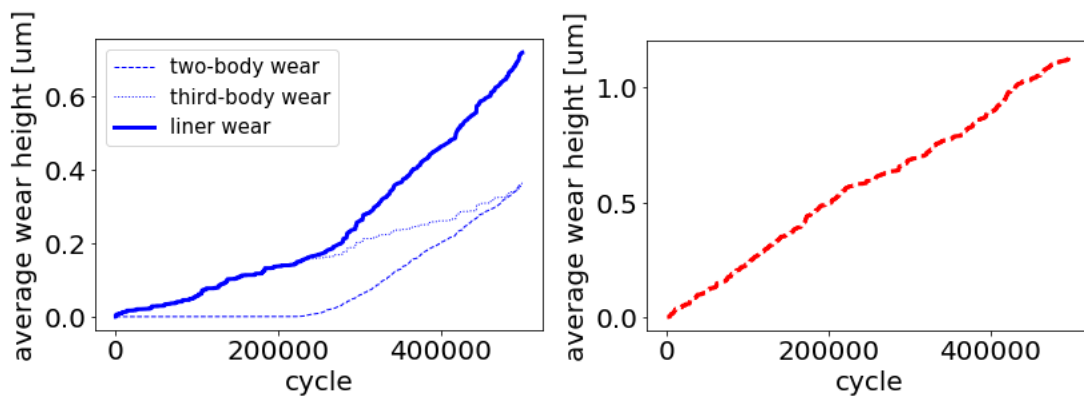
(a)   (b)

Figure 4.27. material wear loss in simulation 1 (a) liner material loss (b) ring coating material loss

Figure 4.27 shows the material loss of simulation 1 as a function of time, for both the ring and the liner. The material volume loss is converted to the average height reduction over the entire simulation area. The ring wear rate does not vary extensively over the simulation period, but the liner wear rate shows an obvious transition. The transition of the liner wear rate mainly comes from the two-body wear mechanism. Because the simulation starts from a smooth ring, it takes time for the ring to get roughened enough to create significant two-body wear loss on the liner side. This transition is largely dependent on the generation of high peaks on the ring surface to perform continuous two-body scratches on the liner.

If the ring coating is switched from PVD to DLC, the ring wear speed shows the same trend in experiment 2, with a different magnitude. The ring wear rate is smaller in simulation 2 because of two effects. As the DLC coating is harder, the critical indentation for pure cutting also increases for the same particle. This makes more particles fall into the plowing abrasion pattern, which does not lead to material loss. The material on the surface is merely pushed to different positions and redistributed. Meanwhile, the average indentation depth decreases for particles with a harder surface, resulting a lower volume loss even in the cutting mode. The trend of liner wear rate of simulation 2 is also different from simulation 1. The liner wear rate transition occurs significantly earlier than the case with PVD ring coating. As the ring gets harder, it becomes easier for particles to get embedded into the liner and slide against the ring. At the same time, particles tend to plow on the ring surface and create pile-ups, which have the potential to generate continuous two-body scratches. Overall, the smoothness of the harder ring coating is damaged quicker for the harder coating. However, eventually, the ring roughness of DLC is smaller than the roughness of the worn PVD surface. As plotted in Figure 4.29, the ring roughness of simulation 2 after wear is smaller than the ring roughness in simulation 1, because of smaller average particle indentations. The DLC roughness eventually is smaller, although it evolves faster to break the smoothness at the beginning.

139
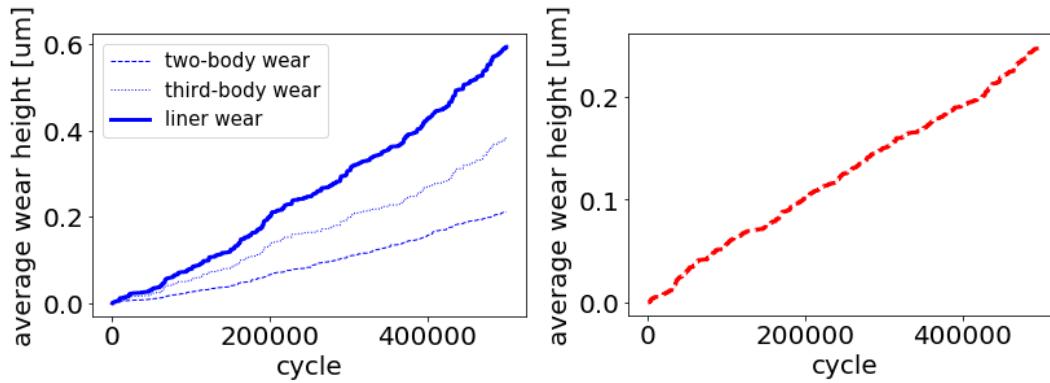
Figure 4.28. material wear loss in simulation 2 (a) liner material loss (b) ring coating material loss
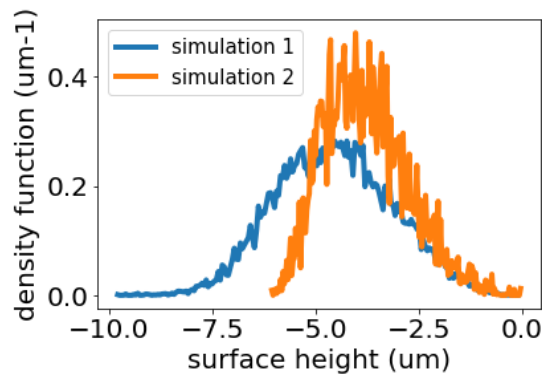
Figure 4.29. ring roughness after wear

