Research Review on Atkinson Cycle Internal Combustion Engines: Technological Development and Optimization Strategies

Jiabei Liu

School of Automobile Service Engineering Shanghai Normal University Shanghai, China

1000530597@smail.shnu.edu.cn

Abstract. This paper provides a comprehensive review of the principles, characteristics, and application prospects of Atkinson Cycle internal combustion engines in hybrid vehicles. It delves into the crucial role of fluid mechanics in optimizing cylinder design for Atkinson cycle engines, emphasizing its importance in improving engine efficiency, fuel economy, and overall performance. The paper contrasts traditional fluid mechanics research methods, such as computational fluid dynamics (CFD) and empirical testing, with machine learning approaches, showcasing the latter's transformative impact on research precision and efficiency. Through an in-depth analysis of machine learning techniques—such as clustering, modal decomposition, neural networks, and ensemble learning—it elaborates on their application methods, strengths, and potential to accelerate research processes in the study of Atkinson cycle engines. The review highlights how machine learning can enhance simulation accuracy, reduce computational costs, and predict complex flow behaviors, making it a valuable tool for advanced engine research. Additionally, it discusses the integration of these approaches into hybrid vehicle design, providing valuable insights for engineers and researchers. This paper serves as a reference for innovation and technological development, offering guidance for future advancements in internal combustion engine research and hybrid powertrain optimization.

Keywords: Atkinson cycle, Internal combustion engine, Fluid mechanics, Machine learning, Hybrid vehicle.

1. Introduction

With increasing global attention to energy efficiency and environmental protection, automotive engine technologies have been constantly evolving. As a technology capable of improving the thermal efficiency of engines, the Atkinson cycle has been widely studied and explored for its application. Against the backdrop of the rapid development of hybrid vehicles, the advantages and potential of Atkinson cycle internal combustion engines have become even more prominent. Meanwhile, fluid mechanics research has provided an important theoretical basis and technical support for the design of the engine cylinder [1]. And the introduction of machine learning methods has brought new ideas and optimization approaches to traditional research, demonstrating unique value and application potential especially in fluid mechanics-related research.

2. Atkinson Cycle

The Atkinson cycle is a form of internal combustion engine that is different from the common Otto cycle. Its core feature is to make the expansion ratio greater than the compression ratio through structural changes or by delaying the closing of the intake valve[2]. The main purpose of this design is to improve the thermal efficiency of the engine and thus enhance its economic performance. However, the Atkinson cycle also has obvious drawbacks, namely, it will reduce the power performance of the engine. Due to the reduced air volume intake, the output power of the engine is limited under high load working conditions. Moreover, the characteristic of the delayed closing of the intake valve makes it difficult to be combined with turbocharging technology, which restricts its wide application in the field of traditional fuel vehicles to some extent. However, with the rise of electric and hybrid vehicles, the Atkinson cycle has ushered in new development opportunities. In hybrid vehicles, the Atkinson cycle engine can work in coordination with the electric motor to make up for the defect of insufficient power and fully utilize its economic advantage of high efficiency and energy saving.

3. Fluid Mechanics in Atkinson Cycle Internal Combustion Engines

Fluid mechanics plays an indispensable role in the research and design of Atkinson Cycle Internal Combustion engines. During the intake process, through the study of the flow characteristics of the air flow in the intake port and by means of computational fluid dynamics (CFD) simulation and other methods, the shape and size of the intake port can be optimized to reduce the intake resistance and improve the intake efficiency [3]. For example, the streamlined intake port designed according to the principles of fluid mechanics can reduce the separation and vortex of the air flow, making the intake smoother. During the combustion process, the air flow movement in the combustion chamber, the interaction between the fuel spray and the air flow, and the flame propagation characteristics are all closely related to fluid mechanics. Reasonable design of the shape of the combustion chamber, the position of the spark plug, and the layout of the fuel injector can promote the formation and combustion of the mixture, improving the combustion efficiency and stability. For example, by generating vortices and tumbling flows of appropriate intensities, the flame propagation can be accelerated, the combustion duration can be reduced, and pollutant emissions can be decreased. During the exhaust process, optimizing the design of the exhaust port can reduce the exhaust resistance, improve the exhaust efficiency and reduce the amount of residual exhaust gas, creating favorable conditions for the next intake cycle. In addition, fluid mechanics research also has important guiding significance for analyzing the flow distribution of the cooling system and the cooling medium of the engine as well as the flow characteristics of the lubricating oil in the lubrication system, ensuring that the engine operates within the normal working temperature range and reducing the friction and wear of components.

