Honda R&D

Technical Review

2026

Vol.38



Comment on the Cover

The cover theme for Vol. 38 is “Layered Canyon.”
Multiple layers of strata in natural canyons form beautiful, harmonious landscapes. Inspired by such landscapes, we created this image to show individual challenges and skills as layered strata that come together into a single, powerful current driving our company forward.
The bright oranges and reds represent the commitment and fire of individual engineers. We produced this cover to express the hope that this passion will grow ever greater and further accelerate Honda’s pursuit of the future.



Design Center
Yusuke KAWAUCHI

Technical Review

2026

Honda R&D

CONTENTS

Contribution

Application of Bayes' Theorem to Statistical Analysis of Small Sample Data . . . . . Shinsuke SAKAI 1

Realization of Strategic Color Lineup . . . . . Masaharu YOSHIMURA 11

Introduction / Case Study

Development of Electric Motocrosser CR ELECTRIC
. . . . . Seiji KATO, Hiroyuki AKIYOSHI, Makoto KITAOKA
Masayuki KINOSHITA, Takuya FURUYAMA 23

Demonstration Experiment of Electric Outboard Motor for Small Boat
. . . . . Takamasa TAKESHIGE, Kenta KATO, Shohma HIGANO 31

Development of Honda E-Clutch System for Motorcycles
. . . . . Junya ONO, Tatsuya RYUZAKI, Masazumi NAITO 38

Technology for Lower Limit Prediction of Lithium-ion Batteries Performance after Degradation
Employing Diffusion Process Model . . . . . Kaoru OMICHI, Kotone SUZUKI
Hideki NAGATSUKA, Michihiro YASUDA 48

Lunar Night Power Supply: Circulative Renewable Energy System Capable of Continuous Power Generation
. . . . . Ryosuke INANISHI, Hayato DAIMON, Kazunori FUKUMA
Takumi KAWANO, Toru HONDA 56

New Transmission Mechanism Technology Utilized in Prosthetic Legs
. . . . . Hiromi ONO, Kei SHIMADA, Yutaka HIKI
Seiji ONOZAWA, Fumihito GOTO 64

Technical Paper

Construction of Turbulent Component Decomposition Method Using Time Filtering with High-speed PIV
and Application to Hybrid Engine Development
. . . . . Yasuhiro OKURA, Makoto SEGAWA, Hiromi ONIMARU
Seisuke SATO, Yasuhiro URATA 72

Numerical Study on Enhancing Power and Thermal Efficiency of Large Motorcycle Gasoline Engine
Using Pre-chamber Jet Combustion Simulated by Combustion CFD
. . . . . Hirokazu ANDO, Takumi TANAKA
Kengo TOMIZAWA, Yosuke INOUE 85

Verification of HC Trap Materials and Catalysts to Reduce Exhaust Emissions during Cold-start in V6 Engine
. . . . . Hiroki TAKEORI, Koji NEMOTO, Yuki OKU
Ryohei IKUTOMO, Keita NAKAO, Toyohiro USUI 97

Integrated Simulation Technology for Design, Manufacturing, and Materials to Reduce Surface Distortion of
Automotive Outer Panels . . . . . Toru FURUSAWA, Yoichi TOYOOKA, Eiichiro YOKOI
Shigeru TAKAMI, Masao UCHIYAMA, Yoshinori TANAKA 109

Development of Evaluation Method for Automotive Outer Panel Surface Distortion
Using Curvature Difference and Application to Body Design
. . . . . Yoichi TOYOOKA, Toru FURUSAWA, Eiichiro YOKOI
Shigeru TAKAMI, Masao UCHIYAMA, Yoshinori TANAKA 117

Short-time Prediction Tool for Flow Fields around Side Mirrors Using Pix2Pix GAN
. . . . . Takuro HASHIMOTO, Koji OKABE
Takeshi YAMATODA, Takafusa KIDO 123

Quantitative Measurement Scale for Effectiveness of Adapted Vehicle
. . . . . Arisa ENDO, Toshihiro KOYAMA
Keiichi SUNAMI, Tetsuya GOMI 131

Retraining-free Customizable End-to-end Automatic Speech Recognition
. . . . . Yui SUDO, Shakeel MUHAMMAD, Naoaki SUMIDA 139

Enhancement of Water Resistance by Methanation Catalysts with LDHs Structure
. . . . . Hideaki YONEDA, Kazuki YANAGISAWA
Yohei IIJIMA, Junpei YOSHIDA 146

Application of Bayes' Theorem to Statistical Analysis of Small Sample Data

Shinsuke SAKAI, Ph.D.*

1. Foreword

In recent years, a probabilistic approach is increasingly needed as mechanical products are expected to provide higher functionality. This is not unrelated to the fact that, as mechanical products provide higher functionality, the consequences of any accident tend to be more serious. A typical example would be an on-board composite hydrogen tank stored under high pressure. In his examination of standards development, for example, G. W. Mair proposes to introduce a probabilistic approach to the consideration of safety margins[(1)]. The present author introduces the concept of a probabilistic approach, which is often discussed in terms of GTR and other standards, as an approach to safety standards for composite hydrogen tanks for fuel cell vehicles[(2)]. The current paper provides an overview from the perspective of reliability design that the author has experienced outside the automotive field. This paper has mentioned the historical background, details of the method, advantages and disadvantages, examples of use in other fields, and the possibility of use in fuel cell vehicles. Following this, an important action would be to start an evidence-based discussion of safety margins for bursting strength, with specific application of a probabilistic approach to the current situation in Japan. An essential part of this process is the organization of statistical data. The basic statistics involved include the mean value and variance of the bursting pressure of hydrogen tanks. However, obtaining basic statistics of hydrogen tank bursting strength experimentally is not easy in terms of cost and labor. Moreover, considering the severity of the results, the need for a higher degree of safety performance means that the probability of bursting is extremely small, which also makes it a challenge to take a large sample size. If a very large amount of data could be obtained, it might not such a challenge to evaluate the basic statistics of the population with a high degree of accuracy based on so-called sample statistics. However, in the case of composite hydrogen tanks, if large sample sizes are not expected, then the application of sample statistics is generally a challenge. In such cases, there is an approach called Bayesian statistics that is effective for evaluating small-sample statistics. For example, Bayes' theorem is used for the statistical processing of maintenance data in risk-based inspections of plant pressure equipment[(3), (4)]. This is related to the fact that one cannot obtain large samples of maintenance data. This paper summarizes the basic ideas of Bayes' theorem as well as examples, applications, and precautions related to applying the theorem.

2. Characteristics of Bayesian Statistics

Before explaining Bayes' theorem in detail, this chapter discusses its characteristics. When an evaluation of the basic statistics involved in the breakage of a mechanical product is required, it is advisable to first consider applying Bayes' theorem in the following situations.

1. The event in question is very infrequent
2. Causing the event by experimentation would require excessive cost and labor
3. The event may be transitory

In such situations, one cannot easily increase the sample size rapidly and it would be challenging to apply sample statistics that assume a large sample size. Additional information follows concerning the third situation above. For example, when considering plant maintenance, the damage status of piping and pressure vessels depends also on operating conditions such as internal pressure. If operating conditions are changed at some point in time, the basic statistics related to the probability of damage may also subsequently change. In addition, design specifications are likely to be modified during the development phase of products like hydrogen tanks. In this case as well, while the basic statistics on damage probability may be predicted to change, the sample size will never increase if sampling restarts each time there is a change in design specifications. Even in challenging situations like these, Bayesian statistics has the following characteristics that allow it to handle such situations to some extent.

1. For statistical processing, it uses not only the information that there was damage, but also information that there was no damage.
2. In identifying the statistical distribution, this method updates the certainty of the parameter from a small sample size.
3. The certainty of the parameter is determined at the starting point by such factors as experienced values, literature, and subjectivity.

In sample statistics, a parameter such as the mean is

* Visiting professor, Yokohama National University

evaluated as a single value by statistical processing from sample data. In Bayesian statistics, however, the parameter is not evaluated as a single value, but is rather treated as something ambiguous; that is, a level of certainty. In extreme cases, certainty may be determined subjectively as a starting point, which is why it is called subjective probability. Although it may seem unwise to allow subjectivity to enter into the statistical process, in Bayesian statistics, the statistics are updated each time sample data is obtained, so that certainty gets dependably closer to the correct answer, and is thus expected to eventually converge to the true value. Even if the event is transitory as described above, Bayesian statistics can track the transition of the event, since processing is accomplished by updating. It is wise, however, to have a good understanding of the principles of Bayesian statistics, because if not applied appropriately, the efficiency of such convergence will not occur.

3. Basic Principles of Bayesian Statistics

3.1. Approach

If the object of interest is a pressure tank, assume that its bursting pressure is normally distributed. The probability density function with a normal distribution is expressed as:

$$f(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (1)$$

where x is the random variable for bursting pressure, and there are two parameters: mean μ and variance σ^2. $f(x|\mu, \sigma)$ is a probability density function with the parameters μ and σ. A normal distribution with mean μ and variance σ^2 is sometimes expressed as $N(\mu, \sigma^2)$. To identify the basic statistics of bursting pressure is to find μ and σ^2. If a sufficiently large sample size n can be obtained, the sample mean $\hat{\mu}$ and unbiased sample variance $\hat{\sigma^2}$ can be evaluated from the following equations.

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (2)$$

$$\hat{\sigma^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 \quad (3)$$

If a larger sample size can be obtained, these parameters can be evaluated with greater accuracy. When this is not expected, however, one should consider using Bayesian statistics. Here, for simplicity, the variance σ^2 is assumed to be known, and only the mean value μ is unknown. The way to estimate this is explained as follows. The approach can be applied even if both parameters are unknown, but that process will be discussed later.

The evaluation using Eq. (2) establishes one sample mean value, and there is no room for subjectivity in the procedure. For that reason, it can be called an objective procedure. Bayesian statistics, in contrast, adds uncertainty to the estimated quantity μ. In other words, it does not assign a single definitive value. If new experimental data is obtained, the result indicating whether a pressure vessel did or did not burst under certain conditions is an objective fact. Bayesian statistics involves updating the uncertainty of an estimated quantity based on a few objective facts. This is called Bayesian updating. While one can say that Bayesian updating does not suddenly yield the correct answer for the parameter, it does approach the correct answer. Uncertainty before updating is called the prior distribution, and uncertainty after updating is called the posterior distribution. Figure 1 illustrates the approach conceptually. The vertical axis is the probability density.

For example, when the left side of Fig. 1 is the prior distribution of uncertainty and μ_0 is the correct answer for the parameter, the figure shows how, based on the results of a single burst test, the posterior distribution of uncertainty gathers near μ_0 based on the Bayesian update results. The greater the number of burst test samples, the better the convergence results that can be expected, and increasing the number of times tests are repeated will yield results asymptotically approaching the correct answer. The first prior distribution to be set is called the initial prior distribution. If the initial prior distribution is off the mark, there is unlikely to be good convergence performance, so it is important to set it carefully. The initial prior distribution can be set based on the results of similar products in the past when available, on the opinions of experienced persons, or on information found in the literature. This initial prior distribution is sometimes called subjective probability, because in any case it is not given definitively but is determined subjectively in a sense. If starting with no information beforehand, the uncertainty is distributed over a wide area, and is called an uninformative prior distribution. In Fig. 1, the prior distribution is set as a continuous distribution, but it can also be set as discrete values, as in Fig. 2. That is, instead of expressing the

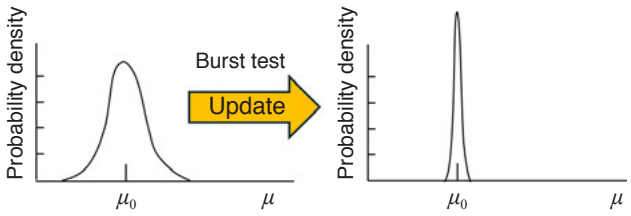


Fig. 1 Principle of Bayesian update for continuous distribution

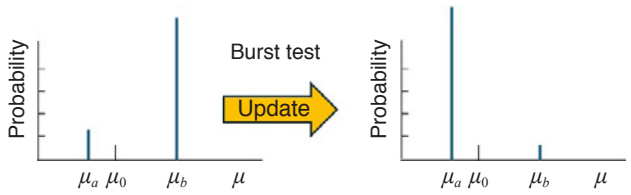


Fig. 2 Principle of Bayesian update for discrete values

uncertainty of the mean as a continuous distribution, it is estimated as a discrete value in the form of either μ_a or μ_b.

In this figure, in the prior distribution shown on the left, uncertainty confidence values for the μ_a and μ_b values estimated are set discretely. In this example, μ_a has a lower confidence level than μ_b, even though μ_a is closer to μ_0, the true value. After sample values are obtained, the Bayesian update results in a higher confidence level for μ_a, and the true value can be determined to be closer to μ_a.

3.2. Basic Principles

3.2.1. Estimating discrete values

Suppose that a parameter one wishes to estimate by Bayes' theorem, such as a population parameter, is θ. When the parameter is θ, the probability density function for an observed event x, such as bursting, is expressed as $f(x|\theta)$. In ordinary sample statistics, the sample value $\hat{\theta}$ is evaluated from sample data by statistical procedures, leaving no room for subjectivity. Bayesian statistics, on the other hand, sets a value for θ subjectively and also assigns a confidence level for the uncertainty of that value. Here, for simplicity, two discrete values (θ_1 and θ_2) have been set. P_1 and P_2 represent the confidence level for each value at the start. Bayes' theorem assumes that observed events necessarily arise based on the value of θ_1 or θ_2. This is to say, the condition $P_1 + P_2 = 1$ needs to be satisfied. For example, if there is a high degree of confidence that the parameter is θ_1, one may set $P_1 = 0.8$ and $P_2 = 0.2$. Observed events, such as bursting, are now written as A. At this time, the logical relationship can be represented by a Venn diagram as in Fig. 3. Here, C_1 indicates that the observed event is caused by θ_1 and C_2 indicates that it is caused by θ_2. From now on, the probability of an event will be expressed as P (event). The probability that the causal event is C_1 and that the event A is observed can be expressed by the following equation using the basic properties of a Venn diagram.

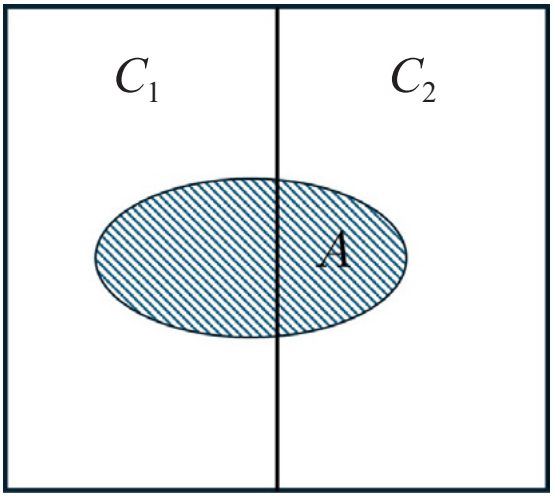


Fig. 3 Venn diagram

$$\begin{aligned} P(C_1 \wedge A) &= P(C_1)P(A|C_1) \\ &= P(A)P(C_1|A) \end{aligned} \tag{4}$$

where $\wedge$ denotes the AND logical relationship and $P(A|C_1)$ denotes the conditional probability that A is observed under the condition of cause C_1. By transforming Eq. (4), one can obtain:

$$P(C_1|A) = \frac{P(C_1)P(A|C_1)}{P(A)} \tag{5}$$

$P(A)$ can be written as:

$$P(A) = P(C_1)P(A|C_1) + P(C_2)P(A|C_2) \tag{6}$$

Therefore, substituting this into Eq. (5) can be expressed as:

$$P(C_1|A) = \frac{P(C_1)P(A|C_1)}{P(C_1)P(A|C_1) + P(C_2)P(A|C_2)} \tag{7}$$

The components on the right side of Eq. (7) are given as below in the above case.

$$\begin{aligned} &P(C_1), P(C_2) \Leftrightarrow P_1, P_2 \\ &P(A|C_1), P(A|C_2) \Leftrightarrow f(x|\theta_1), f(x|\theta_2) \end{aligned} \tag{8}$$

Thus, all quantities on the right side are given quantities. On the other hand, $P(C_1|A)$ on the left side indicates the probability of C_1 under the constraint that there was an observed event A. It therefore indicates the degree of confidence in C_1 after incorporating the observed event. Thus, $P(C_1|A)$ is the posterior probability after the Bayesian update. $P(C_2|A)$ is found similarly:

$$P(C_2|A) = \frac{P(C_2)P(A|C_2)}{P(C_1)P(A|C_1) + P(C_2)P(A|C_2)} \tag{9}$$

$P(C_1|A)$ and $P(C_2|A)$ indicate that the degree of confidence in C_1 and C_2 has been updated from $P(C_1)$ and $P(C_2)$ based on the observed event A. In this manner, Bayesian statistics updates the degree of confidence based on observed events. When an observed event is subsequently obtained, this $P(C_1|A)$ and $P(C_2|A)$ can be used as prior probabilities for the next time, and the same procedure can then be used to continue updating. Thus, an event for which the degree of confidence is higher is used as a hint to estimate θ.

Here, a simple example is considered. This case supposes that the physical quantity to be measured has a normal distribution, with mean $\mu_0 = 0$ and standard deviation $\sigma_0 = 1$. Consider estimating a physical quantity that has this distribution from experimentally measured values. In this case, the Bayes' theorem described above is used to update each measurement. For simplicity, it is assumed here that the standard deviation $\sigma_0 = 1$ is known and the mean is to be estimated. First, the two discrete events in Table 1 are the subject of estimation. The prior probabilities (confidence) are also shown. These two discrete values and the initial prior probabilities are assumed to be set based on the knowledge of people with experience. The sum of the prior probabilities needs to be 1.0. In this case, even though the population mean of 0.0 is correct, a confidence level as low as 0.3 is given to

Table 1 Two discrete events

Symbol	Events	Prior probability
C_1	$\mu = -0.1$	0.3
C_2	$\mu = 1.5$	0.7

-0.1, which is close to the correct answer, while a confidence level of 0.7 is set for 1.5, which is farther from the correct answer. Next is a simulation of how this confidence level is updated by trial experimentation. First, generate 10 $N(0.0, 1.0^2)$ normal random numbers and observe how the Bayesian update is performed with experimental data 10 times.

The ith random number is expressed as $rnd(i)$. The Bayesian update when the first experimental value $rnd(0)$ is obtained is performed using Eq. (7). The right side of this equation is all given quantities, as below.

$$\begin{aligned} P(C_1) &= 0.3 \\ P(A|C_1) &= f(rnd(0)|-0.1, \sigma_0) \\ P(C_2) &= 0.7 \\ P(A|C_2) &= f(rnd(0)|0.3, \sigma_0) \end{aligned} \tag{10}$$

Here, the first observed event A corresponds to obtaining experimental value $rnd(0)$. The probability density function $f(\ )$ is given by Eq. (1). This results in obtaining updated confidence value $P(C_1|A)$. Similarly, the C_2 updated confidence value $P(C_2|A)$ is obtained by Eq. (9). Furthermore, for the next experimental value $rnd(1)$, the next update can be performed as $P(C_1|A) \rightarrow P(C_1)$ and $P(C_2|A) \rightarrow P(C_2)$. The same procedure should be repeated after that. Figure 4 shows the change in posterior probability with Bayesian updating for the 10 experimental values. It shows that the C_1 confidence level, which is close to the true value, has risen, while the C_2 confidence level, which is far from the true value, has fallen. As a result, there is greater confidence that the true value is a C_1 event. Thus, Bayesian statistics is characterized as a method in which predictions and confidence levels are set flexibly and subjectively when there is not a sufficient amount of sample data in the initial stage, after which the confidence level of the parameter is updated each time data is obtained based on a small number of experimental data points.

In the above example, a low confidence level of 0.3 was given for an initial prior distribution of -0.1, which was close to the correct answer, but if this confidence level had been set more appropriately high at 0.7, the results would be as shown in Fig. 5. Comparing this to Fig. 4 shows that the confidence level for the correct answer converges earlier to 1.0. Thus, if the initial prior distribution can be set more appropriately, it will converge to the correct answer more quickly, but if it is off the mark, the degree of convergence will not be as good. As this can result in less efficiency, the key is knowing how to appropriately set the initial prior distribution by referring to experienced people, past case studies, and other factors.

3.2.2. Estimation of continuous distribution when mean is unknown and variance is known

In the example above, instead of estimating the parameters of the burst probability density function to be estimated, a scheme was used in which the burst probability was assumed to be one of two discrete values and the degree of confidence in those discrete values was updated with Bayes' theorem. To extend this to a more general estimation method, the parameters themselves of the burst probability density function of Eq. (1) could be estimated. There are two parameters: the mean μ and the standard deviation σ. This section considers using Bayes' theorem to estimate these parameters. Although the theorem can be used for both parameters, for simplicity of explanation, only the mean μ is considered here. To make parameter estimation versatile enough for general use, it is advisable to use a format that expresses the level of confidence as a continuous distribution and then updates the continuous distribution, rather than use a format that estimates μ as

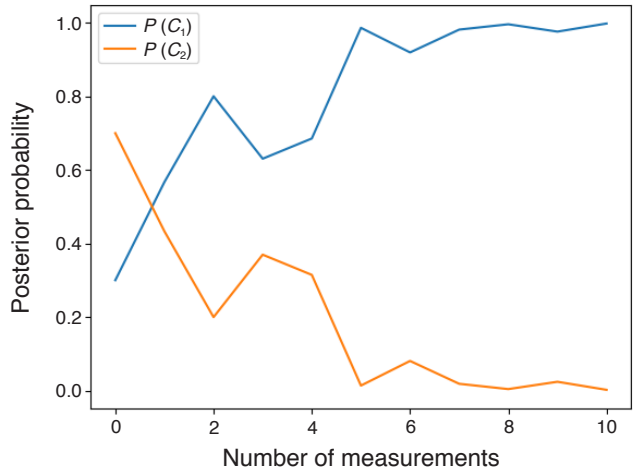


Fig. 4 Variation of posterior probability ($P(C_1)$ = 0.3, $P(C_2)$ = 0.7)

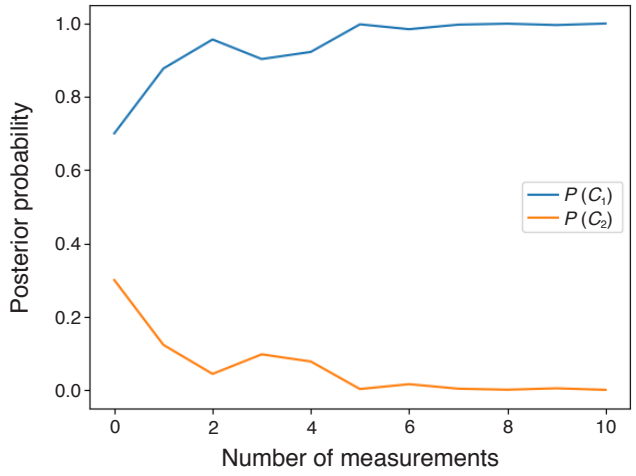


Fig. 5 Variation of posterior probability ($P(C_1)$ = 0.7, $P(C_2)$ = 0.3)

discrete values. That is to say, for a particular mean value μ, the burst probability density function is:

$$f(x|\mu) = \frac{1}{\sqrt{2\pi}\sigma} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \quad (11)$$

This is called a likelihood function. The probability density function $g(\mu)$ is set as a continuous distribution expressing the uncertainty of μ. The approach of setting a continuous distribution will be discussed later, but first, it is set as a prior distribution, based on experience, as in the case of the discrete method. At this time, the posterior probability distribution when the observed value $X = (x_1, x_2, ..., x_n)$ has been obtained can be expressed as follows based on the Bayesian update approach.

$$g'(\mu|X)d\mu = \frac{g(\mu)\prod_{i=1}^{n} f(x_i|\mu)d\mu}{\int_{-\infty}^{\infty} g(\mu)\prod_{i=1}^{n} f(x_i|\mu)d\mu} \quad (12)$$

Then, the posterior probability density function $g'(\mu|X)$ is derived based on Bayes' theorem. If the Bayesian update approach is applied each time an observed value is obtained, the continuous distribution function, which expresses the degree of confidence, can be expected to gradually converge to the true value. However, there is one great challenge with this. If one sets a completely arbitrary function as $g(\mu)$ for the prior distribution, the posterior distribution, $g'(\mu|X)$, would have no relation to the prior distribution. In this case, evaluating Eq. (12) becomes quite difficult. As computational capabilities have increased in recent years, it has become relatively easy to derive probability density functions numerically using the power of computers, even for complex expressions like that shown in Eq. (12). Markov Chain Monte Carlo Simulation (MCMC)[5] is a typical method for doing this. In addition, software like Stan and PyMC3 are readily available to assist in such calculations. One approach is to use such software for arbitrary distributions, but fortunately, by attaching certain conditions to the prior probability, this process is extremely easy. That is the conjugate prior distribution approach. From the sake of general-purpose use and ease of processing, it is considered better to use the conjugate prior distribution approach, which is explained below.

The challenge now is that the prior distribution $g(\mu)$ and the posterior distribution $g'(\mu|X)$ have completely different distribution forms. It is known, however, that if one adopts a particular distribution form as $g(\mu)$ corresponding to the distribution form of $f(x|\mu)$ that is being addressed, the $g'(\mu|X)$ obtained by applying Bayes' theorem will have the same kind of distribution form. However, the parameters would be subject to change. When such a relationship is established, it is said that conjugacy is established, and a prior distribution that was set in this way is called a conjugate prior distribution. It is known that if $f(x|\mu)$ has a normal distribution, the conjugate prior distribution for μ has a normal distribution. In other words, $g'(\mu'|X)$ also has a normal distribution.

The case here assumes that the random variable of interest follows a normal distribution $N(\mu, \sigma^2)$, and that the variance σ^2 is known. A continuous distribution is set for the confidence level for parameter mean μ and a Bayesian update is performed. In this case, one uses the fact that the conjugate prior distribution for the mean is a normal distribution. In other words, if a normal distribution is set as the prior distribution of level of confidence for the mean, the posterior distribution will also be normal. When $N(\lambda_1, \tau_1^2)$ is set as the prior distribution, it is derived that the posterior distribution when a sample of $x = (x_1, x_2, ..., x_n)$ is obtained is as follows[6]:

$$N(\lambda_2, \tau_2^2) \quad (13)$$

where

$$\hat{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\lambda_2 = \frac{n\tau_1^2\hat{x} + \sigma^2\lambda_1}{n\tau_1^2 + \sigma^2} \quad (14)$$

$$\tau_2^2 = \frac{\sigma^2\tau_1^2}{n\tau_1^2 + \sigma^2} \quad (15)$$

In other words, the two parameters of the posterior distribution that is a normal distribution can be easily evaluated from Eqs. (14) and (15). This posterior distribution can be used as the prior distribution the next time and thereafter.

The following is an example in which the correct answer for $f(x|\mu, \sigma)$ is $N(0, 1^2)$, which has a normal distribution. In this case, one starts the initial prior distribution for $N(\lambda, \tau^2)$ from $N(0.5, 1^2)$ and performs Bayesian updating. Assuming that 10 random number observations of $N(0, 1^2)$ are available at the time of measurement, one will see how the prior distribution changes as a result of Bayesian updating. The transition can be easily traced using Eqs. (14) and (15). Figure 6 shows the transition of the prior distribution for two Bayesian updates. It also shows that with each

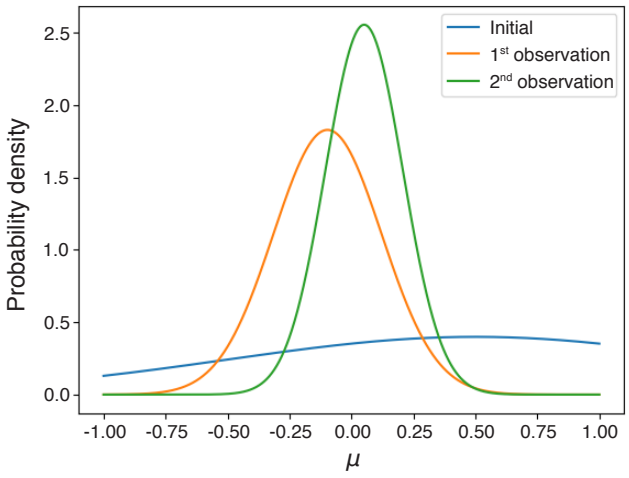


Fig. 6 Variation of prior distribution by Bayesian update

Bayesian update, the predicted prior distribution of the mean converges near the correct answer.

3.2.3. Estimation of continuous distribution when mean and variance are both unknown

This section considers a case in which both the mean μ and standard deviation σ are unknown and are estimated by Bayesian updating. In this case, the posterior probability distribution when the observed value $X = (x_1, x_2, ..., x_n)$ has been obtained is evaluated with the following formula with Bayesian updating.

$$g'(\mu, \sigma|X)d\mu d\sigma = \frac{g(\mu, \sigma)\prod_{i=1}^{n} f(x_i|\mu, \sigma)d\mu d\sigma}{\int_0^{\infty}\int_{-\infty}^{\infty} g(\mu, \sigma)\prod_{i=1}^{n} f(x_i|\mu, \sigma)d\mu d\sigma} \quad (16)$$

Here, $g(\mu, \sigma)$ is a two-dimensional simultaneous probability density function that is a conjugate prior distribution and is found to be expressed as in Eqs. (18) and (19)[7, 8].

$$g(\mu, \sigma) = \pi(\mu|\sigma^2)\pi(\sigma^2) \quad (17)$$

$$\pi(\mu|\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2/k_0}} exp\left[-\frac{(\mu - \hat{\mu_0})^2}{2\sigma^2/k_0}\right] \quad (18)$$

$$\pi(\sigma^2) = \frac{(\hat{\sigma_0}^2 v_0/2)^{v_0/2}}{\Gamma(v_0/2)} \frac{exp\left(\frac{-v_0\hat{\sigma_0}^2}{2\sigma^2}\right)}{(\sigma^2)^{1+v_0/2}} \quad (19)$$

Note that $\hat{\mu_0}$, k_0, v_0, and $\hat{\sigma_0}$ are parameters of the prior distribution, and in Eq. (19), the probability density function is an inverse gamma distribution expressed by the two parameters below.

$$f_{Igamma}(x|\alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}(1/x)^{\alpha+1} exp(-\beta/x) \quad (20)$$

$$\alpha = v_0/2 \quad (21)$$

$$\beta = \hat{\sigma_0}^2 v_0/2 \quad (22)$$

It is derived that the parameters of the conjugate prior distribution are updated by the following calculation:

$$k' = k + n \quad (23)$$

$$m' = [km + n\hat{\mu}]/k' \quad (24)$$

$$v' = v + n \quad (25)$$

$$v(s')^2 = vs^2 + (n-1)S^2 + kn(\hat{\mu} - m)^2/k' \quad (26)$$

where $\hat{\mu}$ is the sample mean and S is the unbiased standard deviation of the sample data.

Here is an example that shows how the continuous distribution of confidence level is updated by Bayes' theorem in a case where the correct answer for the population is $N(7, 2^2)$ and 10 samples are obtained twice by Monte Carlo simulation. First, as the initial prior distribution of the confidence level, one begins far from the correct answers for the parameters, starting with $\hat{\mu_0} = 4.0$, $k_0 = 1.0$, $v_0 = 2.0$, and $\hat{\sigma_0}^2 = 4.0$. Figure 7 shows the simultaneous probability density function for the initial prior distribution of the confidence level. The figure shows that the coordinate values indicating the peak of this distribution are far from the correct answer, $(7, 2^2)$.

Figure 8 shows the result for the simultaneous probability density function obtained by generating 10 random numbers and performing Bayesian updating with Eq. (16). This could be interpreted as meaning that the coordinate values of the peak position are close to $(7, 2^2)$, but the figure shows that the density values are widely distributed around the peak value. Figure 9 shows the simultaneous probability density function obtained by Bayesian updating after generating the next 10 random numbers. The figure shows that the coordinate values of the peak asymptotically approach the correct solution, $(7, 2^2)$, and the density values are concentrated in that area. If Bayesian updating continues to be performed in the same manner each time samples are obtained, it is expected that the simultaneous probability density function values will be concentrated around the correct parameters.

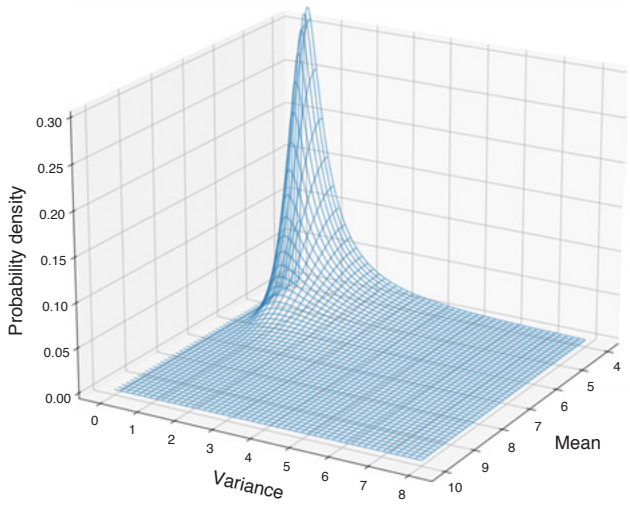


Fig. 7 Initial prior distribution

4. Application for Obtaining Statistical Data on Products with Design Changes

There are cases in which the product of interest is still in the development stage, but the statistical properties of the product are needed and statistical data needs to be obtained. If the product's manufacturing cost is high, it is already a challenge to increase the sample size. And if that product is in the development stage and its strength characteristics may change, there will likely be many challenges in obtaining the statistical properties of the product. Using Bayes' theorem is considered effective precisely in such cases. In other words, Bayes' theorem assumes that the initial prior distribution is as in Fig. 7. Bayesian updating can yield a posterior distribution like that in Fig. 8 in a single test, regardless of sample size. Then, assume that the design of the product will subsequently change. In general, when there are design changes during the development stage, it is considered

that in many cases the changes will not be substantial. In those cases, it is unlikely that the parameter values of the statistical properties will change dramatically. If so, the distribution in Fig. 8 can be used as a prior distribution for estimating the statistical properties of the next product. Then, Bayesian updating can be expected to appropriately track changes in the statistical properties. However, if there is an extreme change in statistical properties, using the Fig. 8 distribution as a prior distribution may not be effective, in which case the distribution is reset, and one can proceed with Bayesian updating.

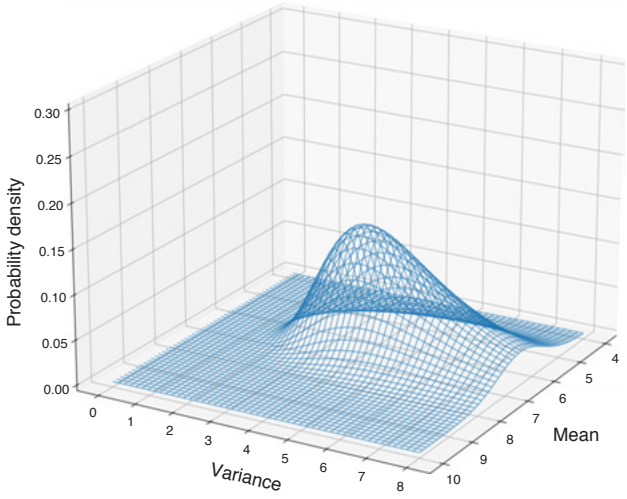


Fig. 8 Second prior distribution

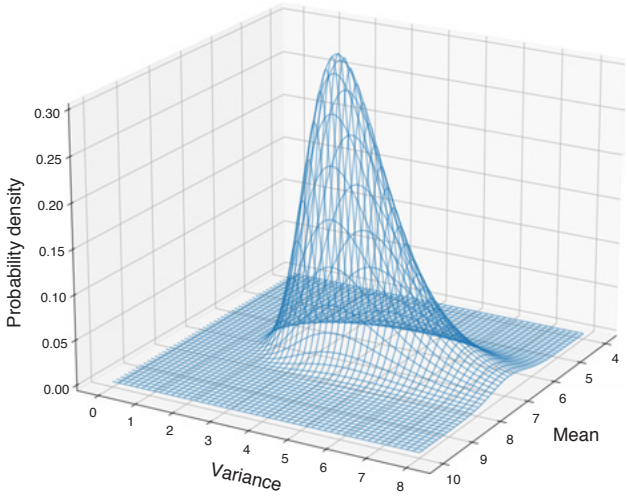


Fig. 9 Third prior distribution

5. Example of Application for Pressure Equipment Maintenance

To give a concrete example of an industrial application of Bayes' theorem, this section presents its application to evaluate the probability of damage to pressure equipment. This information is necessary to perform risk-based inspection of such equipment. Risk-based inspection is one approach to maintaining piping, pressure vessels, and other pressure equipment. The main type of maintenance method in the past was to systematically conduct inspections at a fixed time, in a fixed location, and using fixed inspection methods. To eliminate the irrational aspects of this approach, risk-based inspection continuously reviews inspection plans using risk as an indicator. The first guideline to appear in this area was API581[9] issued by the American Petroleum Institute. The corresponding Japanese standard is published by the High Pressure Institute of Japan[10]. This section presents an example of the application of Bayes' theorem to damage probability evaluation used in API581, while referring to reference[4].

Figure 10 shows the general procedural flow for conducting a risk-based inspection. Since risk values are used as indicators when developing inspection programs, there are steps in which risk assessment is conducted, which are shown inside the dashed line of the figure. In this part, risk values are analyzed, and risk is defined as the consequence multiplied by likelihood (in bold frame) of damage. As Bayes' theorem is used to evaluate the part relating to likelihood of damage, this is introduced below.

Corrosion of metallic materials by internal fluids is one

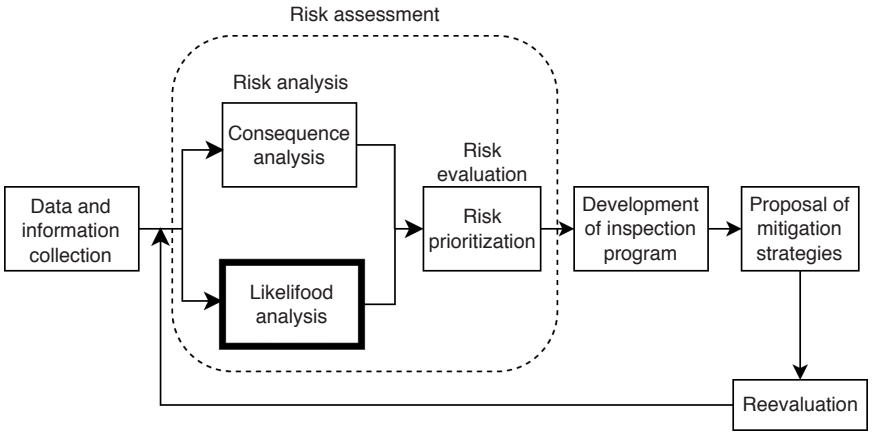


Fig. 10 Procedure of risk based inspection

Table 2 Definitions of damage states

Damage state	Definition
1	Damage does not exceed the predicted value of the damage model
2	Damage is at least equal to the predicted value of the damage rate model and at most twice that value
3	Damage is at least twice and at most four times the predicted value of the damage rate model

Table 4 Relation between inspection effectiveness and confidence

Damage rate state	Poorly effective	Fairly effective	Usually effective	Highly effective
1	0.33	0.5	0.7	0.90
2	0.33	0.3	0.2	0.09
3	0.33	0.2	0.1	0.01

Table 3 Allocation of initial confidence level

Actual damage rate range	Low reliability data	Moderate reliability data	High reliability data
Predicted rate or less	0.5	0.7	0.80
Predicted rate to two times rate	0.3	0.2	0.15
Two to four times predicted rate	0.2	0.1	0.05

of the main damage modes in the maintenance of piping and pressure vessels. As long as equipment is operated under the same operating conditions and environment, the rate of corrosion is expected not to change greatly because corrosion progresses steadily. Damage occurs when corrosion penetrates through the thickness of the plate, allowing internal fluid to leak. Therefore, one can evaluate the likelihood of damage; that is, the damage probability, based on the percentage of plate thickness that remains. If the residual wall thickness is measured by appropriate non-destructive testing methods and one assumes that corrosion progresses steadily, one can also predict future wall thinning. Based on this, one can perform a damage probability evaluation and review inspection plans while referring to the results of the risk assessment. In this case, it is important to accurately evaluate the corrosion rate by appropriate non-destructive testing methods. However, if measurements are made with less accurate non-destructive testing methods, the accuracy of the evaluated corrosion rate will decline correspondingly. In that sense, it is important to evaluate the effectiveness of non-destructive testing methods, which is determined by a combination of the type of damage targeted and the non-destructive testing methods. This is called inspection effectiveness, and the accuracy of the evaluated corrosion rate requires an evaluation of inspection effectiveness. However, inspection effectiveness cannot be expressed as an objective quantity, and inevitably, the evaluation will contain some subjectivity. By applying Bayes' theorem to the evaluation of corrosion rates, one can incorporate an evaluation of an ambiguous quantity like inspection effectiveness and then evaluate the accuracy of corrosion rates. A concrete explanation follows.

Step 1

It is necessary to draw some conclusions from past findings regarding the target damage mechanism. Typical examples include whole-surface corrosion and localized corrosion, but here it is decided to use whole-surface corrosion as an example. Although it is generally a challenge to obtain corrosion rate data in the early stages, if historical inspection data is available, it may be used as a rough estimate. If historical data does not exist, use the predicted wall-thinning rate based on the literature. Here, 0.2 mm/year is used as a possible prediction based on values found in the literature.

Step 2

Determine whether the level of corrosion rate reliability set in step 1 corresponds to low, moderate, or high. If there is complete field data, the confidence level can be classified as high, but if there is no existing data and it is inferred, as in this example, it is classified as low. The API guideline provides detailed classification criteria. What Bayes' theorem seeks to predict is not corrosion loss itself, but the accuracy of the thinning rate currently in use, which it classifies into three states. Perform Bayesian updating of the degree of confidence to determine which of the three states applies. The three accuracy states are shown in Table 2. Table 3 shows the initial degrees of confidence, or prior probabilities, for the three damage states used when applying Bayes' theorem.

Step 3

Perform Bayesian updating of the degree of confidence in the damage conditions defined in Table 2. The damage state indicates the degree of ambiguity of the evaluation result as found by inspection, with larger values indicating a greater degree of ambiguity. The degree of confidence depends on the inspection effectiveness described above. Damage evaluation accuracy depends on the non-destructive testing method used and the control volume ratio. A 100% volume inspection with a highly accurate non-destructive testing method would increase the degree of confidence in damage state 1 in Table 2. This measure of effectiveness is the inspection effectiveness, and the degree of confidence varies with the inspection effectiveness. For each level of inspection effectiveness, the relationship between the damage state and the degree of confidence that it can be accurately measured are given in Table 4. This

table appears to have been developed based on the opinions of experts and the like.

These values signify the conditional probability of each damage state for the subsequent Bayesian update. In other words, the table shows the probability that the predicted rate will be observed, on the condition that the current damage state will be one of the three states.

Step 4

Perform Bayesian updating to determine how the degree of confidence in each damage state changes as a result of performing inspections. That is, update the degree of confidence by using Bayes' theorem based on the results at the time of inspection. Apply Bayes' theorem according to the following equation:

$$P[A_i|B_k] = \frac{P[B_k|A_i]P[A_i]}{\sum_j P[B_k|A_j]P[A_j]} \quad (27)$$

where the symbols have the following meanings:

- $P[A_i]$ Prior probability: Degree of confidence that the damage state is i
- $P[A_i|B_k]$ Posterior probability: The updated degree of confidence when the event at the time of inspection was A_k (denoted as B_k in the equation)
- $P[B_k|A_i]$ Conditional probability: The conditional probability that damage state k will be observed when the current damage state is i. In the API, this is mapped to the inspection effectiveness. Table 4 shows values for $i = 1, 2, 3$; $k = 1$, but there are separate tables for $i = 1, 2, 3$; $k = 2$ and $i = 1, 2, 3$; $k = 3$ in the standard.

Assuming that during the initial inspection the predicted corrosion rate was observed using *Usually effective* non-destructive testing, the Bayesian update procedure would be as follows.

1. Since the prior probabilities are based on data from the literature, use the low reliability data from Table 3 as $P[A_i]$.
2. Use the values from the relevant part of Table 4 based on the fact that the *Usually effective* predicted corrosion rate was confirmed as the inspection effectiveness.
3. Perform evaluation assuming that the results at the time of inspection showed the damage rate to be as predicted, that is, B_1.

Therefore, the correspondence of the variables in Eq. (27) is as follows.

$$P[A_1] = 0.5,\ P[A_2] = 0.3,\ P[A_3] = 0.2$$
$$P[B_1|A_1] = 0.7,\ P[B_1|A_2] = 0.2,\ P[B_1|A_3] = 0.1 \quad (28)$$

Substituting these values into Eq. (27) yields the following:

$$P[A_1|B_1] = \frac{0.7 \times 0.5}{0.7 \times 0.5 + 0.2 \times 0.3 + 0.1 \times 0.2} = 0.814$$
$$P[A_2|B_1] = \frac{0.2 \times 0.3}{0.7 \times 0.5 + 0.2 \times 0.3 + 0.1 \times 0.2} = 0.140$$
$$P[A_3|B_1] = \frac{0.1 \times 0.2}{0.7 \times 0.5 + 0.2 \times 0.3 + 0.1 \times 0.2} = 0.046 \quad (29)$$

That is, Table 5 is obtained.

Table 5 Confidence level after first inspection

Damage rate state	After a usually effective inspection—
1	0.814
2	0.140
3	0.046

The result is that the degree of confidence that the damage state is 1 has increased. Thus, Bayes' theorem updates the degree of confidence in the current damage rate state by incorporating even a subjective measure of inspection effectiveness. If this procedure is followed at each subsequent inspection, it will be possible to determine the current damage rate state.

Step 5

Based on the predicted rate of 0.2 mm/year for damage state 1, a worst-case predicted rate of 0.4 mm/year for state 2, and a worst-case predicted rate of 0.8 mm/year for state 3, the probability of damage at the next inspection can be evaluated for each damage state. The procedure is omitted here because it is detailed, but the overall probability of damage is evaluated based on the probability of damage in each state and its degree of confidence in Table 5. The risk is then assessed by multiplying this and an evaluation of the consequence.

6. Conclusion

This paper has shown that using Bayes' theorem is effective when the objective is to obtain statistical properties of bursting strength for equipment like composite hydrogen tanks. The main reason for this is that it is so challenging to obtain large amounts of bursting strength data for reasons of cost and labor, and thus it is a challenge to evaluate bursting strength by sample statistics. With Bayes' theorem, it is possible to use small sample data to increase the degree of confidence in the parameter by Bayesian updating. However, setting the initial prior distribution, which is to say the initial confidence level, involves arbitrariness. Depending on the circumstances, therefore, there could be aspects that are hard to handle, such as setting the initial prior distribution subjectively. This paper used examples to explain Bayesian statistics, from the basics to application. The author hopes this paper will make at least a small contribution to obtaining small-sample statistical data.

References

(1) G. W. Mair, '*Safety Assessment of Composite Cylinders for Gas Storage by Statistical Methods*', Springer International Publishing AG 2017, doi: 10.1007/978-3-319-49710-5

(2) Sakai, S.: Working Toward Introducing Reliability Design, *Honda R&D Technical Review*, Vol. 37, p. 1-12

(3) Sakai, S.: Application of Bayesian Theorem to the Establishment of Probability of Failure Data for Equipments at Risk Based Inspection Planning, Principle of Bayesian theorem -1-, Journal of High Pressure Institute of Japan, Vol. 42, Issue 5, p. 284-290, (2004) (in Japanese), doi: 10.11181/hpi.42.284

(4) Sakai, S.: Application of Bayesian Theorem to the Establishment of Probability of Failure Data for Equipments at Risk Based Inspection Planning, Application in guidelines -2-, Journal of High Pressure Institute of Japan, Vol. 43, Issue 1, p. 22-27, (2005) (in Japanese), doi: 10.11181/hpi.43.22

(5) Gilks, W. R., Richardson, S., Spiegelhalter, D.: Markov Chain Monte Carlo in Practice, Chapman and Hall/CRC, 512p., (1995), doi: 10.1201/b14835

(6) DataArts, Bayesian statistics: https://www.data-arts.jp/bayesian_statistics/normal_distribution/known_variance.html, (accessed 2024/12/24) (in Japanese)

(7) https://qiita.com/raidofuwa/items/f665b96b0233674727be, (accessed 2024/12/24) (in Japanese)

(8) https://cafeunder.github.io/rosenblock-chainers-blog/2018/03/07/bayes-gaussian-mean-sigma.html, (accessed 2024/12/24) (in Japanese)

(9) American Petroleum Institute. *API publication 581, Risk based inspection: Base resource document, 1st edition*, 2000.

(10) Risk Based Maintenance Handbook, HPIS Z 107-1TR-4TR, High Pressure Institute of Japan, (2010) (in Japanese)

■ Author ■



Shinsuke Sakai

Realization of Strategic Color Lineup

Masaharu YOSHIMURA[1]

ABSTRACT

The coloring of motorcycles is an important element that conveys the concept and appeal of the model to customers and greatly influences their purchasing motivation. During the author's assignment in Vietnam, I devised a "coloring strategy" to highlight the individuality of each model by reviewing the coloring of all models from a comprehensive perspective. Based on this strategy, the upper-grade models set the trend in coloring, and the entry models followed. This enabled us to provide customers with high-satisfaction products. Building on the example in Vietnam, we applied the "coloring strategy" to 50 cc scooters in Japan and reviewed the coloring for each model. This provided each different customer for each model with high-satisfaction color options, while also achieving a product lineup with high business contribution. Although conditions vary by region and market, such as differences in market background and production facility capabilities, the "coloring strategy", which aims to optimize the product lineup by highlighting the character of each model from the "color" perspective, can be applied to the development of various models.

1. Importance of Coloring in Motorcycles

The coloring of motorcycles is an important element in conveying the concept and appeal of each model.

Some customers have found colors that they like on the products that were on display at dealerships and ultimately bought them because they "like the color." Another feature of motorcycles is that customers select a motorcycle based on the image of themselves riding the bike with considerations to how the bike's coloring would work with their outfits. In this age when it is commonplace for people to wear their favorite colors for clothing, bags, and other items, the importance of motorcycle coloring is increasing (Fig. 1).



Fig. 1 Motorcycle coloring and user fashion influencing one another

The first thought that comes to a motorcycle designer's mind when studying the coloring of a product is how and under what circumstances the customer will be riding. Will it be for commuting or for leisure, or for getting around town every day? What will be the riding style? The designer considers the customer's lifestyle as such based on preliminary research and information. After considering the colors that may suit customer lifestyles, the designer carefully selects the color schemes and submit them as a proposal to the sales division. The color lineup for each model is determined through discussions to help ensure the optimal color variation in line with the concept of each model.

The author, a designer in the motorcycle development division, was assigned to Vietnam in the early 2000s. There, I was given the opportunity to be in charge of the coloring of all models sold in Vietnam and developed a bird's-eye view of the entire color lineup. As mentioned earlier, "color lineup" usually refers to the color variations for a single model. However, as I became in charge of the coloring of all models in Vietnam, I should inevitably see the color lineup of all models at the same time.

At that time, Honda had three models being manufactured and sold in Vietnam. Despite each model having several color variations, they all gave similar impressions and each model's unique concept was not fully expressed through their coloring. With this observation, I decided to develop a strategy for coloring schemes that would best showcase the individuality of each model, with the goal of achieving overall optimization. This strategy enabled us to break away from the duplication of colors and

[1] Design Center

images that may had been caused by siloed model planning, and to create a color lineup in which each model has a clearer individuality. This has made it possible to provide customers with the pleasure of choosing colors.

The following is a description of the coloring strategy developed.

2. Show Model that Led to Coloring Strategy Development

2.1. Assignment to Vietnam and sense of crisis

Honda's research and development of motorcycles for The Association of Southeast Asian Nations (ASEAN) region is conducted by the Honda Research & Development Southeast Asia-Thai Office (HRS-T) based in Thailand. The coloring design for the Vietnamese models was handled by HRS-Singapore Office (HRS-S) at the time, however with the increase in attention on Vietnam as an emerging market, demand arose to establish a new research institute in Vietnam. In 2003, the author was assigned to Hanoi with the mission of establishing HRS-Vietnam Office (HRS-V).

I went out on the streets of Hanoi as soon as I arrived at my new assignment, and was struck by the abundance of motorcycles made by Japanese manufacturers, especially Honda, as well as their copies (Fig. 2). What surprised me was that many of the old and new motorcycles on the streets were of the same dark colors as in the past. There were a few brightly colored motorcycles on the road, but the number was limited.



Fig. 2 Street corner in suburbs of Hanoi in 2003

At two of the dealers we visited, we were able to talk to about 10 Honda vehicle users (Fig. 3). We got straight to the point and asked them about the coloring of their motorcycles, which had been of interest to us. Older users responded that Honda's colors are reliable and reassuring, while young users mentioned that Honda was their parents' generation's choice and that they themselves would choose a cool motorcycle made by a Japanese manufacturer other than Honda. We were aware that Honda motorcycles were chosen for their trusted toughness, but we now heard directly from customers that design and color were not among the reasons. In particular, the words of young users made us realize the urgent need to enhance the design and coloring of Honda models to make them more attractive, considering the rejuvenation of the consumers expected to occur in the near future.



Fig. 3 Example of Honda dealerships visited

2.2. Dealer convention and show model

Just as we were beginning to think about how to make the livery more attractive, an opportunity presented itself. I was to meet the president of Honda Vietnam, a Honda motorcycle manufacturing and sales company in Vietnam, and headed for the hotel lounge with the director of HRS-V. In the lounge, the president and sales manager of Honda Vietnam were being given a presentation on the dealer convention that was to take place two months later. The dealer convention is a business meeting to which all Honda dealers in Vietnam are invited to attend, where outstanding managers are awarded and company policies for the next fiscal year are explained. The lounge of the hotel used for the meeting was the very site where the dealers' convention was to be held.