Experimental wear results are also used for comparison [50]. In Figure 4.30, the ring wear losses in engine experiments are plotted with both PVD (chromium nitride) and DLC (hydrogen-free carbon) coatings [48,49]. The wear volume is also evaluated by the height reduction on the ring surface, with different positions measured. The total testing time is 500 hours. By averaging the height reduction at different ring positions, the overall ring wear speeds of the two coatings can be estimated, listed in Table 4.1. It is easy to observe that the calculated ring wear rates are higher than the wear rates obtained from experiments. Going through all the assumptions and approximations, the assumption of rigid spherical third-body particles may lead to an overestimation of the wear rate. In the next section, another assumption is introduced to reduce the estimation error caused by the rigid spherical particle assumption.
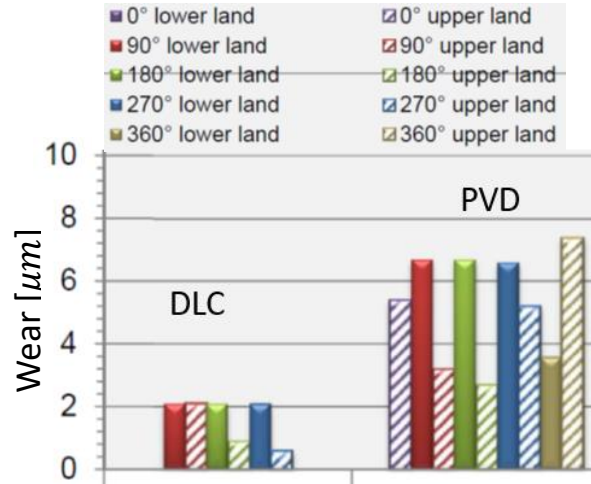
Figure 4.30. experimental results of ring wear [50]

Ring coating	average wear rate by simulations [$\mu m/h$]	average wear rate by experiments [$\mu m/h$]
PVD	0.24	1.15×10^{-2}
DLC	0.06	3.30×10^{-3}

Table 4.1. the average wear rates of ring coatings

4.4.3 One additional assumption for trapped particles

As compared in the previous section, the ring wear rate is, at least one order of magnitude, larger than the ring wear rate obtained from experiments. The rigid spherical particle assumption is considered to be the main contributor to overestimating the wear rate because the soot particles in real engines are agglomerated and random-shaped. Although the primary carbon particles are round and hard, most of the agglomerated soot particles are possible to be deformed if large external forces are exerted on the particles. As shown in Figure 4.31, agglomerated particles

141

usually have loose structures [51]. If a particle is trapped, such structures are not able to create indentation and generate abrasive scars using their nominal diameters. Therefore, a new assumption is made in order to offset the overestimation of wear rate due to non-spherical agglomerated soot particles.
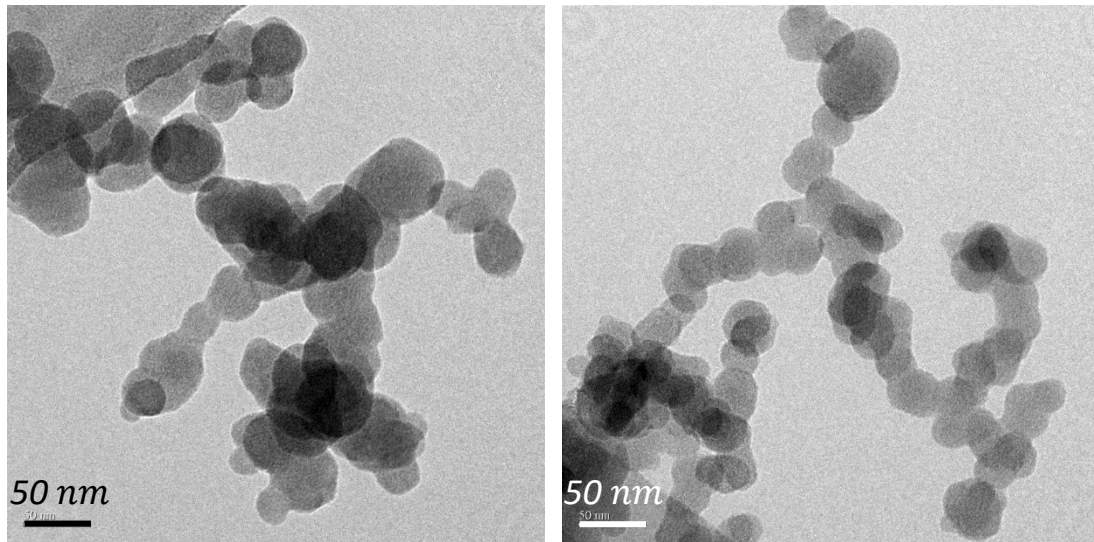
Figure 4.31. agglomerated particles [51]