4. Limitations of Traditional Fluid Mechanics Research Methods

Traditional fluid mechanics research methods, such as computational fluid dynamics (CFD) and visualization techniques like laser-induced fluorescence (LIF), have certain difficulties and limitations when studying Atkinson cycle internal combustion engines.

When using CFD to study Atkinson cycle internal combustion engines, it is necessary to solve complex partial differential equations to describe the fluid flow and heat transfer processes. Due to the complex geometric structure and transient working process of internal combustion engines, CFD calculations are often very time-consuming. For example, performing high-precision three-dimensional CFD simulations for the entire engine working cycle may take several days or even weeks, depending on the number of grids, the complexity of the physical model, and the performance of computer hardware. Moreover, CFD simulations will generate massive amounts of data, and post-processing and in-depth mining of these data to obtain valuable information is a challenging task[3]. For example, extracting the characteristics of the in-cylinder flow field from three-dimensional CFD results, such as the generation and evolution of vortex structures and the interaction between fuel sprays and air flows, requires professional post-processing software and complex analysis techniques, and it is often difficult to directly relate them to the performance indicators of the engine.

In addition, once the settings such as the geometric structure, physical model, and boundary conditions of the CFD model are determined, modifications and adjustments are relatively complicated. For example, when it is necessary to study the impact of a slight change in the shape of the engine intake port on the in-cylinder flow, it is required to re-mesh, set the boundary conditions, and conduct long calculations. And for new engine design concepts, the CFD model may need to be rebuilt.

5. Application of Machine Learning in the Research on Atkinson Cycle Internal Combustion Engines

Clustering is an unsupervised learning method that divides similar data points into different groups (clusters) [4]. It has significant importance in the research on in-cylinder fluid mechanics of Atkinson Cycle internal combustion engines. In the research on engine cylinders, it can be used to classify working states or operating conditions according to the similarity of in-cylinder data. For example, based on multi-dimensional data such as in-cylinder pressure curves, temperature distributions, and fluid velocity fields, the working cycles of the engine can be divided into different categories, such as normal combustion, incomplete combustion, knock, and other different operating condition categories.

The K-Means clustering algorithm divides clusters according to the distance of data points to the cluster centers. In the processing of engine in-cylinder data, for example, when clustering in-cylinder pressure, temperature, and velocity data under multiple different rotational speeds and loads, by setting an appropriate number of clusters (the K value), selecting k samples as cluster centers, and then calculating the distance of each sample to the cluster centers, similar multi-physical field change patterns are grouped into one category, which helps researchers understand the typical working modes of the engine under different operating conditions and the corresponding in-cylinder fluid mechanics characteristics.

$$J = \sum_{i=1}^{K} \sum_{j=1}^{n} ||x_j - \mu_i||^2 \tag{1}$$

Among them, J is the sum of the squared errors within the cluster, K is the number of cluster centers, n is the total number of data points, x_j is the jth data point in the dataset, and μ_i is the clustering center of the ith cluster.