In the presentation by the advertising agency in charge of planning and management of the event, colorfully illustrated characters to be used in the "TOI YEU VIET NAM" ("I love Viet Nam") campaign after the dealer convention were being introduced (Fig. 4).



Fig. 4 Characters of "I Love Viet Nam"

When I introduced myself and told them that I was a designer, they asked for my opinion about these characters. Having been thinking of ways to enhance the appeal of the

coloring, I proposed that real motorcycles with the coloring of each of the characters be produced and displayed at the dealers' convention. I had heard that the convention had consisted mainly of the projection of presentation materials and explanations. The proposal was based on the idea that displaying actual motorcycle models representing the new set of characters would bring out the pizazz of the models and make the appeal more impactful.

While the president and sales manager of Honda Vietnam showed interest and agreed to the proposal, the director of HRS-V opposed to it. The reason for this was that, based on his extensive experience as a motorcycle development manager, he knew the difficulties involved in producing show models, from procurement of materials and prototyping to finishing. They only had less than two months to complete show model for the dealers' convention.

In order to move forward with this proposal, a concrete plan had to be created to satisfy the director of HRS-V and the staff of the vehicle production department at the plant. So I decided to prepare a plan and schedule to complete the show models in two months.

2.3. Request for cooperation to the plant

The day after the meeting, I attended a regular meeting attended by all department managers at the plant. There, I explained the plan for the show model display for the dealers' convention and its production plan. They did not object to the plan since the company president had agreed to it but raised a question about the procurement of paint colors shown in the character illustration. I explained that we, HRS-V, would develop and prepare the paints, and secured their commitment to cooperate.

2.4. Paint development and show model production

In order to procure the necessary paints, I requested a Japanese expat in charge of local procurement to set up a meeting with a paint manufacturer. HANOI SYNTHETIC PAINT COMPANY is a local paint manufacturer that has a technical cooperation partnership with KAWAKAMI PAINT MANUFACTURING CO.,LTD. (hereinafter, KAWAKAMI PAINT), which supplies paints to Honda motorcycle factories in Japan. A meeting was arranged with the Japanese expat dispatched from KAWAKAMI PAINT .

Immediately, we explained the concept of displaying real motorcycles at the dealers' convention and asked for cooperation in the short-term development of paints. However, they immediately responded as impossible to meet the deadline in under two months. New color development usually starts with creating a satisfactory color through repeated color mixing, then tested for performance such as weather resistance, including color fading, and compatibility with the environment of the actual motorcycle usage, mainly with adhesion. Some of the performance checks require time on the order of a year, making it impossible to develop new colors on this schedule.

In response to this, we reevaluated the necessary items for this project and studied how to go about them. I first focused on the fact that most of the time during color mixing is spent on the exchange of paint color samples between paint manufacturers and designers, rather than on the actual process of color mixing. Another fact was that, since it is the sales staff who communicate the color direction desired by the designers, it is difficult for the nuances to be directly conveyed to the colorists. Therefore, I proposed that designers go directly to the manufacturer's color mixing site and judge color tones on the spot. This would shorten the time required for the development of the paints.

In addition, since new color development is based on the assumption of mass production application, it is common to use only pigments whose performance, such as weather resistance, has already been verified for short-term development. The number of pigments available is therefore limited, necessitating the time-consuming color mixing for any new materials. However, since the new colors are limited to show models, they do not need to have the performance required for mass production and we should be able to expect a wider availability of pigments and quicker color mixing.

I assured the paint manufacturer that, should the colors developed for this project be used in mass production, stipulated amount of time will be spent on its development. The manufacturer agreed to the conditions, thus marking the beginning of a development period of under two months.

After the color mixing at the paint manufacturer, the colors need to be checked again on the painting line at the factory, where the painting facilities are different. KAWAKAMI PAINT provided extensive support in this project, including sending a representative at the factory to help us reproduce the gorgeous colors on the painting line for the first time (Fig. 5).

In this way, the paint manufacturer, factory, and research institute worked together to overcome numerous hurdles, and the colorful show models were successfully completed (Fig. 6).

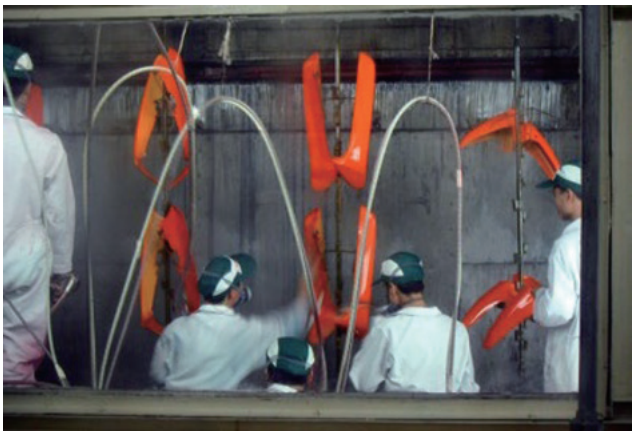


Fig. 5 Paint line at plant



Fig. 6 Completed show models

2.5. Dealer convention

The dealers' convention, where the show models were displayed, was a great success (Fig. 7). All eyes were on the display of motorcycles painted in new, never-before-seen colors. The visitors' reactions showed their strong interest in vibrant colors. This, combined with the sense of urgency I felt in our conversations with young users immediately after my arrival, became the turning point to seriously consider the need for immediate reflection to Honda Vietnam products in the future.

3. Strategic Color Lineup Concept and Mass Production Application

In September 2003 following the dealer convention, I was appointed by HRS-T, the research institute for the ASEAN region based in Thailand, to serve a dual role of color development manager and designer at HRS-V. Up until then, development was led in Thailand while coloring design was done in Singapore; with this instruction it became possible to develop products for Vietnam by the local staff and Japanese expats assigned there. This was indeed the realization of what Honda Vietnam has been yearning for, to have its own development capability for their country.

3.1. Development that leverages "real" Vietnam

To take on the development work for ourselves, we first went back to understanding the lifestyle of motorcycle users in Vietnam.

It turned out that the majority of users use their motorcycles for commuting or for work during the day, then for recreational use with their family at night. The daytime use scenario is common globally, but I could not guess what "recreational use with family at night" meant. So, I went out at night and saw hundreds of motorcycles with two, three, or even four family members, riding happily together (Fig. 8). After arriving at work the next day, I asked a local staff member about the situation. He told me that many families do not have air conditioners or televisions in their homes, and that they ride around town as a leisure activity, enjoying the cool of the evening with their whole family. There is no specific destination, but rather city cruising with family and friends. The motorcycle is both a reliable and sturdy



Fig. 8 Users happily cruising in city at night



Fig. 7 Colorful show models showcased at dealer convention draws large crowd

partner, and a vehicle for leisure that is loved by the whole family.

Looking again at the color lineup of Honda motorcycles from this perspective, we found that they all had a sober, albeit sturdy, appearance. Other brands, however, were also offering vibrant colors fit for leisure. The earlier comment by the young users, wanting a Japanese brand other than Honda for himself, apparently stemmed from the imprinted image created by such observations. If this is the case, we could appeal to families and young people that Honda motorcycles, often used for work, by incorporating stylish and trendy coloring design.

Based on this idea, I carried out a study on the overall color lineup while reconsidering the position of each model in the entire lineup, as summarized in the next section.

3.2. Coloring strategy: Logic for color lineup optimization

At the time, Honda Vietnam offered the following three models with their respective unique characters. Their characteristics are described below (Fig. 9).

- Future (110 cc, USD 1,556): Upper-grade model with a sporty appearance. Its sales have stagnated after 5 years since its launch.
- Wave α (100 cc, USD 856): Entry model with strong sales for its affordable price; many copies in market
- Super Dream (100 cc, USD 1,431): Iconic base model that built the foundation of Honda Vietnam; preferred by users not wanting changes.

Based on the market situation of each model and its ideal position in the product lineup, the following coloring development policy was established after clarifying differences in trend acceptance between urban and rural areas, time differences in the propagation of trends, and differences in disposable income.

(1) Upper-grade model Future is to be positioned as a trend-setting model and will be applied with two-tone colors and large graphics for a bold design expression.
(2) Entry model Wave α is to be positioned as a trend-following model and will be applied with the two-tone coloring and bold graphics similar to Future, a few months after Future's launch
(3) The iconic basic model Super Dream is to be positioned as a solid choice for long-term customers who want a peace of mind; more subtle graphics will reflect the styling trends

The thought here was that the maximum effect of the color lineup could be achieved by creating a system in which Future showcases trends and Wave α follows suit, so that customers can see at a glance what the coloring and graphics have in common.

This system of correlating the coloring of all models and developing, producing, and launching them in an effective order was collectively named the "coloring strategy.

3.3. What is to be gained with coloring strategy

In my thinking, the following was going to become possible through the coloring strategy.

(1) Customers of Future can enjoy the latest trend colors ahead of others.
(2) Customers of Wave α can obtain a motorcycle with latest trend coloring on the upper-grade model a few months after Future.
(3) Sales division can leverage the trends to establish a sales strategy and brand strategy.
(4) Production division can disperse the man-hours required to launch new models, which used to be concentrated at the end of the year, and equalize production by offsetting the launch of new colors for Future and Wave α even by a few months.
(5) Paint manufacturers and purchasing departments can systematically produce and purchase paints, thus stabilizing supply cost.
(6) Development division can benefit from product brand image enhancement through a design-led approach and equalized development man-hours.

3.4. Preliminary coordination to realize coloring strategy

I took this coloring strategy, a system I had developed through simulation, to the president of Honda Vietnam and the head of the sales department. The idea was welcomed:



Fig. 9 Honda's lineup and color variations circulated at 2003 dealer convention

"This is the first time we have seen a strategic presentation of only color changes, and the I love Viet Nam campaign color presentation at the recent dealer convention was a great success, so let's try product development based on this 'coloring strategy.'"

3.5. Challenges and method for mass production of coloring strategy

A development in accordance with the coloring strategy immediately began. The challenge was the development and development period for the new colors.

Although we knew that there were no new colors suitable for trend-setting at the Honda Vietnam plant, we checked again the situation at the paint shop. Nine colors were introduced to the paint line in the six years since the plant's establishment in 1998, six of which were used for the current three models. This was quite a limited number for an emerging market in the ASEAN region, where even a new research institute was being established.

There are two ways to develop a new color: one is to develop an original color exclusively for Vietnam in cooperation with a local paint manufacturer, and the other is to fine-tune a paint that has been mass-produced in Japan for use in Honda Vietnam. The former method has the advantage that designers can develop locally optimized colors that reflect the latest trends, but it requires a sufficient development period promised to the paint manufacturer. The latter method would use the data of paint formulation already proven in Japan and only requires fine-tuning to suit the factory facilities; however, it is undoubtedly more difficult to create colors that best works with the local mindset. Based on my conviction to develop colors in line with the coloring strategy in a short span of time, we chose to go with the latter method.

I immediately contacted a local paint manufacturer and asked them to order color samples of all colors that had been delivered to Honda motorcycle factories in Japan. Several colors were selected from the color samples received, and local color mixing began. Colors that could potentially be trendy in Vietnam were selected, and as with the development of colors for show models, the designers were present at the color mixing site to facilitate rapid color development. Because the original color mixing data made it easy to reproduce the color tones, the color mixing was completed at the fine-tuning level.

3.6. Early coordination with Japan

In parallel with ordering color samples, the materials research department at the motorcycle development center in Japan was consulted about new color development methods in the coloring strategy, as testing and approval by the materials research department are important for paint development. In addition, we explained the development situation to the head of development of our technical partner, a paint manufacturer in Vietnam, and requested further technical support from the Japanese side. The development of paints proceeded smoothly with the full cooperation of both the materials research department and the paint manufacturer, and the mass production of the new paints was on track as planned.

3.7. Realization of the coloring strategy and its result

Future, the first model under the coloring strategy, made its launch in the Vietnamese market with an extensive sales promotion program; it was with much fanfare, astonishing for a color change for a product that was launched six years prior. The prelude to this was a line-off ceremony held at the Honda Vietnam factory by shutting one of its production lines. This too was unprecedented for a minor model change involving only a color change, but I persuaded the president and plant manager that we should show the workers that Honda Vietnam is changing with the implementation of the new coloring strategy. (Fig. 10).

Honda Vietnam's sales manager was also very active in realizing plans such as renting out clubs in Hanoi and Ho Chi Minh City for presentation events and exhibiting in a corner of downtown (Fig. 11).



Fig. 10 Line-off ceremony



Fig. 11 Night club used for launch event; invited first-customers' "Future" parked at storefront

The Future in new colors, which was introduced with such a bang, received a favorable response from young fashion leaders, who thought it was cool and sporty (Fig. 12). However, the actual number of units sold increased only slightly. Sales personnel said that despite the stylishness of the coloring, it was “too new” and did not suit Future, and that they were not ready to buy it out of fear for a resale value risk and the eyes of others. This was an unexpected result, as we had thought that if it was cool, people would buy it. At the same time, there were also hints of Future customers seek to gain a sense of security that comes with the higher product price. It was not just about making it attracting. After hearing this report, the president of Honda Vietnam instructed the company to proceed with the coloring strategy and display Future models at storefront continue with the coloring strategy, so first put the Future on display at dealers. The number of shops displaying Future in their shops gradually increased. This was the start for customers to get accustomed to the new colors.

Six months later, Wave ZX, the second model under the coloring strategy, was launched as planned (Fig. 13).

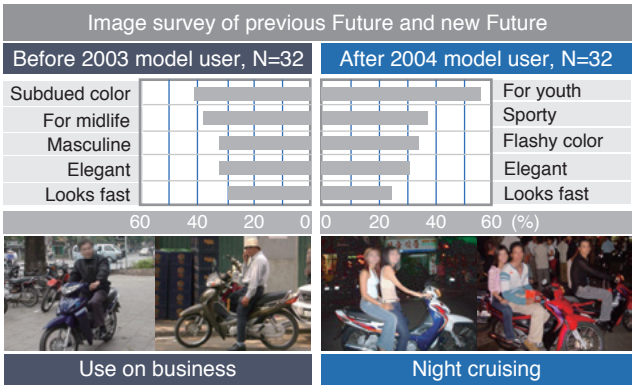


Fig. 12 Image survey of Future at launch

The differences between the existing Wave α and Wave ZX are a change in coloring to the Future image and the adoption of the name ZX to appeal to young users. In addition to the coloring and naming, the Wave ZX adopted front disc brakes, and was recognized as a new model, becoming the fourth pillar of the Wave ZX lineup following Super Dream, Future, and Wave α.



Fig. 13 Wave ZX with coloring inspired by Future

The sales volume of Wave ZX increased without affecting those of Future and Wave α. I believe that this is the result of a coloring strategy that created a virtuous cycle of trends by having the higher-end models initiate the trend and the lower-priced models follow them. At the time of my assignment, I was told by a young user that Honda was the choice of their parents’ generation and that, if they were to purchase one for themselves, they would choose a Japanese manufacturer other than Honda. However, after the implementation of the coloring strategy, the image of Honda motorcycles was greatly rejuvenated. Honda has taken the first step toward becoming a brand that is chosen by people who say Honda’s designs are cool because of their coloring (Fig. 14).

A year and a half had passed since I had been assigned there.

Future, the trend-setting model, underwent a series of model changes, including Future II and Future Neo, accompanied by changes in appearance (Fig. 15).



Fig. 14 Honda dealership with colorful models in 2006



Fig. 15 Future, Future II, and Future Neo achieved both evolution and continuity of graphics

Each model has a continuity in graphics to combine the graphics of trend to combine with a sense of security that the customers could recognize the Future at a glance. As Wave ZX and its evolutionary version, Wave RS, became hits by capturing trends, Future's position as a trend-setting model solidified. Despite its higher price range, Future gradually but steadily established itself in the market, recording 281% sales volume in 2006 compared to 2003. Fig. 16 shows the sales volume of each model.

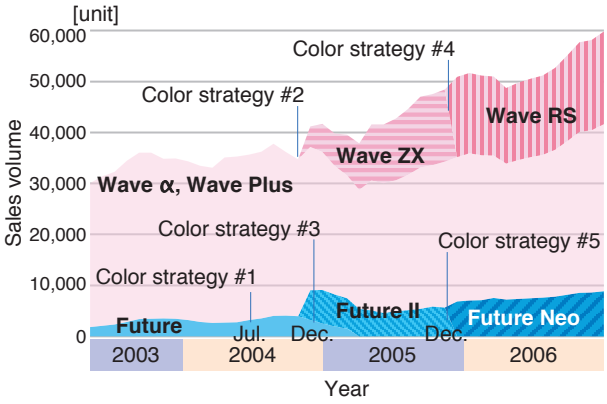


Fig. 16 Sales volume for each model of Honda Vietnam

Incidentally, the names of the model were taken from a provisional name that I used when I first started working on the coloring strategy to link the models, and they became official product names.

4. Applying Evolved Coloring Strategy to 50 cc Scooters in Japan

In 2009, the author returned to the development base in Japan and was appointed as the person in charge of color change development of the 50 cc Dio scooter for the Japanese market.

At the time, the domestic motorcycle market continued to decline due to various factors, such as stricter parking regulations, price changes due to emission regulations, and the global economic downturn[(1)]. The 50 cc scooter market was also affected by the spread and expansion of electrically power assisted bicycles[(2)] after 2007 (Fig. 17). Against this market background, 50 cc scooter manufacturers continued to increase color variations and introduce new colors in a rush to drive up purchase motivation (Fig. 18).

4.1. Five models and 30 colors of 50 cc scooter

The 2009 Honda 50 cc scooter lineup consisted of five models: Dio, a standard model for men; Cesta, a scooter equipped with a shopping basket and designed for housewives; Zoomer, a scooter for street-style youth; Scoopy, a scooter with high fashion appeal; and Today, a basic model. The color variation was 30 types in 23 colors.

The color palette of each model was repeatedly changed based on customer feedback and sales data, and because

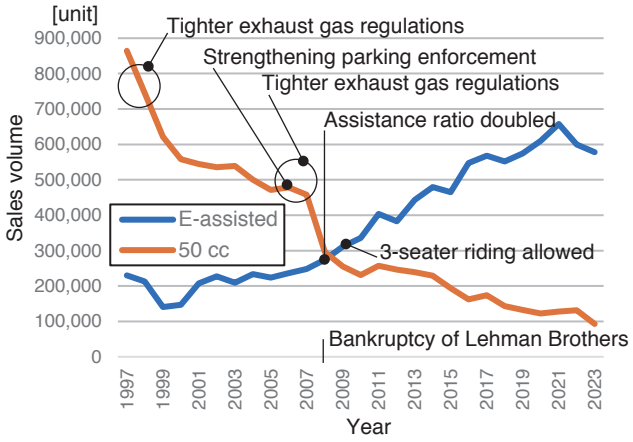


Fig. 17 Sales volume of 50 cc model scooter and domestic production volume of electrically assisted bicycles



Fig. 18 50 cc scooter market in chaos

the styling designers determined the coloring scheme for their respective product based on the trends, all the models had similar color configurations. These may have caused a lack of color differentiation among the models. In addition, retailers were unable to display color variations of all models in their stores, and customers were unable to check the colors on the actual products, causing confusion. Some customers were seen to buy a model other than the one they wanted, saying, "Since it's the same color, I'll just buy this model in the store." It was apparent that the coloring, which was assumed to convey product appeal and the model concept, was not functioning.

Under these circumstances, changing the color of Dio alone would not be sufficient to convey the appeal of the product to customers, and would not solve the confusion that may have caused to customers and the sales floor. In order to get out of this situation, I thought it necessary to evolve and apply the "coloring strategy" rolled out in Vietnam that proved to be immensely successful, look at the entire 50 cc scooter lineup, and change colors to best suit each model. In order to realize the coloring strategy

for the domestic scooters, I thought it necessary to gain the authority of a development manager who could take charge of all five models and took on the responsibility.

4.2. Coloring strategy for Japan from 2009 to 2011

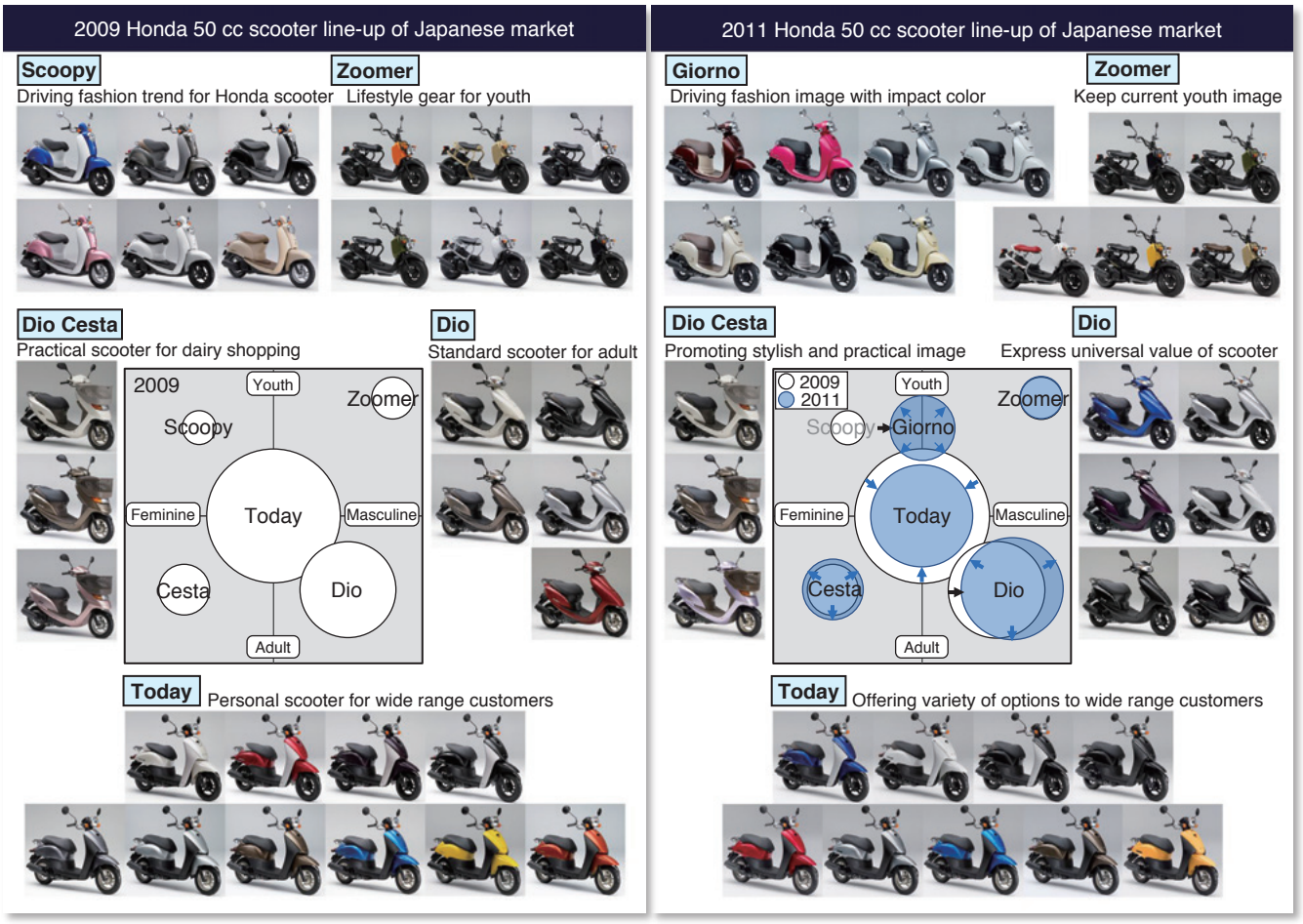
The first step was to create a positioning map as of 2009 [Fig. 19(a)]. A positioning map was not necessary for the Vietnamese market, as the three local models were intuitively recognizable due to their differences in size and character. On the other hand, the Japanese lineup of 50 cc scooters required differentiation among models with the same displacement. A tool was needed to make it easy to recognize the differences in characters at a glance, and so this gave us the idea of creating a positioning map. To make one that is simpler and easier to understand, the vertical axis was set as an image of age (old to young) and the horizontal axis as an image of gender (female/male). Today, which accounts for 56% of Honda 50 cc scooter sales, was placed at the center of the map. Next, Dio, Cesta, Zoomer, and Scoopy were mapped by age and gender according to their concepts.

Looking at the resulting map, it appeared that the models were well differentiated, but I felt a stronger sense of urgency about the fact that the sales of Today accounted for 56% of the total.

It is understandable that Today is selling well because of its attractive balance of specifications and cost. This however meant that the lineup was not functioning: customers were buying Today because the other models were lacking in attractiveness, and because the models offered similar color schemes the customers would simply buy the lower-priced product. I therefore consciously created a wrap-around portion of the map to show the competition among the Honda models. The map also had the function of highlighting the different characterizations of each model, that is, the areas where the character needed to stand out.

Next, I began creating a positioning map not for the 2010 model to be developed this time, but for the 2011 model to be developed the following year [Fig. 19(b)]. This is what was evolved for the coloring strategy applied to Japan: rather than making a one-year fix, it was seen from a birds-eye view to clarify what should be done for a more comprehensive planning purpose for the models for their longer-term success.

The most significant change in the 2011 lineup was the discontinuation of Scoopy, which led the “fashion” line and played a role in attracting female users, and the introduction



(a) 2009 (b) 2011

Fig. 19 Honda’s Japanese 50 cc scooter lineup and positioning map for 2009 and 2011

of Giorno with a unisex image targeting all fashionable young people.

Therefore, Today was kept in the center of the positioning map, but the size of the circle was slightly reduced. The size of the circles for Giorno, as well as the continuing models Dio and Dio-Cesta were enlarged while help eliminating any overlaps with Today.

The goal was to make each model to stand out separately from one another in order to create a lineup where each has its own attractive coloring scheme. To this end, it was decided to offer a minimum number of basic color variations for Today to meet the needs of customers who demand for lower-priced products, while enhancing the entire product lineup by offering other models with their own unique coloring schemes matching their concepts.

There was no plan for Zoomer to change the color scheme for this year, but as 2011 was the 10th year since its launch, I proposed a limited sale of vehicles in a special 10th anniversary livery. Also proposed was a reduction in the number of colors, focusing on the current best-selling colors. Zoomer was to be positioned as a stand-alone model set apart from all other models with the intention to grow it as a city bike for the younger customers.

4.3. Result of coloring strategy for Japan in 2010

After creating the positioning map for 2011, I started the process of creating the positioning map for 2010 (Fig. 20).

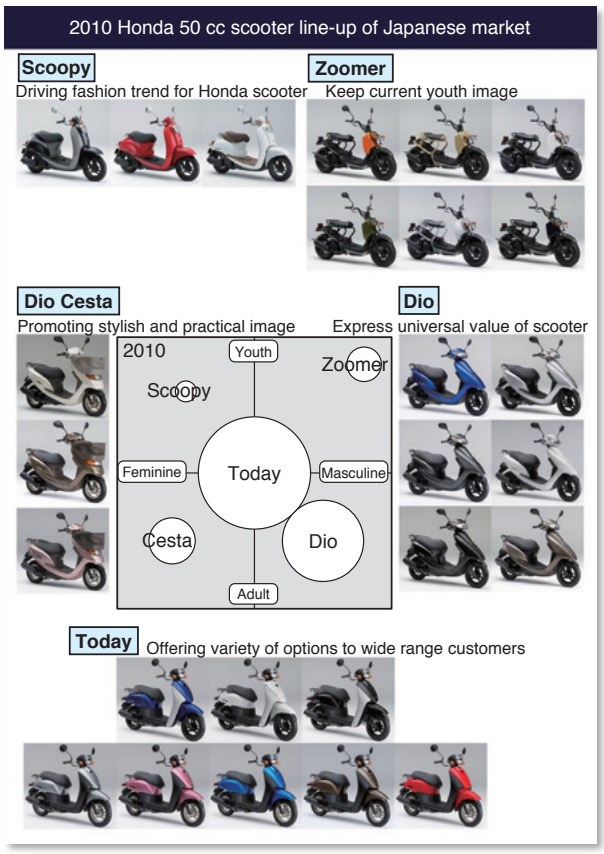


Fig. 20 Honda's Japanese 50 cc scooter lineup and positioning map for 2010

The first thing considered was how to handle Today. The color variation setting was changed from the previous all-round color variation to a setting which takes into account the color variation with keeping the regular colors, but a decision was made to reduce the number of colors as the year progressed, rather than doing it immediately at once.

Dio was given a color lineup of six colors to express a sense of luxury and power and to distinguish it from the Today model, as well as to be competitive against Company A's competing model and Company B's competing model with seven and nine colors, respectively. The color lineup was also intended to encourage customers to switch from Today.

As mentioned above, Scoopy was to smoothly pass the baton to its successor model, Giorno, which was to be launched in 2011 as a new fashion-oriented model. A team of young designers carried out color development of Giorno based on the concept of the return of real fashion scooter. Scoopy's color lineup for 2009 consisted of eight colors in three types (Special, Standard, and Deluxe) to cover a wide range of fashion-oriented preferences. To allow for a smooth linkage from Scoopy, the team narrowed Giorno's colors to three with a unisex fashion image (Fig. 21).



Fig. 21 Color lineup change in Scoopy from 2009 to 2010

4.4. Market reaction to coloring strategy for Japan

In 2011, the newly launched Giorno received a great response for its new concept of "The return of a real the original origin fashion scooter". (Fig. 22).

Giorno sold approximately 16,000 units this year, whereas its predecessor Scoopy sold around 3,000 units per year. In contrast, the sales volume of Today, which previously accounting for the majority of 50 cc scooter



Fig. 22 Color lineup of 2011 Giorno

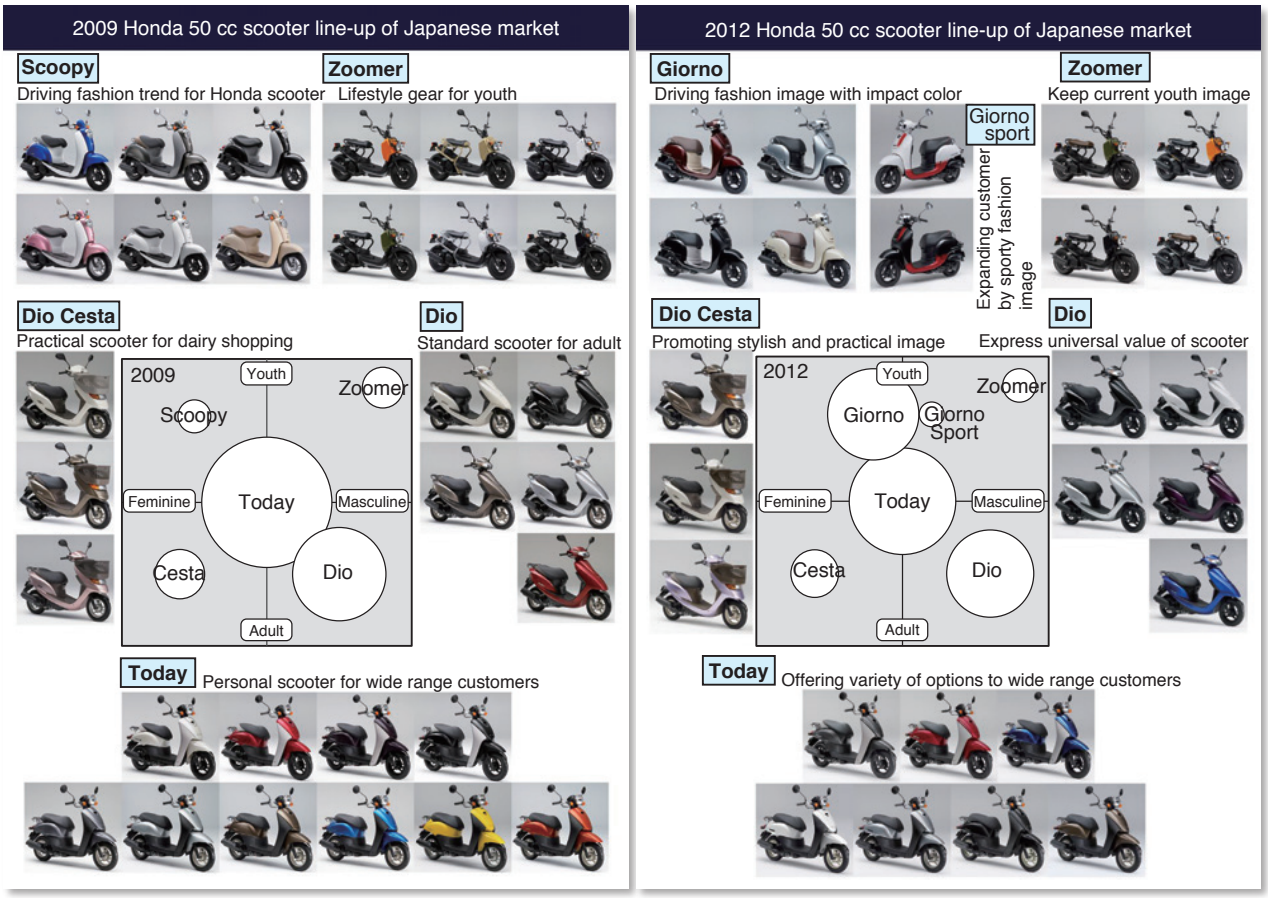
sales, decreased by approximately 9,000 units. These results indicate that the equivalent of those customers shifted to Giorno, and I believe that we were able to achieve one of our coloring strategy goals, which was to differentiate each model by its coloring scheme. On the other hand, the reduction of the number of colors was still insufficient to avoid confusion at retail storefronts, so we decided to continue with the coloring strategy and work on the reduction of the number of colors in 2012. Fig. 23 (a) and Fig. 23 (b) show the lineups for 2009 and 2012, respectively.

Dio, Cesta, Giorno, and Zoomer have settled on the number of colors targeted by the 2011 measures, leaving only Today.

The task of narrowing down the color variations of Today, which were reduced by only one color in 2011 began in earnest.

At this point, the color selection process was going smoothly because other models were already available in a range of unique colors. The sales department was convinced by this proposal that there would be no problem in reducing the number of color variations as long as the color selection was in line with the coloring strategy. This was because the sales performance over the past two years had clearly demonstrated the effectiveness of the coloring strategy.

Comparing the number of models in 2009 and 2012, the number of colors has gone from five models with 30 types and 23 colors to six models with 25 types and 17 colors. The coloring strategy has thus helped to settle the confusion that may have caused to customers and the sales front, as



(a) 2009 (b) 2012

Fig. 23 Honda's Japanese 50 cc scooter lineup and positioning map for 2009 and 2012

mentioned at the beginning of this chapter.

The sales share of Today decreased from 55% to 38%, while that of Giorno, which was a model change from the Scoopy with a 4% share, increased to 27%. The sense of urgency about the sales of 50 cc scooters, which was heavily skewed toward Today, as described by the author at the beginning of this chapter, has been resolved (Fig. 24).

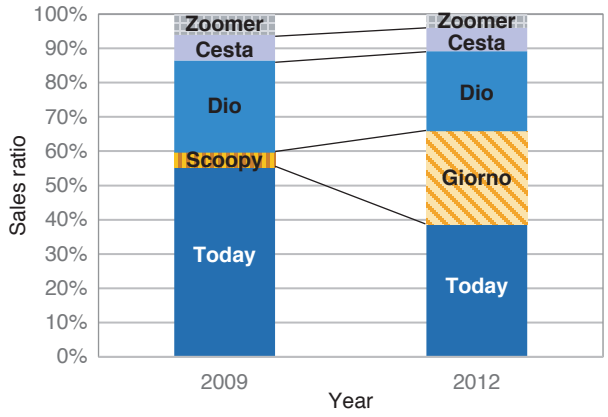


Fig. 24 50 cc scooter sales ratio of Honda

This is the result of adequate color selection based on differentiation and the clarification of the role of each model, and customers were now able to choose a color design that is in line with the individuality of each model. The color lineup that makes the most of each model's concept continues to be effective without blindly increasing the number of colors.

Furthermore, a comparison of sales volume between 2009 and 2012 in the 50 cc scooter market in Japan, where scooters account for 70% of the total, shows that Honda's sales volume increased by about 3% while other companies' sales volume decreased by about 10% (Fig. 25).

With this, I believe we have verified that the "coloring strategy" described here is a development method that can achieve maximal effect with minimal colors.

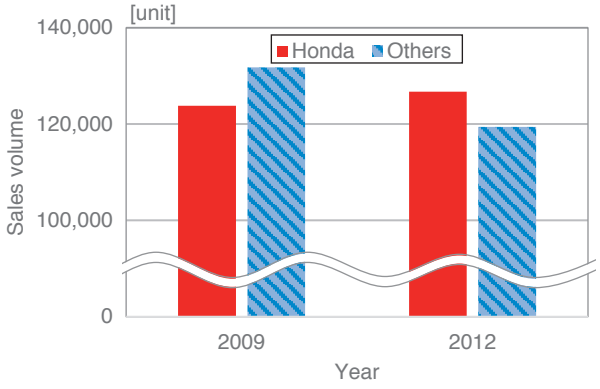


Fig. 25 50 cc model sales of Honda and other manufacturers

5. Summary

Some say that we are only talking about colors, and others say that changing colors of a product at hand is easy.

However, as this article has described, strategic coloring changes can yield maximum effect.

Recently, the world of coloring has been increasingly described by the term CMF (Color, Material, Finish), and Honda has also established a CMF group several years ago to develop attractive colors.

It is my belief that a very important thing in carrying out such development is to have clear objectives and goals: we should ask ourselves, "What do we want to do with the colors, and how?"

Using colors in a specific way may be a way to bring you closer to achieving a goal. Well, it may not be the color, but rather something the reader is in charge of....

I sincerely hope that this article will be of assistance to future generations in their thinking.

References

(1) https://jamaserv.jama.or.jp/newdb/ (accessed 2024/12/26) (in Japanese)
(2) https://jbpi.or.jp/business/statistics/production/electric/ (accessed 2024/12/26) (in Japanese)

■ Author ■



Masaharu YOSHIMURA

Development of Electric Motocrosser CR ELECTRIC

Seiji KATO[1] Hiroyuki AKIYOSHI[1] Makoto KITAOKA[1]
Masayuki KINOSHITA[1] Takuya FURUYAMA[1]

ABSTRACT

In an attempt to offer a new value of EV in the FUN category such as the motorsports, the electric motocrosser CR ELECTRIC has been developed. The employment of launch control and energy management systems upon the selection and control of battery cells enabled realization of the dynamic performance and the range required for the motocross machine. Developing the L-shape battery pack and optimizing the frame body stiffness, the specification of complete vehicle similar to the CRF450R are achieved. Moreover, the development of high-voltage pack as high as 400 volt along with the redesigning of the cooling system to adjust to the electric motocrosser resulted in the downsizing and weight reduction of various component parts. With these measures taken, the maneuverability and stability such as nimbleness, roadability, and turnability have been accomplished.

1. Introduction

The changes of social environment call for realization of carbon neutralization not only in automobiles but also in motorcycles. Honda aims at a carbon neutralization in all products and all corporate activities by 2050. It is Honda's policy to achieve carbon neutrality in all motorcycle products in the 2040s. As a core environmental strategy to achieve the said goal, we are actively promoting electrification of motorcycles[(1)].

In the realization of carbon neutrality in the mobility sector, even the motorsports are not exceptional. In the Formula 1 and MotoGP, introduction of carbon-neutral fuel has already started[(2), (3)]. Regarding the electrification, races exclusively for electric vehicles (EVs) such as the Formula E and MotoE are being held. However, the energy density of lithium-ion battery is approximately 2% of that of gasoline, which is an issue in races requiring frequent use of high-power output. Meanwhile in EVs, the power output is easy to control due to the recent advancement of motor control techniques, having easy-to-control drive-ability. In addition, as the EV is quiet and clean without emissions of exhaust gas, it is an advantage that races can be held in urban cities and in-door arenas.

Honda has offered electric motorcycles to the market as small commuter motorbikes for use in short distances such as delivery and/or means of day-to-day transportation. Adding to the said applications as a creation of new value in FUN category, we have developed the electric motocrosser with a focus on motocross races, which are popular in Europe and North America, and also held in stadiums and the likes.

As a technical validation, the EV motocrosser was raced in the Kawagoe round, the 8th race of 2023 Japan Championship Motocross Series. The results were very positive; 2nd place in Heat 1, and two hole-shots in three heat races. That attracted media attention as Honda's first race participation with the electric motorcycle, resulting in a recognition of EV.

In this paper, we introduce the technical measures employed in the electric drive unit such as the battery, which is an essential part in the electric motocrosser CR ELECTRIC shown in Fig. 1, and the chassis techniques to package such parts.



Fig. 1 External view of CR ELECTRIC

[1] Honda Motor Co., Ltd.
E-mail: seiji_01_kato@jp.honda

2. Aims of Development

2.1. Features of Motocross

The motocross is such a sprint race that contests speed in unpaved, short race courses. Various places such as outdoor flat lands and hills or stadiums are used as the field. The race course has manually-created ups and downs. As Fig. 2 shows[4], some big jumping spots are included to demonstrate three-dimensional actions such as jumping and landing. As the track conditions varies depending on the weather or repetition of races, the vehicle has to deal with a variety of elements.

The race program is such that a heat race lasting from 15 minutes to 30 minutes is held two to three times a day. Base on the results of heat races, a final race is held for overall ranking. With thirty to forty machines line up side by side and all dash out together from the starting gate, the spectacular start scene is one of its charms.

In motocross races held under such conditions, the compactness of vehicle body and the lightness of weight are the keys. A bulky and heavy motocrosser is difficult to ride no matter how powerful the motor is. To travel on slippery off-road surfaces, linear torque delivery and easy-to-handle traction performance are essential although a high maximum power is also important[5].

Due to the race style in that a large number of racers start together, a hole shot (to be the first one to enter turn 1 after starting) is very advantageous to win the race or finish higher.

2.2. Features of EV

2.2.1. Installed energy

The internal combustion engine (hereafter ICE) converts the thermal energy occurring from combustion of gasoline to the mechanical energy to produce driving force. Meanwhile in EVs, the motor converts the electric energy in the battery to the mechanical energy to produce driving force. In current battery technology, the energy density per weight is approximately 2% of that of gasoline. Even if the higher driving force conversion efficiency of electric motor is taken into account, the available energy is approximately 40% of that of an ICE vehicle in the same class. These characteristics result in the short range and heavy vehicle weight of Evs.

While the gasoline-burning ICE loses much energy as waste heat, the electric power conversion efficiency is higher in Evs, and the waste heat is not as much as in the ICE. On the other hand, the battery changes characteristics depending on the temperature. For efficient use of limited energy on the vehicle, it is essential to adjust heat management to the operating conditions and working environment.



Fig. 2 Motocross riders jumping on track

2.2.2. Output characteristics

Figure 3 shows driving force diagrams of EV and internal combustion engine vehicle (ICEV). In the EV, a high torque is attainable at lower revolutions due to the characteristics of the battery and the motor. The torque and the power output occurring from the electricity can be easily modified by the control program. This feature is different from the ICE in that an adequate torque is unattainable until the engine speed reaches to a certain level.

Figure 4 shows difference of power output characteristics in EV and ICEV. While an appropriate gear has to be used in the ICEV depending on the vehicle speed to maintain a higher power output, a flat power output up to higher revolutions is attainable in the EV without changing the gear.

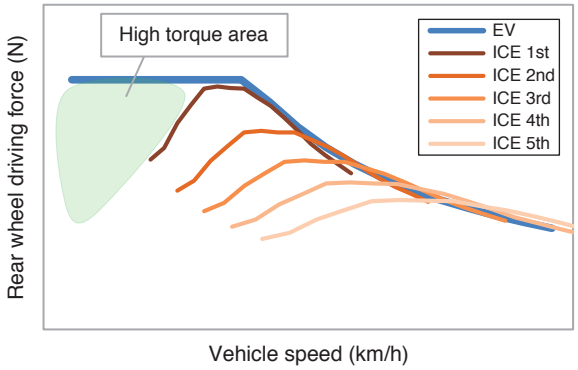


Fig. 3 Comparison of driving force between EV and ICEV

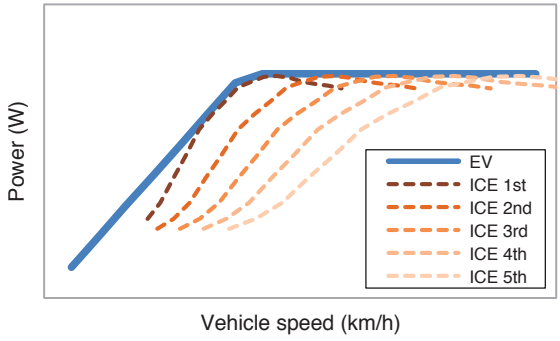


Fig. 4 Comparison of power characteristics between EV and ICEV

Table 1 Specifications of 2021 model CRF450R

Overall length [mm]	2182
Overall width [mm]	827
Overall height [mm]	1267
Ground clearance [mm]	336
Seat height [mm]	965
Caster angle [deg]	27.7
Trail [mm]	114
Wheelbase [mm]	1481
Front brake	Single 260 mm disk
Rear brake	Single 240 mm disk
Front suspension	49 mm fork
Rear suspension	Pro Link
Front tire	80/100-21 51 M
Rear tire	120/80-19 63 M

In other words, the maximum speed is easily reachable only by the throttle operation. It is therefore not necessary to select an appropriate gear depending on the speed, allowing easy control of the vehicle only by a simple operation, which is a feature that makes the vehicle user friendly.

2.3. Development Goals

The power output characteristics of EV mentioned in the preceding chapter can be considered an advantage in the motocross competition where the course conditions keep changing. To take that advantage, the machine has to have the power output and range comparable to the ICEV. We therefore set such goals that the machine has the same level of power as the CRF450R and able to finish a heat race that lasts for 15 minutes. Also, the key factors in the race are turnability and off-road-ability, which are affected by the location of center of gravity (CG) and the frame stiffness of the complete vehicle. We targeted at the specifications of complete vehicle similar to those of the CRF450R as Table 1 shows also in the EV carrying different component parts.

2.4. Issues of EV in Motocross

As Fig. 5 shows, due to the difference of installed energy when the complete vehicle specifications required for the motocrosser are compared, if the installed energy in the EV is simply equalized to the ICEV, the vehicle body becomes bulky and heavy. If a higher power output is continuously used from the limited amount of installed energy, the target range is unattainable. It is therefore necessary to work out a way to extend range while maintaining a higher power output taking advantages of the characteristics of EV. Also, as the energy for use differs, the component parts also differ. Since it is not easy to commonly use conventional parts simply, the major components such as the frame body have to be modified, which is also an issue.

There is also an issue of heat from electric components. In the EV, the power loss that occurs when the power is supplied to drive the motor is converted to the thermal energy, which generates heat. The major components of EV, i.e., the motor, battery, and inverter, generate heat when electric power is supplied. Since such parts may be damaged when the temperature exceeds the limit, a function is incorporated to limit the power output to maintain the temperature under the prescribed level. As the lowering of acceleration and/or torque from the limiting of power negatively affect motocross race performance, a proper thermal management of electric powertrain parts is required.

Introduced from the next chapter on is the elemental techniques employed in this paper to deal with the issues in the EV motocrosser.

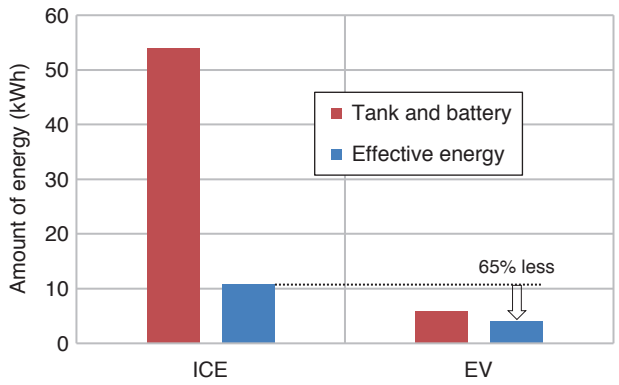


Fig. 5 Comparison of installed energy in motocrossers of EV and ICE

3. Elemental Techniques for EV

3.1. Battery

The battery is a critical part when determining performance and specifications of the complete EV. Described hereafter are the construction of the battery pack (hereafter “pack”) and its system that can be smallest and lightest when realizing the necessary power output and range required by the complete vehicle.

3.1.1. Battery pack construction

The pack is composed of a number of battery cells (hereafter “cells”) and the performance of pack depends on the performance of cells. The performance of cells is evaluated by the gravimetric energy density and the gravimetric power density. It is practically important that both densities are at a light level. The representative cells of various cell suppliers are selected and their gravimetric energy density and gravimetric power density are plotted on the graph in Fig. 6.

The cell performance required for an EV motocrosser can be separated to the gravimetric power density that allows the same performance as the CRF450R and to the

gravimetric energy density that permits finishing of the race. The cell is selected from the separated areas taking into account the vehicle concept and the shape of pack. In this development, the main target was set at finishing of a heat, and the cells having a higher output among those in the required area were selected.

Next, Fig. 7 shows the shapes of general cells. In the pouch type and prismatic type cells, the cuboid shape and the electrodes gathered on a side are the features. It is therefore necessary to adjust the cell shape to the cuboid shape in order to allow efficient layout. On the other hand, in the cylindrical cell, a waste space is less likely even when the external pack shape becomes complex because of the freedom of negative electrode layout and its external shape. In this development, the cylindrical cell is selected from the selection area that satisfies above-mentioned performance requirements, and worked out the L-shape pack shown in Fig. 8. This permits laying of the cells, that are the heavy objects in the pack, in the rear and lower part of the complete vehicle as Fig. 9 shows. With the L-shape pack developed, the target CG of the EV motocrosser, which is the same as the CG of CRF450R, has been attained.

3.1.2. Battery system

In the pack of general EV scooters, the range and the battery life are critical factors. Accordingly, the mainstream is such a method that maximizes the said range and life by increasing the number of parallel connections while minimizing the number of series connections of cells. With this method employed, the output voltage of battery pack often falls in the zone between 50 V and 100 V.

The EV motocrosser requires battery capacity that allows finishing of the race in addition to the higher power output to stay competitive in the race. To satisfy these requirements, the cells having a higher power output explained in the preceding chapter are used, and the number of series connections are maximized to realize a higher power output as a pack. Accordingly developed is the higher voltage pack in the 400 V zone, which is the first in Honda motorcycles. Moreover, the minimum number of parallel connections is set to allow finishing of a heat race.

With the pack voltage set as high as in 400 V zone, the output current of the pack is lowered, which permits reduction of heat generation from the major components existing in the high-voltage wire lines. It also allows downsizing and weight reduction of the motor and the inverter, keeping the increase of vehicle weight of EV motocrosser less than 10% from the CRF450R.

3.2. Control System

Combining three units, i.e., the inverter integrated in the motor, the battery management unit contained in the battery pack, and the vehicle management unit that calculates driving force, the complete vehicle control system has been developed.

Introduced hereafter is the control functions employed to finish higher in the race taking advantages of characteristics of EV.

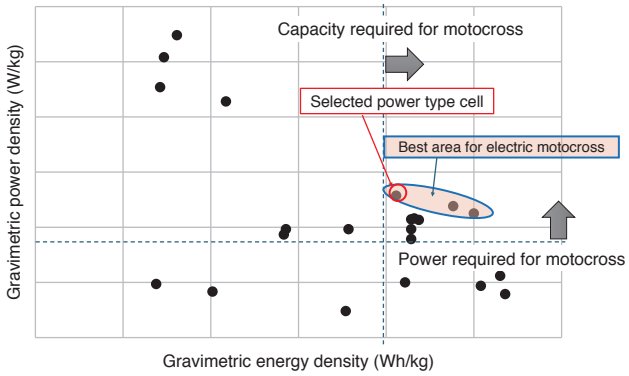


Fig. 6 Comparison of gravimetric energy density and gravimetric power density of various battery cell types

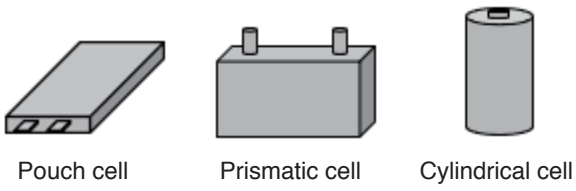


Fig. 7 General cell shapes

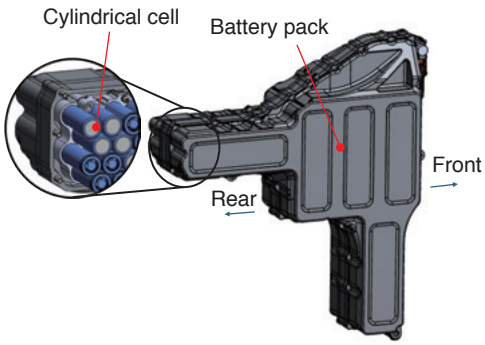


Fig. 8 Battery pack appearance and cells

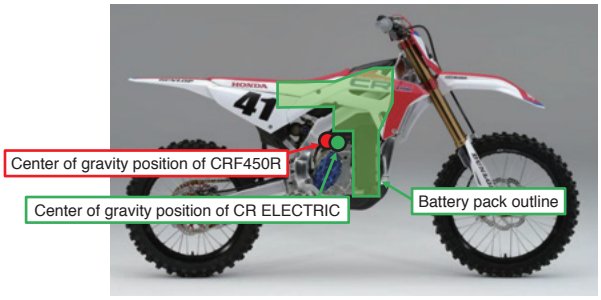


Fig. 9 Battery pack outline and vehicle CG position

3.2.1. Launch controller

To make a hole shot in motocross, it is necessary to maintain a higher driving force and power output from the start to the turn 1. However, if the driving performance specifically tuned for the start is left as it is throughout the race, the vehicle would become too difficult to control to maintain advantage in the race. We have therefore added the drive mode switching function (hereafter "launch controller") that functions only at the time of initial acceleration when starting the race.