As illustrated in Figure 4.32, a trapped particle can experience extensive deformation, being compressed into a more compact structure. With the compact structure, the particle will cause indentation and abrasion on the two surfaces. Considering the possible deformation of particles between the two surfaces, one additional assumption is made. For a particle entering a clearance smaller than its diameter, it is deformed into a more compact sphere, whose diameter is adaptive to the clearance size. The new diameter assigned to the trapped particle is only determined by the clearance size, regardless of its original size. The new diameter of the particle is also assumed to be h+2d, with h the local clearance size and d the diameter of a primary carbon particle. In this work, d is set to be 50 nm. This new assumption is only applied to a trapped particle, while the calculation of the previous entrapment process still uses the original particle diameter.
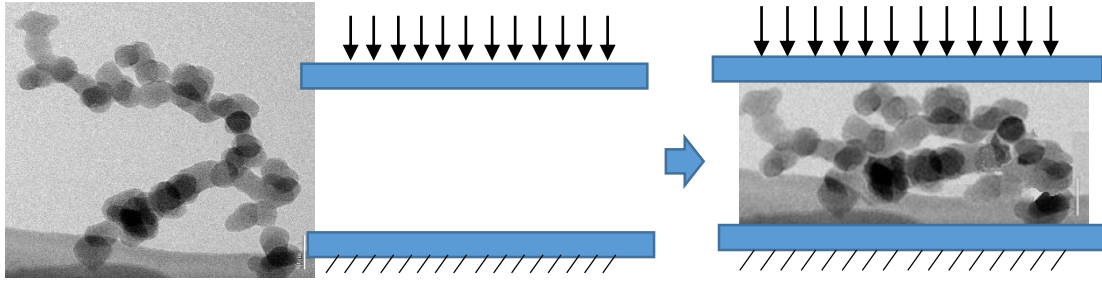
Figure 4.32. possible deformation of a trapped particle
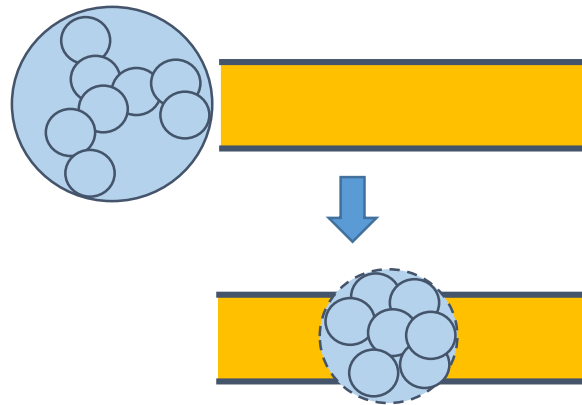
Figure 4.33. the deformed particle can have diameter adaptive to the local clearance

With the same input setup in the previous section, simulations are performed adding the new assumption about trapped particles. Figure 4.34 show the ring coating wear rates as functions of time. The averaged wear rates are summarized in Table 4.2. The predicted wear rates now are comparable to the experimental results, from the order of magnitude. However, the absolute magnitudes of the predicted wear rates are now smaller than the experimental ones, which could be caused by scenarios not covering by the new assumptions. The break-up of a trapped particle is not considered here, with only one smaller deformed particle added into simulation. In fact, a real particle could be squeezed into many parts, grinding the two surfaces together. In addition, the assumed trapped particle size may also bring simulation errors, without considering some large round particles having the potential to cut the surfaces.