Agglomerative hierarchical clustering is also a commonly used method. It classifies data by constructing a hierarchical structure of clusters. When analyzing the correlation between the stability of the combustion process in the engine cylinder and the fluid motion, agglomerative hierarchical clustering can be used to analyze the combustion data and fluid mechanics data in consecutive working cycles [5]. By observing the similarities and differences between the combustion data and fluid field data at different levels, the hierarchical structural characteristics of the in-cylinder fluid motion under stable and unstable combustion states can be identified, providing a basis for optimizing the fluid regulation in the combustion process. To better understand the algorithm principle of agglomerative hierarchical clustering, we will illustrate it with diagrams. According to the algorithm of the average distance in hierarchical clustering, the engine combustion data is processed. The formula is as follows,

$$dist(C1, C2) = \frac{1}{|C1|\cdot|C2|} \sum_{P_i \in C1, P_j \in C2} dist(P_i, P_j) \tag{2}$$

where |C1| and |C2| respectively represent the number of samples in the cluster classes.

A comparative analysis is conducted using both the K-means clustering and hierarchical clustering methods. The K-means clustering determines K clusters through an iterative process and optimizes the center point of each cluster. This method is applicable to large-scale datasets and is easy to visualize. The hierarchical clustering method gradually merges or splits data points by calculating the distances between them until the desired hierarchical structure is achieved, and it can generate a dendrogram, which shows the relationships among different data points.

The advantage of clustering lies in its ability to discover the natural grouping structure in the data without the need to know the category labels of the data in advance, which is helpful for exploratory data analysis. In the research on engine cylinders, it can uncover new operating condition patterns or types of combustion anomalies as well as the related fluid mechanics anomalies[6]. However, the quality of the clustering results depends on the feature selection of the data and the parameter settings of the clustering algorithm. For example, for complex multi-physical field data in the engine cylinder, if the selected features cannot well represent the essential characteristics of the data, it may lead to unreasonable clustering results. Moreover, different clustering algorithms may produce different clustering results for the same data, and they need to be evaluated and selected according to specific problems [5].

Modal decomposition aims to decompose complex data into simple modes and has unique application value in the research on in-cylinder fluid mechanics of Atkinson cycle internal combustion engines. Principal component analysis (PCA)[6], a commonly used modal decomposition method, converts the original in-cylinder fluid data (such as pressure and velocity data at multiple positions) into new uncorrelated variables (principal components) through linear transformation, reduces the data dimension and extracts the main variation patterns. These principal components can characterize the main features of the in-cylinder fluid physical processes, such as the pressure rise and fluid acceleration patterns caused by combustion, and the pressure fluctuation and velocity change patterns resulting from the intake and exhaust processes. Empirical modal decomposition (EMD) and its improved methods (such as ensemble empirical modal decomposition, EEMD) can handle nonlinear and non-stationary in-cylinder fluid signals and decompose them into a series of intrinsic mode functions (IMFs). The core step of EMD is the sifting process, which is used to extract IMFs. The mathematical expression of its algorithm is that given a signal x(t) that needs to be decomposed into n IMFs and a residual term $r_n(t)$.

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \tag{3}$$

Taking non-stationary fluid signals such as in-cylinder pressure and velocity as examples, EMD or EEMD can decompose different frequency IMFs. The high-frequency components.

Taking non-stationary fluid signals such as in-cylinder pressure and velocity as examples, EMD or EEMD can decompose them into intrinsic mode functions (IMFs) with different frequencies[7]. The high-frequency components represent the local fluid disturbances caused by combustion, while the low-frequency components reflect the overall fluid motion resulting from the working cycle. This helps to gain an in-depth understanding of the composition and variation laws of the in-cylinder fluid signals as well as their interaction relationships with combustion and mechanical motions.

The advantage of modal decomposition lies in simplifying complex data and extracting key information. For example, it can separate the fluid components related to combustion and mechanical vibration in the in-cylinder pressure signals, which facilitates the separate exploration of the influence mechanisms of combustion and mechanical motion on in-cylinder fluid mechanics. However, its limitations are that its applicability is restricted by the characteristics of the data. PCA has difficulty capturing the complex variation patterns of in-cylinder fluid data with strong nonlinear relationships. Although EMD and its related methods can handle nonlinear and non-stationary data, modal aliasing may occur during the decomposition process, and further processing is required to ensure the accuracy of the results[6].