The launch controller is enabled from an operation when the vehicle is standing still and set by the rider before the start. Closing the throttle after starting automatically disables it. After being disabled, the ride mode will change to the mode the rider selected before enabling of the launch controller.

When starting the race, it is necessary to drive the rear wheel at the highest-possible power output while preventing wheel spin. Also, to maintain stable acceleration while preventing lifting of front wheel (wheely) when accelerating after the start, the power output characteristics were altered. Figure 10 shows power output characteristics in each mode.

At the time of start, the driving force in the low engine speeds is lowered by approximately 20% rather than merely applying the maximum power. In addition, the maximum power output is set higher by approximately 25% to reach a higher top speed. These parameters can be preset to any given value depending on the track and/or the conditions. This permits selection of optimum characteristics for each rider or race, contributing to maintaining of advantages.

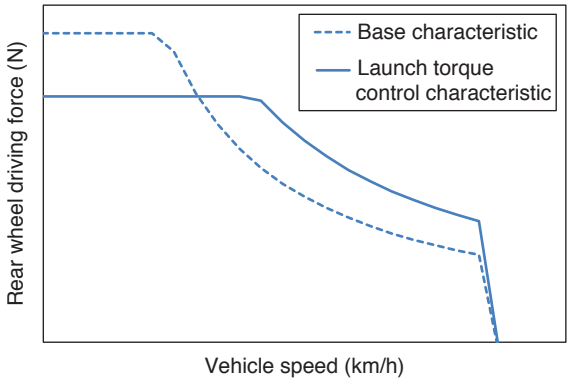


Fig. 10 Launch torque control characteristics under different operating conditions

3.2.2. Energy management

It is necessary to regulate electric power consumption during the race in order to efficiently consume the limited amount of installed energy. As a control application to assist that, the energy management function is added.

As Fig. 11 shows, the following are the methods to regulate electric power consumption.

(1) To attain advantages of EV motocrosser when starting the race, a priority is given to the motor power output without considering electric power consumption when in the launch control mode.

(2) In the race after cancelling the launch controller, the rider rides at such race-oriented output characteristics that make the output such a level that allows full use of electric power within the presumed time length of race.

(3) The changes of electric power consumption during the race are not easy to estimate. To deal with this issue, the adequacy and the inadequacy relative to the power consumption in each lap are calculated. From the outcomes, a compensation is made by the function that adjusts output characteristics (auxiliary application).

To realize that, the output control mode is produced and the electric power consumption is measured at each designated output. Figure 12 shows plotted data indicating the relationship between the output control mode and the power consumption. Based on the data, the race mode was produced so as to set the output at such a level that allows full use of electric power within the presumed time length of race.

Also as an auxiliary application, the allowable power consumption in each lap was estimated from the remaining electric power and the remaining race length after finishing the launch control. Then, the function was added that adjusts real time the motor torque control algorithm from the linear interpolation of pre-measured output and the power consumption.

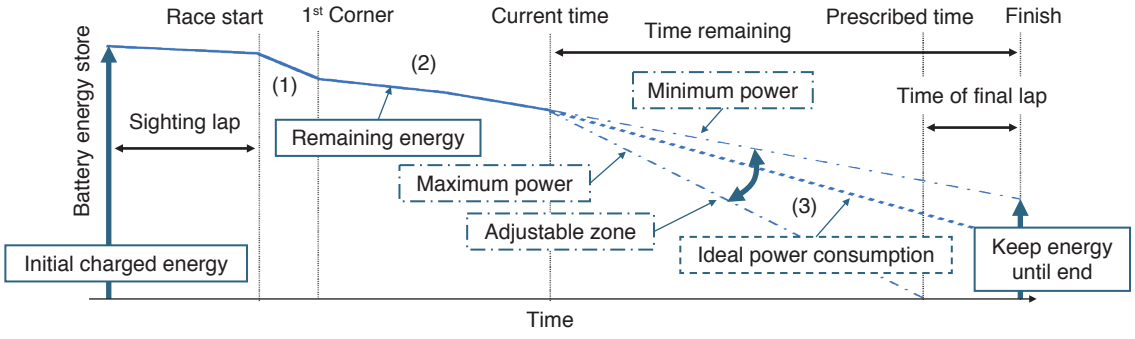


Fig. 11 Concept of battery energy management control during race

3.3. Cooling System Parts

To finish at a position higher than competition in the motocross, it is necessary to complete a heat and another race without reaching to a power restriction due to an increase of temperature. It is therefore necessary to control temperatures of various parts. In this vehicle, cooling is conducted for the motor and the invertor. As the cooling target, the following two requirements were set and the parts specifications were determined accordingly. First, the machine shall allow riding without entering the power restriction zone during a heat. Second, after cooling the parts until the next heat after completion of a heat, the machine shall allow riding throughout a race. Introduced in this chapter are the parts in the cooling system to satisfy the above requirements.

3.3.1. Setting of heat dissipation capacity

To set dissipation capacity of heat from the motor and the inverter under the race conditions of EV motocrosser, test runs were conducted using radiator cores having various surface areas. Taking temperature data of high-voltage parts from the tests, the relationship between the heat dissipation from the radiator and the part temperature during the race were measured.

Figure 13 shows the inverter temperatures when running with various types of radiators. Figure 14 shows motor temperatures when running with various types of radiators. From the test results, the power loss data of the motor and the inverter to use at this time were compared, and the amount of heat generation that determined temperatures of various parts after completing a heat was estimated. Also, from the difference of water temperature depending on the radiator core surface area, the cooling capacity per unit of radiator core surface area under motocross race conditions was estimated.

Consequently, compared to the ICEV, the radiator core surface area was reduced by 64%, and the weight was lowered by 77%. Figure 15 shows overview of the light-weight radiator. This resulted in the setting of optimum cooling performance necessary for a race as an EV motocrosser.

3.3.2. Water pump selection

In many ICE vehicles, a mechanical water pump taking power from the rotational motions of the engine is used. If a mechanical pump is used in the EV in the same manner, the pumping depends on the motor rotation. This means that the coolant circulation stops when the EV requiring no idling is staying still. Also, in the case of mechanical pump, it is necessary to lay parts on the rotational axis or drive the pump with a set of gears. Accordingly, the motor housing becomes wide sideways, causing an increase of body size.

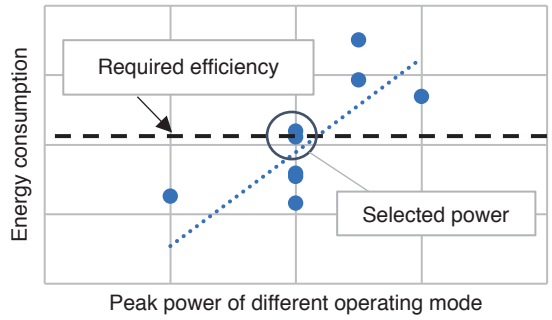


Fig. 12 Relationship between peak power of different control modes and energy consumption

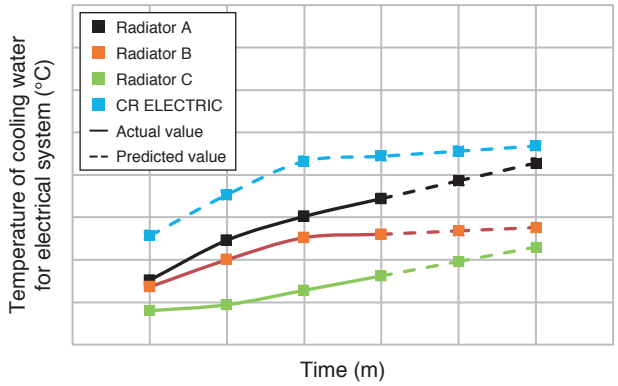


Fig. 13 Heat generation of inverter under varying operating conditions

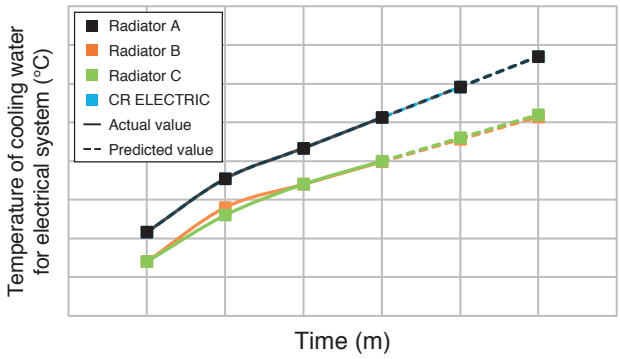
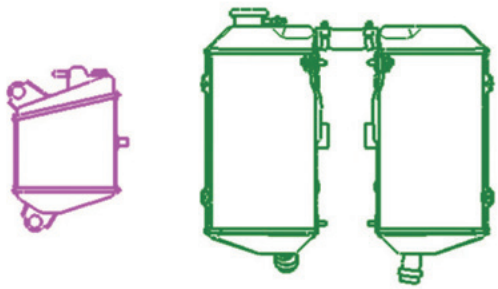


Fig. 14 Heat generation of motor under varying operating conditions



(a) CR ELECTRIC (b) 2024 Model CRF450R

Fig. 15 Radiator size

The selection of an electric water pump to deal with the said issue allows downsizing of the power unit and permits coolant circulation when the vehicle is staying still. Thus, the specifications of cooling system to achieve the target cooling performance of EV motocrosser has been established.

3.4. Complete Vehicle

3.4.1. Specifications of complete vehicle

To allow control of vehicle as per rider's intention, we set a target at the specifications of CRF450R. Setting conditions such as the ground clearance that does not impair roadability and the fore-and-aft weight distribution that can take advantages of traction, the overall layout of electric drive train components was decided as shown by Fig.16.

Based on that, as Fig. 17 shows, the specifications of complete vehicle such as the caster angle, the trail length, the seat height, and the ground clearance were set at the same values as the CRF450R, and the wheelbase similar. Thus realized is the maneuverability necessary for the EV motocrosser.

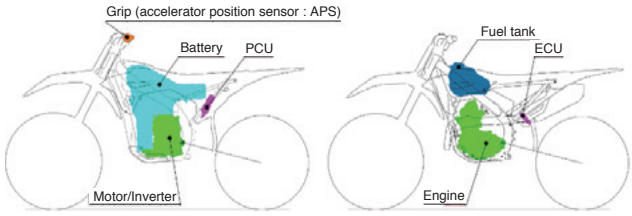
3.4.2. Frame body

To attain required battery output and capacity, it is critical in EV motocrossers to efficiently package the battery in the limited space while minimizing vacancy. If the conventional frame body designed for the engine size of CRF450R is used, it is not easy to hold the battery having a capacity adequate for a race. Meanwhile, the stiffness of frame body affects maneuverability and stability such as nimbleness, roadability, and turnability. Especially, as the motocross races are held on rough surfaces including jumps, the stiffness balance becomes essential. That is, a stiffness balance based on the experiences from the conventional ICEV is required.

That in mind, the aluminum, semi-double-cradle, twin-tube frame employed in the CRF450R was chosen as the basis of the main pipe. For easy installation and removal of the battery and to attain necessary frame stiffness, the cross-sectional shape of the main pipe is modified, and such a layout is employed that straightly connects the main pipes to the steering head pipe.

Figures 18 and 19 show the shapes. Said measures permit to maintain the change of frame stiffness as small as less than 10% of that of the CRF450R, assuring the same level of maneuverability as the conventional ICEV.

Also as Fig. 20 shows, to allow an adequate width for the large capacity battery, the space between the main pipes



(a) CR ELECTRIC (b) CRF450R

Fig. 16 Vehicle layout

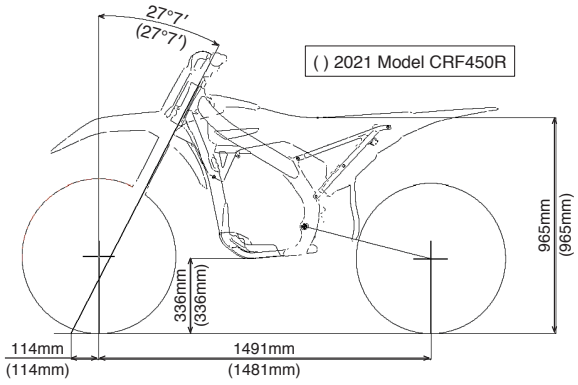


Fig. 17 Vehicle geometry of CR ELECTRIC

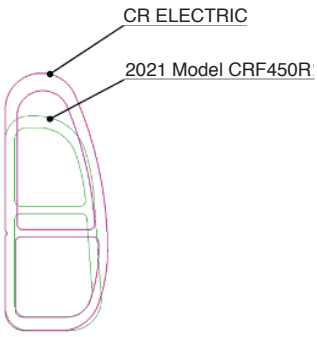
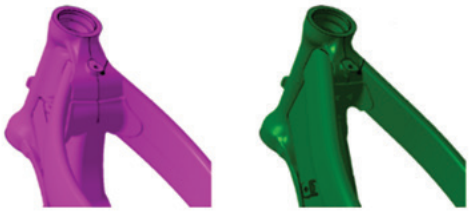


Fig. 18 Cross-sectional shape of main pipe



(a) CR ELECTRIC (b) 2021 Model CRF450R

Fig. 19 Shape of connecting part of main pipe to head pipe

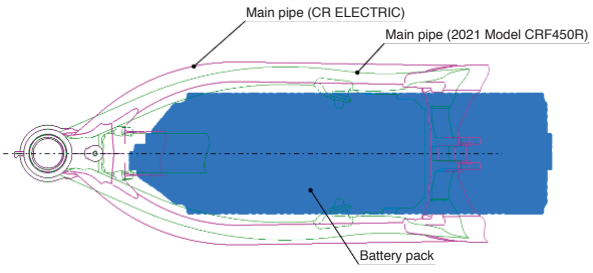


Fig. 20 Route of main pipe

is widened from right behind the steering head pipe. Also as Fig. 21 shows, the upper part of the main pipe is bent in such a manner that follows a vertical line. Accordingly, enabled is smooth installation and removal of the battery in the up and down direction of the vehicle body as Fig. 22 shows.



(a) CR ELECTRIC (b) 2021 Model CRF450R

Fig. 21 Bending shape of main pipe

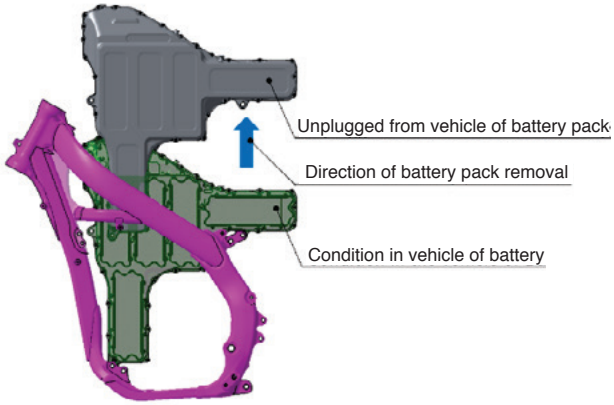


Fig. 22 Direction of battery pack removal

4. Conclusion

Through the development of EV motocrosser CR ELECTRIC, the following have been achieved.

(1) The employment of cylindrical cells for the battery and the creation of the L-shape battery pack enabled to locate the CG of the complete motorcycle at the place similar to that of the CRF450R while attaining the target power output.
(2) The use of battery voltage in the 400 V zone, which is the first in Honda motorcycles, reduced the heat generation from major components, allowing to achieve the light and compact package. Accordingly, enabled is the increase of vehicle weight as small as less than 10% from the CRF450R.
(3) The newly developed ride mode switching control enables retention of power necessary for a hole shot and energy management necessary for finishing of a race as an EV motocrosser.
(4) By estimating and optimizing heat generation in the EV components when used in motocross, we have achieved the light and compact cooling system by reducing the weight by 77% and lowering the core surface area by 64% from the CRF450R.
(5) Despite having the component parts different from the conventional motocrosser, we have achieved the complete vehicle specifications similar to those of the CRF450R, and attained such a level of handling performance that is necessary for motocross.

To deal with a larger number of races, further increase of range and enhancement of battery charging and replacement during the race are called for. With the CR ELECTRIC brought to a reality with an aim at the motor sport in the carbon neutral era, it is now possible to hold races in urban cities and/or indoor arenas, which would enable an unprecedented appeal to users.

References

(1) https://global.honda/en/newsroom/news/2022/c220913aeng.html (accessed 2024/12/24)
(2) https://global.honda/en/tech/motorsports/Formula-1/Powertrain_e-fuel/ (accessed 2024/10/17)
(3) https://www.motogp.com/en/racing-together/fuel-future (accessed 2024/10/17)
(4) https://honda.racing/ja/jmx/post/jmx-2023-rd08-ia1-race (accessed 2025/4/15) (in Japanese)
(5) Hoshi, N., Shimura, Y., Okuma, T., Nomura, A., Gogo, K.: Introduction CRF450R Equipped with 4-Stroke Engine, Honda R&D Technical Review, Vol. 14, No. 1, p. 21-30

■ Author ■


Seiji KATO


Hiroyuki AKIYOSHI


Makoto KITAOKA


Masayuki KINOSHITA


Takuya FURUYAMA

Demonstration Experiment of Electric Outboard Motor for Small Boat

Takamasa TAKESHIGE[1] Kenta KATO[1] Shohma HIGANO[1]

ABSTRACT

The authors conducted a demonstration experiment of a small electric outboard motor system that utilizes an electric scooter power unit and swappable batteries. As the use case, the experiment took place on the "Horikawa Meguri" sightseeing tour in Matsue City, Shimane Prefecture, in which a small boat is used to tour a moat.

The braking distance when berthing was reduced by 30% by using a switch that changes propeller rotation direction and a speed adjustment function that enables control even at very low speeds. In terms of operability, the authors verified how operators used the equipment and tuned the relationship between the throttle position and tour boat speed for an electric outboard motor. Furthermore, the use of swappable batteries enabled quick battery replacement on board. As a result, the performance and feeling of the outboard motor were better than those of internal combustion engine outboard motors.

1. Introduction

Some regions of the United States and Europe are starting to consider zero emissions regulations for small outboard motors of 19 kW or less[(1)]. Torqeedo GmbH is one company expanding its business worldwide, especially in Europe, as an electric water mobility manufacturer[(2)]. In Japan, Yamaha Motor Co., Ltd. began demonstration operations of HARMO, a new boat- maneuvering technique integrated with an electric propulsion unit and steering control system on cruise boats in August 2020[(3)].

In November 2021, Honda announced the concept model of the small electric outboard motor shown in Fig. 1 as a proposal for a next-generation clean outboard motor in a quest to achieve carbon neutrality by 2050[(4)].

In response to this announcement, the Matsue City Tourism Promotion Public Corporation in Shimane Prefecture inquired about using Honda's small electric outboard motor on a tour boat.

Matsue City has announced a long-term goal of achieving net zero CO_2 emissions to become carbon neutral by 2050. As one effort to achieve this goal, the tourism promotion corporation wanted to reduce CO_2 emissions by replacing Honda internal combustion engine (ICE) outboard motors on tour boats operating in the city with electric outboard motors[(5)].

Honda is also working to develop electric outboard motors for mass production, and while it is a challenge to achieve the same long-hour, high-load operation as ICE outboard motors, given battery energy density and cooling performance, Honda also hoped to hear actual customer feedback and clarify the value provided by electric outboard motors, appropriate sales prices, requirements, and



Fig. 1 Concept model of electric outboard motor for small boat

[1] Solution System Development Center
E-mail: takamasa_takeshige@jp.honda

operating methods.

This situation led to the start of a business demonstration project making maximum use of existing resources to build a pilot small electric outboard motor in a short period of time and operate it in Matsue.

Customers on the “Horikawa Meguri” sightseeing tour take a leisurely ride on a small boat through the moat in the Matsue castle town, where they can enjoy the scenery. As shown in Fig. 2, the ride is about 3.7 km long and takes about 50 minutes. During the busy season, more than 20 boats may be on the course at the same time, and at these times a single boat may travel the course up to eight times a day. Boats are about 23 ft long (about 7 m) and have a capacity of up to 13 people. Boats normally operate even in the rain and are in use year-round.

In some places along the course, there are rows of private homes. These are defined as “quiet zones,” where boats are required to travel with the ICE outboard motors running at a low speed of rotation[6].

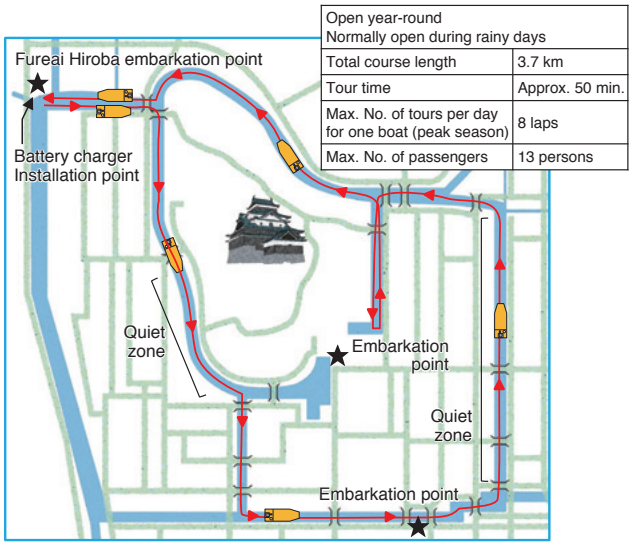


Fig. 2 Tour course overview

2. Specification Review and Target Performance

2.1. Power Performance Requirements

During the tour, boats operate at a speed of about 4 to 6 km/h. To determine the performance requirements of the outboard motors on the boats, the researchers checked the load that they were under during one lap of the course.

As Fig. 3 shows, the power unit used during the tour requires an output of about 1.0 kW for cruising and temporarily up to 3.5 kW when decelerating, such as when berthing at the embarkation points and when stopping. The resulting energy consumption is 1.2 kWh per lap.

2.2. Study of Electric Outboard Motor Power Unit and Motive Power Supply

For the study of a power unit capable of producing this motive power, the researchers looked at the BENLY e:, the business-use electric scooters shown in Fig. 4, which was being sold to corporate customers at the time. BENLY e: is an electric scooter based on the 110 cc class BENLY and powered by the Mobile Power Pack e: (MPP), the swappable battery shown in Fig. 5.

The maximum output of the ICE outboard motor BF9.9 currently used on the tour boats is 7.3 kW. This motor was chosen so boats can run quietly at low rpm. As Fig. 6 shows, the BENLY e: power unit is capable of the range of outputs used during the tour. It was therefore decided to make the tour a use case that could take advantage of the output characteristics of the electric motor, which can generate high torque from the low-rpm range.

2.3. Setting the Target Performance as an Electric Outboard Motor

Based on the use case, the researchers set the following performance targets in terms of motive power and marketability. The purpose was to provide ease of use equal

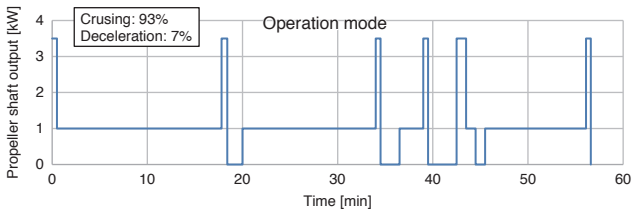


Fig. 3 Mode of used power during one tour lap



Fig. 4 BENLY e:



Fig. 5 Mobile Power Pack e: (MPP)

to or greater than that of ICE outboard motors along with the new value associated with electric motors.

(1) Needs to have braking performance and control system layout capable of berthing a boat
(2) Throttle grip with speed adjustment function needs to allow natural adjustment from very low speed to maximum speed without having to release and re-grip
(3) Needs to be possible to swap batteries while the boat is in operation during rain

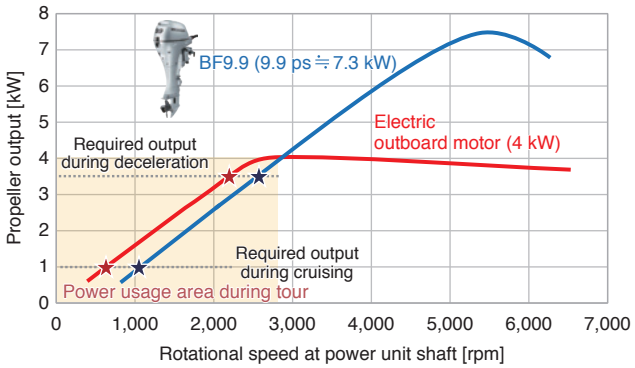


Fig. 6 Comparison of output characteristics between ICE outboard motor and electric motor

3. Electric Outboard Motor System Specification Configuration

3.1. Overview

As Fig. 7 shows, the electric outboard motor system built for this research consists of three components: the outboard motor itself, a battery box, and an electric device box that contains various wiring connectors and devices and fuses for using two MPPs in series. As shown in Fig. 8, the battery box was moved on board for ease of use, and the electric device box was mounted in the space where the fuel tank had been housed to make effective use of the space it had occupied[8), (9)].

3.2. Braking Performance and Control System Layout Specification Configuration

Figure 9 shows a conventional ICE outboard motor. Here, the drive shaft directly connected to the engine can only rotate in a constant direction. To change the propeller rotation direction, it is necessary to operate the clutch and change the engagement of the forward and reverse gears. As shown in Fig. 10, to perform the switching operation, one hand operates the throttle grip while the other is used to move the clutch by physically operating the lever. The electric outboard motor shown in Fig. 11, the use of a motor makes it possible to switch the drive shaft's direction of rotation by operating a control switch. This eliminates the need for a clutch, reverse gear, and a shift lever to operate

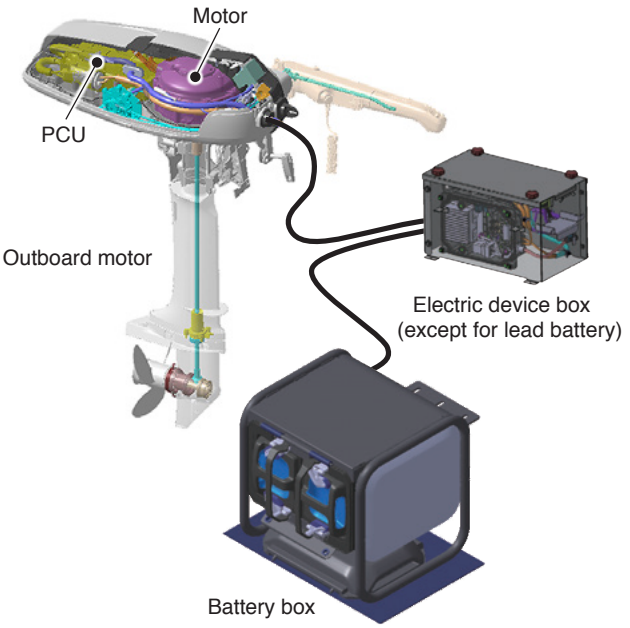


Fig. 7 Configuration of electric outboard motor system

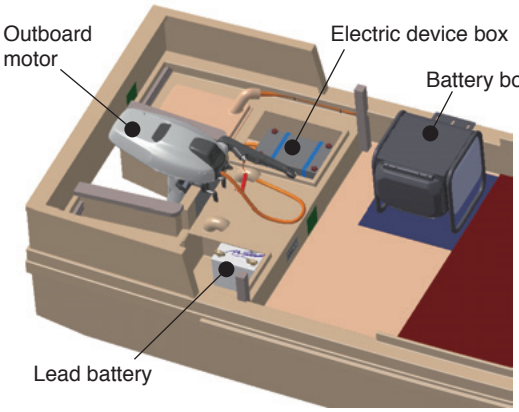


Fig. 8 Image of electric outboard motor system mounted on boat

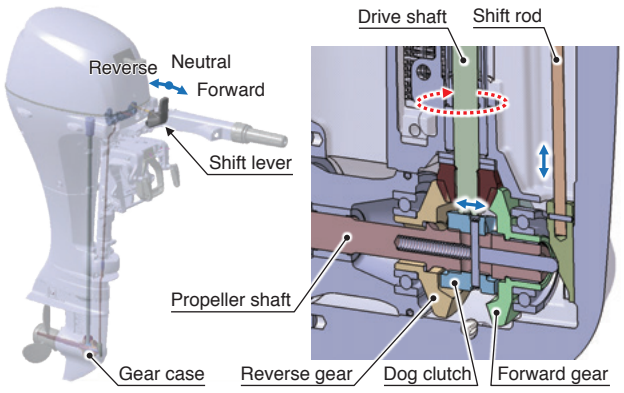


Fig. 9 Internal structure of gear case of ICE outboard motor (BF9.9)

them. The reverse gear and clutch were eliminated based on an existing ICE outboard motor gear case. However, it was necessary to reconsider adjustments for tooth contact among other factors since the gear, which had previously rotated in a fixed direction, would now rotate in forward and reverse. TOHATSU CORPORATION developed the electric outboard motor gear case shown in Fig. 12.

To electrify the control switch, an alternating switch was installed at the end of the throttle grip, as shown in Fig. 11. The switch is positioned to be close to the thumb while the operator is holding the throttle grip. This allows easy one-handed operation, in contrast to the two-handed operation required with ICE outboard motors. During berthing, when the switch is pressed with the throttle open, the power control unit (PCU) outputs an instructional value equal to the output torque multiplied by a negative value. This allows the motor to quickly decelerate, stop, and switch smoothly and quickly to reverse.

This control system simplifies the stopping operation and reduces the distance it takes to brake from a speed of 5 km/h by 30% compared to an ICE outboard motor, enabling the operator to easily berth the boat[10].



Fig. 10 Conventional method of shift switching



Fig. 11 New method of shifting by alternating switch

3.3. Controllability of Speed Adjustment Function

Speed is adjusted by opening and closing the throttle grip. As Fig. 13 shows, the operator operates the throttle by rotating the throttle grip, which turns the accelerator position sensor (APS) through a cable, as in the BENLY e:. The fully open throttle grip angle was set to 100 degrees, the same as that of an ICE outboard motor. To use the BENLY e: power unit for this purpose, the map of power output against throttle position must be tuned specifically to the boat, because the resistance and propulsive force mechanisms at work will differ for a scooter and boat. In addition, while the ICE outboard motor could not turn the propeller below idle speed, the electric motor can turn it at extremely low rotation speeds and generate high torque even at low rotation speeds, making it possible to operate

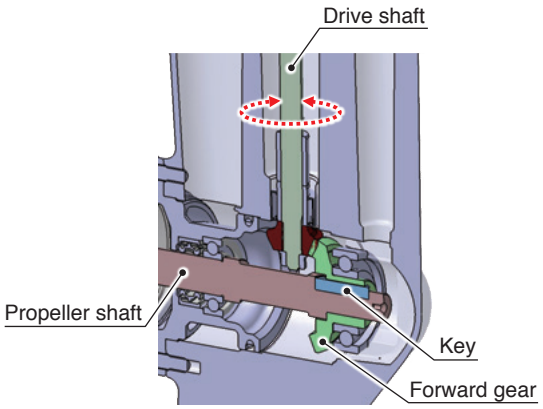


Fig. 12 Internal structure of gear case of electric outboard motor



Fig. 13 APS position in electric outboard motor

the boat even at extremely low speeds. Figure 14 shows the relationship between tour boat speed and the tuned throttle position. When establishing operability, the researchers verified how actual operators used the equipment and took the following into consideration.

(A) Considering the burden on the wrist during grip operations, it needs to be possible to control the boat speed range in practical use without exceeding a 40° throttle position.
(B) It needs to be possible to make adjustments in the cruising speed range (4 to 6 km/h) without much twisting of the wrist.
(C) It needs to be possible to operate the throttle over a wide range of angles when cruising at the extremely low speeds required for subtle maneuvering.

3.4. Adaptation of Swappable Battery to Tour Boat

The battery, the energy source for the motor, is an MPP that can easily be installed and removed, replaced, and transported.

There are many examples of electric outboard motors in the same or greater power range from other manufacturers that have large-capacity batteries mounted inside the vessel. Under that operating approach, a large-capacity battery capable of meeting the needs of the use case for one day is installed, and the battery is recharged over a period of several hours during the night when not in use. The battery is charged from a power source on land, for example, at the pier, installed especially for this purpose[7]. This study used MPPs because they do not require major modifications of existing boats and do not require a land-based power supply, which is expensive to install.

In the case of a boat, MPPs need to be swapped outdoors where there may be precipitation. This is a major difference from the case of an electric scooter, whose MPPs are likely to be swapped indoors or under the eaves of a building. As mentioned previously, these boats run as many as eight times per day, making it necessary to swap MPPs four times. Swapping time is limited to a few minutes while passengers are embarking and disembarking. Another need in the case of tour boats is that the shape and installation layout should not obstruct the view from the operator's or passengers' seats.

There are many points to consider in this use case, such as a box shape for easy replacement and installation layout on board, and water resistance during operation and swapping.

Specific approaches to address these issues include a shape that allows the MPP to be inserted and removed horizontally, as shown in Fig. 15, to help prevent water exposure. Additionally, the battery height is set to be at least 120 mm, which helps prevent water from entering the MPP box in the event of rain accumulation in the boat and makes it easier to work with. The box is located on the deck in front of the steering seat, making it easy for the operator to swap batteries and maneuver the boat.

By doing these operations, the researchers were able to construct a battery system compatible with the existing tour boat navigation use case and hull.

When using the MPP, batteries need to be swapped after every two laps of the tour. To facilitate this operation, a battery charger was installed in a maintenance shed at the "Fureai Hiroba embarkation point" in Fig. 2. Batteries were swapped while passengers were embarking and disembarking.

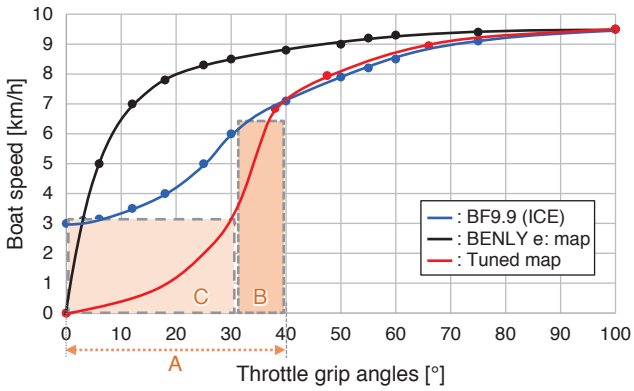


Fig. 14 Boat behavior at each throttle position



Fig. 15 Battery box shape

4. Demonstration Experiment Results

An electric outboard motor equipped with the new control system and power map described above was installed on a tour boat and the operator operated the boat to evaluate how it felt. As Fig. 16 shows, the output characteristics of the electric outboard motor provided better acceleration and responsiveness than the ICE outboard motor. Throttle controllability and ease of deceleration scored higher than those of the ICE outboard motor as a

result of the specification configuration.

The following are two new values to be gained by electrification.

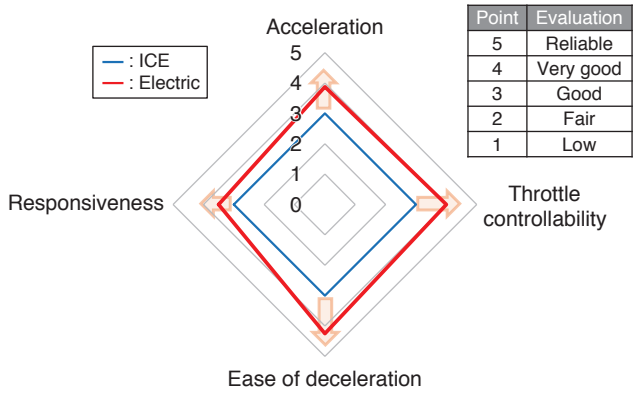
From the point of view of the tour boat operator, it is easy to detect and remove foreign objects tangled in the propeller.

Algae blooms occur in the “Horikawa Meguri” route during the summer. Although the algae is regularly cut, it cannot all be cut away, which is why it becomes tangled in propellers, as shown in Fig. 17, lowering motive power performance and increasing energy consumption. With an electric outboard motor, there is less vibration to the operator’s hand and arm, so even slight changes in this vibration that may be caused by an unbalanced propeller tangled with algae can readily be perceived. As a result, the operator can sense algae entanglement and quickly and easily remove it from the blades by reversing the direction of propeller rotation.

Regarding sound, the research showed that there is even greater value when the viewpoint is expanded to include all the passengers on the tour boat.

Given that there are quiet zones on the tour course, the researchers believed that replacing engines with motors would reduce the impact of noise on the surrounding environment. During tours, some situations require two-way communication, such as when the operator is talking to passengers as their guide and passengers ask questions of the operator. The noise from the ICE outboard engine sometimes made it a challenge for the operator to hear passengers’ questions, making it necessary to reduce the engine speed and to ask them to repeat their question.

With less noise, passengers can better hear the operator talk about the site, and communication with the passengers goes more smoothly. This enhances the value unique to a tour boat, since passengers enjoyed the ride even more.



Point	Evaluation
5	Reliable
4	Very good
3	Good
2	Fair
1	Low

Fig. 16 Comparison of drivability between ICE onboard motor and electric motor



Fig. 17 Tangled algae around propeller

5. Conclusion

The researchers constructed an outboard motor system, described below, that is easier to operate than an ICE outboard motor and is compatible with the existing tour boat navigation use case and hull.

(1) The control system that makes it easy to switch the direction of propeller rotation facilitates stopping, resulting in a 30% reduction in the distance it takes to brake from a speed of 5 km/h compared with a conventional control system.

(2) The tuning of the speed adjustment function was set to reflect the way the operator actually operates the boat, making it easy to operate at extremely low speeds and cruising speeds.

(3) MPPs were used as the motive power source and the battery box shape was optimized for both quick replacement and rain exposure for tour boats, which operate frequently and need to be water-resistant. The result was performance that was satisfactory for this use case.

Afterword

Test rides have already been offered for general passengers on boats operated by “Horikawa Meguri” operators, and many have commented on the quietness of the boat. They note, for example, that the tour is more enjoyable because they can hear birds singing, the sound of water moving on the surface while cruising, and the sounds of birds and insects. The system has also been praised for enabling operators to perform all control with one hand.

As Honda engineers, the researchers were deeply impressed by the opportunity to hear comments from people onsite and be able to develop this project together with the people of Matsue City. They will continue to study the value of electrification over the long term and hope also to study value that can be provided to passengers other than those on this tour boat.

References

(1) https://ww2.arb.ca.gov/news/carb-approves-updated-regulations-requiring-most-new-small-road-engines-be-zero-emission-2024, (accessed 2024/07/17)

(2) https://www.torqeedo.com/en/, (accessed 2024/08/09)
(3) https://global.yamaha-motor.com/news/2020/0731/harmo.html, (accessed 2024/08/09)
(4) https://www.honda.co.jp/marine/prototype-matsue/04/, (accessed 2024/09/10) (in Japanese)
(5) https://www.honda.co.jp/marine/prototype-matsue/01/, (accessed 2024/12/05) (in Japanese)
(6) https://www.matsue-horikawameguri.jp/language/en/course.html, (accessed 2024/12/05)
(7) https://www.torqeedo.com/en/news-and-press/blog/blog-2021-10-15.html, (accessed 2024/12/05)
(8) US20240317377A1
(9) US20240317078A1
(10) JP,2024-049740,A

■ Author ■


Takamasa TAKESHIGE


K e n t a K A T O


Shohma HIGANO

Development of Honda E-Clutch System for Motorcycles

Junya ONO[1] Tatsuya RYUZAKI[1] Masazumi NAITO[1]

ABSTRACT

We have developed the Honda E-Clutch system, which allows starting, gear shifting, and stopping without manually engaging or disengaging the clutch while permitting manual clutch operations when necessary. The clutch actuator added to the existing engine enables controlling the clutch automatically. With the three-split construction employed in the engine-side clutch lever and by the appropriate detection of rider's manual clutch operation, the system permits smooth change over from automatic to manual and vice versa. Equipped with the automatic clutch, the rider can now enjoy sportier and easier riding than before while maintaining an option of manual clutch operations. The system is configured to minimize restrictive factors in mounting the new system to the vehicle so that it is applicable to various types of motorcycles.

1. Introduction

Recently in the FUN category motorcycles (predominantly for the hobby), such drivetrain systems requiring no clutch operations like the dual clutch transmission (DCT) and the quick shifter have been developed and applied to mass-production motorcycles. The DCT is an automatic transmission (hereafter as AT) system composed of two hydraulically-controlled clutches and a specifically-designed transmission gear box, allowing starting, gear shifting, and stopping without manually operating the clutch. Upshifting and downshifting of the gear are also automated. The quick shifter uses the conventional manual transmission (hereafter as MT) gear box, and although manual clutch operations are necessary when starting and stopping, the gear can be shifted just by the pedal operation because the engine is controlled cooperatively for upshift or downshift of the gear. Since rider's controls from the clutch lever are reduced by such a system, the rider can travel at ease and concentrate on riding, which allows even sportier riding. Such a system, therefore, is well accepted by motorcycle users, and its applications have been increasing in recent years. In the meantime, it is necessary to accurately estimate clutch capacity in order to precisely control the clutch automatically by the electronic control system as seen in the DCT. Accordingly, such a system configuration using a hydraulic system is commonly employed. In the system consisting of a hydraulic control system, however, there were issues with an increase of the number of component parts and the difficulty of packaging in the limited space. Unlike automobiles, the motorcycle has a high degree of freedom in the maneuver of vehicle such as banking and wheely, demands are still high for control of vehicle by an instant adjustment of traction by manually operating the clutch. A system like the DCT having no clutch lever cannot deal with such a demand. Accordingly, required is such a system that allows both automatic and manual control of the clutch and fits easily in the limited space.

To deal with such demands, using the compact and easy-to-mount clutch actuator, the Honda E-clutch system has realized starting, gear shifting, and stopping without manually operating the clutch while maintaining the same configuration of existing MT and allowing conventional manual clutch operations. This paper introduces the objectives of the system, the configuration and overview of operations, the automatic control of clutch to realize target functions, and the switching from the manual clutch operation to the automatic clutch control and vice versa.

2. Development Goal

The authors consider that there is a joy of riding and controlling the "FUN" motorcycle using right and left hands, right and left feet and body actions, having the rider's body is augmented, the motorcycle becomes a part of rider's body, rather than a mere means of transportation. Among such actions, even the efforts of clutch operation

[1] Honda Motor Co., Ltd.
E-mail: junya_ono@jp.honda

and/or gear shift are a pleasure and add such a value that is an essence of fun. (An appeal of MT) On the other hand, there are some operations that some riders don't prefer to perform depending on the situation. Examples are listed below.

- Clutch operation when the engine can stall at the time of starting or stopping, or when the rider is afraid of a tip over when standing still.
- Clutch operations in traffic congestion or when the rider is tired after a long-distance travel.
- Clutch operation when pronounced instability may occur in the motorcycle when down-shifting the gear.
- Shifting the gear to neutral when stopping at a traffic light

Such "undesirable operations" can be eliminated by automating clutch operations.

The automated clutch operations can also stabilize motorcycle behavior when sports riding in twisty roads by delicately controlling the clutch just like the expert riders do. This allows the rider to concentrate on the throttle control, braking, gear shifting, and motorcycle maneuvering for "sportier" riding. Thus, the addition of automatic clutch control while maintaining the manual clutch operation makes the FUN motorcycle more user friendly and adds even more "fun."

The Honda E-Clutch system has satisfied the above-mentioned requirements by the simple system configuration applicable to existing engines. This system is aimed at providing a wide variety of riders from the beginners to experienced ones with the opportunity of next level riding (sportier, easier, and more flexibly). Anything the rider can do in the MT can also be performed in the new system which is upward compatible with the MT. Figure 1 compares various functions of MT, quick shifter, Honda E-Clutch, and DCT.

The Honda E-Clutch system requires no manual clutch operation when starting or stopping. Shifting up or down can be performed just by the shift pedal operation. The rider can also engage or disengage the clutch manually from the clutch lever on the handlebar.

Simpler/Flexible
Complex/Inflexible

Clutch type / Operation	Single clutch			Dual clutch
	Conventional manual transmission	Quick shifter	E-Clutch	DCT
Start	Clutch lever and throttle operations	Clutch lever and throttle operations	Throttle operation only	Throttle operation only
Gear change	Clutch lever and shift pedal operations	Shift pedal operation only	Shift pedal operation only (Automatic clutch control)	No operation needed (Automatic gear change / Voluntary operation possible)
Stop	Clutch lever and braking operations	Clutch lever and braking operations	Braking only	Braking only
Manual clutch control	Possible	Possible	Possible	No setting

Fig. 1 Drivetrain system types and operations

3. System Configuration and Functioning

3.1. System Configuration

In the FUN motorcycle, it is necessary to configure the system while minimizing weight increase and its influence on the component layout. That in mind, we added the clutch actuator and the motor control unit (hereafter as MCU) for automatic controls of the clutch actuator to the basic motorcycle having a MT as shown in Fig. 2. The shift sensor similar to the one in the quick shifter is incorporated into the gear shift pedal to detect rider's shifting effort.

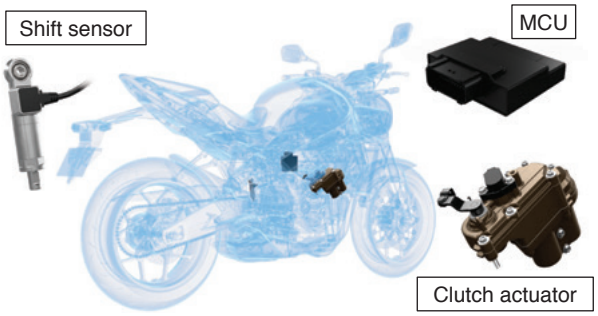
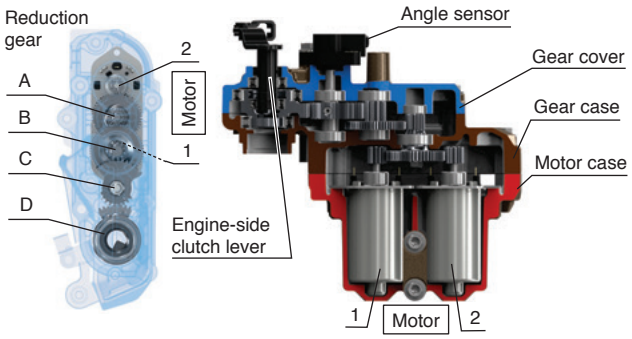


Fig. 2 Components layout

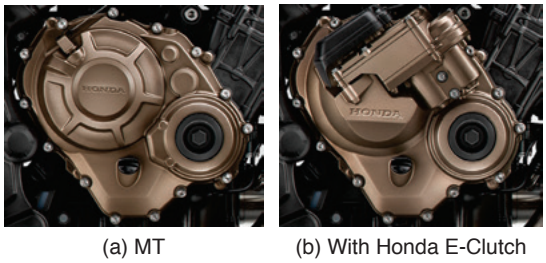


(a) Upper view (b) Cross-sectional view

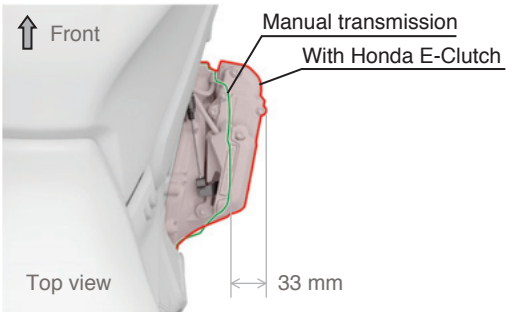
Fig. 4 Internal structure of clutch actuator

Figure 3 compares views of right-side covers of the MT model and with Honda E-Clutch model. As apparent from the picture, the clutch actuator is attached to the engine cover on the right-hand side of the vehicle. The system is configured in such a manner that the motor directly controls the clutch. With this configuration employed, the engine cover is the only part that requires extensive modifications.

Regarding the mounting of clutch actuator, the three-tier construction is employed for the clutch actuator case to accommodate a higher reduction ratio as shown in Fig. 4. Rather than using a single large motor, two small motors are used to narrow down the clutch actuator while maintaining the necessary motor power, which contributes to the increase of width as small as 33 mm to the right-hand side of the engine as shown in Fig. 3 (c)[1].



(a) MT (b) With Honda E-Clutch



(c) Comparative image of engine side cover protruding to the right

Fig. 3 Appearance and size comparison between MT and with Honda E-Clutch

This system requires a changeover of clutch control state from automatic clutch control to manual clutch operation and vice versa.

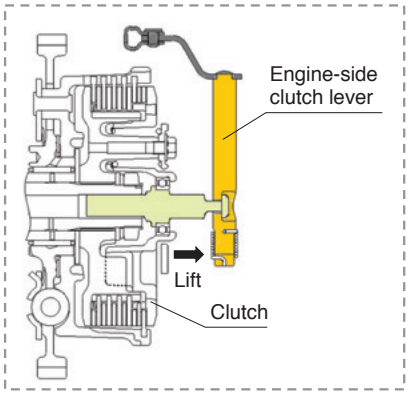
Figure 5 (a) shows the cross-sectional view of a conventional manual clutch. As the rider grabs the clutch lever, the force is transmitted via the clutch cable to the engine-side clutch lever, which turns and lifts the clutch to cut off driving force transmission.

In the Honda E-Clutch, the three-split construction is employed in the engine-side clutch lever of the clutch actuator as shown in Fig. 5 (b), which permits overriding in the direction of disengagement to satisfy the earlier-mentioned requirements[2].

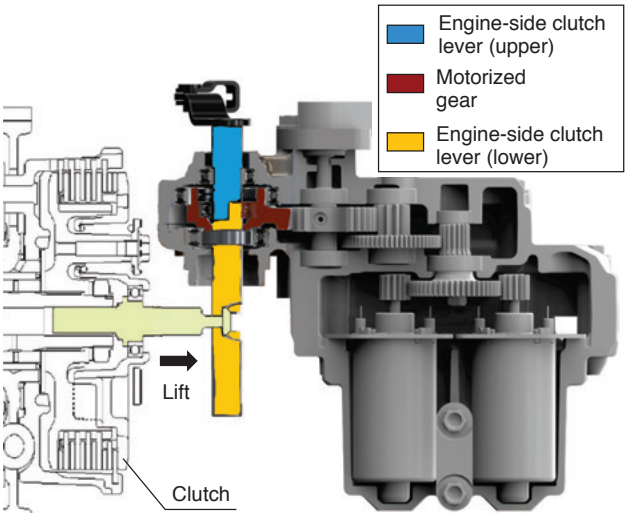
As shown in Fig. 5 (b), the left hand clutch lever is linked via the cable to the engine-side clutch lever (upper side) colored blue.

As the rider grabs the clutch lever, the clutch lever on the engine side turns, which turns engine-side clutch lever (lower side) colored yellow to operate the clutch as shown in Fig. 6 (a).

As shown in Fig. 6 (b), when automatic, the motors turn the gears colored brown, which turns the engine-side clutch lever (lower side) colored yellow to control the clutch. When the rider grabs the clutch lever while the clutch is controlled automatically, the engine-side clutch lever (upper side) drives the engine-side clutch lever (lower side) in the form of overtaking the gear, allowing overriding by the manual clutch operation, as shown in Fig. 6 (c). With this configuration employed, the motorcycle can be ridden as a standard MT vehicle in case the automatic system fails.



(a) Conventional engine-side clutch lever mechanism



(b) Honda E-Clutch lifter

Fig. 5 Engine-side three-split clutch lever mechanism

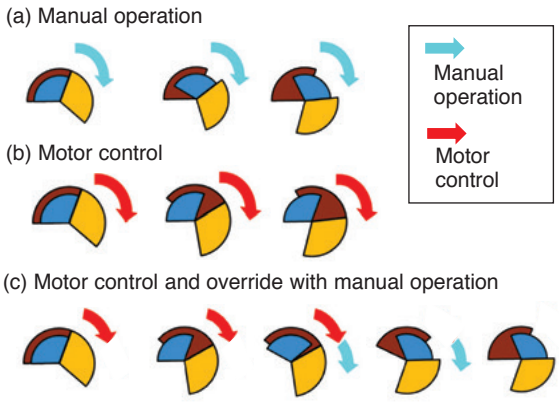


Fig. 6 Cross-sectional view of clutch lever operation diagram for engine-side three-split clutch lever

As shown in Fig. 7, the clutch actuator is controlled by the dedicated controller called MCU. The MCU receives vehicle information from the FI ECU via CAN and appropriately controls the clutch depending on the vehicle status and the clutch actuator conditions (angle sensor output and motor current).

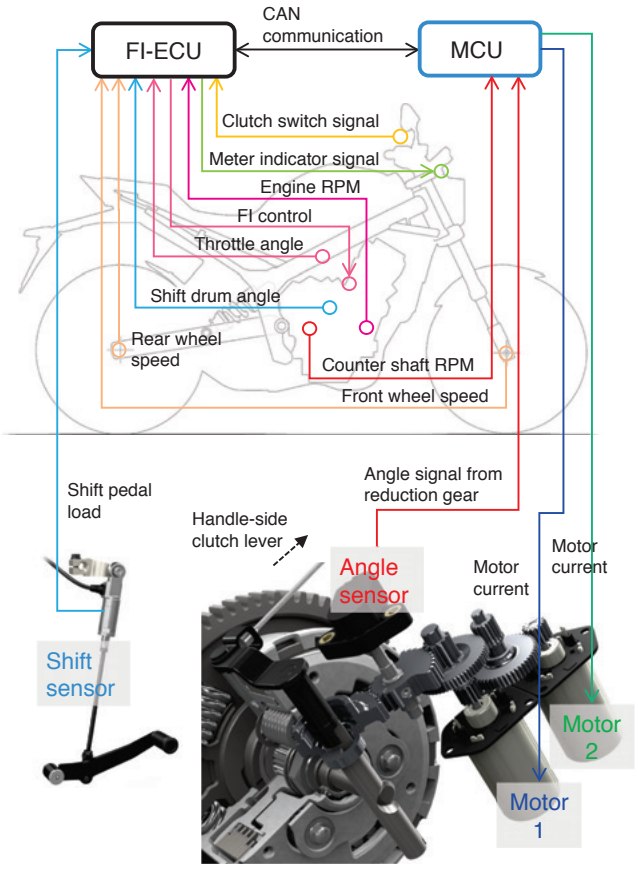


Fig. 7 System configuration overview

The above-mentioned system configuration excels in mount-ability and allows addition to existing engines.