(a)   (b)
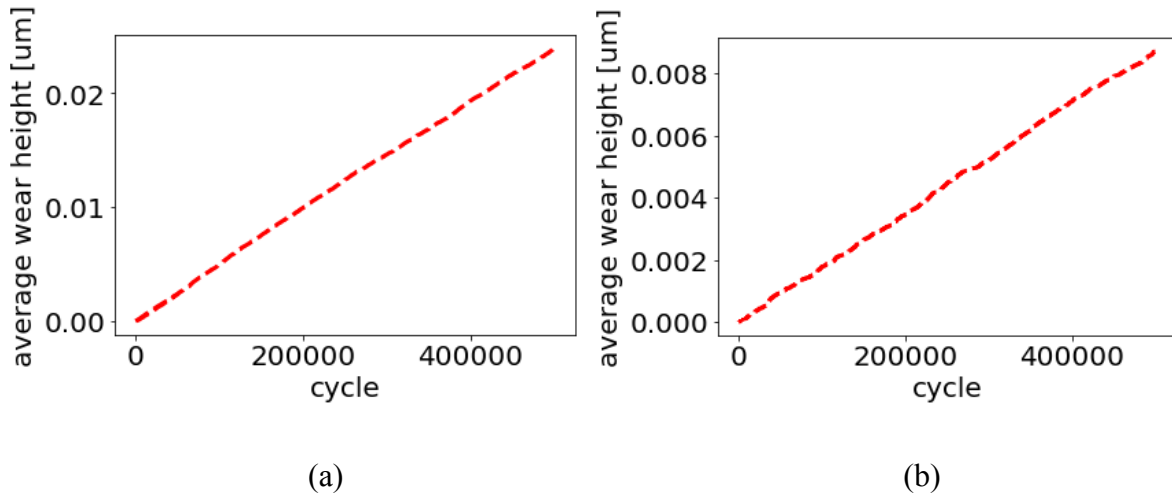
Figure 4.34. the ring coating wear loss (a) PVD coating (b) DLC coating

Ring coating	average wear rate by simulations $[\mu m/h]$	average wear rate by experiments $[\mu m/h]$
PVD	5.51×10^{-3}	1.15×10^{-2}
DLC	1.97×10^{-3}	3.30×10^{-3}

Table 4.2. the average wear rates of ring coatings, with the additional assumption

4.5 Conclusions

In this engine steady-state wear simulation work, two-body abrasive wear and third-body abrasive wear mechanisms are modeled and applied to systematic wear modeling for the piston rings and the liner. The transport and the wear processes of third-body particles play important roles in steady-state engine wear, including particle entrapment, particle movement, particle indentation and particle abrasion. Both short term and long term systematic simulations are performed to investigate the effect of different input parameters. In the short term simulation set, we found the determining factor of third-body abrasive wear is the relative magnitude between

the clearance distribution and the particle size distribution. If the clearance size distribution and particle size distribution have significant overlap, third-body wear rate is high because most of the particles have high chance being trapped between the two surfaces and creating wear scars. The long-term simulations show that the hardness of the ring coating plays an important role. If the ring coating is hard, the wear rate on the ring side will be reduced, but the wear rate of the liner at the beginning stage will be larger. This is because for a harder ring, the smoothness of the ring surface is damaged faster at the beginning due to more particles embedded into the liner surface. Therefore, the tradeoff between the ring wear rate and the liner wear rate at the initial stage should be considered for the design of the ring coating.

Chapter 5. Conclusions and future work

In this chapter, the general conclusions and the potential future work are discussed. The first part summarizes the oil emission modeling and engine wear modeling work, while the second part proposes some potential future work following this thesis.

5.1 Engine oil emission modeling via machine learning

In this work, a complete set of data-driven procedures are defined, including data processing, featuring engineering, model selection, system parameter estimation, prediction and experiments design. A number of general conclusions can be drawn based on this work.

The separation of steady-state oil emission and transient oil emission is an essential data processing procedure to capture the underlying correlation. While steady-state oil emission depends on the engine speed and load, the transient one also depends on the running conditions and oil emission of the past, which either contributes to the oil accumulation in the ring pack of the present time or reflects the oil accumulation above the ring pack that directly emits to the exhaust. The residence time defined here is a critical parameter representing the past influence on the present oil emission. The ML algorithm employed here is effective to identify the residence time from the oil emission measurements and this is a crucial step to extract further transient oil emission information out of engine test results.