Neural networks (as shown in Fig. 1) imitate the structure and function of biological neural networks and have powerful capabilities in the research of in - cylinder fluid mechanics of Atkinson –

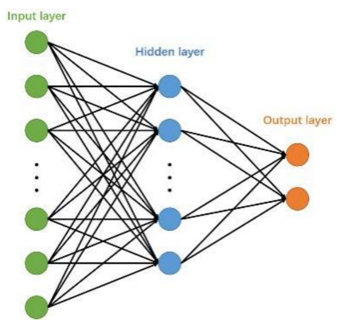

Fig. 1 Neural Networks

cycle internal combustion engines. Composed of a multi - neuron hierarchical structure, they can learn the complex mapping relationships between inputs and output targets. In the research of engine cylinders, they are used to build a variety of prediction models related to fluid mechanics, such as predicting engine performance, combustion status, pollutant generation and other parameters that are closely related to in - cylinder fluid motion. When the multi - layer perceptron (MLP), as a basic neural network structure, is used to build an engine in - cylinder performance prediction model, parameters such as in - cylinder pressure, temperature, fluid velocity and fuel injection quantity are input, and performance indicators such as engine output power and fuel economy are output[6]. Through training to learn the non-linear relationship between input and output, the engine performance under different working conditions can be predicted and the influence of in - cylinder fluid mechanics factors can be considered comprehensively.

$$y = f\left(\sum_{i=1}^{N} W_i x_i\right) \tag{4}$$

Convolutional neural networks (CNN) are used to process grid-structured data, such as in-cylinder image data (images of combustion flames and visualized flow fields). They can automatically extract features like the shape of flames, brightness distribution, vortex structures in the flow field, and streamline patterns, and then judge whether the combustion state is normal or not. Moreover, they can predict key combustion parameters (such as combustion speed, flame propagation angle, fluid turbulence intensity, etc.), enabling an in-depth analysis of the interaction process between combustion and fluid. Recurrent neural networks (RNN) and their variants (long short-term memory networks, LSTM, and gated recurrent units, GRU) are suitable for processing sequential data, such as sequences of pressure, temperature, and fluid velocity changes in the engine cylinder over time. They can capture time-dependent relationships in the data sequences and predict the dynamic changes of in-cylinder parameters (such as the pressure peak in the next working cycle, temperature trends, evolution of the fluid velocity field, etc.), providing a basis for real-time control of the in-cylinder fluid process of the engine.

The advantages of neural networks lie in their ability to learn complex nonlinear relationships and time series patterns, their strong fitting ability, and their capacity to handle multiple types of fluid mechanics-related data and build high-precision prediction models. However, their training requires a large amount of data, the selection of model structures and the adjustment of hyperparameters are complicated, insufficient data is prone to lead to overfitting, and the interpretability of the models is poor, making it difficult to intuitively understand the internal decision-making processes and the meanings of parameters.