The Honda E-Clutch system weighs only 2 kg, which is 1/5 of the DCT which is 10 kg.

3.2. Overview of System Operations

Figure 8 shows overview of operations when the Honda E-Clutch system is enabled (system is ON). When the ignition key is turned ON with the gear position at neutral and the engine is started, the system turns to the automatic control state as indicated in the green frame as shown in Fig. 8. Under these conditions, the rider can start, shift gears, and stop without using the clutch lever. Whether the automatic clutch control is enabled or disabled can be checked by the lighting of the indicator in the meter as shown in the lower right of Fig. 8.

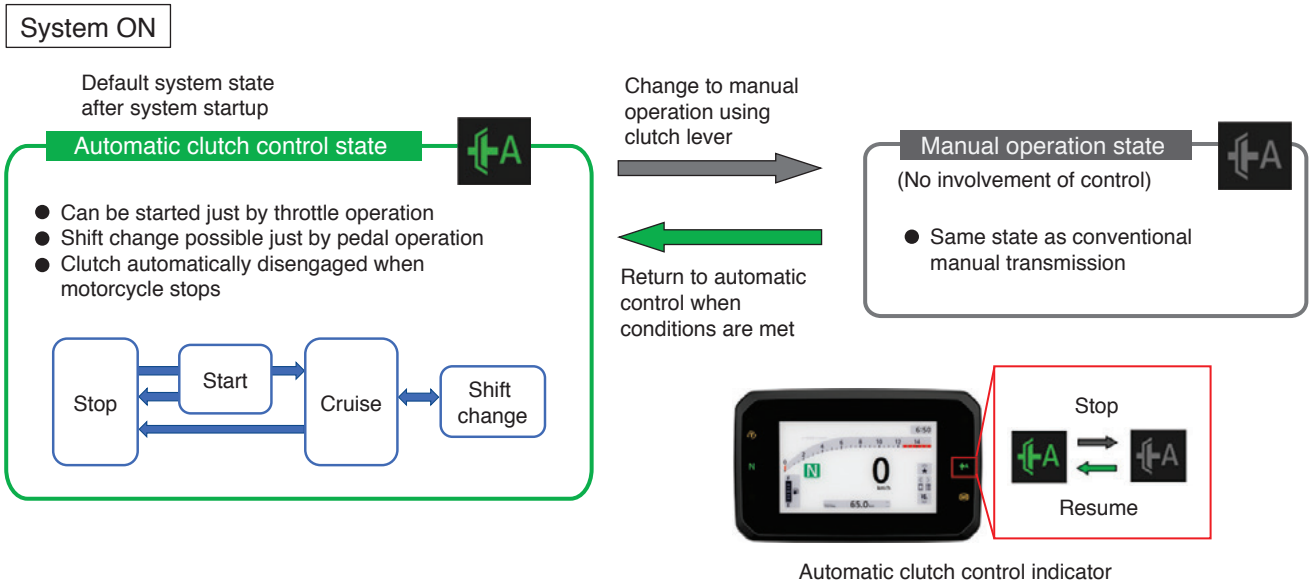


Fig. 8 Operation overview at System ON

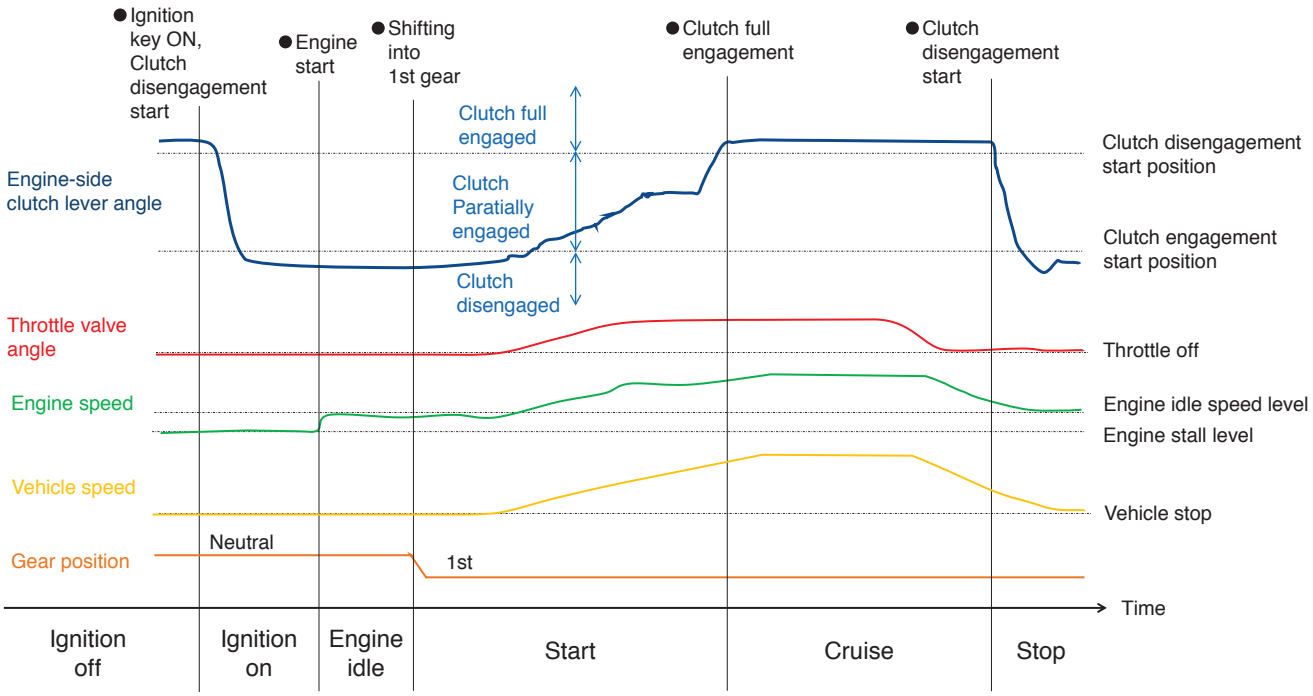
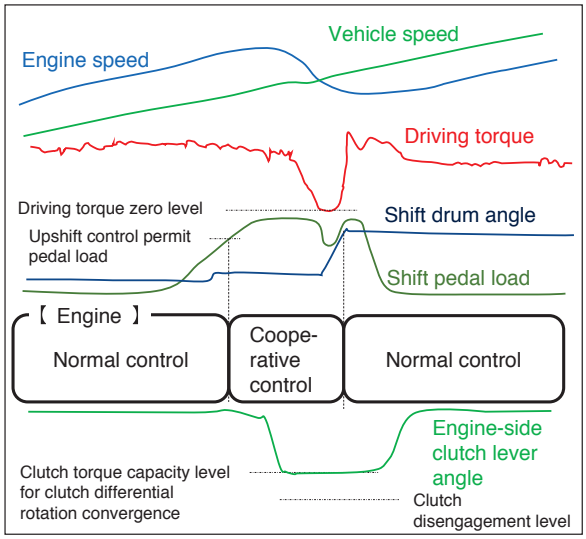


Fig. 9 Transition image of each operation state at ignition on, engine start, starting, cruising, and stopping

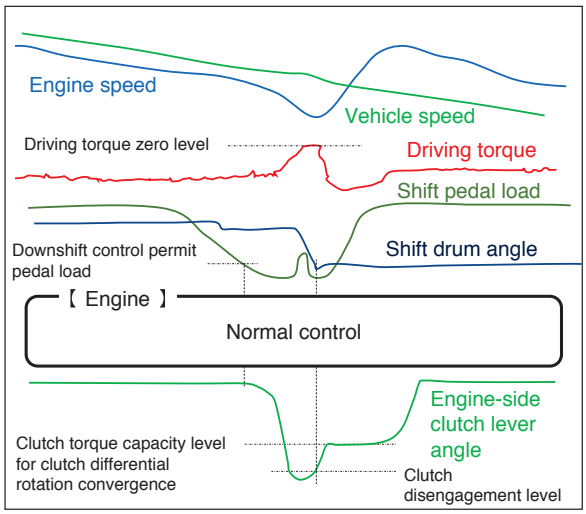
The graphs in Fig. 9 show changes of various parameters when starting, cruising, and stopping using the automatic clutch control.

After starting the engine in neutral and shifting to the first gear using only the shift pedal, the rider can start or stop by regulating the throttle and using the brake without using the clutch lever. The system appropriately controls the clutch as indicated by the changes of the angle of engine-side clutch lever shown by the blue line in Fig. 9 depending on the parameters such as the throttle valve angle, engine speed, and vehicle speed so that the vehicle behavior at the time of starting or stopping becomes linear to rider's input.

The changes of parameters when the gear is shifted only by the shift pedal without using the clutch lever are shown in Fig. 10 taking the shift up during acceleration and the shift down during deceleration as the examples.



(a) Upshifting during acceleration



(b) Downshifting during deceleration

Fig. 10 Operating image at UP, DOWN shifting

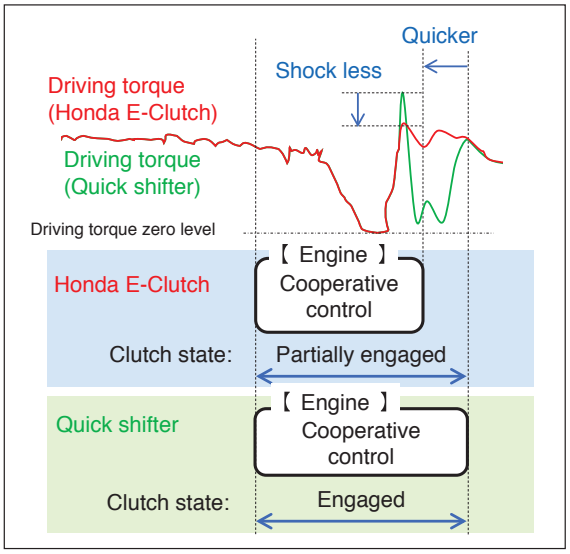


Fig. 11 Comparison of driving force between Honda E-Clutch and quick shifter during upshift

When shifting up during acceleration, the rider's shift pedal effort is detected by the shift sensor just like in the quick shifter, and the system momentarily cuts down the driving force by appropriately controlling the ignition timing and the fuel injection to allow shifting to the next gear. In an ordinary quick shifter, an impact-induced torque (shock) occurs when shifting to the next gear due to the difference of rotating speeds in the transmission gears as Fig. 11 shows. In this system, however, before shifting to the next gear, the clutch is controlled to a partial engagement depending on the engine torque to reduce the torque (shock) occurring at the time of shifting, allowing more smooth shifting than the quick shifter. When shifting down during deceleration, the rider's shift pedal effort is detected by the shift sensor just like when shifting up, and the system disengages the clutch, confirms shifting to the next gear, and synchronizes speeds of the gears by appropriately controlling the partial clutch engagement to allow complete shifting.

Explanations from here on issues rider's manual clutch operations. When the engine is running and one of the gears is selected and the rider grabs the clutch lever, the system detects a manual clutch intervention (rider's manual intervention to the clutch when the automatic clutch control is enabled), and temporarily stops the automatic clutch control. The system identifies a rider's manual clutch intervention from the clutch switch or the load on the actuator motors depending on the conditions of the vehicle. When the automatic clutch control is stopped, the indicator in the meter turns off, and the motorcycle becomes the same as the conventional type equipped with the MT (state of manual intervention). When the operation from the clutch

lever is finished from the said conditions, and the prescribed conditions to be mentioned later are satisfied, or the gear is shifted to the neutral, the automatic mode resumes. Even if the rider unintendedly grabs the clutch lever, and the system turns to the state of manual intervention, the automatic clutch control can be resumed by setting the gear at neutral. With the above-mentioned measures taken, the system permits smooth starting, shifting, and stopping without using the clutch lever as well as allowing resuming to the automatic clutch control requiring no additional operation. As shown in Fig. 12, the motorcycle can be turned to a conventional motorcycle equipped with the MT by disabling the Honda E-Clutch system (System OFF) from the meter.

These functions help the rider to enjoy sportier, easier, and more flexible motorcycle riding.

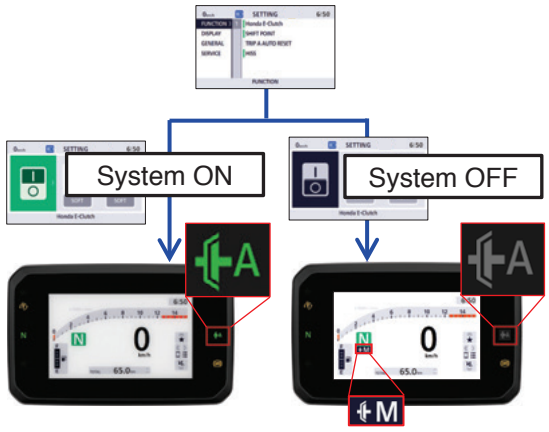


Fig. 12 System ON and System OFF

4. Automatic Clutch Control

The automatic clutch control by directly driving from the motors requires appropriate controls depending on the conditions of clutch. It is also necessary for the automatic clutch control to take into account deviations of various component parts. The clutch control satisfying these two requirements enables smooth starting, shifting, and stopping without using the clutch lever. Explained hereafter is the technical measures for the automatic clutch control.

4.1. Angle Feedback Control and Current Feedback Control Depending on Various Clutch Engagement State

The clutch engagement state can be sorted to the following three types when controlling the clutch by the clutch actuator.

(1) Engaged
(2) Partially engaged
(3) Disengaged

Figure 13 shows the relationship between the engine-side clutch lever angle (control angle), the engine-side clutch lever load (control load), and the torque transmitted by the clutch in each state.

When the clutch is engaged or disengaged, the torque transmitted by the clutch is constant where the changes of engine-side clutch lever load is small. In view of that, the control of position by the clutch lever angle is suitable when the clutch is engaged or disengaged. On the other hand, when the clutch is partially engaged, changes of clutch lever load is large, and as the torque transmitted by the clutch is inversely proportional to the clutch lever load, the control of load by the clutch lever load is suitable. In this system, the clutch lever angle can be calculated by measuring the angle of reduction gear in the actuator using the angle sensor. Accordingly, the control of position using the reduction gear angle (angle feedback control) is employed. The clutch lever load can be estimated by measuring the actuator motor current. We have therefore decided to apply the current feedback control using the motor current for the control of load.

Shifting the above-mentioned control of position to the control of load at the point the clutch engagement state changes such as the disengagement start point and the engagement start point, the clutch control performance has been enhanced[3].

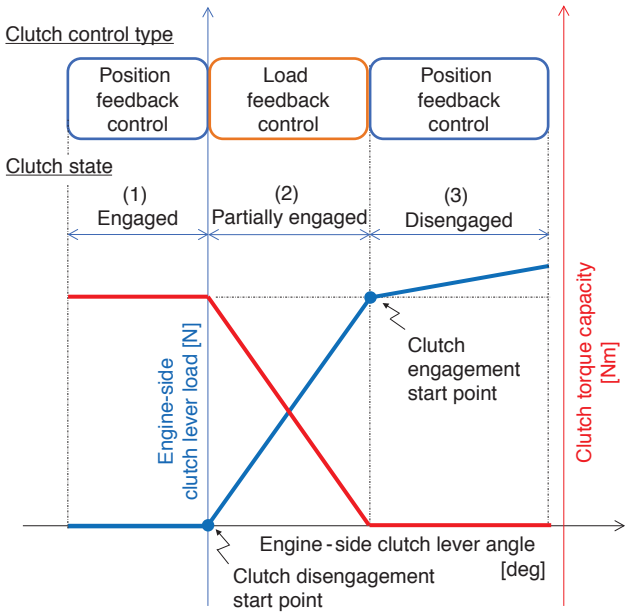


Fig. 13 Relationship between clutch state, engine-side clutch lever angle, and engine-side clutch lever load

4.2. Control of Clutch Considering Deviations of Component Parts

The system needs to control the position and the load

of engine-side clutch lever taking into account the starting positions of disengagement and engagement and the motor current (load on the clutch lever) at these positions. Such a position and current differ due to the deviation of each component part and/or the characteristics of electrical part. To realize appropriate clutch control, it is necessary to compensate for the deviations that differ from a vehicle to another. In this system, the MCU learns in each vehicle the starting positions of disengagement and engagement, the friction current of unloaded motor, and the current at the engagement position, and compensates the deviations. The system learns the starting position of disengagement by detecting the changes of working speed resulting from the changes of load when the actuator turns in the direction of disengagement from the clutch lever angle when in the engaged area. Similarly, the system learns the starting position of engagement by detecting the changes of working speed resulting from the changes of load when the actuator turns in the direction of engagement from the clutch lever angle when in the disengaged area.

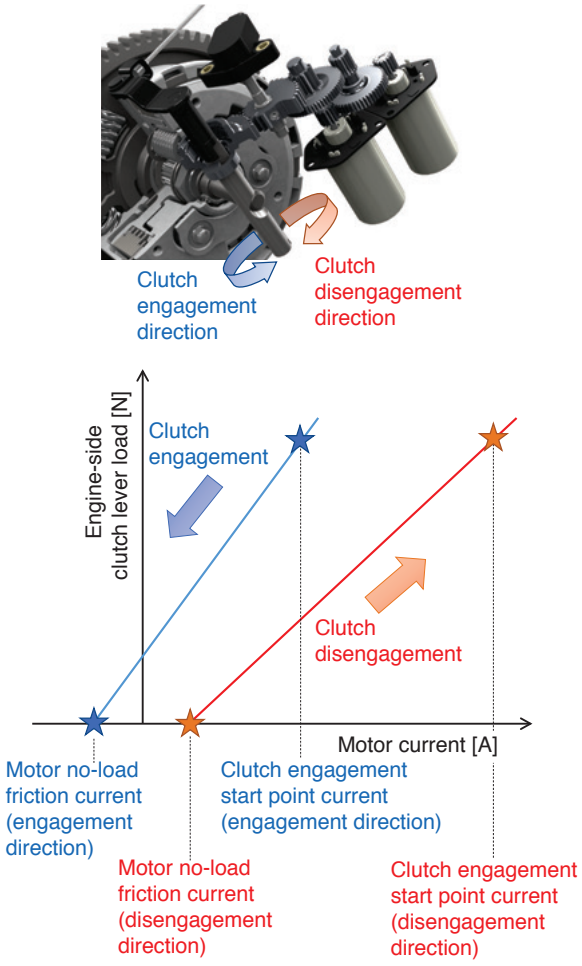


Fig. 14 Relationship between engine-side cultch lever load and motor current

Figure 14 shows the relationship between the engine-side clutch lever load and the motor current.

The relationship between the engine-side clutch lever load and the motor current is not a simple one-to-one correspondence, but differs depending on the direction of actuator motion. That is because the motor brush friction and the direction of friction occurring in the reduction gear differ depending on the direction of motion. We have therefore devised the system to learn the friction current of unloaded motor and the current at the starting position of engagement from the current occurring at each position when the actuator is operated. In the current values that the system has learned, the working directions of engagement and disengagement are takin into account[4), (5)].

By learning the above-mentioned positions and current values, various deviations of component parts in mass-production vehicles are compensated, realizing the automatic clutch control directly driven from the motors.

5. Judgment of Manual Clutch Intervention and Resuming to Automatic Clutch Control

For a temporary discontinuation of automatic clutch control when the clutch is manually operated (intervened), the manual clutch operation needs to be identified accurately. When resuming the automatic clutch control from the manually-intervened condition, the resuming needs to take place without strangeness for the rider. The system appropriately identifies vehicle conditions and rider's use of clutch, allowing the rider to choose the automatic clutch control or the manual clutch operation. How they are identified is explained hereafter.

5.1. Identification of Manual Clutch Intervention

While the automatic clutch control is temporarily discontinued when manually intervened, the clutch may engage abruptly if the intervention is incorrectly identified. It is necessary to reduce abrupt changes of reaction force on the clutch lever or vehicle behavior by an accurate identification of manual clutch intervention. Figure 15 shows how the intervention is identified. When engaged as in (1), as the clutch actuator is not actuating the clutch, there is no change of clutch condition or vehicle behavior even if the clutch control is temporarily discontinued. When in the condition (1), therefore, an intervention is identified by the clutch switch attached to the left hand clutch lever. When partially engaged (2) or disengaged (3), as the clutch is operated by the clutch actuator, even more accurate identification of manual clutch intervention is necessary. The manual clutch intervention during the automatic clutch control is identified by detecting a shift of the load on the engine-side clutch lever to rider's input from the left hand clutch lever from the changes of clutch control angle (reduction gear angle) and motor current.

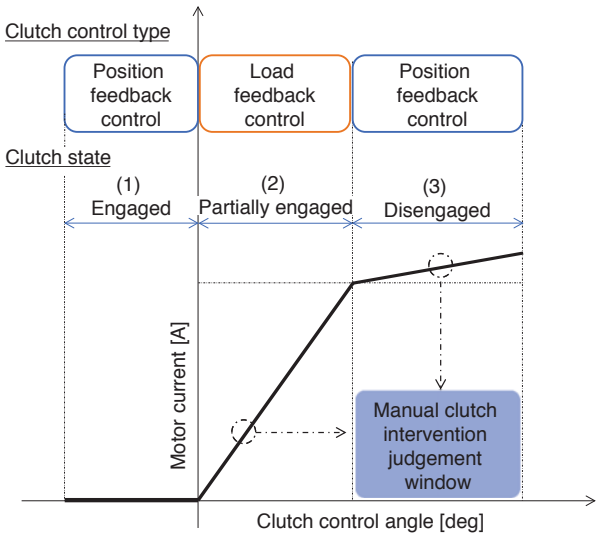


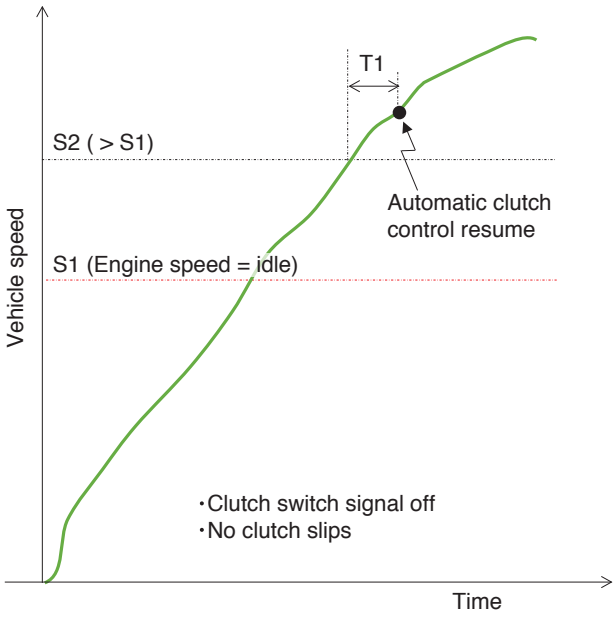
Fig. 15 Manual intervention operation judgment during automatic clutch control

When manually intervened from partially engaged (2) or disengaged (3) condition, the control of load by the actuator turns to rider's input from the left hand. The actuator's clutch control load lowers when manually intervened, and eventually falls in the manual intervention judgement window in Fig. 15. Upon confirmation of the clutch control load surely fallen in the manual intervention judgement window, the system identifies that a manual clutch intervention taking place. Employing the judgement measure for various conditions, we have realized such a method of manual clutch intervention judgement free of strangeness to the rider such as changes of reaction force on the clutch lever or vehicle behavior[6].

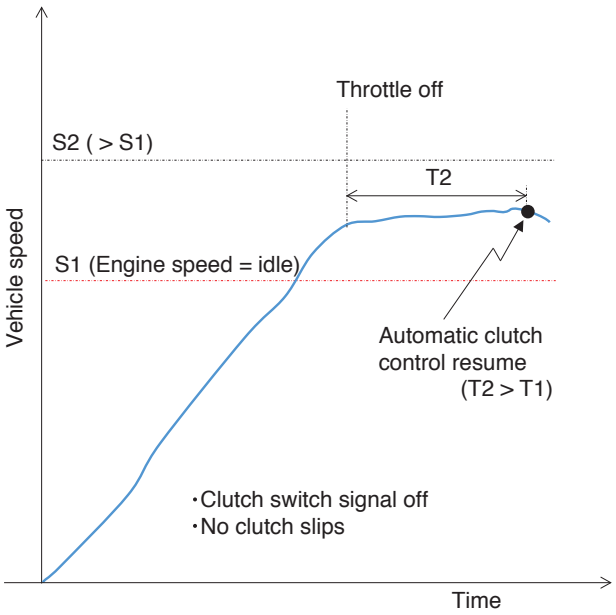
5.2. Resuming Automatic Clutch Control after Manual Clutch Intervention

When the rider finishes manual clutch intervention, it is necessary to resume automatic clutch control at a right timing. When the automatic clutch control is resumed, the system may disengage the clutch depending on the vehicle conditions (such as vehicle speed and engine speed). We have therefore employed the three resuming patterns to adjust to the vehicle conditions as in Fig. 16.

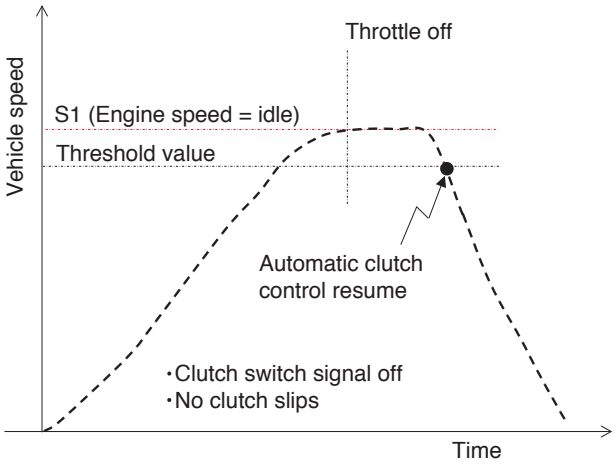
The pattern A shown in Fig. 16 (a) is when the vehicle speed is adequately higher above the idling vehicle speed after completion of the manual clutch operation. (The idling vehicle speed is the speed when the clutch is engaged and the engine speed at idling.) Under these conditions, the clutch is engaged as well when in the automatic clutch control. Since no change occurs in the vehicle behavior under the said conditions, the system resumes the automatic clutch control in a short period of time. In the pattern B shown in Fig. 16 (b), although the vehicle speed is such that the clutch is disengaged when in the automatic clutch control, if the automatic clutch control is resumed instantly under such conditions, the clutch may be disengaged, the traction may change abruptly causing sudden changes of vehicle behavior against rider's expectation. To deal with this issue, a standby period longer than that in pattern A is provided and only after the vehicle conditions have been settled, the system resumes the automatic clutch control. The pattern C shown in Fig. 16 (c) is such a pattern of resuming automatic clutch control that is intended to provide assistance for prevention of engine stall when the brake is operated without grabbing the clutch lever when standing by for resuming of automatic clutch control in the pattern B. When the rider is grabbing the clutch lever and the vehicle speed does not reach the idling vehicle speed as shown in Fig. 16 (d), the system does not resume the automatic clutch control. As such, by employing such a system configuration that resumes automatic clutch control at an appropriate timing depending on the vehicle conditions after completion of a manual clutch operation, we have realized the automatic clutch control free of strangeness and erratic vehicle behavior[7].



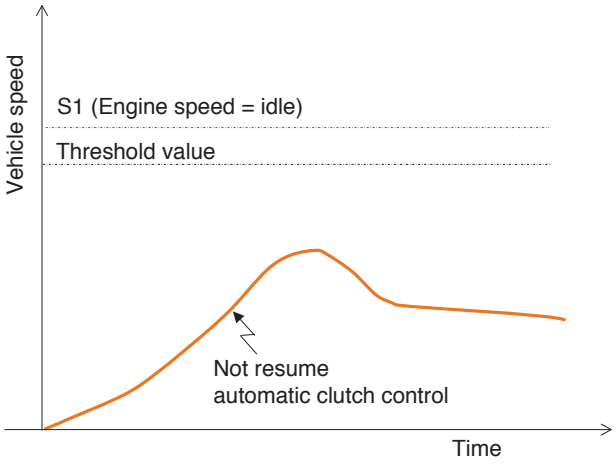
(a) Pattern A



(b) Pattern B



(c) Pattern C



(d) Other condition

Fig. 16 Resume pattern of clutch automatic control state depending on vehicle state

6. Conclusion

The Honda E-Clutch system has been developed by adding the clutch actuator, the MCU, and the shift sensor to the basic motorcycle having a MT. The appropriate clutch control has been realized by the selection of clutch control method according to the clutch engagement state along with the automatic compensation by the learning of system deviations of component parts. The automatic clutch control free of strangeness has been realized by the identification of manual clutch intervention depending on the conditions of clutch and the resuming of automatic clutch control depending on the vehicle conditions after completion of the manual operation.

With this system employed, it is now possible to start, shift the gear, and stop without grabbing the clutch lever, allowing the rider to concentrate more on the riding. At the same time, the rider can choose conventional manual clutch operations when the rider wants it. This allows the rider to enjoy sportier riding by the automatic clutch control without forcing the rider to change their riding habits. It further permits to avoiding operations that could become undesirable in certain situation and makes the riding easier by taking advantages of the AT.

This system has minimal restricting factors for mounting and is applicable to existing engines. As the increase of weight from addition of the new system is only 2 kg and its influence on the complete motorcycle is minimal, this system is applicable to various types of vehicles. We would like to extend the new system so that it becomes a standard of FUN motorcycles and more customers enjoy such products. In the future, we would like to further enhance this technology and contribute to expansion and advancement of entire motorcycling activities.

References

(1) WO2022/209708,A(2022)
(2) WO2022/209678,A(2022)
(3) WO2023/182104,A(2023)
(4) WO2022/209711,A(2022)
(5) WO2022/209632,A(2022)
(6) WO2022/209634,A(2022)
(7) WO2024/201798,A(2024)

■ Author ■


Junya ONO


Tatsuya RYUZAKI


Masazumi NAITO

Technology for Lower Limit Prediction of Lithium-ion Batteries Performance after Degradation Employing Diffusion Process Model

Kaoru OMICHI[1] Kotone SUZUKI[1]
Hideki NAGATSUKA, Ph.D.[2] Michihiro YASUDA[3]

ABSTRACT

A model for predicting lower limit performance including variation terms was constructed to help achieve performance after durability testing of products equipped with lithium-ion batteries. To predict degradation, the model formula includes two types of stochastic variables: random effects multiplied by the battery durability degradation trend term, and a diffusion process in which variation expands proportionally with time. Parameter optimization was conducted by the maximum likelihood method aided by the Markov Chain Monte Carlo method. As a result, the lower limit performance after degradation, which had conventionally been predicted based on 12 months of measured durability data, was successfully computed based on only four months of durability data. It was also demonstrated that prediction of lower limit performance, which previously required the two pieces of information of the predicted typical performance values after durability testing and performance variation during durability, could be achieved with just one piece of information by using the proposed model.

1. Introduction

Lithium-ion batteries (LiB) have high energy density and power density, and their use is expanding from small electronic devices to large mobility applications[(1)-(5)]. LiB for electric vehicles such as hybrid electric vehicles and battery electric vehicles need to provide durability over long periods of 10 years or more in consideration of diverse regions and applications[(6)]. To verify durability under various usages, it is more efficient to utilize simulation models rather than conducting individual tests using actual vehicles[(7), (8)]. The authors previously constructed a model capable of efficiently predicting battery performance after degradation in a short time. In that study, it was found that, by training the model to learn the degradation behavior of the battery under varying load conditions and predicting degradation through accumulation of the degradation amount, nonlinear degradation behavior could be predicted[(9), (10)]. However, this method was able to predict only the average performance value after degradation. On the other hand, the EU Battery Regulation sets minimum performance requirements throughout the lifetime[(6)], making it necessary to verify the LiB lower limit performance in the development phase. The lower limit performance can be predicted by subtracting the variation in Fig. 1(b) from the values predicted by the model in Fig. 1(a), as shown in Fig. 1(c). That is, it needs to be done in two steps: prediction by the degradation model and prediction of lower limit performance. Durability testing is needed in both steps, making this method both costly and time-consuming. This study aimed to address the aforementioned issue efficiently in a single step by employing the model illustrated in Fig. 1(d).

To achieve this goal, it is necessary to introduce terms into the degradation model formula that appropriately express the variation in performance after durability testing. This study employs a degradation model formula incorporating multiple random terms based on factors resulting in LiB performance variance after durability tests. This method has the issues that the number of parameters increases as random terms are introduced, which can lead

[1] Innovative Research Excellence, Power Unit & Energy
E-mail: kaoru_omichi@jp.honda
[2] Chuo University
[3] TechnoPro, Inc. TechnoPro Design Company

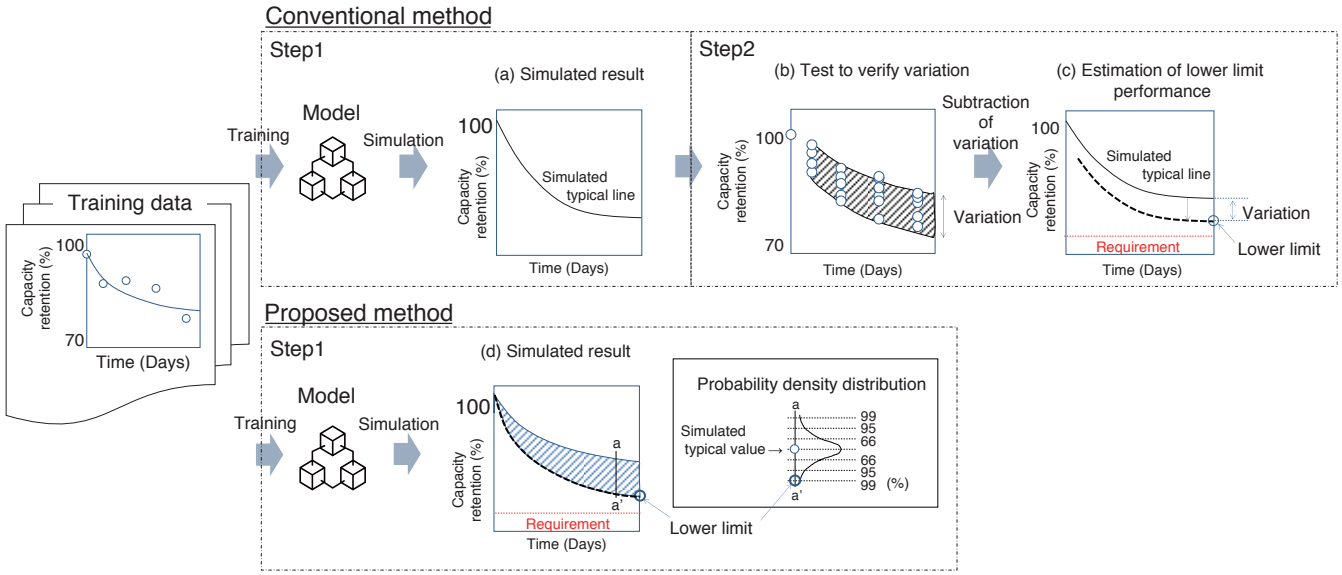


Fig. 1 Estimation of lower limit performance by conventional and proposed methods

to local optimization of parameters, and that the durability testing period needed for model training is unclear. Efforts were therefore made to address these issues.

2. Proposal of Degradation Model for Predicting Lower Limit

2.1. Consideration of Factors Resulting in Performance Variation

The remaining useful capacity after durability testing was analyzed to determine the random terms that represent variation. The battery type and durability test conditions were kept constant, and durability tests were conducted on multiple batteries to measure the changes in capacity over time. 5 Ah-class cylindrical batteries were used as test samples. Figure 2 shows the degradation behavior of 16 batteries charged and discharged from 4.1 V to 3.4 V at an ambient temperature of 55°C. The solid lines represent the behavior of capacity retention rate, and the circle plots indicate the variance of the capacity retention rate at each measurement point. As the variance tends to increase over time, there is individual variation in the degradation speed among batteries. Additionally, even for a single battery, the capacity retention rate fluctuates depending on the measurement timing, so there is measurement variation that is independent of the elapsed time. These two types of variation are present in LiB degradation behavior.

In order to construct a model based on varied performance, the parameters of the model will also include stochastic variation. As an example, Fig. 3 shows the observed degradation speed at 25, 40, and 55°C. The gray-shaded range shows the 95% confidence interval of a degradation model that has been trained on this degradation behavior. That is to say, the results output by a model trained on varied data will also contain stochastic variation, which is referred to hereinafter as model error.

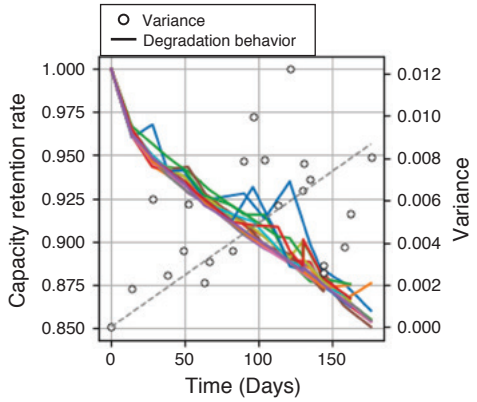


Fig. 2 Degradation behavior of LiB at same loading condition

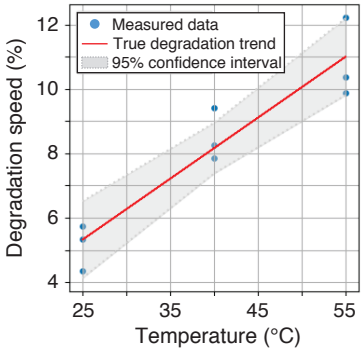


Fig. 3 Observed degradation speed and confidence interval

2.2. Formulation of Degradation Model for Predicting Lower Limit

The measurement variation of the measurement data in the LiB durability tests, the variation of durability performance, and the model error when modeling were reflected in the degradation model formula. The degradation model for predicting the lower limit consists of a deterministic trend term that indicates the degradation trend and a random term that represents variation factors. The model is formulated so that the trend and random terms match the actual phenomena. A parametric model was used so that the model error does not become excessive in case of extrapolation conditions, even with a small amount of training data. Equation (1) shows the basic model formula.

$$dX(t) = \mu(\eta(t), \boldsymbol{\theta}_0)d\eta(t)$$

where

$$\eta(t) = t^b \tag{1}$$

Here, $X(t)$ is the capacity degradation rate (1 - capacity retention rate (%)) at time t; $\boldsymbol{\theta}_0$ is a parameter of the degradation speed; $\mu(t,\boldsymbol{\theta}_0)$ is the partial derivative ($\mu(t,\boldsymbol{\theta}_0) = \frac{\partial}{\partial t}\Lambda(t,\boldsymbol{\theta}_0)$) of the trend term $\Lambda(t,\boldsymbol{\theta}_0)$ with respect to t; $\eta(t)$ is a function that converts the scale of time t[12]; and b is the scale conversion factor.

The random term that represents variation is noted as the stochastic variable ϵ that follows a normal distribution with a mean of 0 and a variance of σ_ϵ^2(both mean and variance constant over time) (Table 1).

For durability variation that expresses behavior that gradually expands over time, a diffusion process was adopted[11), (12)]. $W(t)$ in Table 1 is the standard Wiener process, and σ is the diffusion coefficient.

Model error is a variation term that accumulates in the trend term. This is noted as the stochastic variable λ that follows a normal distribution with a mean of μ_λ and variance of σ_λ^2 (both mean and variance constant over time) and accumulates in the trend term of the performance formula. Note that variables introduced for the purpose of replacing deterministic variables with stochastic variables in the model are generally referred to as random effects[13]. Also, the three random terms above are assumed to be independent of each other.

Table 2 shows multiple candidate models based on combinations of how the above variation terms are incorporated.

In this study, a total of 2^3-1=7 types of model formulas were formulated based on the presence or absence of the three types of error. However, there is no data that includes only model error without measurement variation or degradation variation, so that case was excluded from the candidate models, and the six models M1 to M6 were examined. Note that M1 is the conventional model that considers only measurement error, which is equivalent to a model using the least squares method.

Table 1 Three types of degradation formulas with variation terms

Items	Independence of degradation trends	Time dependence of variation	Formula
Measurement variation	Yes	No	$X(t) = \Lambda(\eta(t), \theta_0) + \varepsilon$ $\epsilon \backsim N(0, \sigma_\varepsilon^2)$
Interindividual variation	Yes	Yes	$dX(t) = \mu(\eta(t), \theta)d\eta(t) + \sigma_B dW(\eta(t))$ $W(\eta(t)) \backsim N(0, \eta(t)))$
Model variation	No	Yes	$dX(t) = \lambda \times \mu(\eta(t), \theta)d\eta(t)$ $\lambda \backsim N(0, \sigma_\lambda^2)$

Table 2 Six types of model formulas with combinations of variation terms

Model No.	Measure-ment variation	Inter-individual variation	SIM variation	Model formula
M1	○			$X(t) = \Lambda(\eta(t), \boldsymbol{\theta}_0) + \boldsymbol{\varepsilon}$
M2		○		$dX(t) = \mu(\eta(t), \boldsymbol{\theta}_0)d\eta(t) + \boldsymbol{\sigma_B dW(\eta(t))}$
M3		○	○	$dX(t) = \boldsymbol{\lambda} \times \mu(\eta(t), \boldsymbol{\theta}_0)d\eta(t) + \boldsymbol{\sigma_B dW(\eta(t))}$
M4	○	○		$X(t) = Y(t) + \boldsymbol{\epsilon}$, $dY(t) = \mu(\eta(t), \boldsymbol{\theta}_0)d\eta(t) + \boldsymbol{\sigma_B dW(\eta(t))}$
M5	○	○	○	$X(t) = Y(t) + \boldsymbol{\epsilon}$, $dY(t) = \boldsymbol{\lambda} \times \mu(\eta(t), \boldsymbol{\theta}_0)d\eta(t) + \boldsymbol{\sigma_B dW(\eta(t))}$
M6	○		○	$X(t) = Y(t) + \boldsymbol{\epsilon}$, $dY(t) = \boldsymbol{\lambda} \times \mu(\eta(t), \boldsymbol{\theta}_0)d\eta(t)$

$\eta(t) = t^b, t > 0, b > 0, \lambda \sim N(\mu_\lambda, \sigma_\lambda^2), \mu_\lambda > 0, \sigma_\lambda > 0, W(\eta(t)) \backsim N(0, \eta(t)), \epsilon \sim N(0, \sigma_\epsilon^2), \sigma_\epsilon > 0$

Table 3 List of trend model formulas

No.	Model	Formula	Number of coefficients
(i)	Linear	$\sum_{i=1}^{5} a_i \times X_i + a_0$	6
(ii)	Linear + interaction	$\sum_{i=1}^{5} a_i \times X_i + \sum \sum_{1 \le i < j \le 5} a_{i,j} \times X_i \times X_j + a_0$	16
(iii)	Quadratic	$\sum_{i=1}^{5} a_{i,i} \times X_i^2 + \sum_{i=1}^{5} a_i \times X_i + a_0$	11
(iv)	Quadratic + interaction	$\sum_{i=2}^{5} a_{i,i} \times X_i^2 + \sum_{i=1}^{5} a_i \times X_i +$ $\sum \sum_{1 \le i < j \le 5} a_{i,j} \times X_i \times X_j + a_0$	21

X_1: SOC; X_2: Temperature; X_3: ΔSOC; X_4: I-charge; X_5: I-discharge

The trend term is expressed as shown in Eq. (2) using the variable $a(\theta_0,t)$ that expresses the degradation speed.

$$\Lambda(\eta(t), \theta_0) = a(\theta_0, t) \times \eta(t) \tag{2}$$

$a(\theta_0,t)$ is a variable indicating the degradation speed composed of the load variables $SOC(t)$, $Temperature(t)$, $\Delta SOC(t)$, $I\text{-}charge(t)$, and $I\text{-}discharge(t)$ and parameters (a_0, a_1, a_2 $\cdots b$). As shown in Table 3, the formulated models include linear models (i) and (ii) with first-order linear factor effects for each degradation factor, quadratic polynomial models (iii) and (iv) that do not correspond to first-order linear equations, and interaction models (ii) and (iv) that have interaction terms between factors. B uses a unique appropriate value depending on the type of battery[10].

3. Construction of Parameter Optimization Process

3.1. Parameter Optimization Method

The trend term parameters (a_0, a_1, a_2 $\cdots b$) and variation term parameters (σ_ϵ, σ_B, σ_λ, μ_λ) included in the model formula were optimized based on actual test data. As these include nonlinear parameters, parameter estimation was performed by the maximum likelihood method with the Log-Likelihood (LL) function as the object variable. Here, let X_i ($t_{i,j}$)) represent the stochastic process that expresses the performance value of cell i at the time $t_{i,j}$. The LL function of model formula M5, which has the largest number of parameters among the proposed models M2 to M6, is given by Eq. (3). To help avoid getting stuck in local solutions in the optimization calculations, the maximum likelihood estimation (MLE) values were obtained using the Markov Chain Monte Carlo (MCMC) method[14), (15)]. Specifically, a normal distribution with large variance was set as the prior distribution of each parameter. The MCMC method employed the Hamiltonian Monte Carlo (HMC) method, which is known to work effectively even with a large number of parameters, and the MLE values were obtained by performing maximum a posteriori (MAP) estimation.

$$l(\theta) = -\frac{\sum_{i=1}^{n} m_i}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\ln|\Sigma_i| - \frac{1}{2}\sum_{i=1}^{n}(\Delta X_i - \mu_\lambda \Delta\Lambda_i(\theta_0))^T \Sigma_i^{-1}(\Delta X_i - \mu_\lambda \Delta\Lambda_i(\theta_0)) \quad (3)$$

However,

$$\{\ \theta = (\theta_0,\ b,\ \mu_\lambda,\ \sigma_\lambda^2,\ \sigma_B^2,\ \sigma_\epsilon^2)\ ,$$

$$\mu_\lambda \in R,$$

$$\sigma_\lambda > 0,$$
$$\sigma_B^2 > 0,$$
$$\sigma_\epsilon^2 > 0,$$

$$\Sigma_i = \sigma_\lambda^2 \Delta\Lambda_i(\theta_0)\Delta\Lambda_i(\theta_0)^T + \Omega_i\ ,$$

$$\Omega_i = \sigma_B^2 \times diag(\Delta\eta_i) + \sigma_\epsilon^2 F,$$

$$F = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 2 & -1 & \cdots & 0 \\ 0 & 0 & 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

$\Delta X = (\Delta X_1, \cdots, \Delta X_n)^T$,
$\Delta X_i = (\Delta X_{i,1}, \cdots, \Delta X_{i,m_i})^T$,
$\Delta X_{i,j} = (X_i(t_{i,j}) - X_i(t_{i,j-1}))$,
$X_i(t_{i,j})$: amount of degradation of cell i at time $t_{i,j}$

$\Delta\Lambda_i(\theta_0) = (\Delta\Lambda_{i,1}(\theta_0), \cdots, \Delta\Lambda_{i,m_i}(\theta_0))'$,
$\Delta\Lambda_{i,j}(\theta_0) = a(\theta_0, t_{i,j})\{\eta(t_{i,j}) - \eta(t_{i,j-1})\}$,
$a(\theta_0, t_{i,j})$: degradation acceleration factor of cell i at time $t_{i,j}$
$\Delta\eta_i = (\Delta\eta_{i,1}, \cdots, \Delta\eta_{i,m_i})'$,
$\Delta\eta_{i,j} = \eta(t_{i,j}) - \eta(t_{i,j-1}) = t_{i,j}^b - t_{i,j-1}^b$
$t_{i,j}$: time at which the amount of degradation of cell i was observed for the jth time.
$1 \le j \le m_i, 1 \le i \le j\ \}$

3.2. Model Formula Selection Method

Selection of the trend formula and variation model formula is determined by Akaike's Information Criterion (AIC) based on the training data. AIC is commonly used as an index enabling selection of the optimal model while helping to prevent overfitting to only the training data. AIC is calculated using Eq. (4) from the final LL value used as the object variable in optimization and the number of parameters in the model (k).

$$AIC = -2LL + 2k \quad (4)$$
k: number of parameters

3.3. Prediction of Lower Limit Performance

The distribution of $\Delta X_{i,1:m_i+1}$ under the given $\Delta X_{i,1:m_i}$ is obtained by the following Eq. (5).

$$\Delta X_{i,\ m_i+1} | \Delta X_{i,\ 1:m_i} \sim N(\mu_{m_i+1|m_i},\ \sigma_{m_i+1|m_i}^2), \quad (5)$$

However,

$$\{\ \mu_{m_i+1|m_i} = \mu_\lambda \Delta\Lambda_{i,\,m_i+1}(\theta_0) + \omega_{i,1}' \Sigma_{i,1:m_i}^{-1}(\Delta X_{i,1:m_i} - \mu_\lambda \Delta\Lambda_{i,1:m_i}(\theta_0))$$
$$\sigma_{m_i+1|m_i}^2 = \sigma_{i,m_i+1}^2 - \omega_{i,1}' \Sigma_{i,1:m_i}^{-1} \omega_{i,1},$$
$\omega_i = (0,\ \cdots, 0\ , -\sigma_\varepsilon^2)'$, ($m_i$-dimensional vector),
$\boldsymbol{\Sigma}_{i,1:m_i} = \sigma_\lambda^2 \Delta\Lambda_{i,1:m_i}(\theta_0)\Delta\Lambda_{i,1:m}(\theta_0)' + \Omega_{i,1:m}$,
$\boldsymbol{\Omega}_{i,1:m_i} = \sigma_B^2 diag(\Delta\eta_{i,1:m_i}) + \sigma_\epsilon^2 \boldsymbol{F}_{obs}$,

$$F_{obs} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 2 & -1 & \cdots & 0 \\ 0 & 0 & 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix},$$

$(m_i \times m_i$ matrix),

$$\Sigma_i^{-1} = \Omega_{i,1:m_i}^{-1} - \frac{\sigma_\lambda^2}{1 + \sigma_\lambda^2 \Delta\Lambda_{i,1:m_i}(\theta_0)' \Omega_{i,1:m_i}^{-1} \Delta\Lambda_{i,1:m_i}(\theta_0)} \times \Omega_{i,1:m_i}^{-1} \Delta\Lambda_{i,1m_i}(\theta_0) \Delta\Lambda_{i,1:m_i}(\theta_0)' \Omega_{i,1:m_i}^{-1},$$
$$\sigma_{i,m_i+1}^2 = \sigma_\lambda^2 \Delta\Lambda_{i,m_i+1}(\theta_0)^2 + \sigma_B^2 \Delta\eta_{i,m_i+1} + 2\sigma_\epsilon^2\ \}$$

Therefore, when the upper α% point of the standard normal distribution is noted as z_α, the 100×(1-α)% confidence interval for the degradation increment $\Delta X_{i,j}$ of the individual cell number i at time t_{i,m_i+1}, under the given $\Delta X_{i,1:m_i}$, is given by Eq. (6).

$$\left[\hat{\mu}_{m_i+1|m_i} - \hat{\sigma}_{m_i+1|m_i} z_{\frac{\alpha}{2}},\ \hat{\mu}_{m_i+1|m_i} + \hat{\sigma}_{m_i+1|m_i} z_{\frac{\alpha}{2}}\right] \quad (6)$$

3.4. Preparation of Training and Validation Data

The results of durability tests conducted under multiple conditions were used as the data set. The tests utilized 4.5 Ah-class cylindrical LiBs as test samples, and were conducted in two modes: storage durability tests and cycle durability tests. The storage durability tests for training data were conducted over a 4-month period under 16 conditions comprising all combinations of four SOC conditions (SOC30, 55, 80, and 95%) and four temperature conditions (25, 45, 50, and 65°C). During the testing period, intermediate performance measurements were conducted at intervals of one to two weeks, and 10 data points were obtained per condition for a total of 160 data points. The cycle durability tests were conducted under 27 conditions planned based on an L27 orthogonal array comprising five cycling control factors: center SOC, operating SOC range, temperature, charge current, and discharge current. Ten data points were obtained over a 4-month period for each condition, for a total of 270 data points. For the validation data, durability tests were conducted over a period of 12 months under 38 conditions separate from the training conditions, and obtained a total of 710 data points.

4. Results and Discussion

4.1. Model Selection Result

Regarding model selection, parameter optimization was performed for 24 types of model formulas comprising combinations of four types of trend term model formulas and six types of variation models (M1 to M6), and the AIC was evaluated. Figure 4 shows the results.

In all cases of M2 to M6 constructed in this study, good results were obtained with an AIC lower than that of the conventional model M1. This demonstrates that model formulas using the diffusion process, random effects, or both can appropriately represent the experimental data.