With a defined residence time, further useful information can be derived, namely, the uncertainty level and the total oil emission with a given engine running cycle. The uncertainty can be caused by a combination of the measurements and the inherent oil emission behavior of the design. Thus, its determination can lead to further analysis of the consistency and robustness of the measurements and design. The interval prediction is found to be adequate for the uncertainty estimation. For the total oil emission prediction for a new test cycle, the methodology developed here by dividing the engine test sequence into chunks based on the residence time proves to be effective. The ability to predict the total oil emission of other driving cycles based on one test enables the development of a test protocol by defining a baseline running cycle that covers,

statistically, all the driving varieties. Albeit unfinished, it is illustrated that an unsupervised learning can be applied for engine oil emission experimental design. Based on the publication, this is the first time ML was used to analyze the engine oil emission measurements. It was found that the algorithms employed here are effective and less data-hungry. Further developments along this direction, such as design new test protocol, connecting to physics modeling, and adding additional constraints in ML based on physics, will help reduce engine test costs and establishing better understanding of oil transport and oil consumption mechanisms.

5.2 Modeling of engine wear

The wear modeling of engine is completed through deterministic approaches by integrating all the critical mechanical processes and sub-models. Both engine break-in and steady-state wear processes are modeled, with different dominant mechanisms, asperity fatigue and abrasive wear, respectively.

The simulation results show an agreement with Archard's wear law that for a specific liner surface finish, the steady-state wear rate is proportional to external nominal load when the nominal load is fully balanced by the asperity contact. With the linear correlation between wear rate and the external pressure, wear coefficients of different liner surfaces can be obtained. Simulations indicate that the maximum steady-state wear rate can be obtained for a specific surface roughness level. In addition, the engine friction evolution due to break-in liner wear calculated by the presented wear model shows the same trend as the experimental measurements. The friction evolution speed during the break-in period depends on the roughness level of the liner plateau part.

The wear simulations of the engine steady-state period are based on abrasive wear, both two-body and third-body abrasion. It is the first time the complete mechanical processes are modeled, from trapping the particles to embedment of the particles on the liner, and subsequent various abrasive wear. With the assumption that the third body particles are have the highest hardness among the three, ring can only be worn down by the third body particles that are embed on the liner. Then, roughening of the ring surface initiates its abrasion to the liner, i.e. two body

wear. In terms of the wear marks on the liner, two body wear tends to create continuous valleys along the sliding direction while the third body is interrupted by the deep honing grooves. The results indicate that two influencing factors are critical for steady-state engine wear, the third-body particle size distribution and the mechanical properties of the ring coating. Abrasive wear rate is high if the particle size distribution has a significant overlap with the clearance height distribution between the two surfaces so that the chances to get the particles trapped are high. This is important for particle filter design to reduce further wear. The material choice of the ring coating is also an important design parameter. For harder ring coatings, the wear rate of the ring will be lower, but the liner would experience a faster transition to significant two-body abrasive wear due to higher embedment rate of the particles on the liner, which in turn causes faster spread of the wear scare on the ring face. However the overall wear effect of ring coating properties on the ring and the liner should be carefully evaluated during the engine component design.

5.3 Potential future work

The work in this thesis is designed to find correlations of the oil emission and provide comprehensive models for engine wear. Some potential future work is proposed in this section to improve the calculation efficiency, prediction accuracy or generalizability of the models.

The methodology of the oil emission modeling is data-driven. As pointed out in Chapter 2, data-driven methods provide advantages to overcoming the modeling difficulties of the complex oil transport and emission process. However, they limit the generalizability and the interpretability of the model, making it rely heavily on the measurement data of each engine. With an increasing understanding about the oil transport and emission process, more physics knowledge can be hybridized with the data-driven methods to create a more interpretable model. There are several approaches to achieve this goal. The most intuitive way for machine learning and physics hybrid is the residual correction. Once a simple physics model is developed, machine learning can be applied to build a mathematical mapping from the output of the physics model to the real measured data. In this case, machine learning is working as a residual calibrator, while the entire

model can have high interpretability and avoid difficulties of modeling the physics details. Another way to connect physics and machine learning is designing physics guided objective functions to guide the optimization direction. Usually, the objective function only has an error term, which is not enough to reflect some important physics properties of the system. Through adding a physics-based term, the optimization algorithm can penalize the regression results that contradict with physics so that the overall training results can fit better with the physics model provided. It is also useful to design an activation function based on physics so that some certain physics constraints can be met. So far, the understanding about the oil emission mechanisms is still limited. With deeper and more robust physics understanding, these hybrid methods can be applied to improve the data-driven modeling.