Ensemble learning integrates multiple machine learning models and combines their prediction results to improve the overall performance, playing a crucial role in the research on in-cylinder fluid mechanics of Atkinson cycle internal combustion engines. In the research on engine cylinders, it can improve the prediction accuracy and stability and reduce the uncertainty of models. Random forest (RF) is constructed based on decision trees, and the comprehensive prediction results of multiple decision trees are used for classification or regression. For example, in engine fault diagnosis, RF integrates the diagnosis results of multiple decision trees for different fault types (such as combustion faults, fluid leakage faults, etc.) to improve the accuracy of fault diagnosis. Each decision tree is trained based on different subsets of in-cylinder data or features (including features related to fluid mechanics), taking advantage of data diversity. Gradient boosting trees (GBT) iteratively trains decision trees, and subsequent decision trees focus on the prediction errors of previous models to gradually improve the overall performance of the model. In the joint prediction of the in-cylinder combustion state and fluid state of the engine, GBT combines multiple weak prediction models (decision trees) into a strong prediction model to accurately classify and predict the combustion state (such as normal combustion, incomplete combustion, knock, etc.) and the fluid state (such as abnormal flow velocity, abnormal pressure fluctuation, etc.). Stacking takes the outputs of different types of machine learning models (neural networks, decision trees, support vector machines, etc.) as new inputs and makes comprehensive predictions through meta-models (linear regression or another neural network). In the research on engine cylinders, stacking combines the ability of neural networks to learn the complex relationships of in-cylinder fluid data and the interpretability advantage of decision trees to improve the comprehensive performance of engine performance and the prediction of combustion and fluid states, providing a reliable decision-making basis for optimizing the in-cylinder fluid mechanic's process. The advantage of ensemble learning lies in effectively improving the performance and stability of models, reducing the risk of overfitting, combining multiple models to take advantage of their complementarity, and accurately predicting and analyzing complex in-cylinder fluid problems of engines. However, its computational cost is high as it requires training in multiple base models, and the selection of base models and combination methods needs experience and experiments. If the performance of base models is poor or their correlations are high, the improvement in the performance of the integrated model may not be significant [8].

6. Advantages of Machine Learning Methods in the Research on Atkinson Cycle Internal Combustion Engines

Compared with traditional CFD methods, machine learning methods have significant advantages in computational efficiency. Clustering can quickly classify a large amount of experimental or preliminary simulation data and group the working states of engines under different operating conditions according to similarity. For example, when dealing with in-cylinder pressure, temperature, and fluid velocity data under different rotational speeds and loads, it can identify typical working patterns in a short time without the need for CFD to conduct detailed fluid mechanics solutions for each operating condition. Modal decomposition extracts the main characteristic patterns of data and reduces the data dimension. When studying in-cylinder flow, it can quickly decompose complex pressure and velocity signals into main modes, focusing on key physical processes and saving time. After training, neural networks can quickly predict engine performance and states. With new input parameters (such as intake valve closing time, fuel injection strategies, etc.), performance prediction only takes a few milliseconds to a few seconds, far exceeding the computational efficiency of CFD, which usually takes a long time. Ensemble learning combines the advantages of multiple models to quickly judge the combustion and fluid states of the engine. For example, it can judge the high-efficiency combustion operating conditions and the optimal fluid flow states, and its computing speed is much faster than that of CFD, as it makes comprehensive evaluations based on the trained sub-models.

Machine learning methods perform excellently in data processing and information mining. Clustering can discover hidden structures and patterns in data, mining similar combustion patterns, abnormal combustion patterns, and corresponding fluid mechanics characteristics of the Atkinson cycle internal combustion engine under different operating conditions. For example, it can find groups of operating conditions with different degrees of delayed closing of the intake valve but similar combustion efficiencies and analyze the similar flow laws of the fluid in the cylinder, providing new ideas for the optimized design of the engine. Modal decomposition simplifies data into easily understandable modal components and reveals the physical mechanisms behind the data. For example, it can separate the fluid components related to combustion and mechanical vibration in the in-cylinder pressure signal and clarify their influences on in-cylinder fluid mechanics, which is difficult to achieve directly through CFD post-processing[7]. Neural networks have a powerful data fitting ability to learn the complex nonlinear relationships between inputs and outputs, mining the comprehensive influences of various parameters on engine performance and emissions, considering the coupling effects of in-cylinder fluid mechanics factors, while CFD can often only analyze the local influences of changes in individual parameters. Ensemble learning combines the information mining capabilities of different models to improve the efficiency of data understanding and utilization. For example, by combining the feature importance evaluation of decision trees and the complex relationship learning of neural networks, it can comprehensively mine useful information from the data of the Atkinson cycle internal combustion engine, including key areas of the flow field, the correlations between fluid parameters and combustion and performance [9].