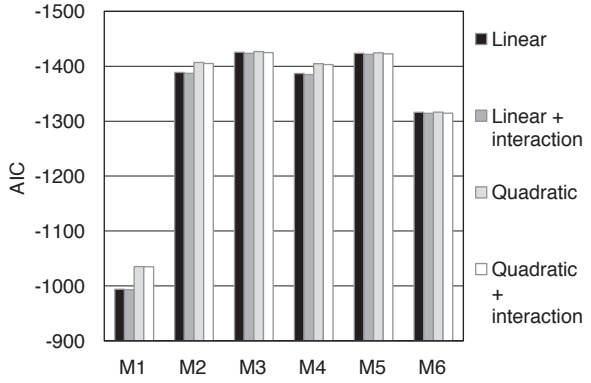
Generally, the accuracy of the model formula increases with the number of parameters, but too many parameters lead to a drop in generalization performance. Therefore, it is appropriate to select the model with the smallest AIC that also balances generalization performance and goodness of fit. Of the degradation model formulas, M3 has fewer parameters than M5 but also a smaller AIC. That is to say, M3 captures the essence of the battery performance degradation obtained from the tests with fewer parameters. Here, the combination of the quadratic (iii) trend term formula and the M3 model formula resulted in the best fit to the experimental data. Also, it was found that enhancement of AIC was not observed with models M4 to M6, which include measurement variation, so there is no need to consider these in the model.

4.2. Prediction of Confidence Interval of Selected Model

Equation (7) shows the model formula and parameters of the prediction model M3 selected in the previous section. In addition, Fig. 5 shows plots of the measurement results from a 12-month durability test conducted at SOC100% and temperature of 45°C, along with an example of the confidence interval prediction result calculated using the optimized M3 degradation model. The confidence interval indicated in gray is the 95% confidence interval at each point in time. The measured plots across the entire range were confirmed to fall within the 95% confidence interval, thus demonstrating the ability to appropriately visualize the confidence interval.

$$dX(t) = \boldsymbol{\lambda} \times \mu(\eta(t),\ \boldsymbol{\theta}_0)d\eta(t) + \boldsymbol{\sigma_B} \boldsymbol{dW}(\boldsymbol{\eta(t)})$$
$$\mu(\eta(t),\ \boldsymbol{\theta}_0)d\eta(t) = (a_1 SOC + a_2 Temp + a_{11} SOC^2 + a_{22} Temp^2 + a_0)d\eta(t)$$
$$\eta(t) = t^b \quad (7)$$

Term	Value	Term	Value
a_0	-5.13	μ_λ	-1.7E-2
a_1	2.6	σ_λ	1.6E-3
a_2	-1.7E-3	σ_B	7.3E-4
a_{11}	-3.2E-1		
a_{22}	1.0E-4		
b	0.55		



			Variation models					
		Parameter	1	2	3	2	4	3
	Para-meter		M1	M2	M3	M4	M5	M6
Model formula of trend term	4	(i) Linear	-993.8	-1388.7	-1425.5	-1386.7	-1423.5	-1316.3
	5	(ii) Linear+ interaction	-992.8	-1387	-1423.5	-1385	-1421.5	-1314.3
	6	(iii) Quadratic	-1035.3	-1406.8	-1426.7	-1404.8	-1424.6	-1316.6
	7	(iv) Quadratic+ interaction	-1034.8	-1405.1	-1424.7	-1403.1	-1422.6	-1314.7

Fig. 4 AIC results of each model

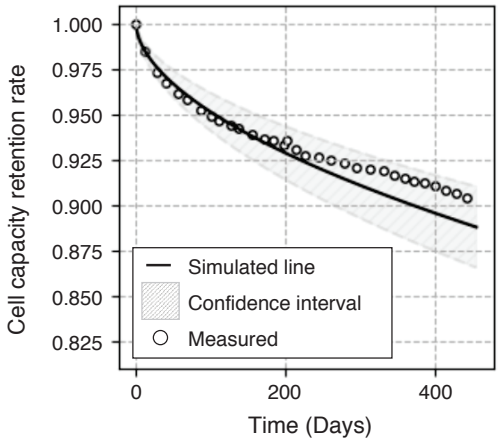
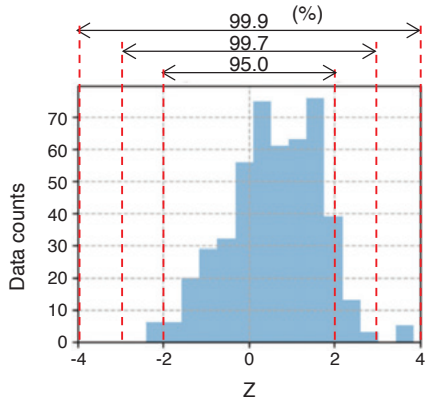


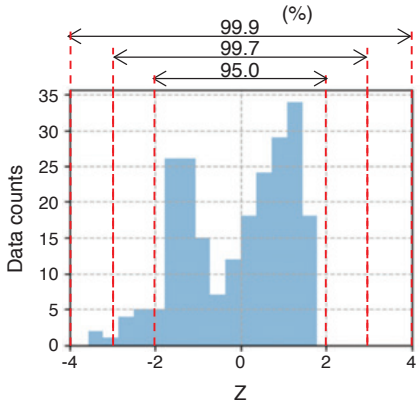
Fig. 5 Confidence interval and measured data points

4.3. Confirmation of Data Distribution

To confirm the likelihood of the model created in this study, the degree of coverage of the predicted performance distribution relative to the measured performance values was analyzed for an amount of data greater than the training data (storage degradation mode data: 24 conditions, 484 points; cycle degradation mode data: 14 conditions, 226 points). Figure 6 shows histograms of the Z values obtained by taking the difference between the typical performance value predicted by the model and the measured value and dividing by the standard deviation at that point. Figure 6(a) is the histogram of the data obtained in storage degradation mode, and Fig. 6(b) is the histogram of the data obtained in cycle degradation mode. The Z values of 4, 3, and 2 correspond respectively to the confidence intervals that theoretically contain 99.9%, 99.7%, and 95%. In this study, these confidence intervals respectively contained 100%, 99.0%, and 93.6% of the storage mode data and 100%, 98.7%, and 94.0% of the cycle mode data, generally matching the theoretical values. It is thought, therefore, that lower limit performance based on confidence can be calculated using this model. The percentage of the measured data contained in the 99.7% confidence interval is referred to herein after as the coverage percentage.



(a) Histogram of data obtained by storage mode degradation

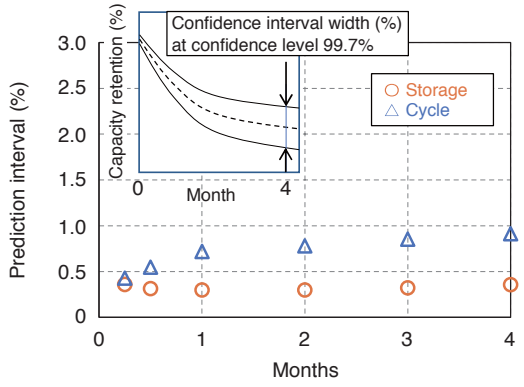


(b) Histogram of data obtained by cycle mode degradation

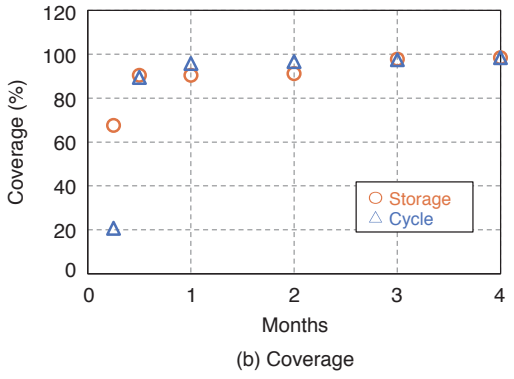
Fig. 6 Histograms of Z values after standardization

4.4. Setting of Training Period

In order to set the period required for model construction, the amount of training data, the width of the confidence interval, and the coverage percentage were analyzed. The width of the confidence interval (%) and the coverage percentage in the 99.7% confidence interval predicted by each model were calculated at the 4-month point, with Fig. 7 showing the results.



(a) Confidence interval width



(b) Coverage

Fig. 7 Confidence interval width and coverage of models trained with data over different periods

A model was constructed using the data up to an arbitrary number of months from four months of training data obtained by storage durability tests under 16 conditions and cycle durability tests under 27 conditions. The coverage percentage was calculated for a total of 380 data points obtained under 38 conditions separate from the training data.

The width of the confidence interval tends to increase slightly as the amount of training data increases. In contrast, the coverage percentage clearly increases, and the coverage percentage of the 99.7% confidence interval increases asymptotically toward 99.7% as theorized. This is presumed to be due to the effect that as the amount of training data increases, the accuracy of the predicted confidence interval is enhanced without greatly impacting the confidence interval result. In this study, a coverage percentage of 99% or more was achieved with a 4-month training period, indicating that the confidence interval can be predicted by training on four or more months of degradation behavior data.

4.5. Confidence Interval in Long-term Prediction

Prediction of the performance degradation of batteries that are used in various ways becomes more uncertain as the prediction period lengthens, which is expected to result in a wider confidence interval.

Figure 8 shows the results of long-term prediction of the confidence interval over 10 years. Figure 8(a) shows the result for the conventional model M1 that considers only measurement variation, and Fig. 8(b) shows the result for the newly proposed optimized prediction model M3.

Whereas the conventional model predicts a uniform confidence interval over a period of 10 years, the model optimized in this study exhibits a confidence interval that expands over time, which is in line with initial expectations. It was confirmed that, compared to the conventional model M1, the model M3 proposed in this study can appropriately model the increase over time of variation appearing in four months of test data.

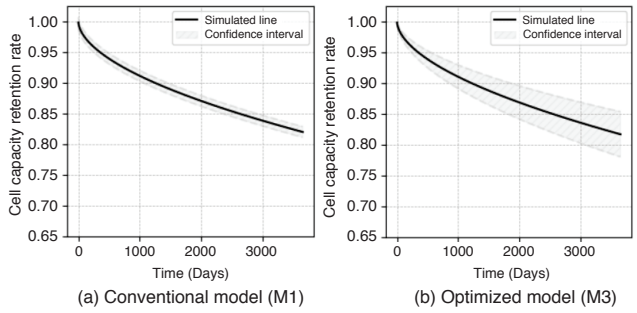


Fig. 8 Simulated confidence interval for 10 years

5. Conclusion

In order to use a model to efficiently predict the lower limit of lithium-ion battery performance after durability testing, a degradation model formula that incorporates random terms was examined. The conclusion drawn from this examination is as follows.

(1) It was found that the variation in lithium-ion battery durability data can be represented by a model incorporating two kinds of stochastic variables: random effects multiplied by the LiB durability degradation trend term, and a diffusion process in which variation expands proportionally with time.
(2) Use of the MLE method aided by MCMC enables optimization of the model without getting stuck in local solutions.
(3) A model was constructed that can appropriately predict the confidence interval over a 12-month period by training on only four months of durability test results.

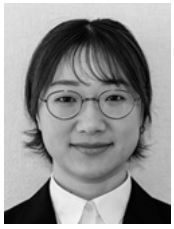
References

(1) New Energy and Industrial Technology Development Organization: Battery Road Map 2013, (2013) (in Japanese), https://www.jstage.jst.go.jp/article/ieejias/134/3/134_NL3_5/_pdf, (accessed 2024/12/24)
(2) Fuji Keizai Group Co., Ltd.: Energi, ogata nijidenchi, zairyo no shorai tenbo 2021 - Dendo jidosha, shasaidenchi bunya hen -, p. 5-12, (2021) (in Japanese)
(3) METI: Ministry of Economy, Trade and Industry, https://www.meti.go.jp/policy/mono_info_service/joho/conference/battery_strategy/0001/03r2.pdf, (accessed 2024/12/24) (in Japanese)
(4) Dunn, B., Kamath, H., Tarascon, J.-M.: Electrical Energy storage for the Grid: A Battery of Choices, Science, Vol. 334, Issue 6058, p. 928-935, (2011), doi: 10.1126/science.1212741
(5) Schmuch, R., Wagner, R., Hörpel, G., Placke, T., Winter, M.: Performance and cost of materials for lithium-based rechargeable automotive batteries, Nature Energy, Vol. 3, p. 267-278, (2018), doi: 10.1038/s41560-018-0107-2
(6) Final Regulation Order, Section 1962.8, Title 13, California Code of Regulations, (2022), https://ww2.arb.ca.gov/sites/default/files/barcu/regact/2022/accii/acciifro1962.8.pdf, (accessed 2024/12/24)
(7) Waag, W., Fleischer, C., Sauer, D. U.: Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles, Journal of Power Sources, Vol. 258, p. 321–339, (2014), doi: 10.1016/j.jpowsour.2014.02.064
(8) LcGrasso, J., Ressler, G., Inguva, S.: MEETING CUSTOMER EXPECTATIONS: Life and Safety of Batteries in Hybrid and Electric Vehicles, SAE Technical Paper TP-AAB-AABC_US_2011_20, (2011)
(9) Tominaga, Y.: Degradation Model for Battery Performance Based on Performance Mechanism, The Electrochemical Society of Japan, Vol. 89, No. 2, p.

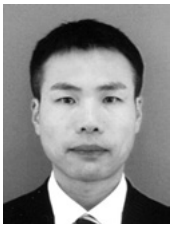
125-132, (2021), doi: 10.5796/denkikagaku.21-FE0014 (in Japanese)
(10) Omichi, K., Kaga, T., Nishiimura, T., Aida, N., Tominaga, Y.: Modeling of Efficient Lifetime Prediction for Lithium-ion Battery Employing Orthogonal Array, Honda R&D Technical Review, Vol. 35, p. 32-38
(11) Xu, X., Tang, S., Yu, C., Xie, J., Han, X., Ouyang, M.: Remaining Useful Life Prediction of Lithium-ion Batteries Based on Wiener Process Under Time-Varying Temperature Condition, Reliability Engineering & System Safety, Vol. 214, (2021), doi: 10.1016/j.ress.2021.107675
(12) Russell, M. B., King, E. M., Parrish, C. A., Wang, P.: Stochastic modeling for tracking and prediction of gradual and transient battery performance degradation, Journal of Manufacturing Systems, Vol. 59, p. 663-674, (2021), doi: 10.1016/j.jmsy.2021.04.008
(13) Wu, C. F. J., Hamada, M. S.: Experiments: Planning, Analysis, and Optimization, John Wiley & Sons, Inc., (2021), doi: 10.1002/9781119470007
(14) Castillo, E., Hadi, A. S.: Modeling Lifetime Data with Application to Fatigue Models, Journal of the American Statistical Association, Vol. 90, Issue 431, p. 1041-1054, (1995), doi: 10.2307/2291341
(15) Jacquier, E., Johannes, M., Polson, N.: MCMC maximum likelihood for latent state models, Journal of Econometrics, Vol. 137, Issue 2, p. 615-640, (2007), doi: 10.1016/j.jeconom.2005.11.017

Author


Kaoru OMICHI


Kotone SUZUKI


Hideki NAGATSUKA


Michihiro YASUDA

Lunar Night Power Supply: Circulative Renewable Energy System Capable of Continuous Power Generation

Ryosuke INANISHI[1] Hayato DAIMON[1] Kazunori FUKUMA[1]
Takumi KAWANO[2] Toru HONDA[1]

ABSTRACT

Development is proceeding on a circulative renewable energy system for lunar night power supply. For proof of principle of the circulative renewable energy system, a system with approximately 1/10 the power of the expected launch unit was constructed, and a cycle test with 39 cycles was conducted, which is equivalent to three years of lunar use.

In the fuel cell stack, the membrane electrode assembly was designed in consideration of changes in the electrolyte membrane properties due to pure oxygen supply. In the electrolysis cell stack, component materials were selected that have reliable flame-retardant properties and oxidation resistance under high-pressure oxygen. In the electrochemical hydrogen compressor stack, discharge of stagnant water on the reaction surface was enhanced by optimizing the flow path structure. In addition, controls and systems were constructed to establish stable water circulation.

Through these initiatives, a cycle test was achieved and demonstrated the feasibility of continuous power generation through water circulation, which is the concept of the circulative renewable energy system and completed the proof of principle.

1. Introduction

Lunar development efforts such as the Artemis program continue to gain momentum[(1)]. The Artemis program aims for astronauts to remain on the lunar surface for long periods of time and carry out sustainable exploration activities. This calls for methods that ensure stable electric power supply while keeping down the cost of lunar transport, which becomes expensive in proportion to the weight. From the perspective of space as a new destination for use of hydrogen technology, development is proceeding on a Circulative Renewable Energy System (CRE system) that combines a fuel cell system[(2)] and a differential-pressure water electrolysis system[(3)] to enable stable lunar power supply. The CRE system has a high weight-energy density that provides advantages in terms of transport cost and achieving large electric power capacity, so further development is expected going forward. This paper introduces the configuration of the CRE system being developed as well as initiatives carried out for proof of principle of its concept, “continuous power generation through water circulation.”

[1] Innovative Research Excellence, Power Unit & Energy
E-mail: ryosuke_inanishi@jp.honda
[2] Innovative Research Excellence

2. CRE System

2.1. Significance of CRE System

A key challenge for long-term lunar stays is stable night power supply. The moon has a rotation period of approximately 27 days and, while it varies depending on the lunar region, night lasts for around 8 to 14 consecutive days. As it is assumed that a single mission will require approximately 2 kW of electric power to operate an unmanned exploration vehicle and other survey equipment, an electric power amount of up to 672 kWh is needed for nights lasting for 14 consecutive days. This amount of electric power needs to be produced by solar power generation during the day and retained to be used during the night. A commonly considered method is to store sufficient electric power for night use in a battery. However, a high-mass battery would result in a higher lunar transport cost, which is said to be 100 million yen or more per 1 kg[(4)]. Also, to help prevent battery performance from deteriorating in the low-temperature lunar environment, it is necessary to store electric power to heat the battery itself, which would further increase the weight.

The CRE system in development uses electricity obtained by solar power generation to electrolyze water and supplies the obtained hydrogen and oxygen to the fuel cell stack (FC stack) at night to generate power. Furthermore, the concept is to continuously regenerate power by electrolyzing the water produced in power generation. Table 1 shows the expected mass of each component of the future CRE system, assuming electric power generation of 672 kWh. The total mass of the entire system is 1,911 kg, and the weight energy density is 352 Wh/kg. In addition, as shown in Fig. 1, the weight energy density is greater than that of a Li-ion battery[5), (6)], which is thought to be advantageous in terms of helping reduce transport cost and achieving large electric power capacity.

2.2. Overview of CRE System

Figure 2 shows an overview of the CRE system. First, water is supplied to the electrolysis cell stack (EC stack), which electrolyzes the water to produce hydrogen and oxygen. The EC stack has a structure similar to that of the differential-pressure water electrolysis stack used to produce high-pressure hydrogen in the Smart Hydrogen Station[(7)], and produces high-pressure oxygen. Meanwhile, the produced hydrogen moves together with the water, separates from the water, and is then supplied to the electrochemical hydrogen compressor stack (EHC stack). In the EHC stack, high-pressure hydrogen is produced by pressurization through an electrochemical reaction. The produced high-pressure hydrogen and oxygen are stored in tanks and then supplied to the FC stack to generate power when electric power is needed. In addition, since water is produced in the reaction in the FC stack, continuous power generation can be achieved by resupplying this water to the EC stack and circulating it.

2.3. CRE System Development Challenges

There are two challenges for lunar use of the CRE system.

The first is the challenge of stable use in space, which calls for addressing various external factors that differ from those on earth. Examples include changes in components over time due to radiation from the sun and other sources, changes in fluid behavior due to low gravity, and launch vibrations in lunar transport.

The second challenge is to realize continuous power generation. Various measures are needed to realize this through water circulation, e.g. it is necessary to balance the amount of gases produced and consumed between connected stacks. When larger amounts are produced, the resulting increase in internal pressure may result in component deformation, and when smaller amounts are produced, the voltage may rise. The CRE system, however, is intended for use in a lunar environment where resources are not easily obtained, rendering the method of balancing by fluid discharge to outside the system unable to achieve continuous power generation. This calls for realization of a system configuration and control methods that achieve smooth circulation within the system.

This initiative focused on the second challenge, namely the realization of continuous power generation. The contents of initiatives and the demonstration test conducted for proof of principle of the concept are introduced below.

Table 1 Expected mass of each component of CRE system

Items	Mass (kg)
CRE system	727
Storage system	814
Fuel water	370
Total mass	1911

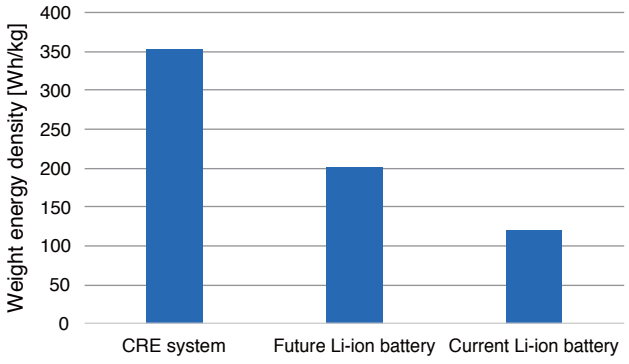


Fig. 1 Comparison of energy density between CRE system and Li-ion battery

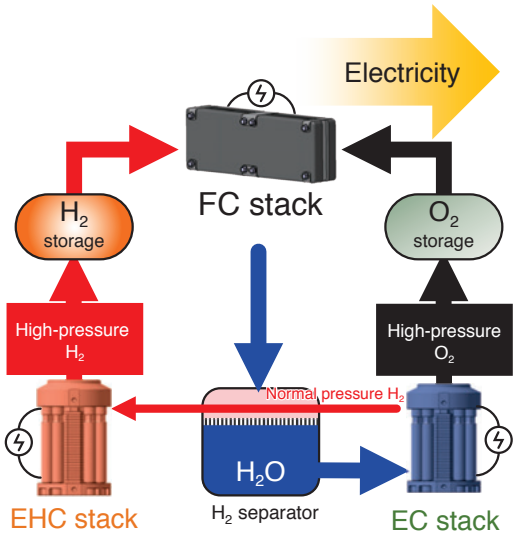


Fig. 2 Schematic diagram of CRE system

3. Initiatives for Proof of Principle

There are three main stages of development before actual lunar use of the CRE system. The first stage is proof of principle of the CRE system. The second stage is evaluation in a lunar equivalent environment using a system with nearly the same specifications as the expected launch unit. The third stage is evaluation of the actual unit that will be launched into space, and mission use.

This time, as the first stage of development, a demonstration test was conducted for proof of principle of continuous power generation through water circulation. In the demonstration test, circulation was performed for a number of cycles assuming actual operation in space and confirmed the feasibility of continuous power generation.

First, the changes to the specifications of each stack comprising the CRE system that were made for the demonstration test, and the contents of system control- and system configuration-related initiatives are described below.

3.1. Control in Gas Pressurization

The hydrogen produced by the EC stack is supplied via the H_2 separator to the EHC stack as shown in Fig. 3. When the H_2 separator pressure drops, the reaction rate slows because less hydrogen is supplied, and the operating voltage of the EHC stack rises. Conversely, when the H_2 separator pressure rises, it applies an extra load to the components. Thus, the operating current of the EHC stack is determined by feedback control using the difference between the H_2 separator pressure and the target pressure, and the amount of reaction hydrogen is adjusted.

An event that lowers the H_2 separator pressure is when oxygen mixes into the supplied hydrogen. Since this oxygen is transformed into water by reacting with hydrogen in the oxygen eliminator installed before the EHC stack, the amount of hydrogen and the pressure become lower. Events that increase the H_2 separator pressure are when high-pressure hydrogen passes though the electrolyte membrane in the EHC stack and moves to the gas flow path on the low-pressure side, and when water generated by the FC stack is supplied from storage to the H_2 separator. In these cases, controlling the operating current of the EHC stack enables operation within the component strength reliability range while also helping to prevent a rise in the operating voltage, even if the pressure inside the H_2 separator fluctuates.

3.2. System Configuration for Stable Depressurization

The EC and EHC stacks have a high-pressure side on one side and a low-pressure side on the other side of the electrolyte membranes inside the cells, which produces a differential pressure. Therefore, when stopping the system or when a tank has reached the specified pressure, it is necessary to depressurize the remaining high-pressure gas in order to lessen the load on the internal components of the stack. However, since external discharge would waste the produced hydrogen and oxygen, this CRE system performs depressurization by suppling the excess gas directly to the FC stack and reactivating it. Hydrogen and oxygen react in the FC stack at a molar ratio of 2:1, but the amount of gas remaining in each stack differs depending on the situation. Therefore, the case can be assumed where either the EC stack or the EHC stack completes depressurization before depressurization of both stacks has been completed.

The set pressures of the depressurizing valves (regulators) shown in Fig. 3 were determined in a manner to address this. The set pressures of regulators A and C between the EC and EHC stacks and the FC stack were set to the depressurization completed pressure. In addition, the set pressures of regulators B and D between each tank and the FC stack were set to a pressure several tens of kPa higher than the depressurization completed pressure.

As a result, the gas remaining on the EC and EHC stack sides is preferentially consumed. When the EC or EHC stack reaches the depressurization completed pressure,

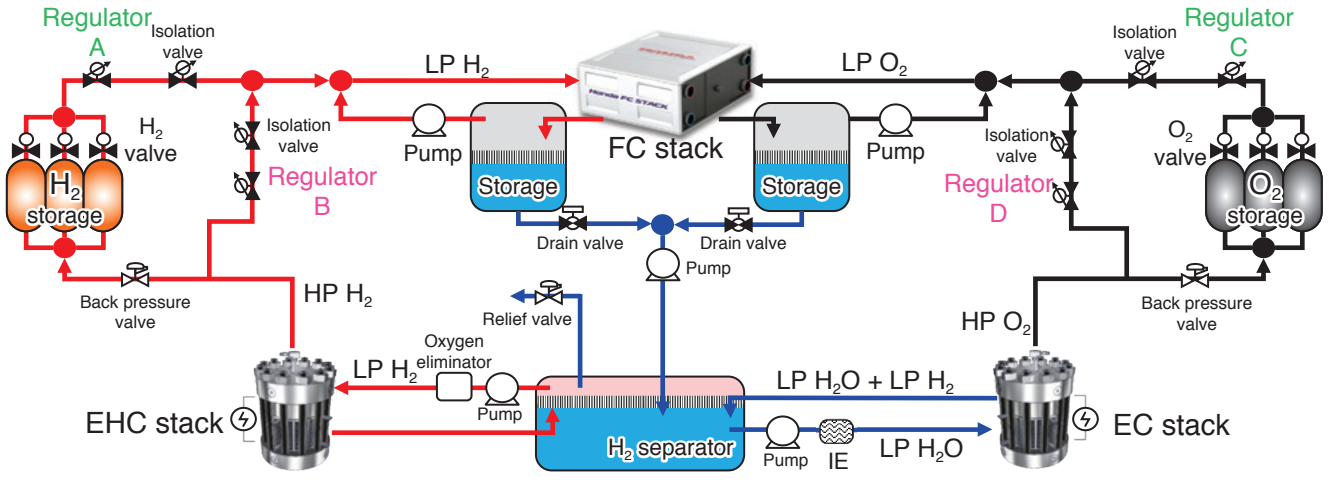


Fig. 3 CRE system configuration

the gas inside the tank that reached the depressurization pressure starts to be consumed. This enables the reaction in the FC stack to continue and depressurize both the EC and EHC stacks by power generation.

3.3. FC Stack

The FC stack is a device that converts chemical energy into electrical energy by the following electrochemical reaction of hydrogen and oxygen in the electrolyte membranes and has a structure that stacks multiple cells that have reaction parts.

$$\text{Anode: } H_2 \rightarrow 2H^+ + 2e^-$$

$$\text{Cathode: } \frac{1}{2}O_2 + 2H^+ + 2e^- \rightarrow H_2O$$

Figure 4 shows the FC stack used in the demonstration test. Although in an FC stack for terrestrial use, the oxygen contained in the air is used for power generation, in the CRE system, pure oxygen manufactured by the EC stack is supplied. One effect of pure oxygen is that the changes in the electrolyte membrane properties over time are more pronounced than those for air. An electrolyte membrane degradation model[8] was constructed by conducting a unit durability test with pure oxygen supply and quantifying the reduction in the membrane molecular weight after durability. This model was then used to design a membrane electrode assembly (MEA) that was expected to exhibit a decrease in performance in the demonstration test.

3.4. EC Stack

The EC stack is a device that performs water electrolysis to produce high-pressure oxygen and low-pressure hydrogen. The goal of the EC stack used in the demonstration test was to produce high-pressure oxygen of 14.7 MPa. Figure 5 shows the EC stack used in the demonstration test, and Fig. 6 shows the schematic diagram of the EC stack. Like the FC stack, the EC stack has a structure of stacked multiple cells. Water is supplied to the electrolyte membranes inside the cells, and oxygen and hydrogen ions (H^+) are generated by applying voltage to the stack. The oxygen outlet is sealed, so the pressure rises over time, enabling high-pressure oxygen to be obtained. Voltage is applied to both ends of the electrolyte membrane, and the oxygen and hydrogen ions generated by water electrolysis pass through the electrolyte membrane and move to the cathode side and are separated. The hydrogen ions are then converted into hydrogen, discharged from the hydrogen outlet of the EC stack, and supplied to the EHC stack.

$$\text{Anode: } H_2O \rightarrow \frac{1}{2}O_2 + 2H^+ + 2e^-$$

$$\text{Cathode: } 2H^+ + 2e^- \rightarrow H_2$$

Oxidation of components present under high-pressure oxygen is promoted, and due to a lower ignition temperature and an increase in the combustion rate, even materials that are non-combustible in the atmosphere may become combustible. Thus, material selection for the EC stack components around the electrolyte membrane on the anode side was performed from the perspective of helping to prevent oxidation and combustion.

3.5. EHC Stack

The EHC stack is a device that electrochemically pressurizes the supplied hydrogen. The goal of the EHC stack used in the demonstration test was to produce high-

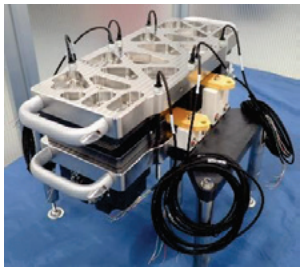


Fig. 4 FC stack



Fig. 5 EC stack

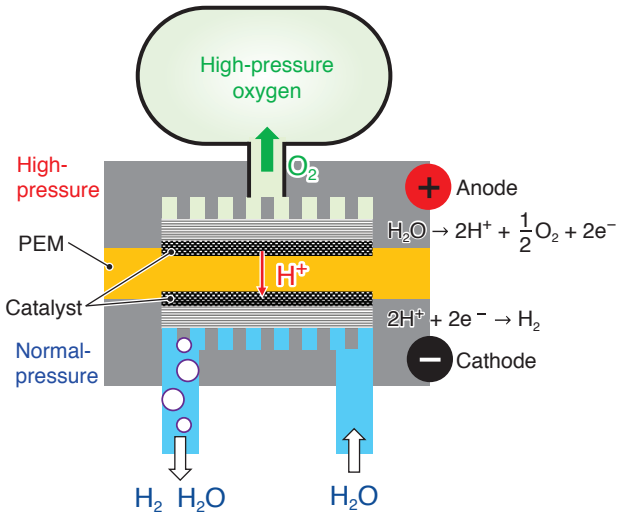


Fig. 6 Schematic diagram of EC

pressure hydrogen of 70 MPa. Figure 7 shows the EHC stack used in the demonstration test, and Fig. 8 shows the schematic diagram of the EHC stack. The EHC stack also has a structure of stacked multiple cells, and when voltage is applied, hydrogen ions generated by the following reaction in the electrolyte membrane inside each cell move to the cathode side that leads to the high-pressure tank. Since the tank is a closed space, the pressure rises over time.

Anode: $H_2 \rightarrow 2H^+ + 2e^-$

Cathode: $2H^+ + 2e^- \rightarrow H_2$

As drying of the electrolyte membrane leads to lower conductivity, wet hydrogen is supplied to the EHC stack. The water in the hydrogen can permeate the membrane and move to the high-pressure side, but the greater the differential pressure via the membrane, the more the water is pushed back to the low-pressure side. As shown in Fig. 8, there is the issue that the water pushed back remains in the membrane reaction part with the result that hydrogen supply is not smooth and the voltage rises. Especially in the low-gravity lunar environment, water detaches less easily from the reaction membrane. To help prevent water retention, it is necessary to suppress local drops in the flow velocity of the hydrogen supplied to the reaction surface that contributes to water discharge. The hydrogen flow path structure was thus optimized to achieve a more uniform hydrogen flow velocity in the reaction part.

Figure 9 shows the flow of hydrogen in the EHC stack. The hydrogen supplied to the EHC stack originates from the inlet frame to each cell, then passes through the flow paths cut into the bottom surface of the inlet frame and flows to the spacer and the electrode. The electrode is a component located at the bottom of the electrolyte membrane, and the flow velocity distribution in this component was equalized to help prevent water retention. An L27 orthogonal table was prepared using 13 design factors such as the diameter and length of the flow paths cut into the bottom surface of the inlet frame, and variation in the flow velocity under each condition was obtained by fluid analysis. A response surface was created based on the analysis results, and the flow path structure was adapted to the shape that achieved the lowest flow velocity variation in the analysis. Figure 10 shows the hydrogen flow velocity for the base shape, and Fig. 11 shows the hydrogen flow velocity for the optimized structure. The gas flows from left to right in the figures.

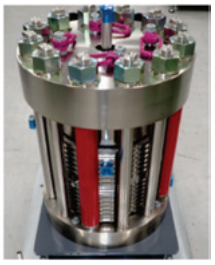


Fig. 7 EHC stack

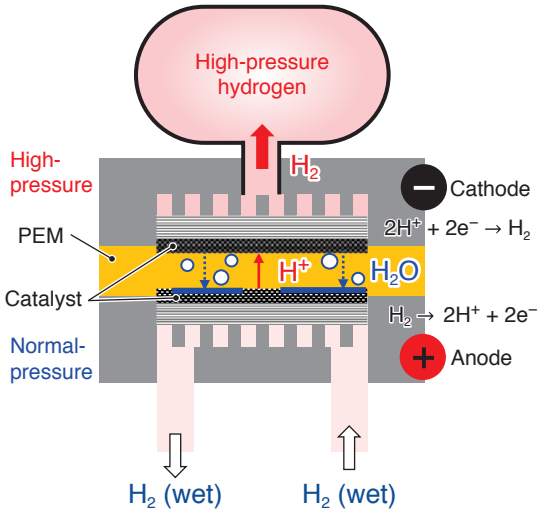


Fig. 8 Schematic diagram of EHC

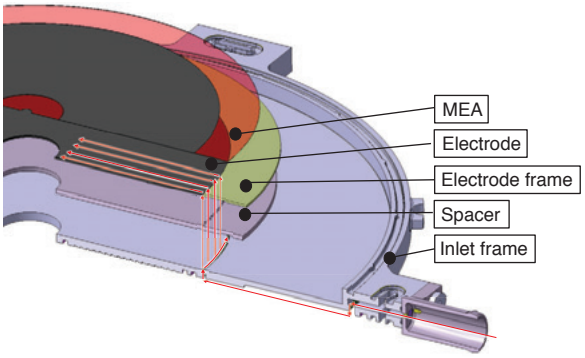


Fig. 9 Flow of hydrogen in EHC

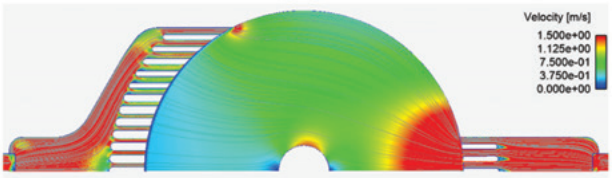


Fig. 10 Velocity distribution in base flow path structure of EHC stack

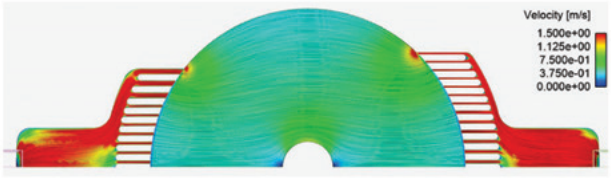


Fig. 11 Velocity distribution in optimized flow path structure of EHC stack

While the semicircular area in the center indicates the electrode, the number and width of flow paths outside the electrode have been changed, resulting in a more uniform flow velocity inside the electrode compared to before the structural change. The cell component with the altered flow path structure was used in the demonstration test to help prevent operational termination due to water retention.

4. Proof of Principle of CRE System

A demonstration test was conducted for proof of principle of the CRE system.

A CRE system with about 1/10 the power of the expected launch unit was created, and a cycle test was conducted as a demonstration test. There are 13 night-crossings per year on the moon. In addition, while the CRE system assumes continuous lunar use of several to 10 or more years, here, the specified number of cycles for the demonstration test was 39 cycles, which is the number of day/night cycles equivalent to three consecutive years, or the expected number of years for a single lunar mission[9]. The total amount of water present in the system at the start of the test was 1,900 ml, whereas the amount of water used for water electrolysis in one cycle is approximately 400 ml. Assuming circulation is established, therefore, proof of principle can be achieved by completing power generation for the specified number of cycles without any event that leads to termination of the test. In addition, this test was conducted using general grid power instead of electric power obtained by solar power generation.

Figure 12 shows the appearance of the created CRE system, and Table 2 shows the specifications. The operation of one cycle of the CRE system will also be introduced. Figure 13 shows the transitions in the operating current and outlet pressure of the EC and EHC stacks, the FC stack power, and the pressure of each tank. The timing of one cycle differs slightly for the EC and EHC stacks, but it can be divided into the four modes of 1) pressurization, 2) storage, 3) depressurization, and 4) generation.

In pressurization mode, gases are produced and pressurization within the stacks begins with the start of the electrochemical reaction. Fluctuations in the supplied hydrogen pressure are addressed by controlling the operating current of the EHC stack as mentioned in Section 3.1. Once the predetermined pressure is reached, operation transitions to storage mode.

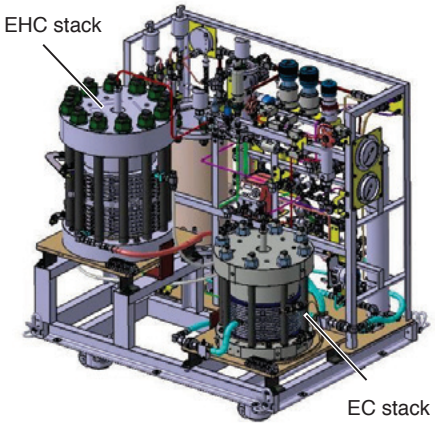
In storage mode, gas supply to the tanks starts and the tank pressures begin to rise. After the tank pressures reach the target pressure, operation transitions to depressurization mode.

In depressurization mode, gas supply to the tanks stops and the pressure inside the tanks is retained. The pressure inside the EC and EHC stacks is decreased by supplying the extra gas remaining in the EC and EHC stacks directly to the FC stack and consuming it for power generation. As shown in the bottom graph of Fig. 13, the FC stack power is not constant in depressurization mode. This is due to adjustment of the FC stack power to help prevent rapid depressurization of the EC and EHC stacks and blistering of the rubber seals and other EC and EHC stack internal components.

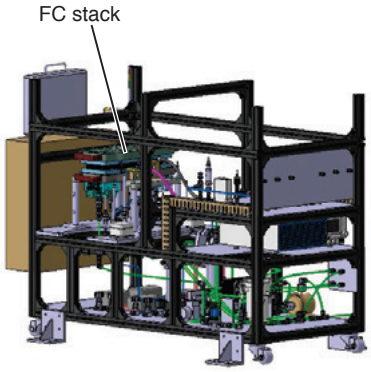
Table 2 Specifications of CRE system for demonstration test

CRE system	System weight [kg]	2900
FC stack	Average power (generation mode) [W]	221.6
	Stack weight [kg]	50
	Stack dimensions [mm]	570 × 266 × 250
EC stack	Pressure of generated oxygen [MPa]	14.7
	Average oxygen production rate [g/h]	179.5
	Stack weight [kg]	260
	Stack dimensions [mm]	Φ393 × 447
EHC stack	Hydrogen pressure increase [MPa]	70.0
	Average hydrogen pressurization rate [g/h]	22.5
	Stack weight [kg]	470
	Stack dimensions [mm]	Φ460 × 623

Then, in generation mode, the tank pressures drop because the gas stored in the tanks is supplied to the FC stack.



(a) EC and EHC systems



(b) FC system

Fig. 12 CRE system for demonstration test

There are two things to consider when setting the cycle test operation method: the first is that in the pressurization and depressurization modes, the operating current and the control times using the depressurizing valves, etc. may vary due to changes in the system specifications. However, these variations due to longer or shorter times can be addressed by maintaining the balance of fluid circulation.

The second is shorter storage mode and generation mode times. Both the capacity and storage pressure of the gas storage tanks are set lower than for actual operation, and one cycle is completed in about five hours. In actual operation, however, the operating time increases due to an increase in the amount of electric power needed. This is expected to have an impact on the durability of each stack and will be confirmed by individual durability in the future. As the purpose of this demonstration test is to confirm the circulation of the entire CRE system, verification can be achieved by conducting a cycle test with the set operation method.

Figure 14 shows the FC power generation amount in each cycle. As a result of this series of initiatives, it was confirmed that power could be generated through 39 cycles without terminating the test, and proof of principle was completed.

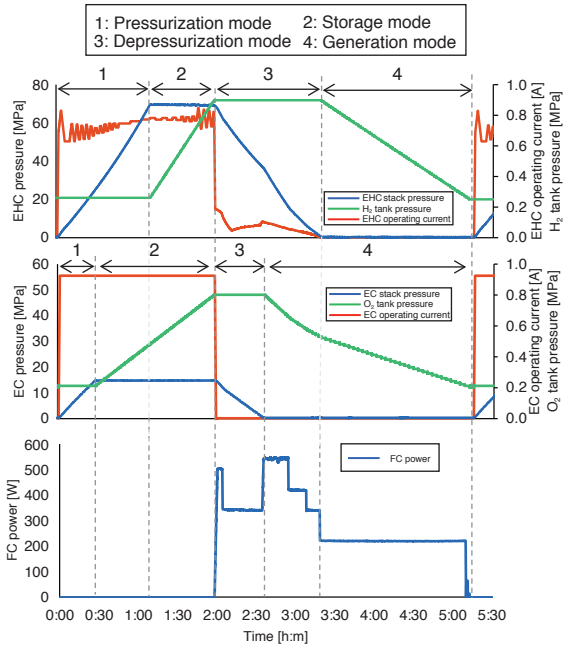


Fig. 13 Operation transition of CRE system

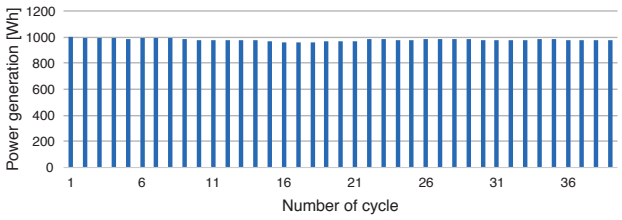


Fig. 14 Power generation in each cycle

5. Initiatives for Stable Use in Space

As mentioned in Chapter 3, evaluation will be conducted in a lunar equivalent environment going forward. To achieve stable lunar use of the CRE system, it is necessary to consider various environments and disturbances that differ from those on earth. This paper introduces initiatives that should be taken going forward to address radiation, low gravity, transport vibrations, and weight reduction to achieve a lower transport cost.

5.1. Radiation

The moon is shielded by a weaker magnetic field than the earth, and the amount of radiation from the sun and other sources is higher in the lunar environment than on earth. Radiation exposure tests were conducted on the stack components to understand the property changes due to radiation. The results demonstrated that, among the stack components, changes in the electrolyte membrane properties were especially noticeable. As shown in Fig. 15, the molecular weight and tensile strength of the electrolyte membrane drop as the radiation-absorbed dose increases. However, as equipment such as presses and cranes are needed to assemble the stacks, disassembly and replacement of the electrolyte membranes cannot be performed on the lunar surface. Therefore, the specifications need to provide durability enabling complete operation, making it necessary to clarify the radiation-absorbed dose and the functional deterioration of each component, and to design a shielding structure that does not produce functional loss over the operation period.

5.2. Low Gravity

Lunar gravity is approximately 1/6 that on earth. The lower the gravity of the environment, the less the buoyancy of air bubbles acts. This means that in the EC stack, the produced gas is not expected to separate from the electrode

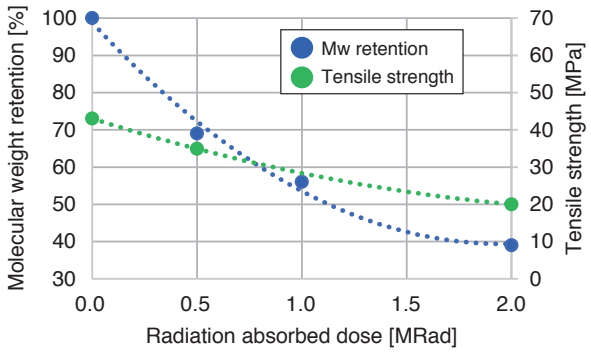


Fig. 15 Relationship between radiation absorption and membrane properties

surface, with the result that mass transport is hindered, and the electrolytic performance deteriorates. In future, tests will be conducted in low-gravity environments by means such as drop towers and parabolic flights on aircraft in order to clarify the influence of gravity and to design water supply conditions and structures for reliable discharge of produced gases.

5.3. Transport Vibrations

Transport rockets used to carry goods into outer space experience severe vibrations during launch. Therefore, the CRE system and each stack require structures that can withstand these vibrations without loss of functionality. Stabilizing with resonance suppression jigs is a generally effective vibration countermeasure, but from the viewpoint of transport costs to space, increases in weight for the purpose of suppressing vibration should be avoided as much as possible. A shaker is currently being used to produce vibrations at the same level as in actual transport, and vibration characteristics, such as the resonance frequencies of the stacks, are being evaluated. Going forward, these results will be fed back to CAE to enhance design estimation technology and utilized in system design.

5.4. Weight Reduction

In order to achieve a lower lunar transport cost, weight reduction is needed for each stack comprising the CRE system. Since the EC and EHC stacks have especially high weights, it is necessary to optimize the reaction part diameters. By increasing the reaction part diameters and the amount of hydrogen produced per cell, fewer cells need to be stacked, which enables shortening of the total length of the bolts securing the stacks. However, increasing the reaction part diameters also increases the diameter of the pressure-resistant components around the outer periphery of the reaction part, which is a factor that increases the weight. Future efforts will be made to optimize the reaction part diameters and achieve weight reduction within constraints such as space in transport.

6. Conclusion

(1) The MEA of the FC stack was designed in consideration of membrane changes over time due to supply of pure oxygen.
(2) Materials with reliable oxidation and flame resistance that can be used even in a high-pressure oxygen environment were selected for the EC stack.
(3) Optimization of the flow-path structures of the EHC stack enhanced discharge performance of stagnant water in the reaction part.
(4) Valve and stack operation control were established to realize stable water circulation within the CRE system.
(5) 39 cycles of operation were achieved using the constructed CRE system, and proof of principle was completed.

References

(1) https://humans-in-space.jaxa.jp/en/future/, (accessed 2024/07/25)
(2) https://global.honda/en/newsroom/news/2023/c230202eng.html, (accessed 2024/07/25)
(3) https://global.honda/jp/news/2018/c181121.html, (accessed 2024/07/25) (in Japanese)
(4) https://www.astrobotic.com/astrobotic-awarded-79-5-million-contract-to-deliver-14-nasa-payloads-to-the-moon/, (accessed 2024/10/17)
(5) Oka, T., Marutani, K., Hamada, S., Osako, Y.: Large Capacity Lithium Ion Battery for Space Use, Mitsubishi Electric Technical Report, Vol. 94, No. 2, p. 41-45, (2020), https://www.giho.mitsubishielectric.co.jp/giho/pdf/2020/2002202.pdf, (accessed 2024/07/25) (in Japanese)
(6) Naito, H.: Energy Storage Devices for Space Exploration, GS Yuasa Technical Report, Vol. 13, No. 2, p. 1-7, (2016), https://www.gs-yuasa.com/jp/technology/technical_report/pdf/vol13_2/13_02_001.pdf, (accessed 2024/07/25) (in Japanese)
(7) Ishikawa, H., Haryu, E., Kawasaki, N., Daimon, H.: Development of 70 MPa Differential-pressure Water Electrolysis Stack, Honda R&D Technical Review, Vol. 28, No. 1, p. 86-93
(8) Kato, T., Nakahara, A., Ohkura, T., Yamamoto, M., Terada, S., Noda, A., Kanaoka, N.: Development of High-durability and Low-cost UEA for Fuel Cell Stack, Honda R&D Technical Review, Vol. 37, p. 27-33
(9) Uehara, S., Ohama, N., Tagai, R., Takahashi, M., Murata, S.: Study Status of Regenerative Fuel Cell System for Pressurized Rover, 68th Space Sciences and Technology Conference, 3A09, (2024) (in Japanese)

■ Author ■


Ryosuke INANISHI


Hayato DAIMON


Kazunori FUKUMA


Takumi KAWANO


Toru HONDA

New Transmission Mechanism Technology Utilized in Prosthetic Legs

Hiromi ONO[1] Kei SHIMADA[1] Yutaka HIKI[1]
Seiji ONOZAWA[1] Fumihito GOTO[2]

ABSTRACT

The researchers developed a transmission mechanism that combines one-way clutches capable of rotating in both directions and can change speeds by switching the connection state of the clutch.

This makes the clutch smaller and lighter than an electromagnetic clutch and makes it possible to transmit driving force for both forward and reverse rotation. Other features of the mechanism are that it consumes power only for switching, not for maintaining the rotational state, and there is no limit to the number of speed ratio levels that can be set.

When this technology is applied to motor-driven prosthetic legs, the walking speed is independent of motor performance, which means that the motor can be made smaller and the prosthetic leg can be made lighter due to the power savings. In addition, the swinging of the leg while walking can be made quieter, which expands the life possibilities for users of prosthetic legs.

1. Development Background and Goals

1.1. Knee Joint Types and Functions

As Fig. 1 shows, there are various types of prosthetic legs, depending on the part of the body that is missing(1). Among these, a person without a natural knee joint uses a prosthetic leg with an artificial knee joint to replace the natural one.

These need to function similarly to the human knee. This includes supporting the weight of the body and flexing and extending the knee, among others. Among these, the role of missing muscles continues to evolve and can be replaced by various means.

The function introduced in this paper is primarily a knee joint yielding function using a damper(2). The yielding function is the ability to flex the knee while applying moderate flexion resistance. This has become a major function of knee joints in recent years to help prevent unintentional sudden buckling of the knee when body weight is applied to the prosthetic leg and enable the user to go downstairs with natural alternating steps (Fig. 2). Recently, knee joints have become available that use a motor to replace the role of the muscles and flex and extend the knee. Both types of knee joints in recent years use knee angle sensors, load sensors, gyro sensors, and other sensors to detect the knee's status and operate dampers, motors, and other devices.

Of these, damper-type knee joints with a yielding function have been the mainstream in recent years, but motor-driven knee joints are also expected to become more common in future.

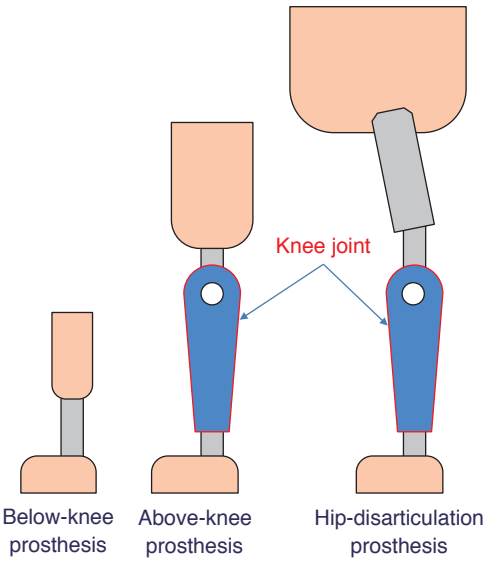


Fig. 1 Types of prosthetic leg

[1] Solution System Development Center
E-mail: hiromi_ono@jp.honda
[2] RAPID ENGINEERING CO.,LTD.

1.2. Background to Development

Damper-type knee joints and motor-driven knee joints each have their strong points and issues.

In a damper-type knee joint, the damper provides an appropriate yielding function during the period when body weight is applied (the “stance phase”), which helps to prevent sudden buckling of the knee. During the period when no body weight is applied (the “swing phase”), the knee is passively flexed and extended in time with the swinging of the thigh. This allows the user to freely operate and walk, which makes the damper-type knee joint easy to handle. However, when the person is going upstairs, the knee joint does not exert force to extend the knee, so going upstairs with alternating steps of the two legs is not easy.

Motor-driven knee joints are driven by a motor in both the stance phase and swing phase. During the swing phase, the knee flexes and extends, which makes walking easier. During the stance phase, power from the motor makes it possible to perform movements that require power, such as going upstairs and standing up, which are not so easy with a damper-type knee joint. For that reason, the researchers believed that addressing the following three issues with motor-driven knee joints could make them easier to handle.

The first issue is that motor-driven knee joints consume considerable power because the knee joint is constantly being driven even while walking, which is how the knee joint is most frequently used. Therefore, the battery for driving the motor needs to have enough capacity for the frequency of use in daily life. This could increase the size and mass of the battery. Secondly, the motor is constantly driving the knee joint while walking, so there is motor drive noise at all times. Thirdly, the speed while walking depends on motor performance. This is because the speed of flexion and extension of swinging increases as walking speed increases. Another factor is that the motor needs to be larger to operate without being restricted by the walking cadence (the number of steps per minute).