For the work of engine wear modeling, the approach is deterministic. However, deterministic modeling is expensive, both in time and computation resource occupation. One potential future work is abstracting some important sub-models with statistical approaches to increase the calculation efficiency. For example, the asperity fatigue calculation currently is based on simulations of propagation of each crack. This process can be greatly simplified with additional assumptions. Here, one possible approach is proposed to increase the calculation efficiency.
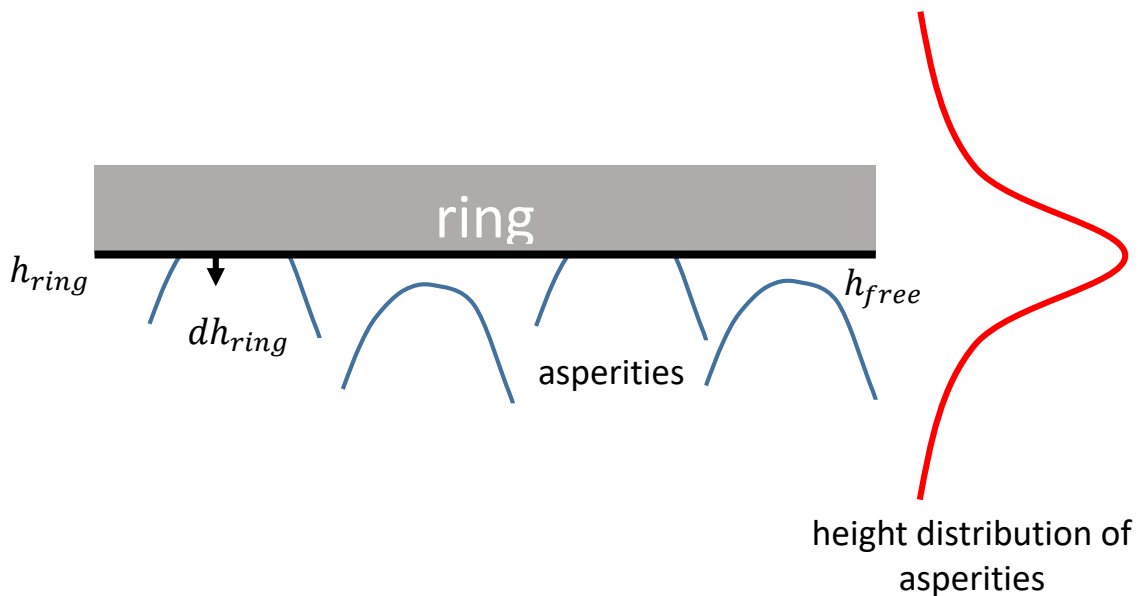

Figure 5.1. asperities with the same size, distributed on the liner surface

The assumption required for the calculation simplification is that all the asperities on the liner surface have the same size and shape. As illustrated in Figure 5.1, these liner asperities, with the same size, are distributed at different heights. Under such an assumption, once the fracture speed of an asperity is determined, the correlation can be applied to all the surface asperities. There are existing models to convert any liner surface roughness to a new surface having the same sized asperities, with the equivalent contacting properties [52,53].

The fracture of each asperity is now considered as a gradual chopping-off of the highest peaks, instead of removing an entire asperity due to crack propagation. For each asperity, h_{free} describes the free surface, above which the asperity material is removed due to wear. Combining the asperity fatigue mechanism during the break-in, the Paris law of crack propagation can be generalized to:

$$\frac{dw}{dt} = Paris\big(h_{free}, h_{ring}, Cof\big)$$

Where w is the crack length and h_{ring} is the actual ring height. Thus, it is possible to calculate the time needed to remove the asperities at any h_{free}. With some additional coefficients in the generalized Paris law and a fixed asperity height distribution $P(h)$, some simplification of asperity fatigue process can be done. The generalized Paris law may need finite element calculation to derive the correlation. However, as the asperity size is assumed to be constant, it can be applied to all asperities once the correlation is obtained. The calculation amount is still significantly reduced compared to simulating the fracture behavior of each asperity.

For the steady-state engine wear, the biggest error source comes from the assumption of rigid and spherical particles. This assumption would remarkably increase the estimated wear rate as most of the particles are agglomerated particles, consisting of many small primary carbon particles. They could be deformed and smashed. Although one additional assumption is introduced to reduce the wear rate to the same magnitude level as the experimental results, the detailed particle transport and trapped process is still not included in the current model. The potential future work can include the morphology of third-body particles and the shape change during the transport process. This will require more detailed discussions about the chemical bonding among primary particles and the morphology distribution of agglomerated particles.

References

[1] Schmidt, T., Gaertner, F., & Kreye, H. (2006). New developments in cold spray based on higher gas and particle temperatures. Journal of Thermal Spray Technology, 15(4), 488-494.

[2] Chen, H. (2011). Modeling the lubrication of the piston ring pack in internal combustion engines using the deterministic method (Doctoral dissertation, Massachusetts Institute of Technology).

[3] Paoloni, F., Burette, G., Gohl, M., Lensch-Franzen, C., & Holzmüller, J. (2020). Beyond Noack Volatility: New Perspective for Engine Lubricant Oil and Particle Emissions (No. 2020-01-2161).