7. Conclusion

The Atkinson cycle internal combustion engine has broad application prospects in the field of hybrid vehicles, and its high efficiency and energy saving are in line with the development trend of automobiles. Fluid mechanics provides theoretical and technical support for its in-cylinder design, but traditional fluid mechanics research methods have limitations. Machine learning methods bring new opportunities and optimization approaches to the research on Atkinson cycle internal combustion engines. Technologies such as clustering, modal decomposition, neural networks, and ensemble learning have significant advantages in computational efficiency, data processing and information mining, model adaptability [10], and flexibility, effectively making up for the deficiencies of traditional methods and providing powerful tools and means for engine performance improvement, optimized design, and fault diagnosis. With the continuous development and improvement of machine learning technologies[11], their applications in the research on Atkinson cycle internal combustion engines will be deeper and more extensive and are expected to promote further innovation and development of this technology.

References

[1] Z. Wang, G. Zhang, T. Sun, et al., "Development and application of a fluid mechanics analysis framework based on complex network theory," *Computer Methods in Applied Mechanics and Engineering*, vol. 435, pp. 117677, 2025.

[2] G. Zheng, G. Yanlin, C. Lingen, et al., "multi-objective optimization of an endoreversible closed Atkinson cycle," *Journal of Non-Equilibrium Thermodynamics*, vol. 49, no. 1, pp. 73-88, 2023.

[3] Z. Yao, C. Chen, T. Wang, et al., "Effects of gas distributor on hydrodynamics in gas–liquid bubble column by visual experiments and CFD simulations," *Chemical Engineering Journal*, vol. 504, pp. 158476, 2025.

[4] K. L. Narayana, V. K. Reddy, and S. Venkataramana, "Analysis of Land Use and Land Cover Using Machine Learning Algorithms on Google Earth Engine for Munneru River Basin, India," *Sustainability*, vol. 13, no. 24, pp. 13758-13758, 2021.

[5] T. G. Gwak, J. U. Hwang, and H. J. Kim, "Clustering of shoulder movement patterns using K-means algorithm based on the shoulder range of motion," *Journal of Bodywork & Movement Therapies*, vol. 41, pp. 64-170, 2025.

[6] Z. Fengnian and L. S. D. H., "Applications of machine learning to the analysis of engine in-cylinder flow and thermal process: A review and outlook," *Applied Thermal Engineering*, vol. 220, 2023.

[7] Y. Sun, H. Song, H. Gao, et al., "Interpretable tool wear monitoring: Architecture with large-scale CNN and adaptive EMD," *Journal of Manufacturing Systems*, vol. 78, pp. 294-307, 2025.

[8] P. R. Vasquez, A. A. A. Lasserre, M. R. Palacios, et al., "Artificial neural network (ANN) in forecasting of poverty line and economic-energetic efficiencies into the maize-based agroecosystems," *Archives of Agronomy and Soil Science*, vol. 70, no. 1, pp. 1-17, 2024.

[9] S. Vikas, J. S. Preet, H. Kamal, et al., "Machine learning approach for predicting the machining characteristics of additive manufactured ABS material," *Proceedings of the Institution of Mechanical Engineers*, vol. 239, no. 1, pp. 129-144, 2025.

[10] J. Jian, R. Z. Lu, J. Liu, et al., "De-mixing variational mode decomposition and its application on operational modal analysis in the presence of closely spaced modes," *Measurement*, vol. 242, no. PC, pp. 116057, 2025.

[11] H. Qureshi, U. S. Arfeen, A. W. Khan, et al., "Python-based machine learning procedure for radiative Sutterby multiple-hybrid nanofluid flow comprising features of chemical processes," *Journal of Radiation Research and Applied Sciences*, vol. 18, no. 1, pp. 101258, 2025.