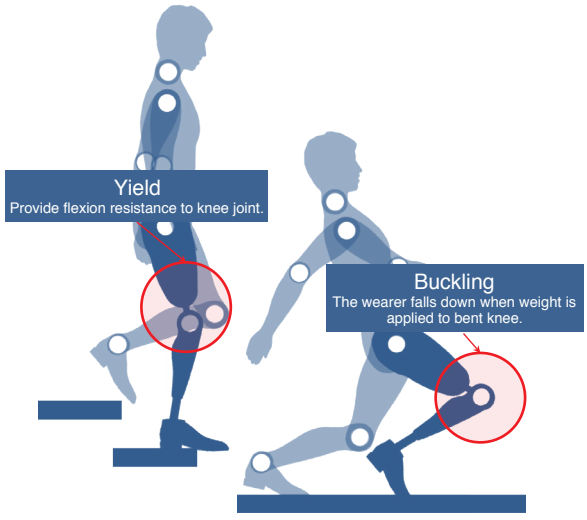


Fig. 2 Yielding and buckling

Table 1 Issues addressed by this research

No.	Issue
1	Power saving while walking on level ground (battery downsizing)
2	Reduction of motor drive noise while walking on level ground
3	Walking cadence independent of motor performance (motor downsizing)

In this paper, to address the three issues listed in Table 1 and below, based on a motor-driven knee joint, the researchers developed not an entire prosthetic leg system but a transmission, which is a technical element of such a system.

- Power saving while walking on level ground (battery downsizing)
- Reduction of motor drive noise while walking on level ground
- Walking cadence while walking on level ground independent of motor performance (motor downsizing)

2. Goals

2.1. Ideal Behavior of Knee Joints

The greatest advantage of motor-driven knee joints is their potential to expand the range of the user’s mobility and activity, for example by enabling them to go upstairs. The researchers believed that this advantage could be used to help realize Honda’s 2030 Vision, which is to “Serve people worldwide with the ‘joy of expanding their life’s potential’[3].

When the user is going upstairs, it is necessary to provide force for the knee joint to extend. By contrast, movements such as walking on level ground do not require power. The researchers believed an effective way to address issues 1–3 in Table 1 would be to aim for behavior similar to that of a damper-type knee joint without providing power.

First, the researchers separately considered walking on level ground and going upstairs, based on whether there is a need to provide extension force to the knee joint.

For walking on level ground, as previously mentioned, the aim was movements like those shown in Fig. 3, which refers to the ease of handling of the damper-type knee joint and the sense of assurance with the yielding. Specifically, during the swing phase, the connection between the motor and the knee axis is cut to allow free rotation of the knee axis so that the user can freely handle the prosthetic leg. During the stance phase, the knee axis is connected to the motor to provide flexion resistance by yielding. The back drive torque of the motor is used as the flexion resistance at this time. As a result, the only power required when walking on level ground is for the switching operation, thereby saving power. In addition, as the knee axis is not connected to the motor during the swing phase, this can help

prevent motor noise. Moreover, the flexion and extension movements are not motor driven during the swing phase, and thus in this concept, walking speed is not limited and all three of the issues listed in Table 1 can be achieved. The motor provides the yielding function, and thus development can aim for a sense of assurance and ease of handling equivalent to those of a damper-type knee joint.

Next, for going upstairs, the motor's driving force is used for extension and flexion of the knee axis, aiming for movements such as those shown in Fig. 4.

Specifically, during the stance phase, the motor forcefully extends the knee to support the body weight. During the swing phase, the motor quickly flexes the knee to help avoid differences in level.

The speed ratio between stance and swing at this time was set to approximately 1:2. The reason is that, in the human gait, the flexion movement of the swing phase is about twice as fast as that of the stance phase, as shown in Fig. 5[4].

In addition, the knee joint needs to be capable of both extension and flexion movements during both the stance phase and swing phase.

2.2. Functional Requirements and Limiting Conditions

To achieve the goals laid out in Section 2.1., the following functional requirements and limiting conditions were set for knee joints.

(1) Functional requirements

While walking on level ground:

During the swing phase, the transmission shall be in neutral state and the user shall be able to rotate the knee axis freely.

During the stance phase, the knee axis shall be connected to the motor and the user shall be able to apply resistance to the knee axis. In addition, the switch from free rotation to yielding shall be instantaneous.

Power shall be consumed only for the switching operation.

When going upstairs:

It shall be possible to connect to the motor at two speed ratios, and the user shall be able to get assist power while high torque is being output during the stance phase and high rotation speed during the swing phase.

The knee joint shall be able to provide driving force in both the extension and flexion directions of the knee axis.

(2) Limiting conditions

The device shall be small enough to fit within the silhouette of the knee.

As it was considered necessary technology for achieving the above functional requirements and limiting conditions, a compact transmission technology was developed that can instantly switch the connection state with the motor between neutral and two speed ratio levels in synchronization with the state transition between the stance phase and swing phase.

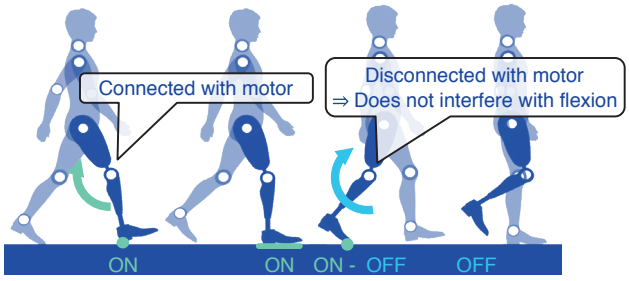


Fig. 3 Aim of walking on level ground and going downstairs

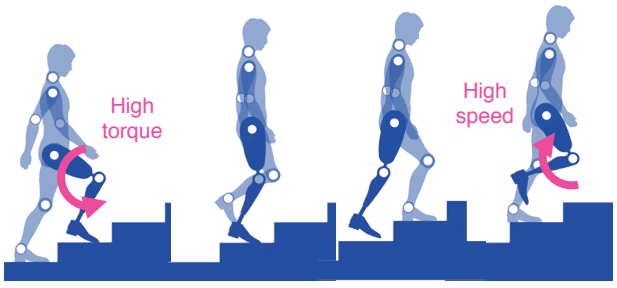


Fig. 4 Aim of going upstairs

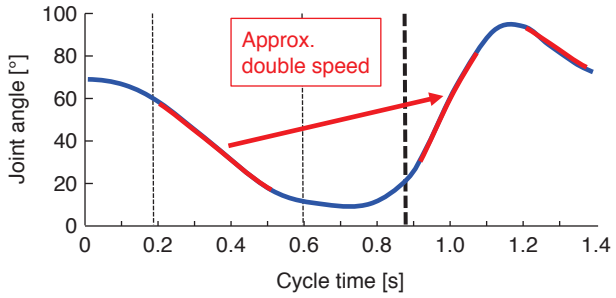


Fig. 5 Speed difference between stance phase and swing phase when going upstairs

3. Basic Structure of Transmission

3.1. Overall Structure of Transmission

The overall structure of the transmission and the names of the transmission's main functional parts are shown in Figs. 6 and 7, respectively.

Three parts, the gear, shaft, and roller, transmit the power. Whether or not power is transmitted between the gear and shaft changes depending on whether the roller is engaged between the shaft and gear. The retainer holds the roller in position. Gears A and B each have a different number of teeth, etc., based on the speed ratio, but the structure on the inside is the same.

To switch the state of power transmission, the rod actuator operates, making the rod move up and down via

a rack-and-pinion mechanism. The rod is cam-shaped, and the cam moves the pin. The pin pushes the pin guide, which moves the retainer and changes the status of the roller.

Grease is applied inside the shaft and to the sliding parts of the clutch.

Figure 7(a) shows the rod in axial cross section. The connection state of gears A and B changes as the rod in the shaft slides along the shaft axis. The rod has two cam shapes, one with a large diameter and one with a small diameter. The clutch is OFF when the part with the large diameter is positioned in the center of gears A and B, and ON when the part with the small diameter is positioned there. Details are provided in Sections 3.2. and 3.3. The motor and rack-and-pinion mechanism shown in the lower part of Fig. 7(a) make the rod slide. The back drive torque of the motor holds the rod's position.

3.2. Switching Clutch from OFF to ON

The state transition during switching is shown in Fig. 8. This section focuses on the change in gear B. Parts of Fig. 8 are shown in enlarged form or in cross section in Figs. 9, 10, and 11.

When the rod is slid downwards from the clutch OFF state, the state of contact between the pin and rod changes and a gap is created (Fig. 9). As Fig. 10 shows, when rotational force is applied to the gear at this time, the retainer rotates while the pin is pushed back by the cam of the pin guide. During this process, the gear and retainer co-rotate due to friction of the grease that was applied between them. The roller is sandwiched between the gear and shaft and transmits rotational force to the shaft (Fig. 11). Torque is transmitted by this type of mechanism. Moreover, because the gear and retainer co-rotate, even if the gear reverses, the roller is sandwiched again, on the opposite side of the shaft, and rotational force is transmitted to the shaft.

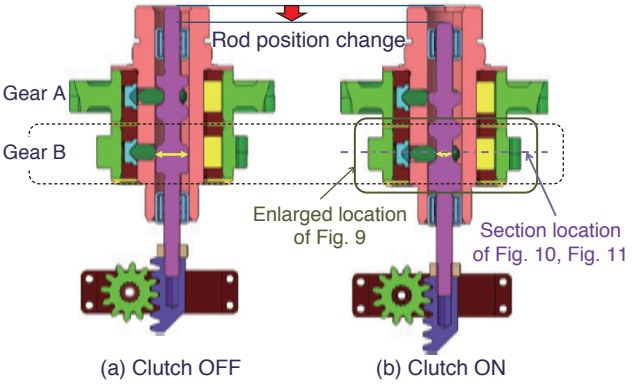


Fig. 8 Switching clutch from OFF to ON

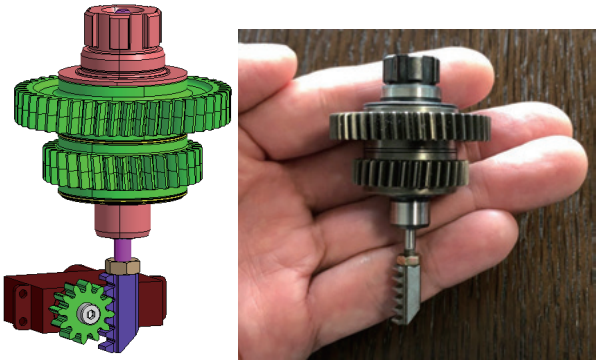


Fig. 6 Overall structure of transmission

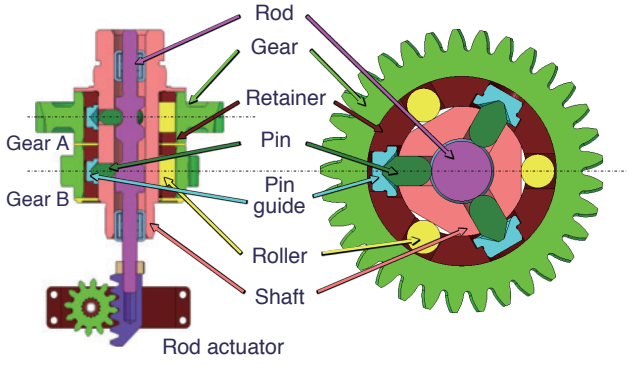


Fig. 7 Structure of transmission

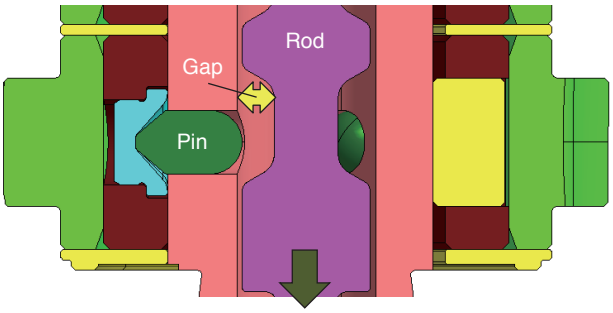


Fig. 9 Rod state when switching clutch from OFF to ON

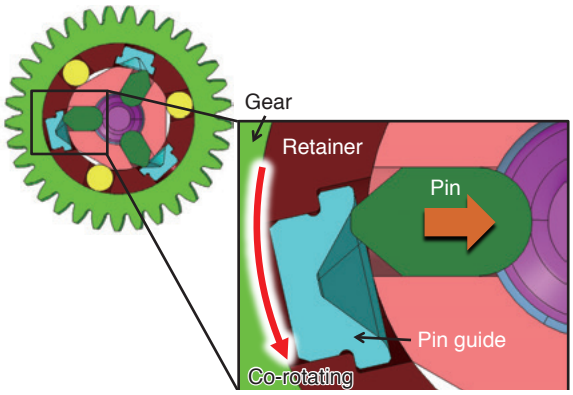


Fig. 10 Pin state when switching clutch from OFF to ON

3.3. Switching Clutch from ON to OFF

When the rod is slid upwards again from the clutch ON state, the pin is pushed outward by the cam on the rod (Fig. 12). During this process, the pin pushes up the pin guide and the retainer and roller are fixed in place (Fig. 13). In that position, the gear and shaft do not come into contact with the roller at the same time, and thus there is idling (Fig. 14).

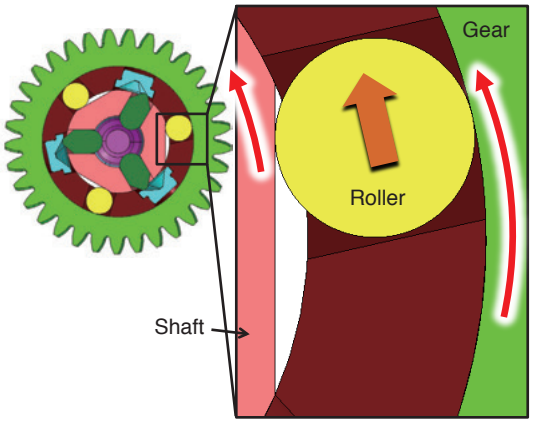


Fig. 11 Roller state when switching clutch from OFF to ON

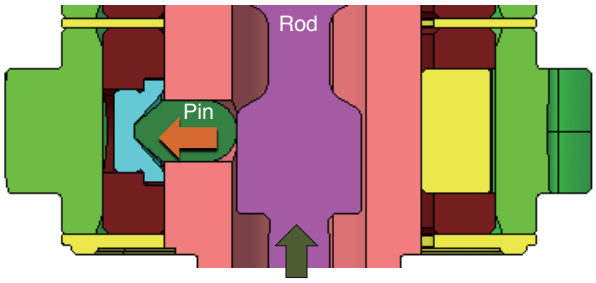


Fig. 12 Rod and pin state when switching clutch from ON to OFF

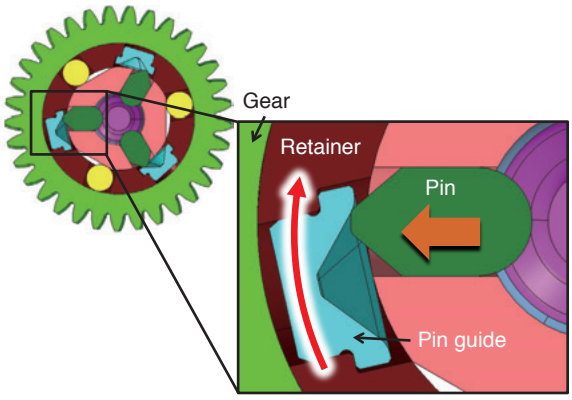


Fig. 13 Pin and retainer state when switching clutch from ON to OFF

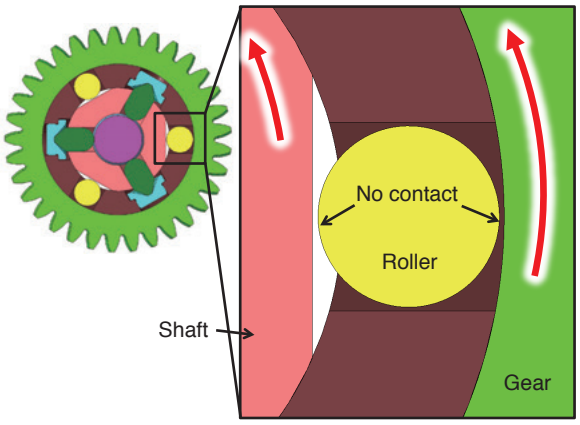


Fig. 14 Roller state when switching clutch from ON to OFF

4. Operation in a Prosthetic Leg

4.1. Incorporation into a Prosthetic Leg

Figure 15 shows an example of the system structure when the transmission is incorporated into a prosthetic leg. Figure 16 shows an example of how the mechanical components are structured. Torque from the motor enters the knee joint via the transmission. The torque exerted by the person from the knee joint side also enters the motor via the transmission in the same way. In this structure, three sensors measure the state of the knee. A load sensor measures the loading and unloading of body weight. An IMU sensor measures the stance of the knee joint. The knee angle sensor measures the angle between the thigh and lower leg components.

4.2. Operation in a Prosthetic Leg

The transmission is made to operate in a prosthetic leg

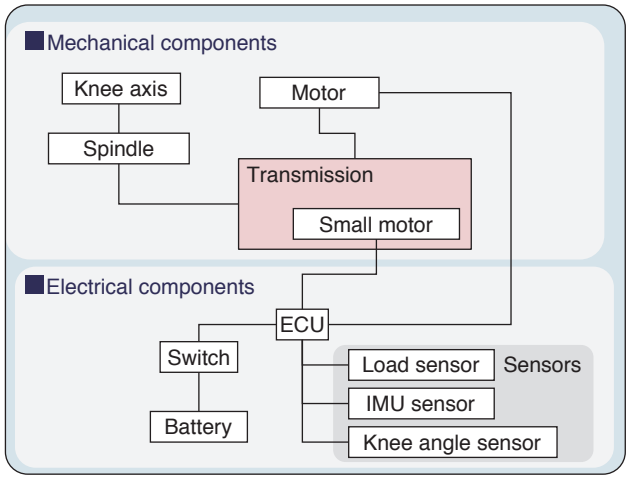


Fig. 15 System structure of knee joint

as shown in Figs. 17 and 18.

For the swing phase of walking on level ground, both gears are set to neutral, as shown in Fig. 17(a). This is the clutch OFF state for both gear A and gear B. Free flexion and extension movements, due to the swinging of the user's thigh, are possible.

For the stance phase, gear A is set to clutch OFF and gear B (the high rotation gear) is set to clutch ON, as shown in Fig. 17(b). The rotation of the knee joint is transmitted to the motor, and proper yielding is made to function by utilizing the back drive torque of the motor. This can help to prevent unintentional sudden buckling of the knee during the stance phase and give a sense of assurance to the user.

Next, this section discusses going upstairs. During the stance phase, high torque is required to support the user's body weight, but the rotational speed is slow. However, flexion and extension during the swing phase require only below-knee inertial force, which means that the high torque is not necessary, but a high rotational speed is required. The transmission mechanism with two gear ratios effectively addresses these conflicting requirements.

For the swing phase, the clutch for gear B (the high rotation gear) is set to ON and the clutch for gear A is set to OFF, as shown in Fig. 18(a). Using the motor to flex and extend the knee joint at high speed enables the user to step onto the next step without touching the staircase risers.

For the stance phase, the clutch for gear A (the high-torque gear) is set to ON and the clutch for gear B is set to OFF, as shown in Fig. 18(b). The output of the motor is transmitted to the knee joint with high torque to assist knee extension, enabling the user to go upstairs.

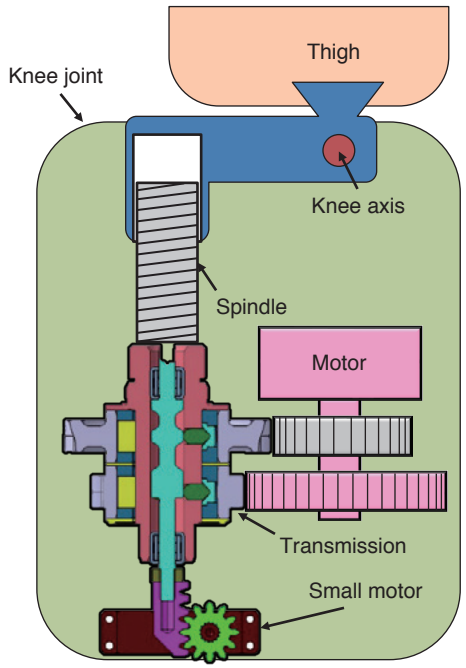


Fig. 16 Mechanical system structure of knee joint without electrical parts

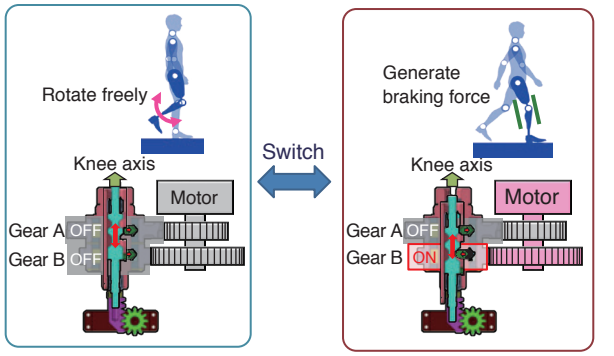
Table 2 Phases, modes, and their goals

Situation	Phase	Transmission mode	Aim
Walking on level ground/going downstairs	Swing	Neutral	Easy to handle
	Stance	High speed	Yielding
Going upstairs	Swing	High speed	Avoid steps
	Stance	High torque	Lift weight

Table 2 summarizes the above transmission modes and their goals during the walking cycle.

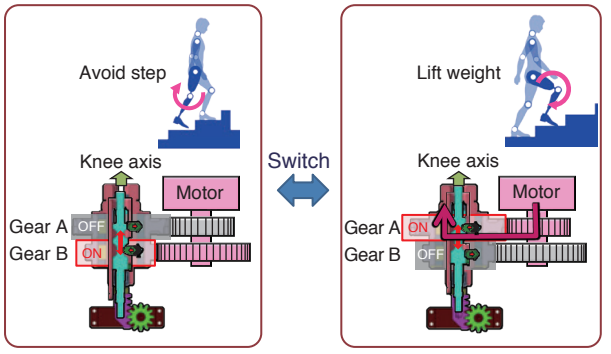
4.3. Check of Operation When Used in a Prosthetic Leg

Log data of walking on level ground is shown in Fig. 19. A force sensor and IMU sensor mounted on the prosthetic leg are used to determine whether the knee joint is in the stance phase or swing phase. During the stance phase, the rod position switches to the high torque mode before load is applied, which enables the resistance of the motor to help prevent buckling of the knee. In the swing phase, moreover, the transmission switches to neutral before the knee angle increases. It was confirmed that, as a result,



(a) Swing phase: neutral (b) Stance phase: high speed

Fig.17 Transmission switching when walking on level ground or going downstairs



(a) Swing phase: high speed (b) Stance phase: high torque

Fig. 18 Transmission switching when going upstairs

the knee angle can be flexed quickly without interference from the motor's resistance.

Log data of going upstairs is shown in Fig. 20. While the user is going upstairs, the force sensor determines whether body weight is being applied to the prosthetic leg, and based on that, switches between the stance phase and the swing phase. The numbers for load (shown in light blue) show how load changes with body weight. They show that when no body weight is applied, the rod position switches from the stance phase to the swing phase, and when body weight is applied, it switches from the swing phase to the stance phase.

4.4. Summary of Issues, Countermeasures, and Results for Motor-Driven Knee Joints

Table 3 summarizes countermeasures and results for the three issues previously listed in Table 1.

Regarding the first issue, power saving while walking on level ground (which accounts for much of the time the prosthetic leg is in use) and downsizing the battery, the transmission lowered power consumption by going into a neutral state during the swing phase. Even yielding using the back drive torque of the motor during the stance phase does not consume power. As power is consumed only during switching and only to move the rod up and down, the capacity of the onboard battery can be reduced. Another advantage is that this transmission does not consume power to maintain either state, as friction-type clutches do.

Regarding the second issue, reducing motor drive noise while walking on level ground, when this transmission is used, no noise is generated because there is no connection with the motor during the swing phase. During the stance phase, noise is generated because of the connection to the motor in situations where yielding is used. However, when walking on level ground, yielding is only used when the user starts to fall and the knee buckles, and basically there is no noise during the stance phase as well. During switching, the small motor used for switching makes very little noise, about the same as other electronically controlled knee joints. The above indicates that a structure and control have been realized that do not generate operating noise when walking on level ground.

The third issue is that the walking speed should be independent of motor performance. This is so that the neutral function of the transmission will allow it to be moved by the user's stump (the part of the thigh remaining after amputation), like a damper-type knee joint.

Table 3 Issues addressed, countermeasures, and results of this research

No.	Issue	Countermeasure	Result
1	Power saving while walking on level ground (battery downsizing)	During the swing phase while walking on level ground, do not perform extension and flexion by the motor, only switch the state of the knee joint.	Power consumption only for switching, except for activities that require power (e.g. going upstairs)
2	Reduction of motor drive noise while walking on level ground		No motor noise except for activities that require power (e.g. going upstairs)
3	Walking speed independent of motor performance (motor downsizing)		Walking speed independent of motor performance because knees are not moved by motors during walking

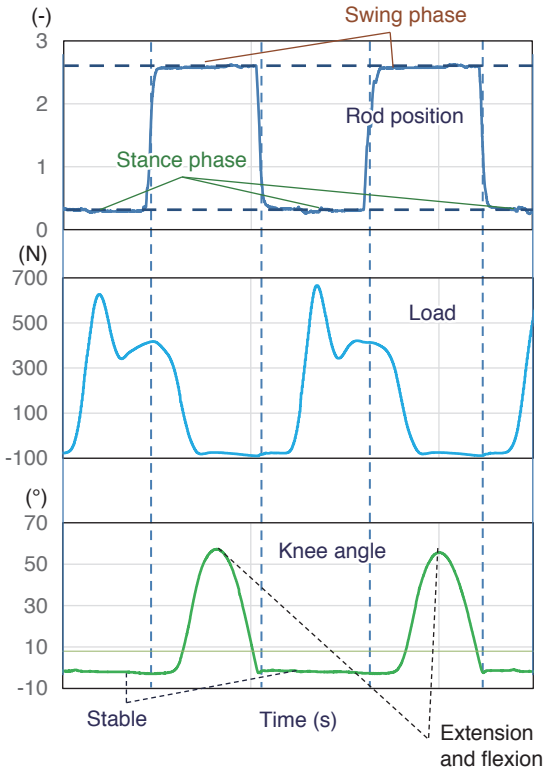


Fig. 19 Log data of walking on level ground

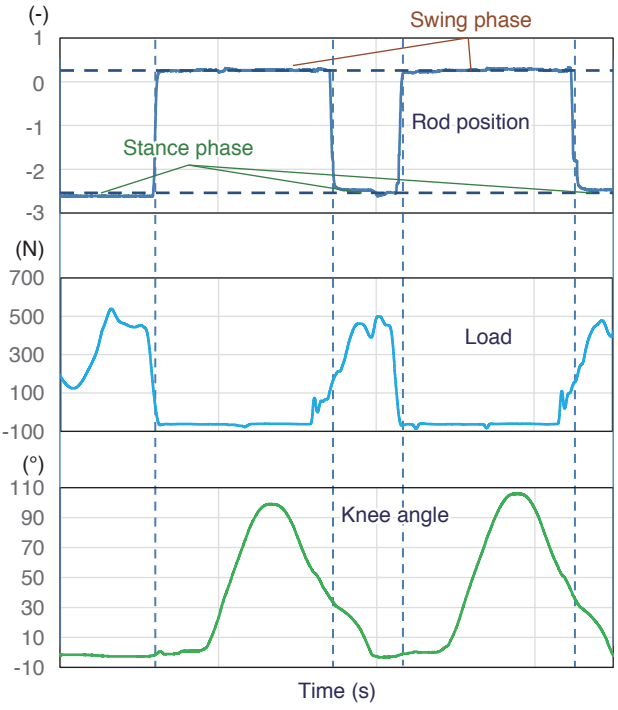


Fig. 20 Log data of going upstairs

Table 4 Performance comparison of roller friction and electromagnetic clutch

Features	Roller friction type	Electromagnetic clutch type		
		Company A	Company B	Company C
Outside shape (mm)	ϕ 25	ϕ 106	ϕ 75	ϕ 90
Mass (kg)	0.2	1.5	1	1
Static friction torque (N·m)	15-20			

In addition to addressing the three issues above, the following functions can be added when applying this technology to prosthetic legs.

- The prosthetic leg is easy to handle, with ability to walk freely on level ground.

By setting the roller clutch connection state in both gears to neutral, the lower leg can be swung out by the user's thigh. The user can walk freely without any restrictions on walking speed or cadence.

- The ability to provide bidirectional rotary drive encourages a sense of assurance for the user.

Because rotary driving force can be transmitted in both directions, it is possible to drive in opposite directions, such as flexion and extension, without switching. This encourages a sense of assurance for the user. The reason is that the knee joint can continue to support the body weight without fear that the clutch will disengage, even if the rotation unexpectedly goes the opposite direction.

- It is possible to switch between multiple reduction ratios.

The knee joint can have multiple reduction ratios, and different output characteristics can be achieved with a single motor, so that the knee joint can provide optimal characteristics during the stance and swing phases. Finding points for greater efficiency, and adjusting the amount of current are expected to contribute to the downsizing of motors and batteries.

4.5. Advantages over Electromagnetic Clutches

Table 4 lays out the characteristics of compactness and light weight that the new technology achieved compared to electromagnetic clutches. The results suggest that the technology could be effectively used not only for prosthetic legs but also for other products.

5. Conclusion

This development started with the goal of developing elemental technology to take advantage of the merits of motorized knee joints, and a prototype knee joint incorporating a new transmission mechanism was confirmed.

When the user is walking on level ground, the motor is not connected during the swing phase, but rather is connected during the stance phase. Power is saved because power is only consumed for switching. In addition, during the swing phase, there is no connection with the motor, which enables walking with less noise. Moreover, the mechanism can be disconnected from the motor and moved by the thigh, making it possible to walk without depending on motor performance.

During the stance phase, yielding functions similarly to a damper-type knee joint, enabling the user to walk with a sense of assurance.

Also, when going upstairs, the speed ratio can be switched according to the characteristics required for the stance phase and swing phase, which helps to make the motor smaller.

This transmission has many merits compared to electromagnetic clutches, including a compact and lightweight mechanism that can transmit torque, the ability to provide bidirectional rotary drive, no limit to the number of speed ratios, and no power consumption to maintain the state during torque transmission. As such, it is an effective means for achieving the objectives of a knee joint in a prosthetic leg.

As the features of this technology are expected to function effectively in other devices, the researchers hope it will be used in devices that can take advantage of its unique merits.

References

(1) https://www.ottobock.com/ja-jp/prosthetic_le/info/structure, (accessed 2024/7/10) (in Japanese)
(2) https://www.ottobock.com/en-us/product/3R80, (accessed 2024/7/10)
(3) https://global.honda/en/about/vision/, (accessed 2024/7/10)
(4) Riener, R., Rabuffetti, M., Frigo, C.: Stair ascent and descent at different inclinations, Gait & Posture, Vol. 15, Issue 1, p. 32-44, (2002), doi: 10.1016/S0966-6362(01)00162-X

■ Author ■


Hiromi ONO


Kei SHIMADA


Yutaka HIKI


Seiji ONOZAWA


Fumihito GOTO

Construction of Turbulent Component Decomposition Method Using Time Filtering with High-speed PIV and Application to Hybrid Engine Development

Yasuhiro OKURA, Ph.D.[1] Makoto SEGAWA[1] Hiromi ONIMARU[2]
Seisuke SATO[3] Yasuhiro URATA[1]

ABSTRACT

A time filtering method was constructed as a means of decomposing the turbulent component that contributes to promoting engine combustion. This time filtering method used the frequency at which the frequency characteristic of the in-cylinder flow changes as the cut-off frequency, and this frequency was confirmed to correspond to the time scale of the relatively large in-cylinder flow induced by the combustion chamber, intake port, and piston shapes. It was also clarified that the cut-off frequency can be uniquely determined by using the inverse of the integral time scale of in-cylinder flow as the cut-off frequency. The time filtering method was applied to a flow field equipped with a tumble control valve, which is an intake switching device. The obtained turbulence was highly correlated with the combustion duration of a metal engine, which confirmed the reliability of the decomposition method. The time filtering method was implemented in commercial analysis software and applied to hybrid engine development. Turbulence was enhanced by the in-cylinder flow enhancement specification that changed the intake port and piston shape from the early development specification and verified the in-cylinder flow enhancement concept for achieving rapid combustion.

1. Introduction

Social attention is increasingly focused on environmental issues such as global warming due to carbon dioxide (CO_2) emitted by human economic activities. In FY2022, CO_2 emissions from Japan's transportation sector accounted for approximately 18.5% of the country's total, and of that passenger cars accounted for approximately 44.9%[(1)]. To reduce CO_2 emissions from passenger cars, measures such as banning the sale of internal combustion engines and shifting to electrification using batteries or fuel cells as power units are progressing on a global scale. However, batteries require rare metals such as cobalt, the stable supply of which is thought to be an issue due to limited areas for extraction and ore production[(2)]. Furthermore, life-cycle assessments that take into account CO_2 emissions produced by battery manufacturing and thermal power generation and other factors suggest that CO_2 emissions from battery electric vehicles may be higher than those of vehicles equipped with gasoline engines in some cases[(3)]. Given that background, efforts are underway to develop fuel compositions suited to gasoline engines in order to achieve carbon neutrality in automotive internal combustion engines[(4), (5)] and to research synthetic fuels derived from renewable electric power (Erneuerbarer strom - fuel: e-fuel) that assume drop-in enabling use of existing infrastructure as is[(6)].

Combustion concepts using lean-burn combustion[(7), (8)] and massive exhaust gas recirculation (EGR)[(9)] in spark ignition (SI) gasoline engines are being studied for the purpose of enhancing thermal efficiency. These combustion concepts add extra air and EGR gas to the air-fuel mixture to reduce pumping loss and heat loss. On the other hand, efforts are being made to enhance the in-cylinder flow by measures such as optimizing the intake port shape to compensate for the decrease in the combustion rate due to a lean or diluted air-fuel mixture[(10)]. Such a method, however, may lead to lower power due to increased intake resistance in the mid-to-high engine speed range. Research and development of intake switching devices is underway to address this issue with the aim of achieving both a higher

[1] Honda Motor Co., Ltd.
E-mail: yasuhiro_okura@jp.honda
[2] Innovative Research Excellence, Power unit & Energy
[3] Honda Racing Corporation

tumble ratio and lower intake resistance according to the engine operating conditions[11), (12)].

2. Turbulent Component Decomposition Method

2.1. Issues with Conventional Methods

Turbulent enhancement is effective for promoting flame propagation combustion in SI gasoline engines, and the cycle average method has conventionally been used to evaluate turbulence in engine cylinders. The cycle average method defines the difference between the cycle-averaged velocity and the instantaneous velocity as the turbulent component[13)-(15)]. It has been pointed out, however, that this method also includes the relatively low-frequency cycle-by-cycle fluctuation of the mean flow component in the turbulent component, so it may overestimate the turbulence. Studies have attempted decomposition using a time filter as another decomposition method. This method used a steady-state time-averaging method to decompose the mean flow component and the turbulent component every cycle based on the velocity near the compression stroke top dead center (TDC). It was assumed that the gas flow can be regarded as a steady flow for a short duration in each cycle near the compression stroke TDC, but a clear physical definition of the averaged time considered to be steady flow is not given[16), (17)]. It was also proposed to use the frequency at which the integrated power spectrum of the cycle-averaged velocity was 90% or more as the cut-off frequency, however the cut-off frequency was considered arbitrary[18), (19)]. In this way, various methods have been proposed for turbulent component decomposition, but ambiguity remains in the cut-off frequency setting. Therefore, for example, when designing an intake port using the measured turbulent intensity as the criteria, it may be necessary to rework the design because the turbulence is overestimated.

2.2. Time Filtering Method

The authors have proposed a time filtering method utilizing the high time resolution of high-speed particle image velocimetry (PIV), which is capable of measuring at high speeds, as a new turbulent component decomposition method[20)]. Figure 1 shows the velocity u and the frequency characteristic of the in-cylinder flow from the intake stroke to the compression stroke. (Refer to Fig. 6 hereafter for the velocity direction.) The velocity of an arbitrary cycle fluctuates relative to the cycle-averaged velocity, and the conventional cycle average method treated this fluctuation as the turbulent component. The time filtering method obtains the frequency characteristic by the discrete Fourier transform every cycle, defines the frequency at which the slope of the cycle-averaged characteristic changes as the cut-off frequency f_{c_slope}, and then decomposes the component with a higher frequency than f_{c_slope} as the turbulent component. Here, $\hat{u}$ and $\hat{v}$ are the discrete Fourier coefficients of the u and v velocity components. Figure 2 compares the turbulent kinetic energy k obtained by the cycle average method and the time filtering method. This shows that the turbulent kinetic energy k_c with the cycle average method is overestimated overall compared to the turbulent kinetic energy k_f with the time filtering method. In particular, the turbulent kinetic energy in the latter half of the compression stroke directly affects combustion, so the combustion characteristics may be overestimated. From the above, the time filtering method is characterized by treating the frequency component higher than the cut-off frequency as the turbulent component, so it does not include the cycle variation of the mean flow component (low-frequency component), which is thought to not contribute to promoting combustion, in the turbulent component.

The object of this study is to construct a turbulent analysis method that can be applied to engine in-cylinder flow. First, the effect of the time resolution needed for the time filtering method was confirmed, and the cut-off frequency setting for decomposing the mean flow component and the turbulent component was investigated. Next, the time filtering method was applied to a tumble

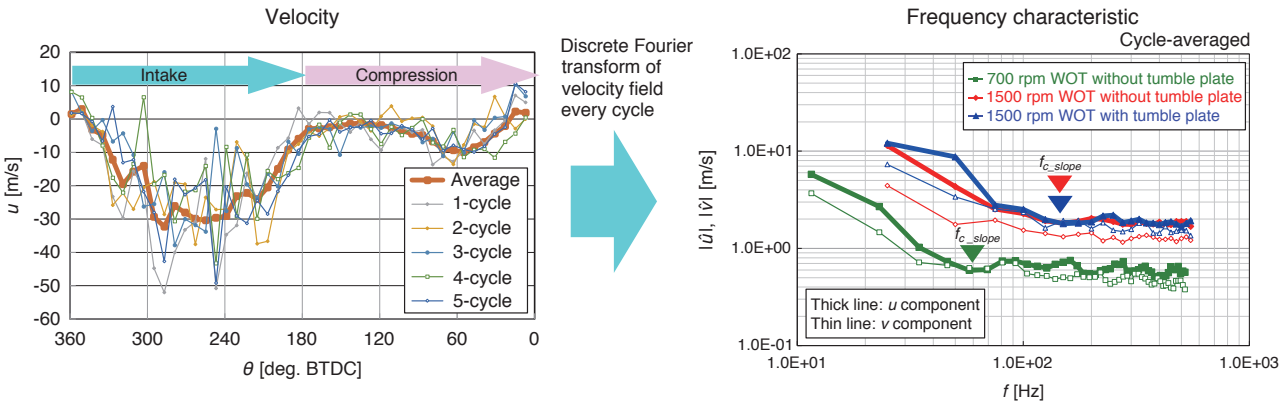


Fig. 1 Velocity and frequency characteristic of in-cylinder flow

control valve (TCV), which is an intake switching device, and the reliability of the turbulent analysis method was verified from the correlation between the obtained turbulence and the combustion characteristics. Finally, the time filtering method was implemented in a commercial PIV auto analysis tool, and the concept of in-cylinder flow enhancement in a hybrid engine was verified. This paper reports these details.

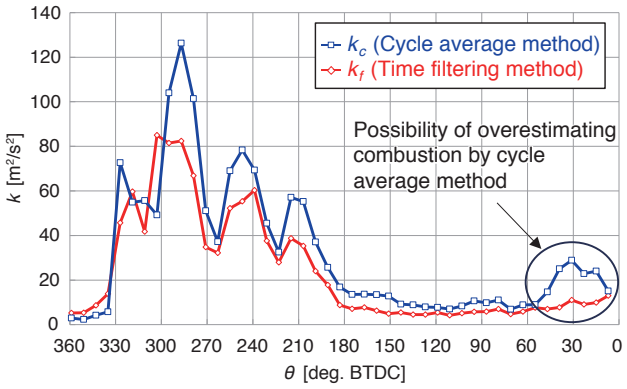


Fig. 2 Comparison of turbulent kinetic energy obtained by time filtering method and cycle average method

3. Construction of Time Filtering Method

3.1. Experimental Technique

3.1.1. Engine specifications

Table 1 shows the main engine specifications, and Fig. 3 shows the appearance of the engine. The engine is an optical full-stroke single-cylinder engine made by Koyama Garage Co., Ltd. It has a bore diameter of 73 mm and a stroke of 89.4 mm and is equipped with two intake and two exhaust valves. A quartz sleeve is installed between the cylinder head and the cylinder block, providing a side view of the inside of the cylinder from the intake stroke to the compression stroke. The top shape of the extended piston can be changed by parts replacement. For optical access from below the piston, it can be changed to a flat shape with a quartz window that provides a view via a 45-degree mirror installed below the extended piston.

3.1.2. Experimental setup

The PIV used for velocity measurement uses a laser or other light source to illuminate tracer particles mixed into the airstream and obtains the velocity from the amount of particle movement in the images captured by a camera. A high-repetition double-pulse Nd:YAG laser (LDP-50MQG made by Lee Laser, Inc.) was used as the light source, and the wavelength of the laser beam was 532 nm. A high-speed CMOS camera (APX-RS made by PHOTRON LIMITED) was used to capture particle images. SiO_2 (Godd Ball B-6C made by SUZUKIYUSHI INDUSTRIAL CORPORATION) with an average particle diameter of 3 μm was used as the tracer and was supplied into the engine from the intake chamber located immediately after the throttle. The air pressure of the tracer supply device was adjusted to achieve a tracer particle concentration that enabled imaging of a few particles in the 16 × 16 pixel inspection area for the PIV analysis described hereafter.

Table 1 Engine specifications

Engine	Optical single-cylinder engine with quartz sleeve
Bore × Stroke	73 mm × 89.4 mm
Displacement	374 cm³
Valve train	DOHC 4 valves

Figure 4 shows the crank angle range over which measurement was performed. The intake valve lift curve,

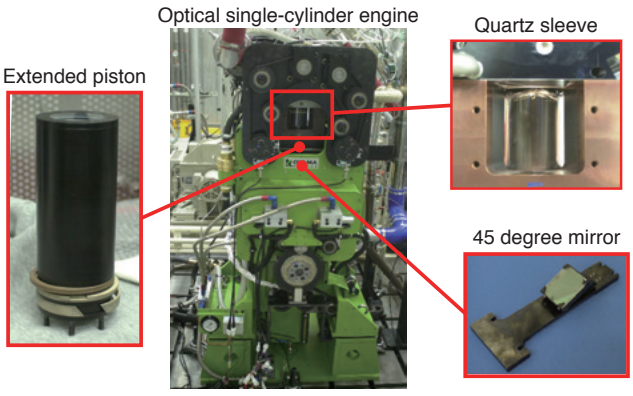


Fig. 3 Appearance of optical single-cylinder engine

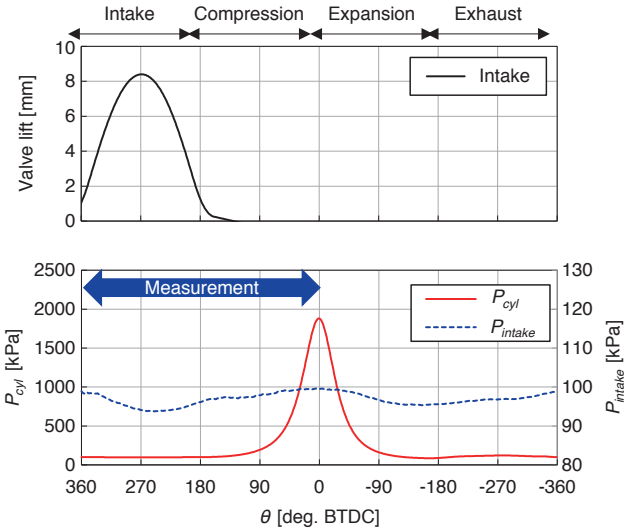


Fig. 4 Measurement range of in-cylinder flow

in-cylinder pressure P_{cyl}, and the intake manifold pressure P_{intake} are added to the figure for reference. The measured crank angle was from the intake stroke TDC crank angle θ = 360 deg. BTDC to the compression stroke TDC crank angle θ = 0 deg. BTDC, and this was measured for multiple cycles. The obtained particle images were analyzed using the commercial PIV analysis software Koncerto ver1.0 made by Seika Digital Image CORPORATION. The analysis algorithm used the recursive FFT cross-correlation method, and the inspection area was varied in three steps from 64 × 64 pixels to 16 × 16 pixels, with overlap of 50%. Three test conditions were used: *Ne* = 700 rpm, 1500 rpm, and 2000 rpm, all at wide open throttle (WOT).

3.2. Examination of Cut-off Frequency Setting Method

In this section, the time resolution needed for the time filtering method is clarified, and the method for setting the cut-off frequency and its physical meaning are examined.

Figure 5 shows the layout of the PIV system used in the test, and Fig. 6 shows the measurement region and analysis area. The laser beam was stretched into a 0.5 mm thick sheet by cylindrical lenses, and then irradiated from the side of the engine into the cylinder through a transparent sleeve. The measurement position was the bore center. Imaging was performed with a frame speed of 20 kilo frame per second (kfps), an imaging range of horizontal 40 mm × vertical 18 mm, and a spatial resolution of 0.077 mm/pixel (512 × 256 pixel). In addition, a horizontal 36 mm × vertical 6 mm range at the top of the cylinder was used as the analysis area to enable turbulent analysis up to the vicinity of the compression stroke TDC.

3.2.1. Comparison of time filtering method and cycle average method

Figure 7 compares the cycle-averaged velocity distribution and instantaneous velocity distribution using the in-cylinder flow at θ = 300.1 deg. BTDC in the early stage of the intake stroke as an example. This showed that the instantaneous velocity distribution and the tumble vortex center (red dot) differed in each of the arbitrary six cycles shown in Fig. 7(b). In addition, the cycle-averaged velocity distribution shown in Fig. 7(a) also differed from the velocity distribution in any cycle.

Fig. 8(a) shows the difference in the energy decomposed by the time filtering method and the cycle average method. The energy E on the vertical axis is the mean square value of the u and v velocity components at the analysis point. The time filtering method obtained the frequency characteristic by the discrete Fourier transform every cycle, and the cycle-averaged values were indicated in blue. The frequency f_{c_slope} at which the slope changes

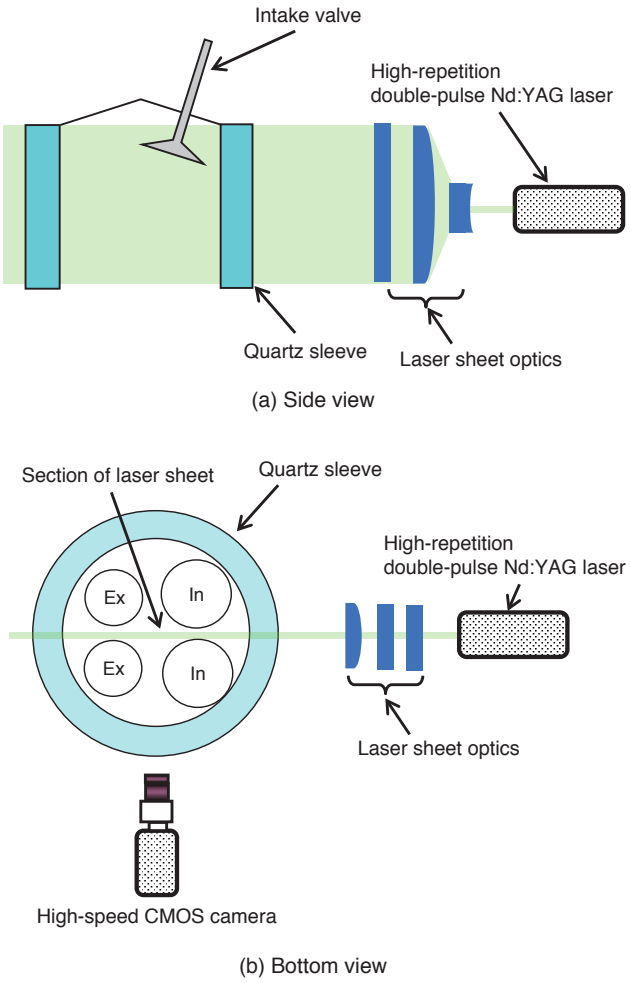


Fig. 5 Layout of PIV system for construction of turbulent analysis method

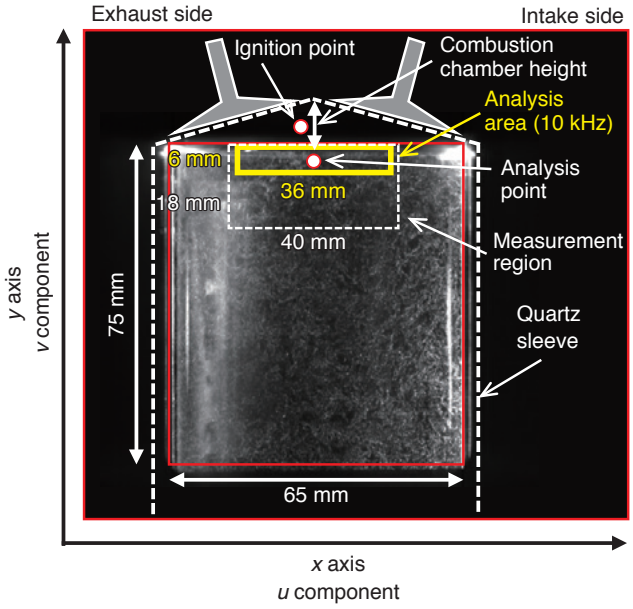
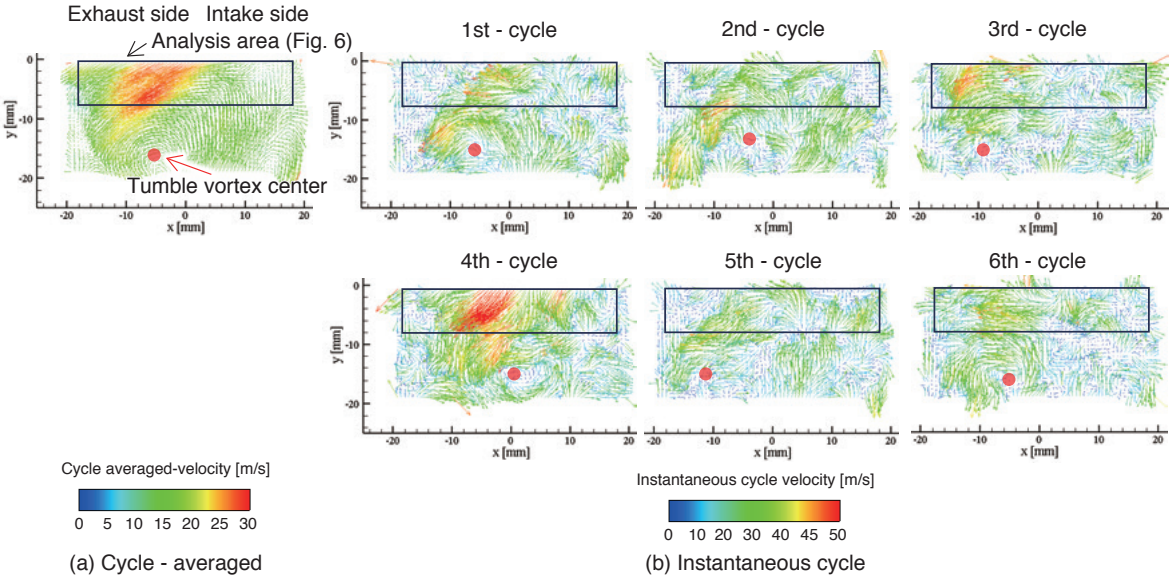


Fig. 6 Measurement region and analysis area for construction of turbulent analysis method

was 360 Hz, and the frequency band higher than that was the turbulent component (blue painted area). This turbulent component corresponds to the mean square value of the mean flow and instantaneous velocity difference shown in Fig. 8(b). The frequency characteristic of the cycle average method was obtained by the discrete Fourier transform from the cycle-averaged velocity and was indicated by the solid red line in Fig. 8(a). The turbulent component of the cycle average method was the red painted area surrounded by the spectra of the time filtering method and the cycle average method, and the area surrounded by the black dotted line was overestimated compared to the time filtering method. This overestimated area is the mean square value of the velocity difference between the time filtering method and the cycle average method in Fig. 8(b); that is to say, the cycle variation of the mean flow component.