[4] Gu, C., Wang, R., & Tian, T. (2019). Modeling the fatigue wear of the cylinder liner in internal combustion engines during the break-in period and its impact on piston ring lubrication. Lubricants, 7(10), 89.

[5] Truhan, J. J., Qu, J., & Blau, P. J. (2005). The effect of lubricating oil condition on the friction and wear of piston ring and cylinder liner materials in a reciprocating bench test. Wear, 259(7-12), 1048-1055.

[6] Papadopoulos, P., Priest, M., & Rainforth, W. M. (2007). Investigation of fundamental wear mechanisms at the piston ring and cylinder wall interface in internal combustion engines. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 221(3), 333-343.

[7] Rolls-Royce Power Systems AG. (2019) Ring/liner wear from EGR-Engine.

[8] Suh, N. P. (1973). The delamination theory of wear. Wear, 25(1), 111-124.

[9] Frommer, A., Beeckmann, A., Freier, R., & Pierotti, R. (2012). Online Oil Consumption Measurement—Steady-state vs. Dynamic. In Conference SIMEA, Brazil.

[10] Frommer, A., Deuß, T., Ehnis, H., & Künzel, R. (2016). Design of the Tangential Force of the Oil Control Ring—Oil Emission Versus Friction. MTZ worldwide, 77(1), 48-53.

[11] Adelmann, J., Becker, S., Rabute, R., Bruno, R. (2018). Optimized oil control ring design for emission reduction. Zylinderlaufbahn, Kolben, Pleuel, 91-103

[12] Yilmaz, E., Thirouard, B., Tian, T., Wong, V. W., Heywood, J. B., & Lee, N. (2001). Analysis of oil consumption behavior during ramp transients in a production spark ignition engine. SAE Transactions, 1828-1837.

[13] Frommer, A., Beeckmann, A., Freier, R., & Künzel, R. (2013). Analysis of Lube Oil Consumption in Transient Engine Operation. MTZ worldwide, 74(1), 26-33.

[14] Chiu, Y. P., Tallian, T. E., McCool, J. I., & Martin, J. A. (1969). A mathematical model of spalling fatigue failure in rolling contact. ASLE transactions, 12(2), 106-116.

[15] Chiu, Y. P., Tallian, T. E., & McCool, J. I. (1971). An engineering model of spalling fatigue failure in rolling contact: I. The subsurface model. Wear, 17(5-6), 433-446.

[16] Tallian, T. E., & McCool, J. I. (1971). An engineering model of spalling fatigue failure in rolling contact: II. The surface model. Wear, 17(5-6), 447-461.

[17] Kudish, I. I., & Burris, K. W. (2000). Modern State of Experimentation and Modeling in Contact Fatigue Phenomenon: Part II—Analysis of the Existing Statistical Mathematical Models of Bearing and Gear Fatigue Life. New Statistical Model of Contact Fatigue. Tribology transactions, 43(2), 293-301.

[18] Ma, Z., Henein, N. A., & Bryzik, W. (2006). A model for wear and friction in cylinder liners and piston rings. Tribology transactions, 49(3), 315-327.

[19] Giorgio, M., Guida, M., & Pulcini, G. (2010). A state-dependent wear model with an application to marine engine cylinder liners. Technometrics, 52(2), 172-187.

[20] Mezghani, S., Demirci, I., Yousfi, M., & El Mansori, M. (2013). Running-in wear modeling of honed surface for combustion engine cylinderliners. Wear, 302(1-2), 1360-1369.

[21] Zum Gahr, K. H. (1987). Microstructure and wear of materials (Vol. 10). Elsevier.

[22] Fang, L., Liu, W., Du, D., Zhang, X., & Xue, Q. (2004). Predicting three-body abrasive wear using Monte Carlo methods. Wear, 256(7-8), 685-694.

[23] Reizer, R., Pawlus, P., Galda, L., Grabon, W., & Dzierwa, A. (2012). Modeling of worn surface topography formed in a low wear process. Wear, 278, 94-100.

[24] Przesmitzki, S. S. V. (2008). Characterization of oil transport in the power cylinder of internal combustion engines during steady state and transient operation (Doctoral dissertation, Massachusetts Institute of Technology).

[25] Hamilton, J. D. (2020). Time series analysis. Princeton university press.

[26] Tealab, A. (2018). Time series forecasting using artificial neural networks methodologies: A systematic review. Future Computing and Informatics Journal, 3(2), 334-340.

[27] Mason, L., Baxter, J., Bartlett, P., & Frean, M. (1999, May). Boosting algorithms as gradient descent in function space. In Proc. NIPS (Vol. 12, pp. 512-518).