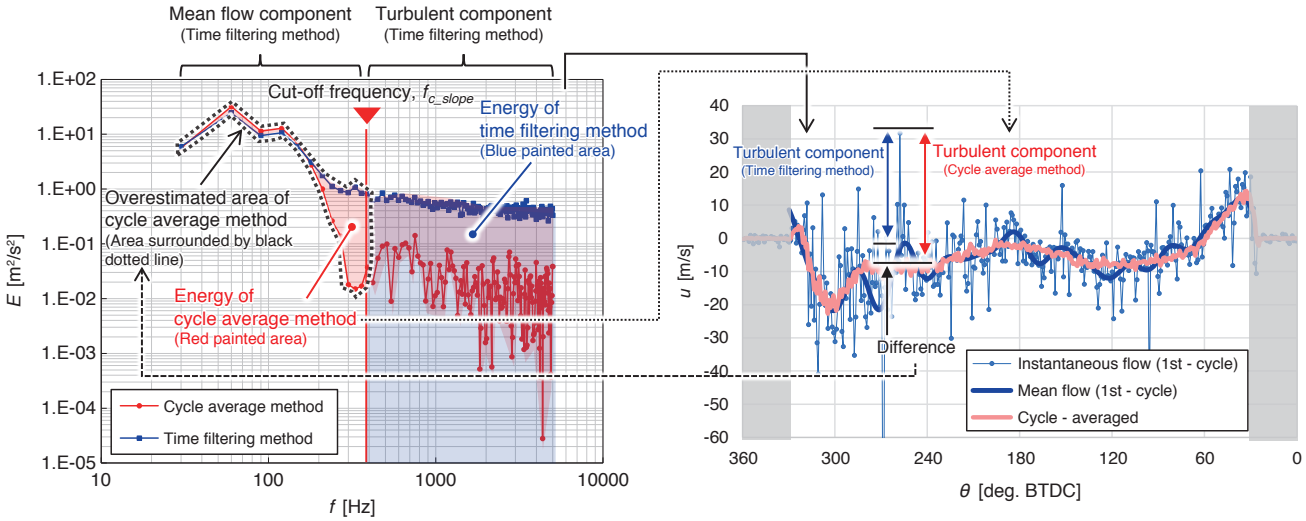
3.2.2. Effect of time resolution

The time resolution is thought to be important with the time filtering method, so it was analyzed. Figure 9 shows



(a) Cycle - averaged (b) Instantaneous cycle

Fig. 7 Comparison of cycle-averaged in-cylinder flow and instantaneous in-cylinder flow for 6 cycles in measurement region, *Ne* = 1500 rpm WOT at θ = 300.1 deg. BTDC



(a) Energy spectrum (b) In-cylinder velocity

Fig. 8 Comparison of energy obtained by time filtering method and cycle average method at analysis point, *Ne* = 1500 rpm WOT

the effect of time resolution on the energy spectrum E_u of the u velocity component at the analysis point. The time resolution was artificially reduced by thinning of the same measurement data in the crank angle direction. When the time resolution was between around 10 to 2.5 kHz, this did not produce a noticeably large effect on the frequency where the slope changes. The discrete Fourier transform is performed in this analysis, so when a lower time resolution is used, frequencies that differ by an integer multiple of the Nyquist frequency f_N that is half-frequency of the true spectrum and the time resolution are viewed as the same due to aliasing. That is to say, the spectra for each f_N interval overlap the true spectrum, resulting in deviation due to discretization[21]. In this flow field, even when the time resolution is lowered from 10 kHz to 5 kHz, the reduction in energy of 360 Hz or more is approximately 5%, so the effect of aliasing on f_{c_slope} is thought to be small, assuming a time resolution of around 5 kHz.

3.2.3. Cut-off frequency setting method

In this section, the integral time scale of in-cylinder flow τ, the physical meaning of f_{c_slope}, and the cut-off frequency setting are examined.

Figure 10 compares the spatial-averaged integral time scale of in-cylinder flow in the analysis area and the characteristic time of the piston speed τ_p. τ_u was obtained by using Eq. (1) to calculate the autocorrelation coefficient R_{uu} from the instantaneous velocity u, and then integrating with Eq. (2) until the autocorrelation coefficient R_{uu} was 0[22]. The integral time scale of in-cylinder flow τ_v of the v velocity component was also calculated in the same manner.

$$R_{uu}(\Delta\theta) = \frac{1}{u(\theta)^2}\sum u(\theta)u(\theta+\Delta\theta) \quad (1)$$

$$\tau_u = \sum R_{uu}(\Delta\theta) \quad (2)$$

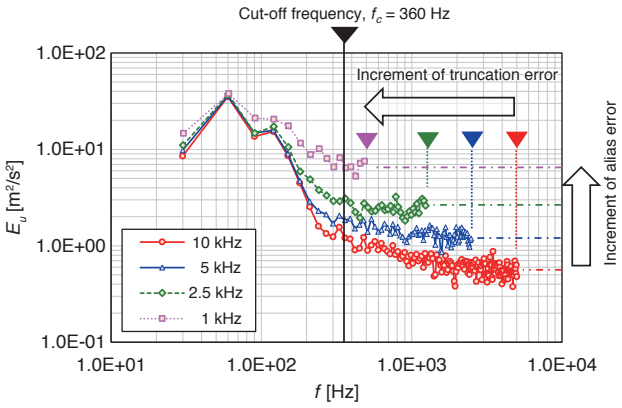


Fig. 9 Effect of time resolution on cycle-averaged in-cylinder energy spectrum at analysis point, *Ne* = 1500 rpm WOT

τ_p was defined as the value obtained by dividing the height of the combustion chamber h_p from the top of the pent-roof combustion chamber to the piston top at bore center by the instantaneous piston speed u_p. h_p at the TDC was 15 mm. Thus, on the integral time scale of in-cylinder flow, particularly the change in τ_v relative to the crank angle exhibited tendencies close to τ_p. In this relationship, the integral time scale of in-cylinder flow is thought to correspond to the relatively large-scale in-cylinder flow induced by the combustion chamber and intake port shapes and the piston speed. The average of τ_u and τ_v at θ = 270 deg. BTDC, which is in the middle of the intake stroke, is used as the integral time scale of in-cylinder flow τ, and its inverse $f_{c\_\theta=270}$ and f_{c_slope} are compared in Fig. 11. The data of Chapter 4 is also plotted in the figure. As a result, the correlation coefficient between $f_{c\_\theta=270}$ and f_{c_slope} was 0.936 in Chapter 3, and 0.918 overall with the Chapter 4 data

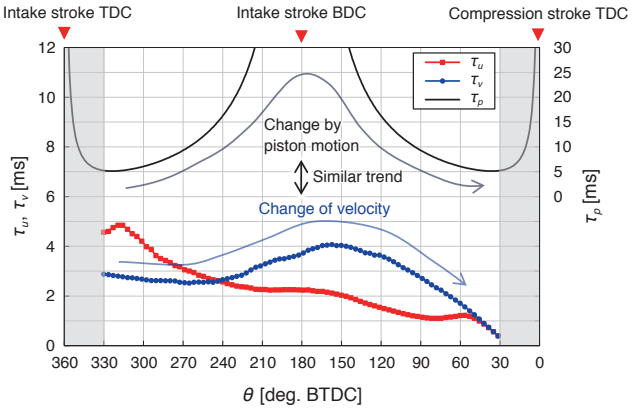


Fig. 10 Comparison of spatial-averaged integral time scale of in-cylinder flow, τ_u, τ_v in analysis area and time scale of piston speed, τ_p. Test condition is *Ne* = 1500 rpm WOT.

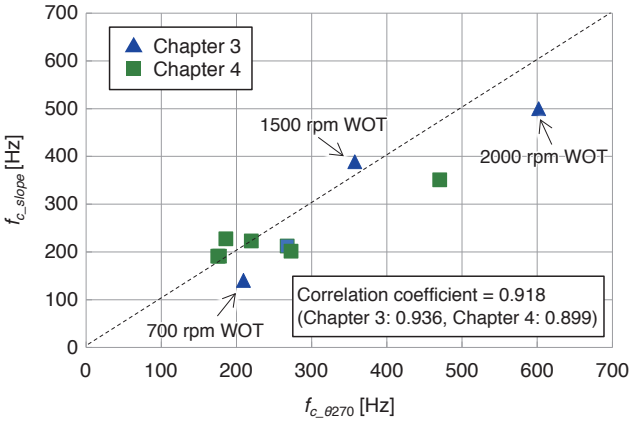


Fig. 11 Relationship between cut-off frequency at changed slope of in-cylinder flow frequency characteristic f_{c_slope} and cut-off frequency obtained from autocorrelation $f_{c\_\theta270}$

added, indicating a high correlation.

From the above, the frequency f_{c_slope} at which the slope of the frequency characteristic of the in-cylinder flow changes is thought to be correlated with the integral time scale of the in-cylinder flow τ and corresponds to the relatively large-scale in-cylinder flow induced by the combustion chamber and intake port shapes and the piston speed. In addition, the cut-off frequency can be uniquely determined using the inverse of integral time scale of in-cylinder flow, $f_{c\_\theta 270}$ as the cut-off frequency. Unless stated otherwise, $f_{c\_\theta 270}$ is used as the cut-off frequency f_c hereafter.

4. Verification of Reliability of Time Filtering Method

In this chapter, the time filtering method was applied to the intake switching device, TCV, and the reliability of the time filtering method, which is a turbulent component decomposition method, was verified based on the correlation between the obtained turbulence and the combustion characteristics.

Table 2 shows the TCV specifications, and Fig. 12 shows the schematic diagram. When the lower passage of the tumble plate is closed by a flap, air flows through the upper passage of the tumble plate, so this specification is called upper flow and is noted as UF. Conversely, when the upper passage of the tumble plate is closed by a flap, air flows through the lower passage of the tumble plate, so this specification is called lower flow and is noted as LF. The longer the length of the tumble plate L_{Plate}, the closer the tip position of the tumble plate to the intake valve stem, and the flow into the combustion chamber is thought to have higher directivity. Therefore, the distance G_{Plate} from the tip of the tumble plate to the intake valve stem is used as the tumble plate parameter, and the three specifications of 2 mm, 25 mm, and 50 mm are provided. The flap and tumble plate specifications are noted together hereafter. For example, the case of upper flow and G_{Plate} = 2 mm is noted as UF 2 mm.

Figure 13 shows the measurement region and the analysis area. The laser sheet beam was reflected upward by the 45-degree mirror and irradiated from the underside of the engine through the piston window and into the combustion chamber. The measurement position was the bore center, and the measurement range was a horizontal 46.6 mm × vertical 24 mm region in the combustion chamber. In order to analyze turbulence from ignition by the spark plug to the early stages of combustion, the spark plug was removed and a dummy plug without the outer and center electrodes was installed. The ignition point of the spark plug was hidden by the intake valve in the intake stroke, so a position offset 4.6 mm to the exhaust side (analysis point) was used for the analysis. In addition, when performing spatial averaging, the target was the horizontal 8 mm × vertical 4 mm range (analysis area) where intake valve lift did not affect the analysis results. The test condition was Ne = 1500 rpm WOT.

Table 2 Specifications of tumble control valve

Specifications	Flow part	G_{plate} [mm]	L_{Plate} [mm]
UF 2 mm	Upper	2	93
UF 25 mm	Upper	25	70
UF 50 mm	Upper	50	45
LF 2 mm	Lower	2	93
LF 25 mm	Lower	25	70
LF 50 mm	Lower	50	45

4.1. Turbulent analysis of TCV

Figure 14 shows the spatial distributions of the mean flow component and the turbulent kinetic energy k_f for the cases without TCV, TCV UF 2 mm (upper flow), and TCV LF 2 mm (lower flow). The cut-off frequency f_c was 267 Hz, 219 Hz and 470 Hz, respectively. The turbulent kinetic energy k_f was obtained by Eq. (3).

$$k_f = \frac{1}{2}(u'^2_f + v'^2_f) \tag{3}$$

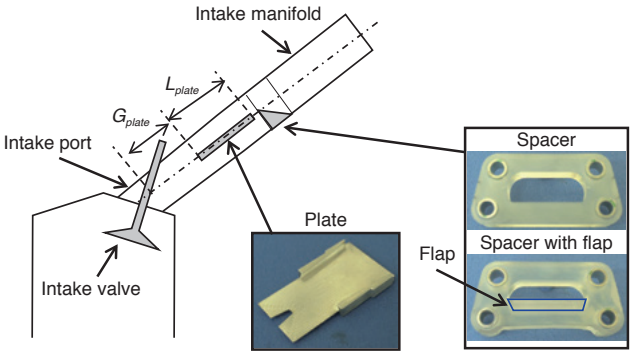


Fig. 12 Schematic diagram of tumble control valve

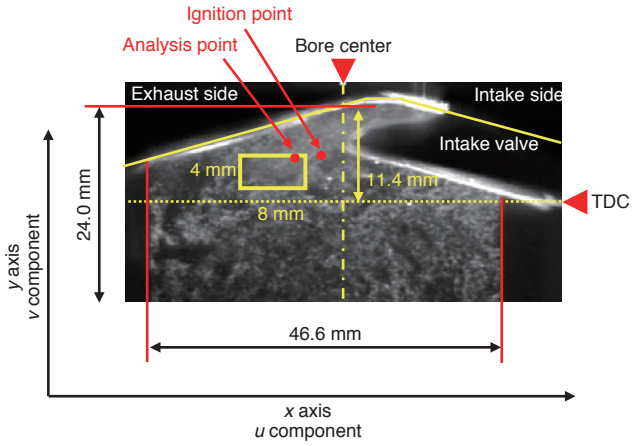


Fig. 13 Measurement region and analysis area for turbulent analysis with tumble control valve

Here, u'_f and v'_f are the turbulent velocity components decomposed by the time filtering method. As shown, the mean flow velocity of TCV UF 2 mm was faster than without TCV, and the vortex center was near the center of the combustion chamber. The mean flow velocity of TCV LF 2 mm indicated almost the same values as without TCV, and there was also no significant difference in the distribution. The k_f of TCV UF 2 mm indicated higher values than without TCV, and the k_f near the center of the combustion chamber where the vortex center of the tumble flow was located indicated particularly high values. On the other hand, the k_f of TCV LF 2 mm indicated almost the same values as without TCV, and there was also no difference in the distribution. The reason for this is that, as shown in the schematic diagram in Fig. 14, the flow into the cylinder from the lower side of the intake port is directed to the exhaust side by the intake valve and does not contribute to enhancing the tumble flow.

4.2. Estimation of turbulent flame structure

The flame structure with a TCV mounted was estimated from the turbulent premixed combustion diagram[23]. Figure 15 shows the results at θ = 0.4 deg. BTDC, which is near the compression stroke TDC. The turbulent intensity u' on the vertical axis used the square root of the turbulent kinetic energy k_f measured in monitoring with the optical engine. The laminar burning velocity S_L was obtained by

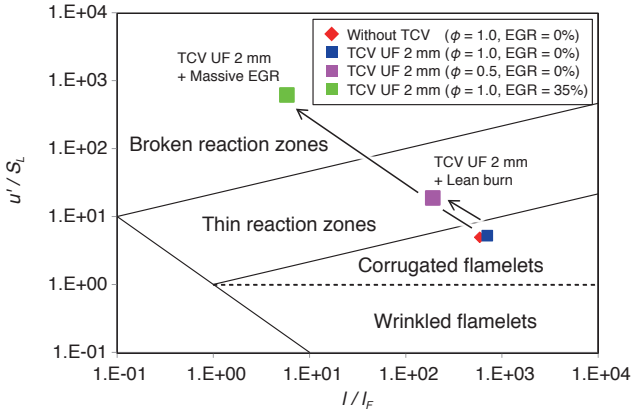


Fig. 15 Effect of enhancement of in-cylinder flow by TCV in analysis area, *Ne* = 1500 rpm WOT at θ = 0.4 deg. BTDC

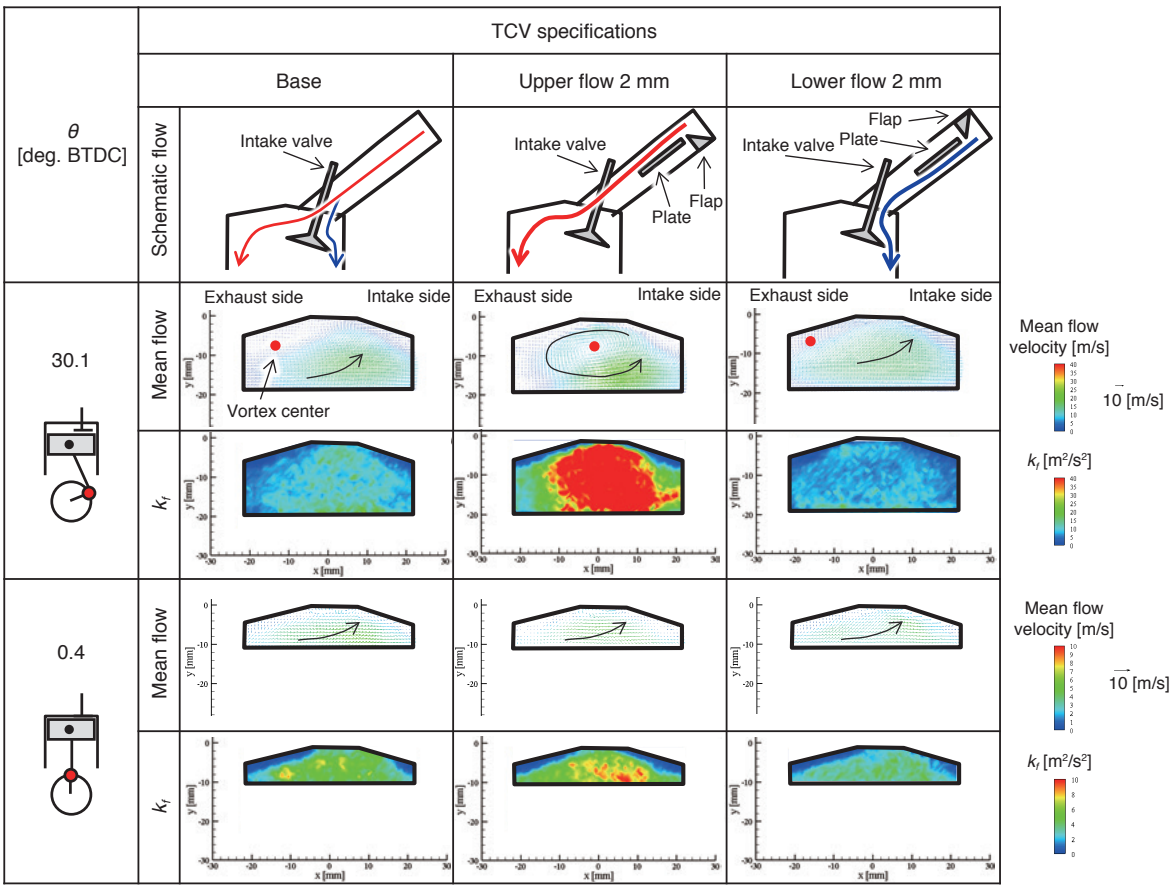


Fig. 14 Distribution of mean flow velocity and turbulent kinetic energy k_f obtained by time filtering method, *Ne* = 1500 rpm WOT in without TCV, upper flow 2 mm and lower flow 2 mm

elementary reaction calculations, assuming that the physical properties of the fluid were an air-fuel mixture with an equivalent ratio of φ = 1.0 and an EGR ratio of 0%[24]. The fuel composition was surrogate gasoline (primary reference fuel: PRF), and the conditions were in-cylinder pressure P_{cyl} = 1270 kPa and in-cylinder gas temperature T_{cyl} = 532 K. The turbulent integral length scale l and the flame thickness l_F on the horizontal axis were respectively obtained by Eq. (4) and Eq. (5).

$$l = u' \times \tau_l \quad (4)$$

$$l_F = \nu_F / S_L \quad (5)$$

Here, τ_l is the integral time scale of the turbulent intensity u', ν_F was the dynamic viscosity coefficient of the air-fuel mixture, and the fuel composition was obtained by the physical property estimation method[25].

As shown, the flame structures of without TCV and TCV UF 2 mm were both classified as corrugated flamelets. The u' and l of TCV UF 2 mm were larger than without TCV, so it was plotted slightly to the upper right of without TCV. In the corrugated flamelets region, wrinkles form on the flame surface, but turbulence is thought to not affect the internal structure of the flame, and the wrinkles on the flame surface are more pronounced than in the wrinkled flamelets region[26]. That is to say, turbulence is enhanced and the increase in the flame surface area is expected to promote combustion. In anticipation of further enhancement of combustion efficiency, the turbulent flame structures were also calculated for lean burn and massive EGR. When the air-fuel mixture was diluted from an equivalent ratio of φ = 1.0 to φ = 0.5, S_L decreased from approximately 42.4 cm/s to 11.5 cm/s, and the flame structure changed to the thin reaction zones. When the EGR ratio increased to 35% at φ = 1.0, S_L decreased to 0.35 cm/s and the flame structure shifted to the broken reaction zones. These flame structures are thought to have local flame quenching due to turbulence. That is to say, in the case of these flame structures, combustion may not be promoted simply by enhancing the in-cylinder flow. Additional measures are therefore needed, such as enhancement of ignitability by suppressing cycle fluctuation of the mean flow around the spark plug, etc.

4.3. Relationship between turbulence and combustion characteristics

The effect of the TCV on combustion was confirmed using a metal single-cylinder engine with the same specifications (hereafter, metal engine). Figure 16 shows the relationship between the turbulent kinetic energy k_f obtained in monitoring with the optical engine and the combustion duration obtained in metal engine firing. The load of the metal engine was the indicated mean effective pressure high-loop (IMEP.H) from the compression stroke to the expansion stroke, and was 1000 kPa. The duration from ignition to a mass fraction burnt (MFB) of 10% was defined as the early combustion duration θ_{ig}-MFB10, and k_f used the value at the ignition timing θ_{ig}. As shown, when k_f increased by the TCV, θ_{ig}-MFB10 decreased accordingly. The most effective TCV specifications compared to without TCV were UF 2 mm and UF 25 mm. In contrast, the LF 2 mm specification increased θ_{ig}-MFB10. The correlation coefficient between the turbulent kinetic energy k_f and θ_{ig}-MFB10 was 0.875, and the k_f obtained by the time filtering method exhibited correlation with combustion, which confirmed the reliability of the turbulent analysis method.

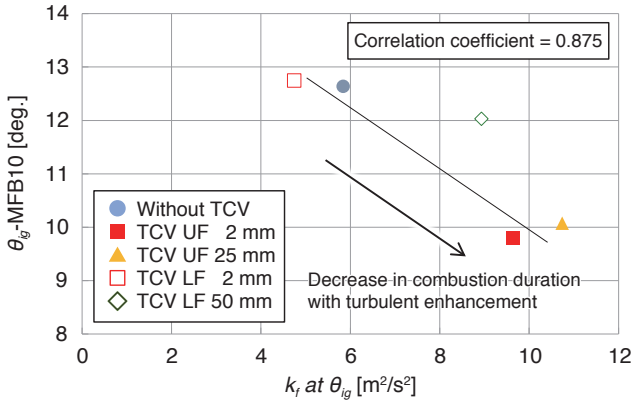


Fig. 16 Relationship between combustion duration on metal single-cylinder engine and turbulent kinetic energy, k_f obtained by time filtering method at ignition timing on optical single-cylinder engine, *Ne* = 1500 rpm IMEP.H = 1000 kPa

5. Application to Hybrid Engine Development

5.1. Standardization of Analysis Tools

To enhance ease of use, the constructed time filtering method was implemented as an in-cylinder flow analysis function in the commercial PIV auto analysis tool FtrPIV-Cruise made by Flowtech Research Inc. This tool features the ability to perform a series of processes such as PIV measurement setup, PIV analysis, statistical processing (turbulent analysis, etc.), and report creation using the same tool[27]. A function was added to automatically perform turbulent analysis and create reports in web browser format simply by selecting the time filtering method using the graphical user interface (GUI). Simple setup and automation with a GUI allow shortening of the work time needed for turbulent analysis and report creation by approximately 30%, which enables timely feedback of verification results to engine development.

5.2. Verification of In-cylinder Flow Enhancement Concept

The engine specification is an inline 4-cylinder direct

injection engine with displacement of 2.0 L for a 2-motor hybrid. In order to achieve class-leading fuel efficiency, this engine has a compression ratio of 13.9 and is equipped with a cooled EGR system. To achieve quick combustion, the tumble flow is enhanced by the intake port, and a bowl-in piston is adopted that smoothly converts the tumble flow to an upward flow by the piston[28].

The intake port and piston shapes were examined using computational fluid dynamics (CFD) to enhance the in-cylinder flow. In order to verify this in-cylinder flow enhancement concept, the in-cylinder flow with early development and in-cylinder flow enhancement specifications were analyzed by PIV. Measurement was performed for 60 cycles at Ne = 2000 rpm WOT with a time resolution of 3 deg. BTDC (frame speed 8 kfps). CFD analysis used the commercial software VECTIS made by Ricardo Software, and the standard k-ε model was applied to the turbulence model. Figure 17 compares the cycle-averaged in-cylinder flow between the early development and in-cylinder flow enhancement specifications. The velocity of the in-cylinder flow enhancement specification obtained by PIV at θ = 180 deg. BTDC was higher than that of the early development specification, and the intake flow from the intake port was smoothly converted to an upward flow. Even at θ = 30 deg. BTDC in the latter half of the compression stroke, the velocity of the in-cylinder flow enhancement specification obtained by PIV was fast, and these tendencies were similar to the CFD analysis results.

Figure 18 compares the turbulent kinetic energy k_f with early development and in-cylinder flow enhancement specifications obtained by PIV. Compared to the early development specification, the k_f with the in-cylinder flow enhancement specification indicated a large value

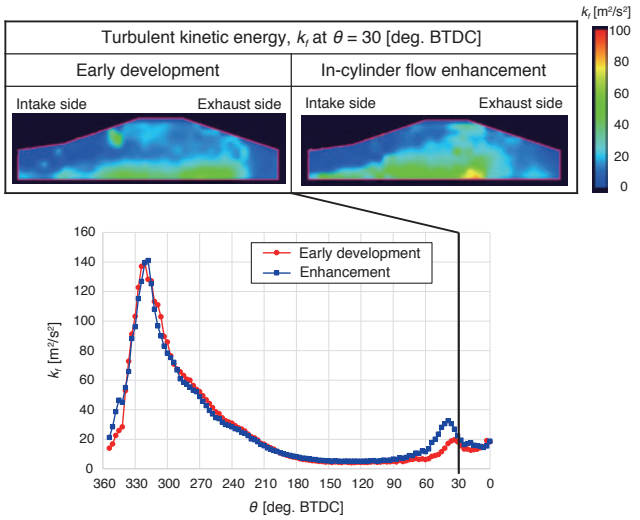


Fig. 18 Comparison of turbulent kinetic energy k_f with early development and in-cylinder flow enhancement specification obtained by PIV, Ne = 2000 rpm WOT

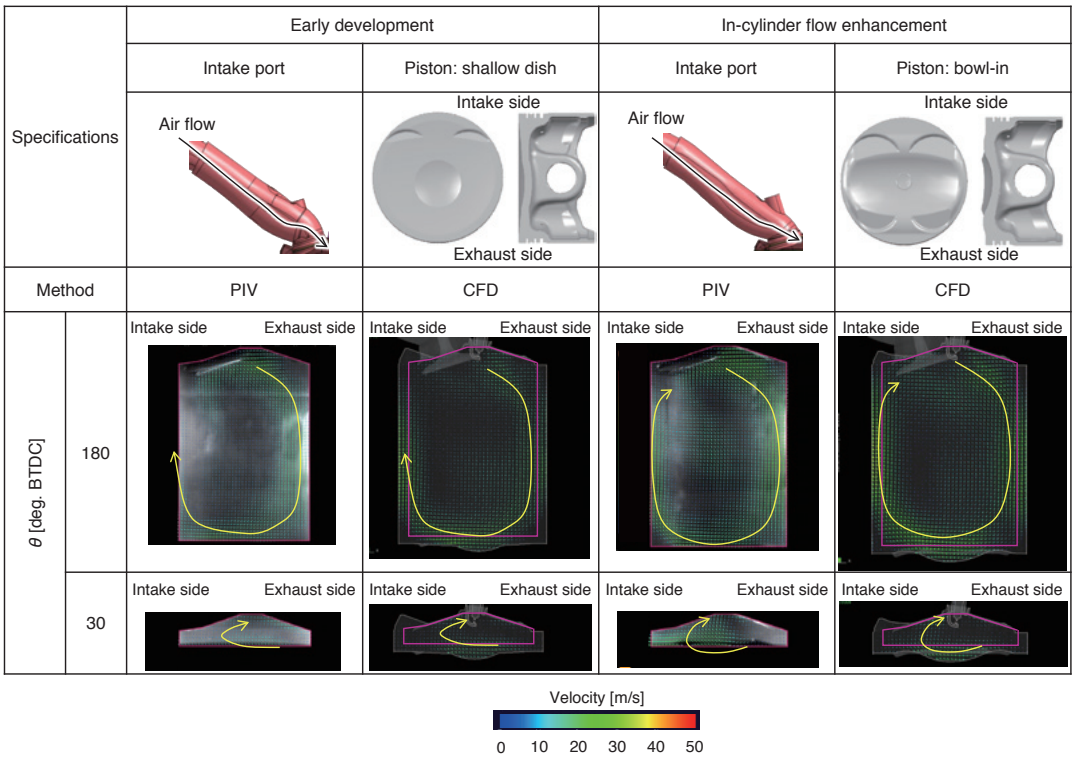


Fig. 17 Comparison of in-cylinder flow with early development and in-cylinder flow enhancement specification obtained by PIV and CFD, Ne = 2000 rpm WOT

from θ = 60 deg. BTDC onward. In the k_f distribution at θ = 30 deg. BTDC, the k_f with the in-cylinder flow enhancement specification tended to be higher than that with the early development specification near the center of the combustion chamber.

Figure 19 compares the turbulent kinetic energy k_f with the time filtering method and the cycle average method. With the time filtering method, changing from the early development specification to the in-cylinder flow enhancement specification increased k_f by approximately 21.9%. However, this increase was 24.9% with the cycle average method, so the cycle average method estimated k_f approximately 3% larger.

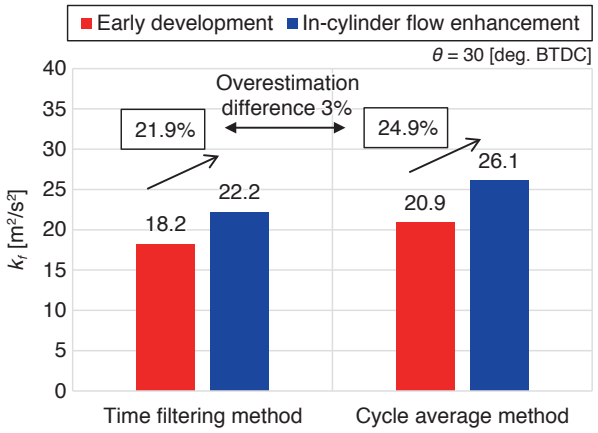


Fig. 19 Effect of time filtering method, *Ne* = 2000 rpm WOT at θ = 30 deg. BTDC

6. Conclusion

A time filtering method that makes use of high time resolution by high-speed PIV was proposed as a new turbulent component decomposition method, and application to engine in-cylinder flow was examined. In addition, the time filtering method was implemented in a commercial PIV auto analysis tool, and the following conclusions were obtained through application to hybrid engine development.

(1) The time resolution was artificially varied, and the turbulent kinetic energy characteristics were verified. It was found that when the time resolution was 5 kHz or more, there was no great effect on the cut-off frequency used to decompose the mean flow component and the turbulent component in this flow field.

(2) The frequency at which the slope of the frequency characteristic of the in-cylinder flow changes was found to correspond to the relatively large-scale in-cylinder flow induced by the combustion chamber and intake port shapes and the piston speed. Furthermore, the cut-off frequency can be uniquely determined by using the inverse of the integral time scale of in-cylinder flow as the cut-off frequency.

(3) The time filtering method was used to analyze the turbulent characteristics of a tumble control valve, which is an intake switching device. The results confirmed that the flame structure was classified as corrugated flamelets, and combustion promoting effects can be obtained by turbulence. There was high correlation between the turbulent kinetic energy and the metal engine combustion duration, which confirmed the reliability of the turbulent analysis method.

(4) The time filtering method was implemented in a commercial PIV auto analysis tool, and the analysis tool was standardized to enhance ease of use. This tool was then applied to hybrid engine development and verified the concept of in-cylinder flow enhancement for realizing quick combustion.

References

(1) Ministry of Land, Infrastructure, Transport and Tourism, https://www.mlit.go.jp/sogoseisaku/environment/sosei_environment_tk_000007.html, (accessed 2024/07/17) (in Japanese)

(2) Ministry of Economy, Trade and Industry, https://www.meti.go.jp/shingikai/mono_info_service/jidosha_shinjidai/pdf/20180831_01.pdf, (accessed 2024/07/23) (in Japanese)

(3) Kimura, S., Matsuura, H., Kikuchi, T., Tsuchiya, K.: Examination of AICE's Technical Scenario for Carbon Neutrality in Vehicles Equipped with Internal Combustion Engines, 2021 JSAE Annual Congress (Autumn) proceedings, No. 121-21, 20216181, (2021) (in Japanese)

(4) Matsubara, N., Kitano, K., Yokoo, N., Nakata, K., Miyoshi, A., Obata, K., Nomura, M., Fujimoto, T.: Development of Internal Combustion Engine and Fuel Composition to Achieve Early Low-Carbonization of Automobiles, 2023 JSAE Annual Congress (Spring) proceedings, No. 20-23, 20235090, (2023) (in Japanese)

(5) Kaneko, K., Matsubara, N., Kitano, K., Yokoo, N., Nakata, N., Yasutake, Y., Naiki, T., Watanabe, M., Obata, K.: Effects of Fuel Components on Thermal Efficiency in Super Lean Burn S.I. Engine, 2023 JSAE Annual Congress (Spring) proceedings, No. 20-23, 20235091, (2023) (in Japanese)

(6) Nakayama, T.: e-fuel ni kansuru gijyutsutyousa to jissyoujikken, Journal of the Society of Mechanical Engineers, Vol. 127, Issue 1266, p. 40-43, (2024) (in Japanese), doi: 10.1299/jsmemag.127.1266_40

(7) Kimura, K., Sakai, H., Omura, T., Takahashi, D.: Development of Super Lean Burn Engine with 50% Thermal Efficiency to Achieve Carbon Neutrality, 2023 JSAE Annual Congress (Spring) proceedings, No. 32-23, 20235145, (2023) (in Japanese)

(8) Hanabusa, H., Kondo, T., Hashimoto, K., Sono, H., Furutani, M.: Study on Homogeneous Lean Charge Spark Ignition Combustion, SAE Technical paper 2013-01-2562, (2013), doi: 10.4271/2013-01-2562
(9) Tagishi, R., Ikeya, K., Takazawa, M., Yamada, T.: Pursuit of Thermal Efficiency of Gasoline Engine, 2014 JSAE Annual Congress (Autumn) Proceedings, No. 121-14, 20145712, (2014) (in Japanese)
(10) Takahashi, D., Nakata, K., Yoshihara, Y., Ohta, Y., Nishiura, H.: Combustion Development to Achieve Engine Thermal Efficiency of 40% for Hybrid Vehicles, 2015 JSAE Annual Congress (Spring) Proceedings, No. 1-15S, p. 15-20, 20155004, (2015) (in Japanese)
(11) Nogawa, S., Nakata, K., Kanda, M.: Improvement of Thermal Efficiency by Lean Boosted Engine, 21th Internal Combustion Engine Symposium Proceedings, A7-4, p. 533-538, (2010) (in Japanese)
(12) Heiduk, T., Kuhn, M., Stichlmeir, M., Unselt, F.: The new 1.8l TFSI Engine from Audi Part 2: Mixture Formation, Combustion Method and Turbocharging, MTZ, Vol. 72, p. 58-64, (2011), doi: 10.1365/s38313-011-0078-1
(13) Kido, H., Wakuri, Y., Murase, E., Fujimoto, K., Wang, Z., Tomita, O.: Variations of Spatial Scales of Turbulence in an Engine Cylinder during the Compression Stroke, Transactions of the Japan Society of Mechanical Engineers Series B, Vol. 50, Issue 452, p. 1114-1121, (1984) (in Japanese), doi: 10.1299/kikaib.50.1114
(14) Jeong, H., Furui, T., Nishiyama, A., Ikeda, Y.: Turbulence characteristics measurement near spark plug with PIV and LDV, 22nd Internal Combustion Engine Symposium Proceedings, 20117085, (2011)
(15) Müller, S. H. R., Böhm, B., Gleißner, M., Grzeszik, R., Arndt, S., Dreizler, A.: Flow field measurements in an optically accessible, direct-injection spray-guided internal combustion engine using high-speed PIV, Experiments in Fluids, Vol. 48, p. 281-290, (2010), doi: 10.1007/s00348-009-0742-2
(16) Wakisaka, T., Hamamoto, Y., Kinoshita, S.: Turbulence Characteristics in Internal Combustion Engines, Transactions of the Japan Society of Mechanical Engineers Series B, Vol. 48, Issue 430, p. 1198-1205, (1982) (in Japanese), doi: 10.1299/kikaib.48.1198
(17) Hamamoto, Y., Tomita, E., Miwa, H.: Measurement of Turbulence of Air Swirl in a 4-Stroke Cycle Engine Cylinder, Transactions of the Japan Society of Mechanical Engineers Series B, Vol. 53, Issue 491, p. 2226-2232, (1987) (in Japanese), doi: 10.1299/kikaib.53.2226
(18) Kamimoto, T., Yagita, M., Moriyoshi, Y., Kobayashi, H., Morita, H.: AN EXPERIMENTAL STUDY OF IN-CYLINDER AIR FLOW WITH A TRANSPARENT CYLINDER ENGINE, Transactions of the Japan Society of Mechanical Engineers Series B, Vol. 53, Issue 492, p. 2686-2693, (1987) (in Japanese), doi: 10.1299/kikaib.53.2686
(19) Ohtani, H., Moriyoshi, Y., Yagita, M., Kamimoto, T.: The Effect of Swirl on the Decay and Generation of In-Cylinder Turbulence during the Compression Stroke: LDV Measurements of Turbulence Length Scales Using the Spatial Correlation Method, Transactions of the Japan Society of Mechanical Engineers Series B, Vol. 56, Issue 530, p. 3173-3180, (1990) (in Japanese), doi: 10.1299/kikaib.56.3173
(20) Okura, Y., Higuchi, K., Urata, Y., Someya, S., Tanahashi, M.: Measurement of In-Cylinder Turbulence in An Internal Combustion Engine Using High Speed Particle Image Velocimetry, Transactions of the Japan Society of Mechanical Engineers Series B, Vol. 79, Issue 806, p. 2193-2206, (2013) (in Japanese), doi: 10.1299/kikaib.79.2193
(21) Hino, M.: Risanka ni tomonau gosa, Supekutoru kaiseki, Asakura publishing, p. 175-177, (1997) (in Japanese)
(22) Hino, M.: Jikosoukankansuu, Spekutoru kaiseki, Asakura publishing, p. 25-39, (1997) (in Japanese)
(23) Peters, N.: Regimes in Premixed Turbulent Combustion, Turbulent Combustion, Fourth printing with corrections, CAMBIDGE UNIVERSITY PRESS, p. 78-86, (2006)
(24) Ohashi, T., Yang, X., Takabayashi, T., Urata, Y., Kubota, S., Katsuyama, H.: Ignition and Combustion Simulation in HCCI Engines, SAE technical paper 2006-01-1522, (2006), doi: 10.4271/2006-01-1522
(25) Ohe, S.: Kitai no nendo no suisanhou, Busseisuisanhou, Databook shuppansha, p. 228-236, (2002) (in Japanese)
(26) Tanahashi, M.: Turbulent Premixed Combustion I — Reference Scales of Turbulence and Premixed Flame —, Journal of the Combustion Society of Japan, Vol. 51, Issue 158, p. 295-302, (2009) (In Japanese), doi: 10.20619/jcombsj.51.158_295
(27) FLOWTECH RESEARCH: https://www.ft-r.jp/soft/cruise.html, (accessed 2024/07/18) (in Japanese)
(28) Kasajima, Y., Egawa, T., Ushio, N., Kondo, T., Yamaguchi, R., Ikeya, K.: Development of 2.0L Inline-4-Cylinder Gasoline Direct Injection Engine, 2022 JSAE Annual Congress (Autumn) Proceedings, No. 111-22, 20226128, (2022) (in Japanese)

List of Symbols

E	Mean square of u and v velocity components [m^2/s^2]
E_u	Energy of u velocity component [m^2/s^2]
f	Frequency [Hz]
f_c	Cut-off frequency [Hz]
f_{c_slope}	Cut-off frequency at where spectrum slope changes [Hz]
$f_{c\_\theta=270}$	Cut-off frequency which was calculated from

	inverse of integral time scale of the in-cylinder flow at crank angle 270 deg. BTDC [Hz]
f_N	Nyquist frequency which is half frequency of time resolution [Hz]
G_{Plate}	Gap from tip of tumble plate to intake valve stem [mm]
h_p	Height of combustion chamber from top of pent-roof to piston surface top at bore center [mm]
k	Turbulent kinetic energy [m^2/s^2] (Subscripts f: time filtering method, c: cycle average method)
l	Turbulent integral length scale [mm]
l_F	Flame thickness [mm]
L_{Plate}	Length of tumble plate [mm]
Ne	Engine speed [rpm]
P_{cyl}	In-cylinder pressure [kPa]
P_{intake}	Intake manifold pressure [kPa]
R_{uu}	Auto correlation coefficient of u velocity component
S_L	Laminar burning velocity [m/s]
T_{cyl}	In-cylinder gas temperature [K]
u	Instantaneous velocity component in x direction [m/s]
u'	Turbulent intensity [m/s]
u'_f	Turbulent velocity component in x direction decomposed by time filtering method [m/s]
u_p	Instantaneous piston speed [m/s]
v	Instantaneous velocity component in y direction [m/s]
v'_f	Turbulent velocity component in y direction decomposed by time filtering method [m/s]
x	Coordinate to bore direction [mm]
y	Coordinate to stroke direction [mm]
θ	Crank angle [deg.]
θ_{ig}	Ignition timing [deg. BTDC]
v_F	Dynamic viscosity of mixture [m^2/s]
τ	Integral time scale of in-cylinder flow [ms] (Subscripts u: x direction, v: y direction)
τ_l	Turbulent integral time scale [ms]
τ_p	Characteristic time of piston speed [ms]
φ	Equivalent ratio [-]

List of Abbreviations

BDC	Bottom Dead Center
CFD	Computational Fluid Dynamics
e-fuel	Erneuerbarer strom - fuel
EGR	Exhaust Gas Recirculation
fps	frame per second
GUI	Graphical User Interface
IMEP.H	Indicated Mean Effective Pressure High-loop
MFB	Mass Fraction Burnt
LF	Lower Flow
PIV	Particle Image Velocimetry
PRF	Primary Reference Fuel
SI	Spark Ignition
TCV	Tumble Control Valve
TDC	Top Dead Center (Before TDC: BTDC)
UF	Upper Flow
WOT	Wide Open Throttle

Author


Yasuhiro OKURA


Makoto SEGAWA


Hiromi ONIMARU


Seisuke SATO


Yasuhiro URATA

Numerical Study on Enhancing Power and Thermal Efficiency of Large Motorcycle Gasoline Engine Using Pre-chamber Jet Combustion Simulated by Combustion CFD

Hirokazu ANDO[1] Takumi TANAKA[1]
Kengo TOMIZAWA[1] Yosuke INOUE[1]

ABSTRACT

A study of combustion methods was conducted using 3D combustion simulation with the aim of enhancing power at full load and thermal efficiency at partial load for a big-bore spark ignition gasoline engine for large motorcycles. The effect of passive pre-chamber jet combustion on power was confirmed at full load. It was further confirmed that a jet sprayed from the pre-chamber into the main chamber caused an increase in turbulent kinetic energy that sped up combustion, and that increasing the compression ratio from 10.1 to 12.1 resulted in an enhancement of 3.6% in indicated work at full load compared to conventional SI combustion. At partial load, the effect of pre-chamber jet 2plug combustion, achieved by ignition in the pre-chamber after the ignition of the main-chamber side-plug with the aim of enhancing combustion stability, was confirmed. The jet sprayed from the pre-chamber facilitates ignition by coming into contact with the flame surface generated by the earlier ignition of the main-chamber side-plug. It was confirmed that this resulted in an increase in indicated thermal efficiency of 1.7 points at partial load compared to conventional SI combustion.

1. Introduction

In order to control global warming, under a scenario in which increases in temperature caused by climate change are kept below 1.5ºC, it will be necessary[(1)] to reduce emissions of the greenhouse gas CO_2 by at least 45% from 2010 levels by 2030. The transportation sector accounts for approximately 20% of total domestic CO_2 emissions, and although motorcycles account for only 0.4% of that, this amounts to CO_2 emissions of 780,000 tons[(2)], and it is important that enhancements are also made to internal combustion engines for motorcycles in order to reduce CO_2.

In addition to the demand to enhance riding performance by increasing maximum power at full load, spark ignition (SI) gasoline engines for large motorcycles are also being required to reduce CO_2 by improving fuel economy through enhancements in thermal efficiency at partial load. Here, the relationship between maximum engine power and fuel economy for the mass-produced motorcycle models of Honda is presented. Figure 1 shows the relationship between maximum engine power and fuel economy in Worldwide-harmonized Motorcycle Test Cycle (WMTC) mode. The graph plots data for motorcycles in WMTC mode class 3-2, which applies to those models with a top speed of at least 140 km/h. As shown by the graph, maximum power and mode fuel economy are in conflict with each other. Increases in volumetric efficiency and higher engine speed are generally used to enhance maximum power. Increasing volumetric efficiency requires intake and exhaust resistance to be lowered, and it is essential to increase the diameter of the intake and exhaust valves, and to increase the bore diameter to permit this. Since increasing engine speed raises piston speed, it is important to shorten the stroke to keep piston speed under control. Accordingly, the tendency for an engine for which the aim is to enhance maximum engine power is to specify a big bore and a short stroke. However, while specifying a big bore and a short stroke is effective in increasing maximum power, at partial load this results in noticeable reductions in both intake air speed and tumble flows formed in the cylinder. This leads to insufficient air-fuel mixture formation and reduced turbulence at combustion timing, resulting in reductions in combustion speed. These factors cause a lowering in thermal efficiency at partial load. This means that enhancing both maximum

[1] Innovative Research Excellence, Power Unit & Energy
E-mail: hirokazu_ando@jp.honda

power and partial-load thermal efficiency in a big-bore, short-stroke engine requires increases in combustion speed to be achieved by techniques other than in-cylinder flow.

The relationship between displacement per cylinder and bore for motorcycle models currently in production at Honda is shown in Fig. 2. As shown in the graph, this research targeted motorcycle models currently in production with a bore diameter of 92 mm, which are classified as having big-bore engines. These models had achieved[3] increases in combustion speed by performing ignition at two plugs, one at the side and one in the center of each cylinder, and this research sought to realize further increases in combustion speed.

A passive pre-chamber was used in prior research examples[4, 5] with the aim of increasing combustion speed in SI engines for large motorcycles. A passive pre-chamber is one that is not equipped with a fuel injection nozzle within the pre-chamber and that draws in the air-fuel mixture from the main-chamber side. In the examples of prior research[4, 5], increases in full-load output power were achieved, but there were no reports of enhancements in partial-load thermal efficiency. In addition, although those research studies used 3D combustion simulation (hereinafter, "combustion CFD"), performance evaluations did not take into account the knocking that can be an issue with actual engines.

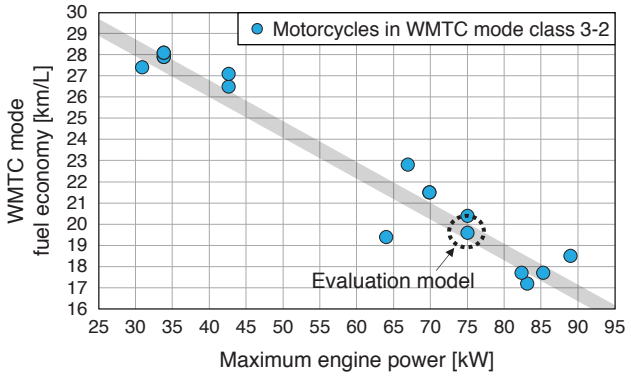


Fig. 1 Maximum engine power and fuel economy of motorcycles in WMTC mode class 3-2

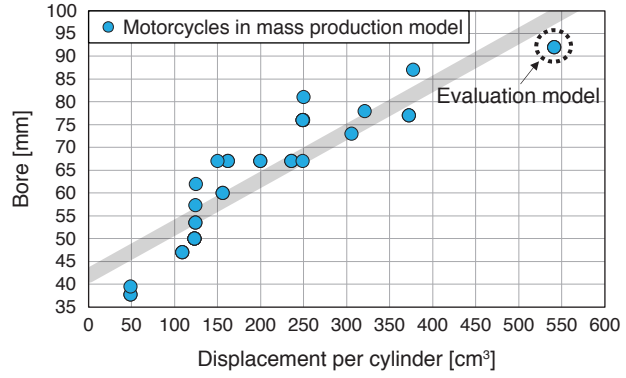


Fig. 2 Displacement per cylinder and bore of motorcycles in mass production model

In this research, a combustion CFD was built for SI engines used in large motorcycles to enable knocking detection through a detailed elementary reaction. Furthermore, combustion CFD was used to analyze the combustion mechanism of flame propagation combustion for a variety of approaches, such as conventional SI combustion, pre-chamber jet combustion, and pre-chamber jet 2plug combustion. Performance evaluations of maximum power and partial-load thermal efficiency were also conducted. This enabled the determination of combustion methods that simultaneously enable enhanced maximum power at full load and higher partial-load thermal efficiency for the reduction of CO_2 in conventional SI engines for large motorcycles, and the results are discussed in this paper.

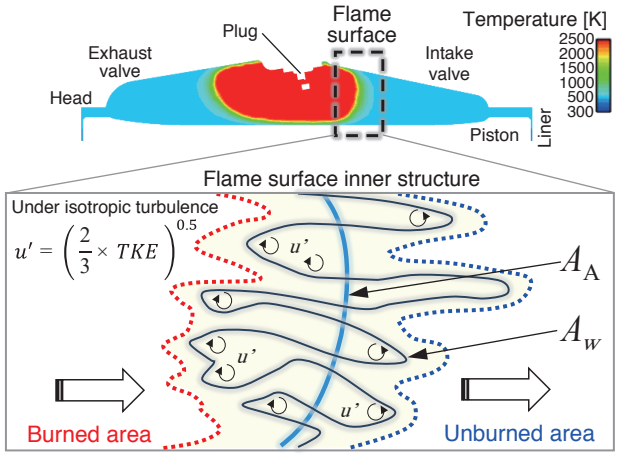
2. Issues with Large Motorcycle SI Combustion and Increases in Combustion Speed

When seeking to enhance thermal efficiency in an SI engine, it is generally thought necessary to increase theoretical efficiency by raising the compression ratio. Because this facilitates the occurrence of knocking, the ignition timing needs to be retarded more than the Minimum advanced for Best Torque (MBT), which increases time loss. Accordingly, to increase thermal efficiency, it is important to mitigate knocking and reduce time loss.

Knocking is closely related to auto-ignition of end gases[6], and because the distance to the bore-end gas portion is particularly large in big-bore SI engines, knocking is easy to induce. To mitigate knocking, it is necessary to complete combustion before auto-ignition occurs, and raising combustion speed is an effective means of achieving this[7, 8]. A summary of flame surface behavior in flame propagation combustion and mechanisms related to combustion speed is shown here in Fig. 3. According to the concept of wrinkled laminar flame proposed by Damköhler et al.[9-11], combustion speed is closely related to the surface area of the wrinkled laminar flame, which is amplified by turbulence. Consequently, Turbulent Kinetic Energy (TKE) within the combustion chamber needs to be increased to increase combustion speed. One method of increasing combustion speed in SI engines that has been attracting attention is pre-chamber jet combustion, in which a jet sprayed from the pre-chamber can amplify TKE on the main-chamber flame surface and increase the wrinkled laminar flame surface area to speed up combustion. This has the result of mitigating knocking, and enabling the compression ratio to be raised and thermal efficiency to be increased[7, 8].

In this research, the effect of pre-chamber jet combustion[7, 8] using a passive pre-chamber was confirmed. In the passive type of pre-chamber jet combustion, fuel is brought into the pre-chamber by drawing in air-fuel mixture from the main chamber, which means that the pre-chamber

can easily run short of fuel[12), 13)]. Moreover, because the temperature falls when burned gases in the pre-chamber pass through the hole, this creates issues for combustion stability at partial load in particular[12)]. For this reason, it is challenging to guarantee enhanced performance over a wide operating range from full load to partial load with a single combustion specification[13)], creating an issue for the popularization of pre-chamber jet combustion[12), 13)].



TKE : Turbulent kinetic energy
u' : Turbulence
A_A : Average flame surface area
A_w : Wrinkled laminar flame surface area

Fig. 3 Schematic diagram of wrinkled laminar flame

3. Combustion Analysis Techniques

3.1. Evaluation Load and Specifications

Figure 4 shows full-load power characteristics for the target models, with operating points in WMTC mode class 3-2 plotted by engine speed and power, and the bubble size indicating the frequency of occurrence.

Figure 4 also shows the evaluation load point conditions for this research. For the evaluation point for maximum power at full load, engine speed was set to 7500 rpm, throttle opening to 100%, and excess air ratio to 0.85. For the evaluation point for thermal efficiency at partial load, engine speed was set to 4000 rpm, which is frequently used in WMTC mode, throttle opening to 26%, and excess air ratio to 1.0.

Figure 5 shows schematic diagrams of the combustion chamber and specifications of the engine used for analysis with the aim of increasing combustion speed. Figure 5(a) shows the parallel twin SI engine for large motorcycles (hereinafter, "conventional SI combustion") that forms the basis of this study. This engine has two plugs, center and side, for each cylinder, providing simultaneous ignition as well as phased ignition timing control[3)]. Figure 5(b) is based on conventional SI combustion but specified with a passive pre-chamber to replace the center plug, with ignition performed only in the pre-chamber (pre-chamber jet combustion). Figure 5(c) shows a pre-chamber jet combustion specification with an additional side plug between the exhaust valves, with ignition performed both in the pre-chamber and by the side plug (pre-chamber jet 2plug combustion). The pre-chamber specifications used in this research and shown in Fig. 5(d) refer to those obtained by the authors in previous research[7), 8)] that had the effect of enhancing thermal efficiency through knocking mitigation, and were set to the same values of 10 holes and a hole diameter of 0.5 mm. On the other hand, because the bore diameter was increased from 73 mm in the previous research to 92 mm for this research, the volume was increased from 0.5 cm^3 to 0.8 cm^3 to increase the reach of the jet.