[28] Friedman, J., Hastie, T., & Tibshirani, R. (2001). The elements of statistical learning (Vol. 1, No. 10). New York: Springer series in statistics.

[29] L. Sabri, S.Mezghani, M. El Mansori, H. Zahouani, "Multiscale Study of Finish-Honing Process in Mass Production of Cylinder Liner", Wear, March 2010.

[30] D. Lawrence, "An Accurate and Robust Method for the Honing Angle Evaluation of Cylinder Liner Surface Using Machine Vision", Int. J. Adv. Manuf. Technol., 2011.

[31] Zhao, Y., Maietta, D. M., & Chang, L. (2000). An asperity microcontact model incorporating the transition from elastic deformation to fully plastic flow. Journal of tribology, 122(1), 86-93.

[32] Andersson, Sören, Anders Söderberg, and Ulf Olofsson. "A random wear model for the interaction between a rough and a smooth surface." Wear, 2008.

[33] Johnson, Kenneth Langstreth, and Kenneth Langstreth Johnson. "Contact mechanics" Cambridge university press, 1987.

[34] Paris, P., & Erdogan, F. (1963). A critical analysis of crack propagation laws. Journal of basic engineering, 85(4), 528-533.

[35] Zach Westerfield, "A Study of the Friction of the Power Cylinder System in Internal Combustion Engines using a Floating Liner Engine", Master Thesis, Massachusetts Institute of Technology, 2015.

[36] Mercedez-Benz, "Honen von Zylinder-Laufflachen," 2008.

[37] Wang, R., Gu, C., & Tian, T. (2019). Reliable Processes of Simulating Liner Roughness and Its Lubrication Properties (No. 2019-01-0178). SAE Technical Paper.

[38] Archard, J. (1953). Contact and rubbing of flat surfaces. Journal of applied physics, 24(8), 981-988

[39] Archard, J. F., and W. Hirst. "The wear of metals under unlubricated conditions." In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 1956.

[40] Ajayi, L., Lorenzo-Martin, C. (2017) Volvo engine ring and liner wear mechanisms analysis.

[41] Liang, X., Wang, Y., Dong, L., Shu, G., & Liu, C. (2014). Experimental study of the influence of a pour point depressant additive on particle size distributions in diesel engine exhausts. Particuology, 13, 73-81.

[42] Wang, Y., Liang, X., Shu, G., Wang, X., Sun, X., & Liu, C. (2014). Effect of lubricant oil additive on size distribution, morphology, and nanostructure of diesel particulate matter. Applied energy, 130, 33-40.

[43] Nikas, G. K. (2006). A mechanistic model of spherical particle entrapment in elliptical contacts. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 220(6), 507-522.

[44] Fang, L., Kong, X. L., Su, J. Y., & Zhou, Q. D. (1993). Movement patterns of abrasive particles in three-body abrasion. Wear, 162, 782-789.

[45] Goddard, J., & Wilman, H. (1962). A theory of friction and wear during the abrasion of metals. Wear, 5(2), 114-135.

[46] Xu, H. (2012). Asperity-scale surface mechanics-Implications to adhesive contacts and microscale deformation behavior of rough surfaces (Doctoral dissertation, UC Berkeley).

[47] Ring Product Development Div , Riken Corporation (2021). Piston Ring Technologies for Natural Gas Engines

[48] Olaya, J. J., Wei, G., Rodil, S. E., Muhl, S., & Bhushan, B. (2007). Influence of the ion–atom flux ratio on the mechanical properties of chromium nitride thin films. Vacuum, 81(5), 610-618.

[49] Martınez, E., Andújar, J. L., Polo, M. C., Esteve, J., Robertson, J., & Milne, W. I. (2001). Study of the mechanical properties of tetrahedral amorphous carbon films by nanoindentation and nanowear measurements. Diamond and Related Materials, 10(2), 145-152.

[50] Li, F. (2019). Preliminary Results of Weichai H1 WP9H336E60 Euro VI Rings and Cylinder Liners After 500 Hours Durability Test

[51] Wang, Y., Liang, X., Shu, G., Dong, L., Sun, X., Yu, H., & Wang, Y. (2015). Effects of an anti-wear oil additive on the size distribution, morphology, and nanostructure of diesel exhaust particles. Tribology International, 92, 379-386.

[52] Zhao, Q. (2014). Modeling of contact between liner finish and piston ring in internal combustion engines based on 3D measured surface (Master thesis, Massachusetts Institute of Technology).

[53] Greenwood, J. A., & Williamson, J. P. (1966). Contact of nominally flat surfaces. Proceedings of the royal society of London. Series A. Mathematical and physical sciences, 295(1442), 300-319.