3.2. Computational Techniques and Reproducibility in Actual Engine

The combustion CFD software used in this research was Convergent Science's CONVERGE v3.0.19, and an overview of the model is provided in Table 1. The flame propagation combustion model used in this research is a detailed chemical reaction (SAGE)[14)], which reproduces combustion through the interaction of detailed elementary reactions and turbulent diffusion. While this has a high calculation load, because the direct combustion chemical reaction is solved using an elementary reaction, it is likely that the dependence on the model coefficients can be reduced even in cases where the specifications and forms of combustion are changed. Furthermore, the use of an elementary reaction calculation enables the reproduction of knocking caused by auto-ignition phenomena driven by low-temperature oxidation reactions, which has been reported in examples of previous research in which it was applied to SI combustion and pre-chamber combustion[4), 15), 16)]. In addition, an SIP-Gr2.0-S5R[17), 18)] model that has been confirmed to reproduce regular gasoline in elementary reactions with high accuracy was used for the fuel elementary reaction formula.

Using this combustion model, the combustion CFD was compared to an actual engine. Figure 6 shows a comparison of combustion behavior between the combustion CFD and

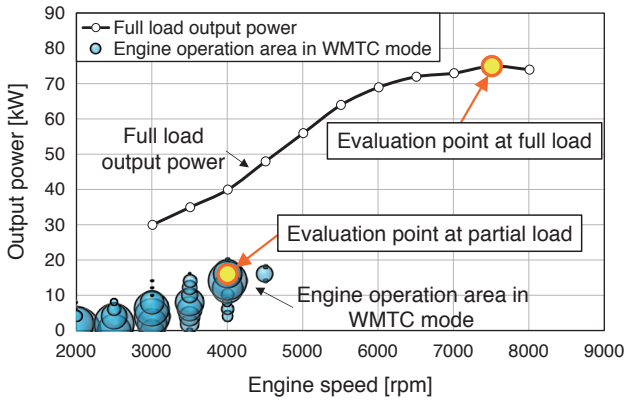


Fig. 4 Schematic diagram of evaluation load conditions

the actual engine, for conventional SI combustion under test specifications at the evaluation point for maximum power at full load and at the evaluation point for thermal efficiency at partial load. For combustion CFD using identical model coefficients, it was confirmed that in-cylinder pressure and heat release rate under full and partial load were within the range of cyclic variation in an actual engine.

3.3. Knocking Detection Techniques and Flame Surface Analysis Techniques

During the analysis of a variety of flame propagation combustion types, techniques using combustion CFD for knocking detection were studied. In previous research, it was concluded that knocking detection can be performed[19), (20)] by using the absolute minimum derivative values for heat release rate shown in equation (1) to detect sharp declines in heat release rate caused by auto-ignition. In this research, the value in equation (1) is referred to as the Knocking Detection Index (KDI) and knocking detection was verified in combustion CFD.

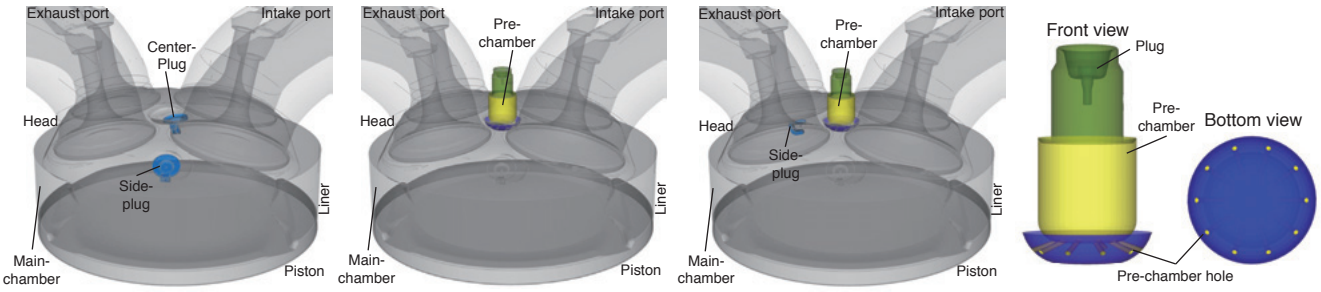
$$Knocking\ detection\ index\ (KDI) = \left| min \frac{d}{d\theta}\left(\frac{dQ}{d\theta}\right)\right| \quad (1)$$

Here, θ is the crank angle [deg.], and Q is heat release [J].

Table 1 Calculation software and models

CFD software	CONVERGE v3.0.19
Turbulence model	RANS ζ-f
Spray model	Rosin-Rammler and Rayleigh-Taylor
Combustion model	SAGE detailed chemical reaction
Ignition model	Add energy source to ignition point
Fuel chemical reaction mechanism	Regular gasoline SIP-Gr2.0-S5R
Mesh size [mm]	0.125-1.0 (adaptive mesh refinement)
Model parameter	Turbulent schmidt number: 0.5 Reaction multiplier: 1.15

Figure 7 shows combustion CFD heat release rates and derivative values for heat release rates under conditions in which auto-ignition occurs for conventional SI combustion at the evaluation point for maximum power at full load. At the time that auto-ignition occurred, it was observed that the heat release rate fell back after first rising, and the derivative values of the heat release rate fell sharply immediately after auto-ignition compared to before auto-ignition. Furthermore, Fig. 8 shows KDI values when changes were made to ignition timing. When combustion center of gravity timing (Mass Fraction Burned 50%: MFB 50%) was advanced gradually from 20 deg. ATDC, the KDI



(a) Conventional SI combustion ε10.1 (b) Pre-chamber jet combustion ε12.1 (c) Pre-chamber jet 2plug combustion ε12.1 (for side-plug between exhaust valves) (d) Schematic diagram of passive pre-chamber

	(a) Conventional SI combustion ε10.1	(b) Pre-chamber jet combustion ε12.1	(c) Pre-chamber jet 2plug combustion ε12.1 (for side-plug between exhaust valves)
Displacement	1082 [cm³]		
Number of cylinders	2		
Bore	92 [mm]		
Stroke	81.4 [mm]		
Compression ratio: ε	10.1	12.1	
Maximum power	75 [kW] / 7500 [rpm] (base)		
Valve system	SOHC 4-valve		
Fuel injection system	Port injection (injection pressure: 343 [kPa])		
Fuel	Japanese regular gasoline (research octane number: 91)		
Ignition system	2plug: Center-plug: conventional Side-plug: conventional	1plug: Center-plug: pre-chamber Side-plug: no-ignition	2plug: Center-plug: pre-chamber Side-plug: conventional (side-plug placement between exhaust valves)
(d) Pre-chamber specifications	-	Hole number: 10	
		Hole diameter: 0.5 [mm]	
		Volume: 0.8 [cm³]	

Fig. 5 Schematic diagrams and specifications of engine

value increased at a constant slope, with the advanced side sloping from MFB 50% at 11 deg. ATDC. This shows that from MFB 50% at 11 deg. ATDC, a spike in heat release occurs due to auto-ignition combustion on the advanced side and the KDI value increases. This technique is interpreted as resulting in clear changes in combustion behavior before and after auto-ignition and, as shown in Fig. 8, calculating the intersection enables detection of knocking limit MFB 50%. From this point, these techniques were applied after taking knocking into account, and performance evaluations were conducted for each specification.

In addition, when analyzing flame surface behavior, because it has the characteristic of peaking at the flame front of excessive oxidants of HO_2 radicals and the like, techniques in which it is used as a marker of flame surface position have been proposed[15), (21)]. In those research studies, the flame surface position was observed using a similar technique.

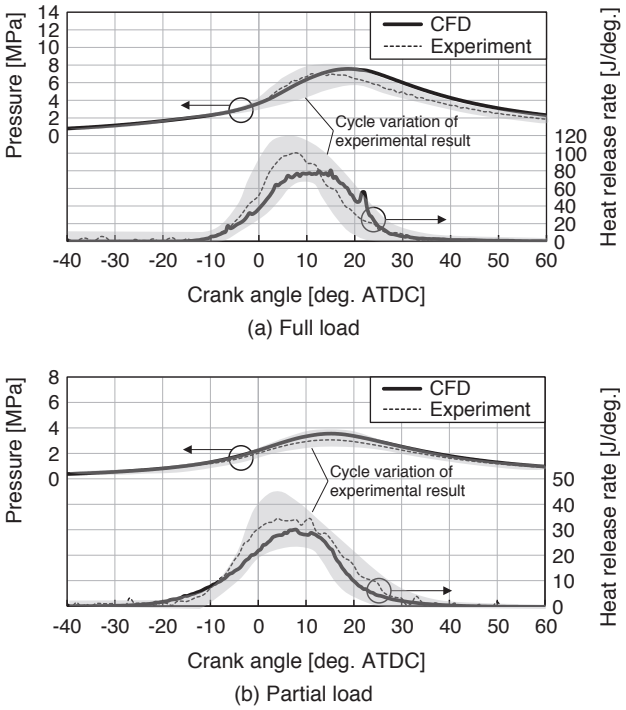


Fig. 6 Comparison of in-cylinder pressure and heat release rate between experiment and CFD

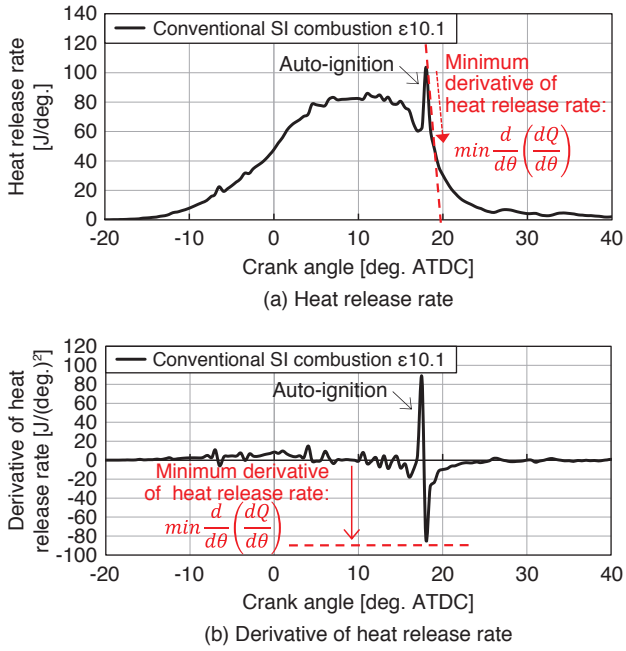


Fig. 7 Heat release rate and its derivative under auto-ignition condition

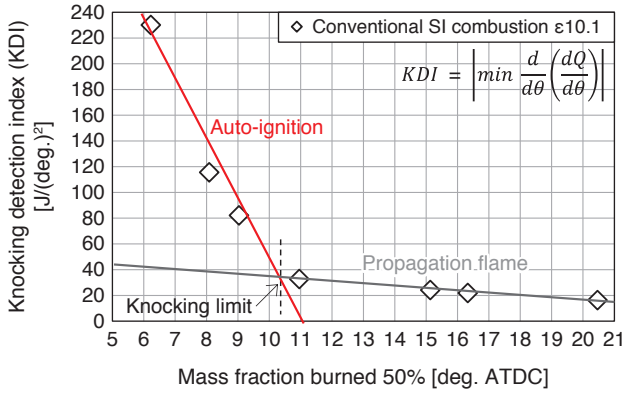


Fig. 8 Knocking judgement using knocking detection index

4. Results of Combustion CFD Analysis at Full-load Power

The specifications used for analysis at full-load power were the conventional SI combustion of Fig. 5(a) and the pre-chamber jet combustion of Fig. 5(b). In previous research by the authors, pre-chamber jet combustion led to increased combustion speed that had the effect of mitigating knocking, enabling an increase in compression ratio of 1.8, and achieving enhancements in thermal efficiency[7), (8)]. This research references those results, specifying that the compression ratio was increased from 10.1 in conventional SI combustion to 12.1. In addition, for conventional SI combustion, the phased ignition of the side plug from the center plug was set to 4 deg., the same as for a mass-produced model.

Figure 9 shows the effects of changes in ignition timing on knocking detection using KDI values under the respective specifications. For both conventional SI combustion and pre-chamber jet combustion, because advances in MFB 50% timing resulted in auto-ignition occurring and significant changes in the KDI, this was used to perform knocking detection. For pre-chamber jet combustion, the slope of changes in the KDI value was

significant, and the knocking limit MFB 50% timing was retarded by 2.4 deg.

Figure 10 shows the results of analysis of combustion duration and indicated work based on changes in ignition timing, with MFB 50% timing on the horizontal axis. The conditions under which maximum indicated work was obtained are shown by the line with circle markers. For this research, combustion duration was defined as combustion duration required for Mass Fraction Burned to progress from 10% to 85%.

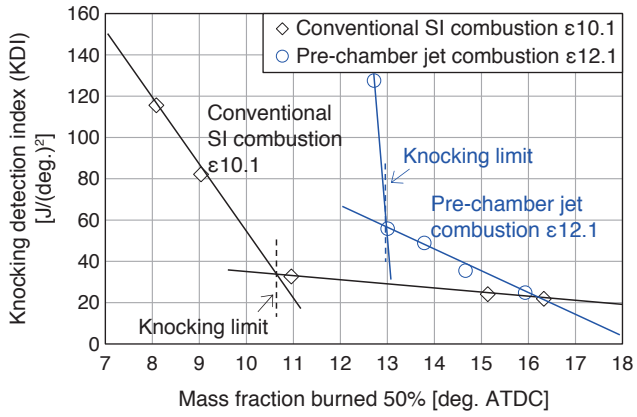
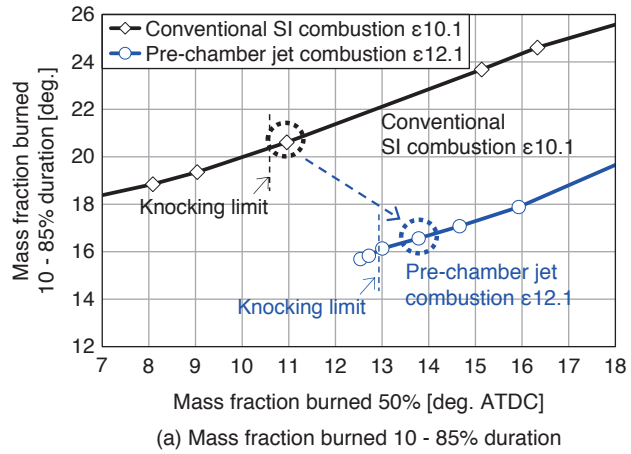
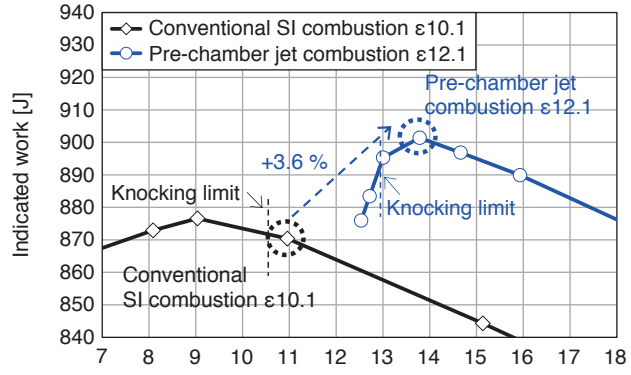


Fig. 9 Knocking limit comparison using knocking detection index



(a) Mass fraction burned 10 - 85% duration



(b) Indicated work

Fig. 10 Combustion duration and indicated work

In Fig. 10(a), when compared at the same MFB 50% timing of 14 deg. ATDC, combustion duration for pre-chamber jet combustion was shortened by 6.2 deg. Furthermore, in Fig. 10(b), indicated work for pre-chamber jet combustion increased by 5.9% when compared at the same MFB 50% timing of 14 deg. ATDC. In addition, for conventional SI combustion, advancing ignition timing to the knocking limit obtained the maximum indicated work. By comparison, pre-chamber jet combustion maximum indicated work was obtained before ignition timing at the knocking limit. When compared under the conditions at which maximum indicated work was obtained, not only was combustion duration for pre-chamber jet combustion shortened by 4.1 deg., but higher theoretical thermal efficiency caused by increases in the compression ratio also achieved enhancements in indicated work at full load of 3.6%.

Figure 11 shows in-cylinder pressure and heat release rate under the conditions at which maximum indicated work was obtained. The graph confirms that, compared to conventional SI combustion, pre-chamber jet combustion has high heat release rates and rapid combustion at or after TDC. In addition, it was confirmed that, as shown in Fig. 12 by the temperature iso-surface around adiabatic flame temperature and the cross-sections of the HO_2 mass fraction, although for conventional SI combustion the flame surface advances gradually on the surface of a sphere centered on the spark plug, in pre-chamber jet combustion the jet sprayed into the main chamber causes the flame surface to spread outwards.

Furthermore, Fig. 13 shows average in-cylinder TKE behavior. In pre-chamber jet combustion, the sprayed jet increased turbulent kinetic energy in the main chamber. This resulted in rapid flame propagation combustion, enabling an increase in indicated work at full load.

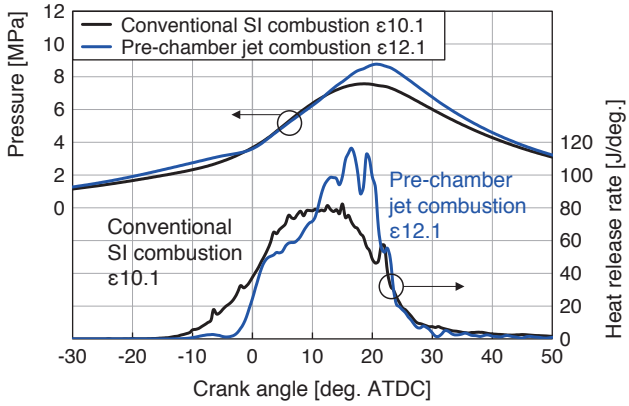


Fig. 11 In-cylinder pressure and heat release rate

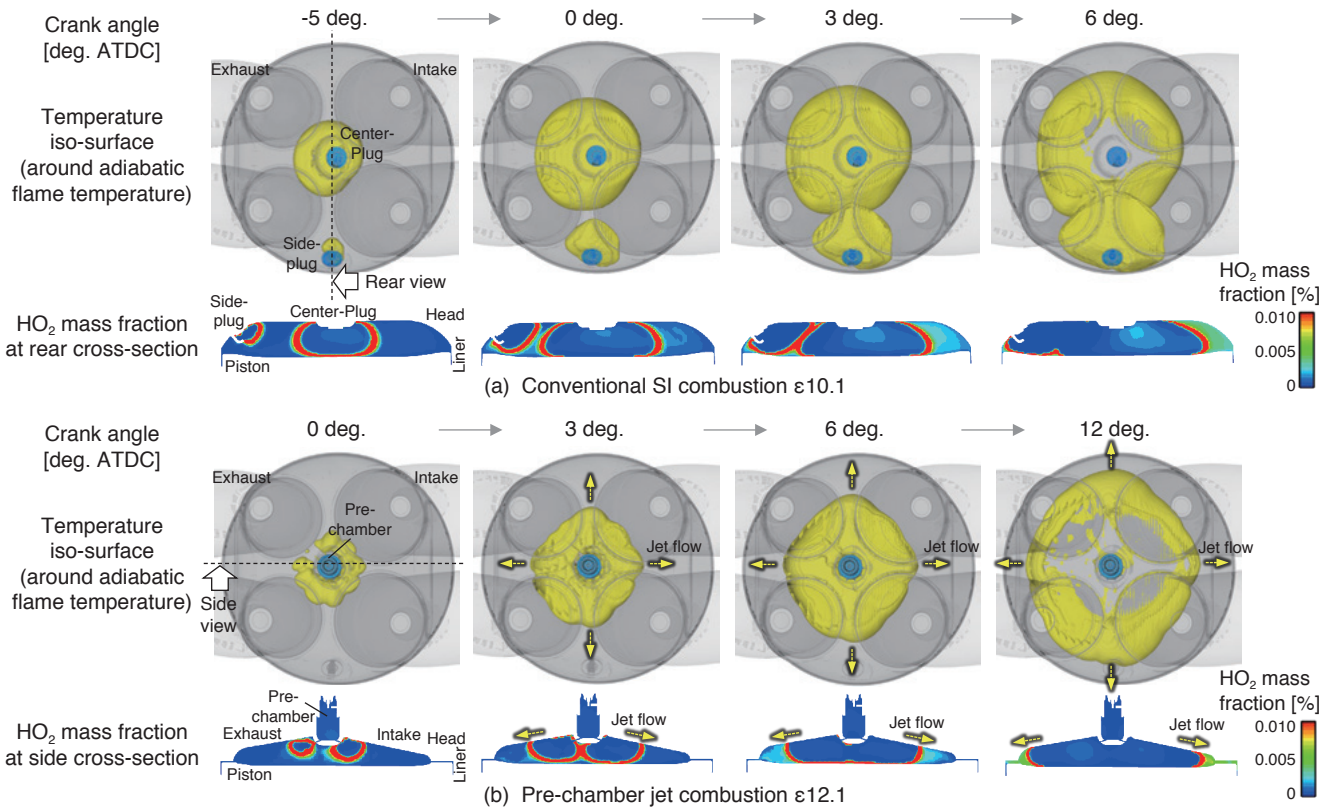


Fig. 12 Flame surface behavior

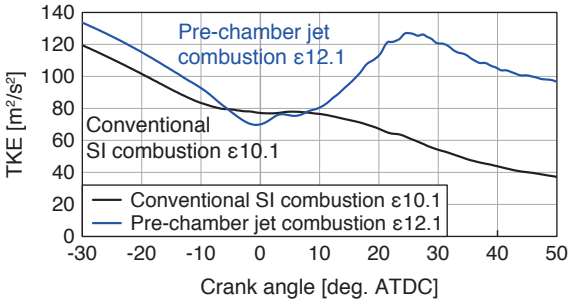


Fig. 13 In-cylinder turbulent kinetic energy

5. Results of CFD Analysis of Combustion at Partial Load Point

5.1. Pre-chamber jet combustion

Figure 14 shows the behavior of in-cylinder pressure and heat release rate when ignition timing is changed during pre-chamber jet combustion using ignition only in the pre-chamber, comprising the specification shown in Fig. 5(b). The graph confirms that for pre-chamber jet combustion at partial load, heat release rates were insignificant immediately after ignition for all ignition timings, and misfires subsequently occurred.

Here, the factors leading to misfires for pre-chamber jet combustion are analyzed. Figure 15 shows residual gas rates and their distribution in the pre-chamber directly before ignition, both under full-load conditions where ignition was possible and under partial-load conditions that led to misfires. As shown in Fig. 15(a), under partial-load conditions, residual gas in the pre-chamber increased from 12.8% to 23.5%. As shown in Fig. 15(b), residual CO_2 increased throughout the pre-chamber. Regarding the cause, it is thought that because intake flow rates and flows are reduced under partial-load conditions, this facilitates increases in residual gas in the main chamber, and also results in insufficient pre-chamber scavenging.

Furthermore, Fig. 16 shows gas temperatures at the inlets and outlets of the pre-chamber holes, and in addition to the previously mentioned reductions in combustion temperature caused by increases in residual gas under partial-load conditions, reductions in temperature caused by heat losses when passing through the holes also increased. Accordingly, while the maximum temperature for the sprayed jet under full-load conditions was 2480 K, under partial-load conditions it fell to 2170 K.

In addition, Fig. 17 compares main-chamber gas density and the distance reached by the tip of the jet originating in the hole outlet (hereinafter, “penetration”). Penetration was taken to be the distance reached by HO_2. As shown by Fig. 17(a), density in the main chamber at the time of jet spray, which occurred at -5 deg. ATDC, was 14.8 kg/m^3

under full-load conditions, while this decreased to 6.9 kg/m^3 under partial-load conditions of low volumetric efficiency. As can be seen from Fig. 17(b), it was confirmed that initial penetration was greater under partial-load conditions than it was under full-load conditions. This is thought to be due to the low gas density of the main chamber facilitating the jet spray.

Additionally, Fig. 18 shows in-cylinder temperature distributions after jet spray. Under full-load conditions, initial penetration was small, with high-temperature gas tending to accumulate in the area around the hole outlet, leading to ignition. Conversely, under partial-load conditions, the large initial penetration facilitated the blending of the jet with the air-fuel mixture in the main chamber immediately after jet spray, resulting in the temperature of the jet falling during the spreading of the jet through the main chamber, and in ignition not being achieved. In previous research that analyzed the factors leading to a decline in ignitability in pre-chamber jet combustion, it was shown that in the case of a strong jet, jet temperatures fall due to larger heat losses in the area of the holes[22]. The phenomenon of the jet spray facilitating the drawing-in of main-chamber premix leading to a reduction in jet temperatures and to ignition instability was also confirmed[22]. A similar phenomenon to that of previous research was also observed in this research, suggesting that reductions in ignitability were caused by these factors.

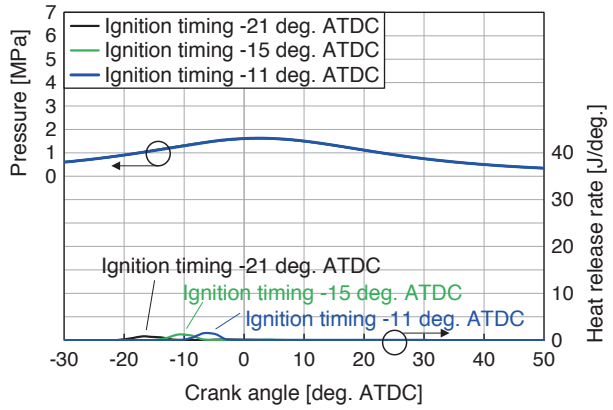
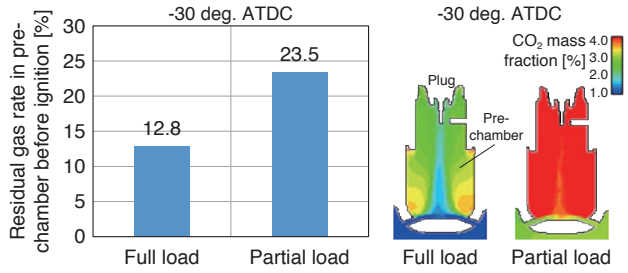
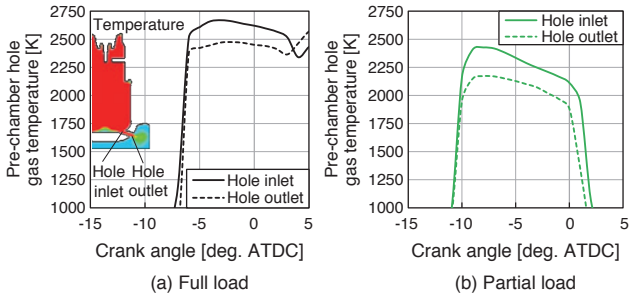


Fig. 14 In-cylinder pressure and heat release rate



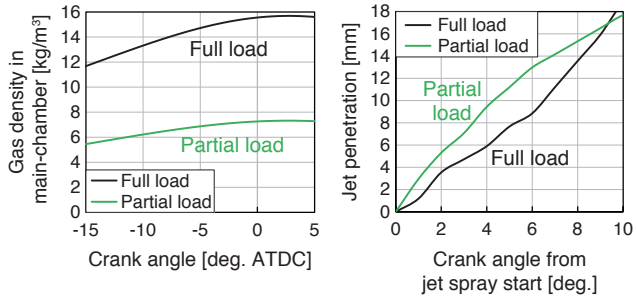
(a) Residual gas rate in pre-chamber before ignition (b) CO_2 mass fraction before ignition

Fig. 15 Residual gas rate and distribution in pre-chamber



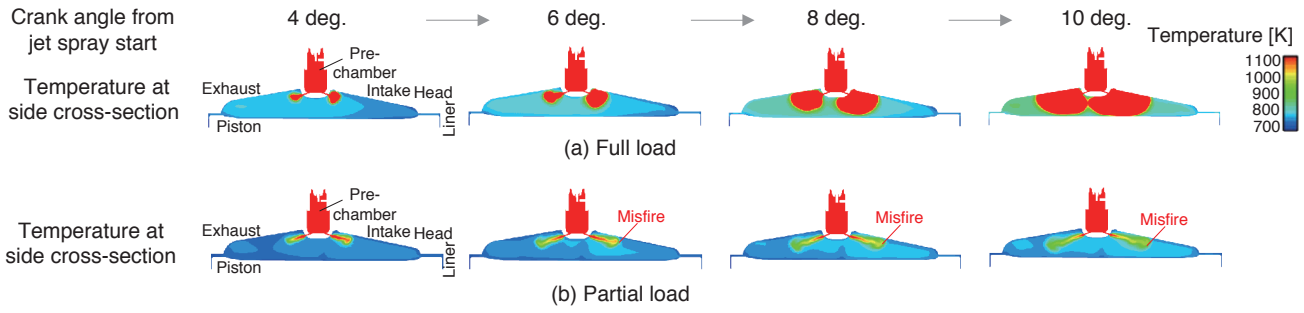
(a) Full load (b) Partial load

Fig. 16 Pre-chamber hole inlet and outlet gas temperature



(a) Gas density in main-chamber (b) Jet penetration in main-chamber

Fig. 17 Gas density and jet penetration in main-chamber



(a) Full load (b) Partial load

Fig. 18 Jet temperature behavior

5.2. Pre-chamber Jet 2plug Combustion

In the previous section, it was clearly shown that, for partial loads, pre-chamber jet combustion solely by pre-chamber ignition under these specifications caused jet temperatures to decline and resulted in misfires. To address this, techniques to facilitate ignition by first igniting the main-chamber side-plug and using main-chamber combustion to raise the temperature of in-cylinder gas, then performing pre-chamber ignition and spraying the jet were studied.

As shown in Fig. 5(c), the specification for this study involved the installation of a side-plug in the pre-chamber jet combustion main chamber. Regarding the location of the main-chamber side-plug at the bore end, there were concerns that, due to the use of main-chamber-first ignition, the unburned area at the opposite end of the bore would experience increased heating time, inducing knocking.

Accordingly, the performance evaluation of pre-chamber jet 2plug combustion was performed after changing the location of the main-chamber side-plug to a position between the exhaust valves, towards the center of the bore.

For phased ignition, conventional SI combustion was set to simultaneous ignition using a center-plug and a side-plug in the same way as for mass-produced models. For pre-chamber jet 2plug combustion, advance studies using phased ignition as the parameter were conducted, and phased ignition was set with the aim of facilitating jet combustion by initiating the jet spray at the point at which the flame resulting from the earlier ignition of the main-chamber side-plug reached the area of the pre-chamber. For the load evaluation of the targeted engine, pre-chamber plug ignition was set to take place at 10 deg. after ignition of the main-chamber side-plug.

The results of knocking detection when changing ignition timing for conventional SI combustion and pre-chamber jet 2plug combustion are shown in Fig. 19. Because the advance of MFB 50% timing resulted in significant changes in the KDI value for both, this was used to perform knocking detection. For pre-chamber jet 2plug combustion, the slope of changes in the KDI value was significant, and the knocking limit MFB 50% timing was retarded by 3.9 deg.

Figure 20 shows the results of performance evaluation at partial load point. Figures 20(a) and (b) show the results of analysis of combustion duration and indicated thermal efficiency, with MFB 50% timing on the horizontal axis. The conditions under which maximum indicated thermal efficiency was obtained for each specification are shown by the line with circle markers. When pre-chamber jet 2plug

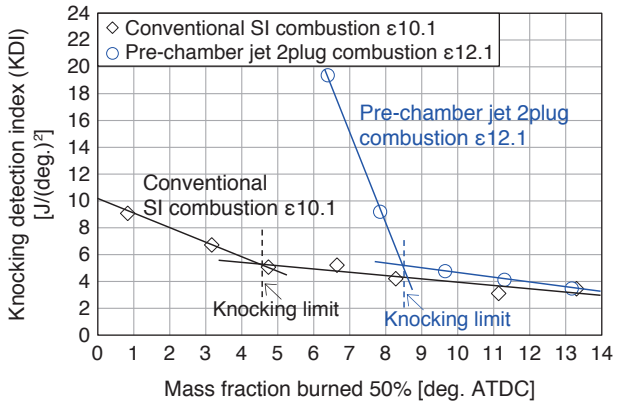


Fig. 19 Knocking limit comparison using knocking detection index

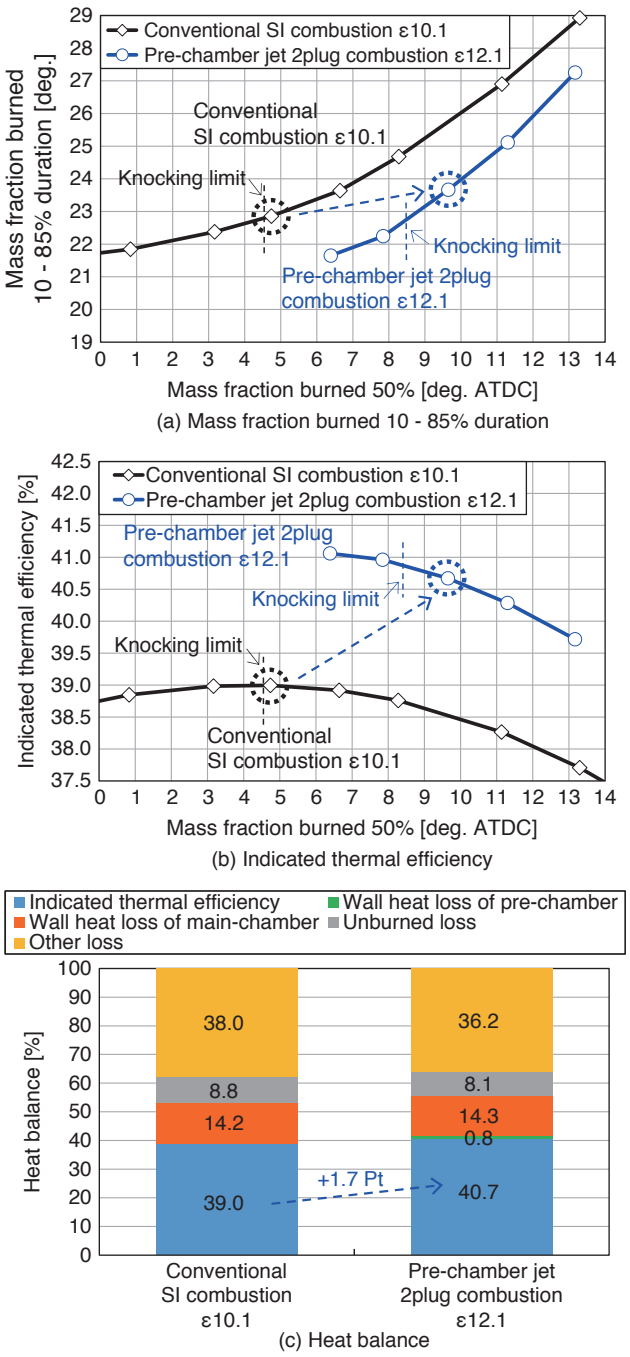


Fig. 20 Combustion duration, indicated thermal efficiency, and heat balance

combustion was compared to conventional SI combustion at the same MFB 50% timing of 10 deg. ATDC, it was confirmed that combustion duration was shortened by 2 deg. and indicated thermal efficiency was increased by 2.1 points.

Also, for the heat balance under the conditions in which maximum indicated thermal efficiency was obtained shown in Fig. 20(c), heat losses increased for pre-chamber jet 2plug combustion when compared to conventional SI combustion, but the effects of raising the compression ratio from 10.1 to 12.1 achieved a 1.7-point increase in indicated thermal efficiency.

Figure 21 shows the behavior of in-cylinder pressure and heat release rate under the conditions at which maximum indicated thermal efficiency was achieved. As shown in the graph, misfires do not occur in pre-chamber jet 2plug combustion, and it was confirmed that the ignition of the jet into the main chamber was followed by a sharp rise in heat release.

In addition, as shown in Fig. 22 by the adiabatic flame temperature and the cross-sections of the HO_2 mass fraction, it was confirmed that for partial loads in conventional SI combustion the flame surface advances on the surface of a sphere centered on the spark plug in the same way as for full loads. On the other hand, in pre-chamber jet 2plug combustion, it was confirmed that using pre-chamber ignition to spray the jet when flame propagation reaches the area of the pre-chamber following the earlier ignition in the main chamber resulted in strong flame propagation originating in the jet.

Furthermore, with a focus on behavior up to the point at which the jet ignites in the main chamber, Fig. 23 shows the distribution of in-cylinder temperatures. The figure confirms that the mechanism driving ignition is that of the jet sprayed from the pre-chamber contacting the high-temperature flame surface caused by the earlier ignition of the main-chamber side-plug.

In addition, Fig. 24 shows average in-cylinder TKE behavior. In conventional SI, in-cylinder TKE declined

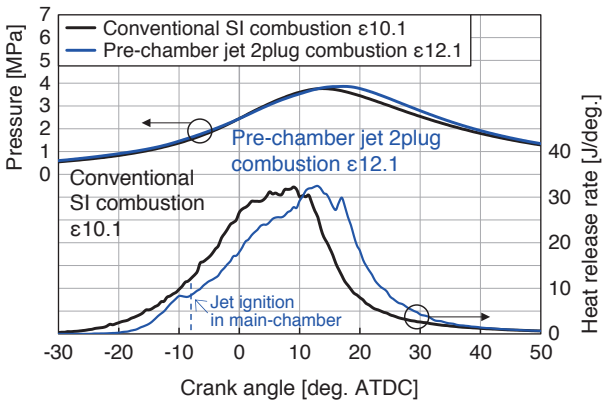


Fig. 21 In-cylinder pressure and heat release rate

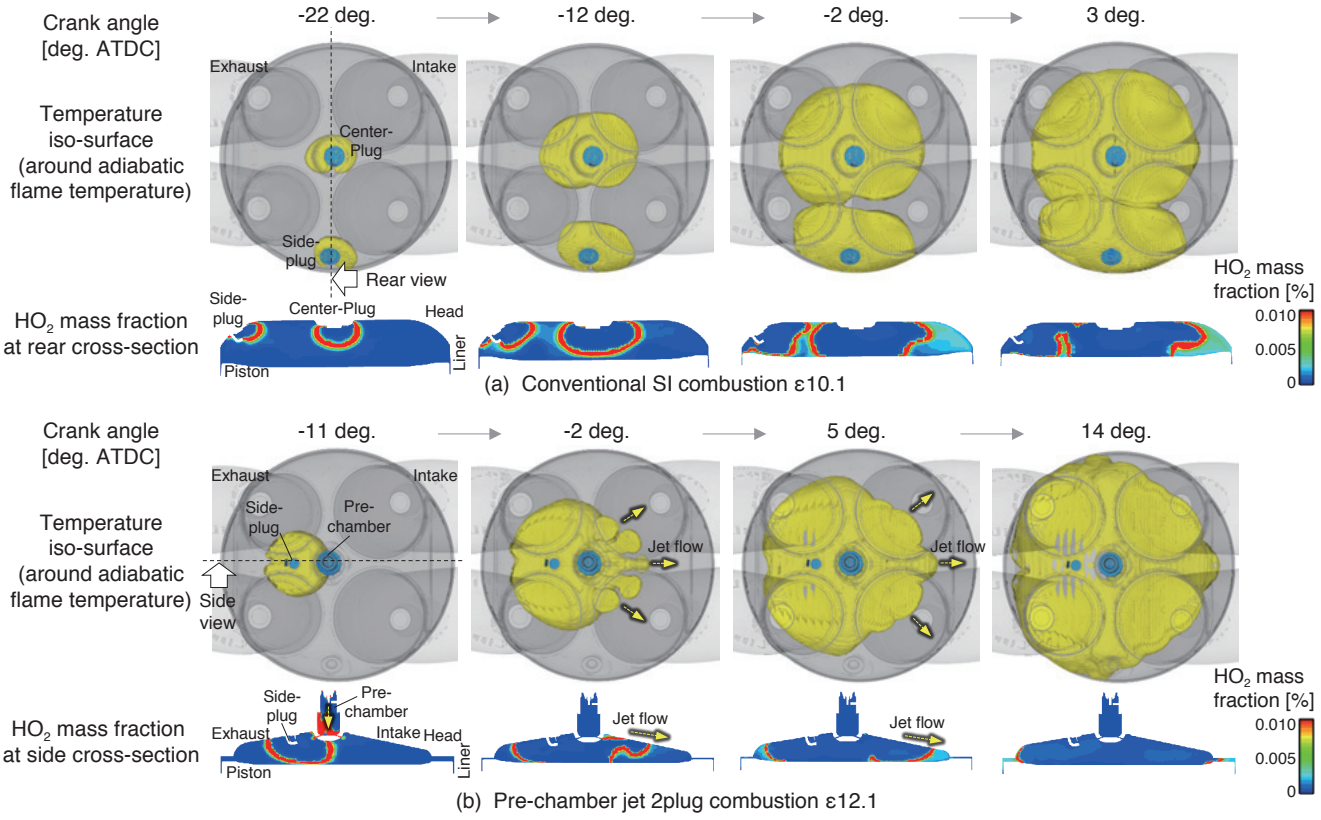


Fig. 22 Flame surface behavior

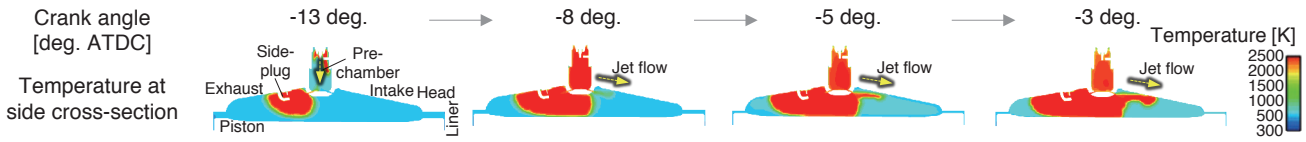


Fig. 23 Temperature behavior in pre-chamber jet 2plug combustion ε12.1

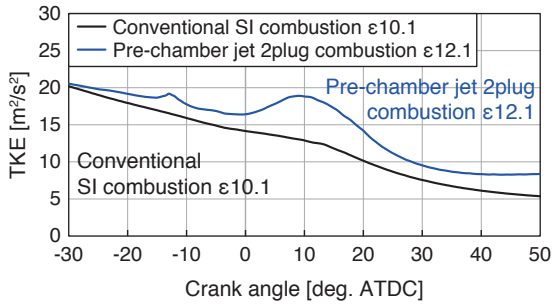


Fig. 24 In-cylinder turbulent kinetic energy

over time, while in pre-chamber jet 2plug combustion, it was confirmed that spraying the pre-chamber jet after main-chamber ignition led to higher in-cylinder TKE. This resulted in a sharp rise in heat release after the jet ignited, increasing indicated thermal efficiency.

6. Conclusion

A study of full-load maximum power and partial-load thermal efficiency in conventional SI engines for large motorcycles was conducted using combustion CFD incorporating knocking detection. The conclusions reached are as follows.

(1) Using detailed chemical reaction models, a combustion CFD was built that incorporates knocking detection techniques that use absolute minimum derivative values for heat release rate. This enabled performance evaluations of maximum power and evaluation of thermal efficiency related to engine combustion specifications to be conducted as a simulation-based study.

(2) At the evaluation point for maximum power at full load, in pre-chamber jet combustion, a jet sprayed from the pre-chamber increased turbulent kinetic energy in the main chamber and sped up combustion.

(3) At the evaluation point for maximum power at full load, the use of pre-chamber jet combustion reduced combustion duration and mitigated knocking in comparison to conventional SI combustion. This enabled the compression ratio to be raised from 10.1 to 12.1, shortening combustion duration by 4.1 deg. and achieving an enhancement of 3.6% in indicated work at full load.

(4) At the evaluation point for thermal efficiency at partial load, it was confirmed that pre-chamber jet combustion using ignition only in the pre-chamber helped prevent flame propagation from being maintained in the main chamber and resulted in misfires. This indicated that it was the result not only of combustion temperature declines caused by residual gas in the pre-chamber, and jet temperature declines caused by heat losses in the area of the holes, but also by temperature declines resulting from the spreading of the jet in the main chamber.

(5) At the evaluation point for thermal efficiency at partial load, for the specification in which compression ratio was raised from 10.1 to 12.1, the effects of pre-chamber jet 2plug combustion were confirmed by relocating the main-chamber side-plug to a position between the exhaust valves, towards the center of the bore, and igniting the pre-chamber after main-chamber side-plug ignition. It was confirmed that ignition of the jet that is sprayed from the pre-chamber is made possible by contact with the flame surface caused by earlier ignition in the main chamber, and that it facilitates combustion in the main chamber.

(6) At the evaluation point for thermal efficiency at partial load, it was confirmed that pre-chamber jet 2plug combustion resulted in an increase in indicated thermal efficiency of 1.7 points compared to conventional SI combustion.

References

(1) Intergovernmental Panel on Climate Change: Global Warming of 1.5℃ Presentation to the wrap-up of the Talanoa Dialogue preparatory phase, COP24 Presentation, (2018), https://apps.ipcc.ch/outreach/documents/446/1544108257.pdf, (accessed 2025/1/1)

(2) Ministry of Land, Infrastructure, Transport and Tourism, https://www.mlit.go.jp/sogoseisaku/environment/sosei_environment_tk_000007.html, (accessed 2025/1/1) (in Japanese)

(3) Morita, K., Okada, N., Okamura, H., Suzuki, S.: Development of 2020 CRF1100L Africa Twin, Honda R&D Technical Review, Vol. 32, No. 2, p. 17-26

(4) Bigalli, S., Catalani, I., Balduzzi, F., Matteazzi, N., Agostinelli, L., De Luca, M., Ferrara, G.: Numerical Investigation on the Performance of a 4-Stroke Engine with Different Passive Pre-Chamber Geometries Using a Detailed Chemistry Solver, Energies, Vol. 15,

No. 4968, (2022), doi: 10.3390/en15144968
(5) Nodi, A., Sforza, L., Lucchini, T., Onorati, A., Buttitta, M., Marmorini, L.: CFD Modeling of Conventional and Pre-Chamber Ignition of a High-Performance Naturally Aspirated Engine, SAE Technical Paper 2024-01-2102, (2024), doi: 10.4271/2024-01-2102
(6) Terashima, H.: Strong Pressure Wave Generations Induced by Inhomogeneous End-Gas Autoignition, Journal of the Combustion Society of Japan, Vol. 57, Issue 181, p. 206-213, (2015) (in Japanese), doi: 10.20619/jcombsj.57.181_206
(7) Ando, H., Shintani, Y., Kobayashi, H., Shiina, R., Kimura, N.: Knocking Mitigation of Gasoline Engine by Reduction of Unburned Area Using Pre-chamber Jet Combustion, Honda R&D Technical Review, Vol. 35, p. 46-56
(8) Ando, H., Shintani, Y., Kobayashi, H., Shiina, R., Kimura, N.: Study of Knocking Mitigation and Thermal Efficiency Enhancement of Pre-chamber Jet Combustion in Stoichiometric Gasoline Engine, SAE Technical Paper 2023-32-0006, (2023), doi: 10.4271/2023-32-0006
(9) Damköhler, G.: Der Einfluss der Turbulenz auf die Flammengeschwindigkeit in Gasgemischen, Z. f. Elektrochem, Vol. 46, Issue 11, p. 601-626, (1940) (in German), doi: 10.1002/bbpc.19400461102
(10) Glassman, I., Yetter, R. A.: COMBUSTION FOURTH EDITION, Academic Press, p. 231-235, (2008)
(11) Kobayashi, H.: Ranryuu to kaen no kanshou ni kansuru jikken to kousatsu, Journal of Japan Society of Fluid Mechanics, Vol. 16, Issue 5, p. 398-403, (1997) (in Japanese), doi: 10.11426/nagare1982.16.398
(12) Zhou, H., Meng, S., Han, Z.: Combustion characteristics and misfire mechanism of a passive pre-chamber direct-injection gasoline engine, Fuel, Vol. 352, No. 129067, (2023), doi: 10.1016/j.fuel.2023.129067
(13) Ono, R., Harada, Y., Nagatsu, K., Suzuki, H., Uchida. K., Iida, S., Fujikawa, T.: Study on the pre-chamber technology application to gasoline engine combustion, Proceedings of the 32nd Internal Combustion Engine Symposium, No. 81, (2021) (in Japanese)
(14) Senecal, P. K., Pomraning, E., Richards, K. J., Briggs, T. E., Choi, C. Y., MacDavid, R. M., Patterson, M. A.: Multi-Dimensional Modeling of Direct-Injection Diesel Spray Liquid Length and Flame Lift-off Length using CFD and Parallel Detailed Chemistry, SAE Technical Paper 2003-01-1043, (2003), doi: 10.4271/2003-01-1043
(15) Hokimoto, S., Kuboyama, T., Moriyoshi, Y., Yamada, T.: Combustion analysis in a natural gas engine with pre-chamber by three-dimensional numerical simulation, Transactions of the JSME, Vol. 81, Issue 830, p. 15-00154, (2015) (in Japanese), doi: 10.1299/transjsme.15-00154
(16) Maas, R., Bekdemir, C., Somers, B.: Numerical Study on the Design of a Passive Pre-Chamber for a Heavy-Duty Hydrogen Combustion Engine, SAE Technical Paper 2024-01-2112, (2024), doi: 10.4271/2024-01-2112
(17) Sakai, Y., Miyoshi, A.: Development of Reduced Chemical Kinetics Mechanism of Gasoline Surrogate Fuel, Proceedings of the 28th Internal Combustion Engine Symposium, No. 28, (2017) (in Japanese)
(18) Sakai, Y., Hasegawa, K., Miyoshi, A.: Development of Reduced Chemical Kinetics Mechanism of Gasoline Surrogate Fuel with Oxygenated Compounds, Proceedings of the 29th Internal Combustion Engine Symposium, No. 34, (2018) (in Japanese)
(19) Ando, H., Takemura, J., Koujina, E.: A Knock Anticipating Strategy Basing on the Real-Time Combustion Mode Analysis, SAE Technical Paper 890882, (1989), doi: 10.4271/890882
(20) Ando, H., Nishiyama, A., Wachi, Y., Kuwahara, K., Sakai, Y., Ohta, T.: Heat Release Rate and Cylinder Gas Pressure Oscillation in Low and High Speed Knock, SAE Technical Paper 2015-01-1880, (2015), doi: 10.4271/2015-01-1880
(21) Shao, J., Rutland, C.: Modeling Investigation of Auto-ignition and Engine Knock by HO_2, SAE Technical Paper 2014-01-1221, (2014), doi: 10.4271/2014-01-1221
(22) Ishii, T., Furukawa, S., Sakaida. S, Tanaka, K., Konno, M., Yoneya, N., Mitsufuji, K.: Study on effect of orifice diameter on turbulent jet by a visualized pre-chamber, Proceedings of the 32nd Internal Combustion Engine Symposium, No. 82 (2021) (in Japanese)

Author


Hirokazu ANDO


Takumi TANAKA


Kengo TOMIZAWA


Yosuke INOUE

Verification of HC Trap Materials and Catalysts to Reduce Exhaust Emissions during Cold-start in V6 Engine

Hiroki TAKEORI[1] Koji NEMOTO[2] Yuki OKU[1]
Ryohei IKUTOMO[1] Keita NAKAO[3] Toyohiro USUI[3]

ABSTRACT

In order to reduce emissions immediately after engine cold-start, which is a challenge for conventional three-way catalyst systems, an HC trap catalyst was developed that combines new HC trap materials and an HC purification catalyst for efficiently purifying the trapped HC. Ag-CHA and Cu-FAU were developed as zeolite HC trap materials that have pore sizes close to the HC molecular sizes and can adsorb HC up to higher temperatures by increasing the stability of the exchanged ions that promote chemisorption. It was also clarified that arranging the HC purification catalyst layer in the upper layer of the HC trap materials is key to effectively purifying the desorbed HC, and that adding an oxygen absorbing and releasing material consisting of CeO_2 and ZrO_2 in a 1:1 ratio to the HC purification catalyst enables purification of the desorbed HC even under stoichiometric conditions. A V6 engine with this HC trap catalyst in an underfloor arrangement was able to reduce HC emissions during engine cold-start by approximately 50% in the LA#4 mode emission test.

1. Introduction

Automotive exhaust gas regulations, which began in the United States in the 1960s, have been gradually tightened to the present, and this trend is expected to continue going forward. This has been addressed in gasoline engines by the invention and evolution of the three-way catalyst (TWC). However, as TWC does not manifest purification functions until reaching the light-off temperature (the temperature at which the purification rate reaches 90%), the exhaust gas is emitted without purification from immediately after engine cold-start until the TWC light-off temperature is reached. As it generally takes several tens of seconds for TWC to reach the light-off temperature, technology for purifying these emissions during cold-start is key for coping with ever tighter emission regulations.

Although lowering the TWC light-off temperature has been studied for many years[(1)-(4)], it is not easy and there is a limit to what TWC can do looking ahead to further tightening of regulations, so innovative technologies with different approaches will be needed. Numerous examples of such research have been reported, including an electrically heated catalyst (EHC) that promotes early light-off by using electric power to heat the catalyst before engine cold-start[(5)], and HC trap catalysts that adsorb (trap) the hydrocarbon (HC) components in exhaust gas in materials with pore sizes close to the HC molecular sizes such as zeolite at the low temperatures during engine cold-start, and then desorb and purify the HC after reaching the catalyst light-off temperature[(6)-(11)]. This HC trap catalyst (HCTC) consists of trap materials and an HC purification catalyst as shown in Fig. 1, and has a simple structure with the trap layers additionally coated on a TWC. This has the advantage of not requiring changes to the existing exhaust layout or installation of additional systems.

While there have been numerous reports on trap materials, the performance is insufficient, and further enhancement is needed in anticipation of tighter regulations going forward. The authors therefore developed an HCTC that can reduce HC emissions from an inline 4-cylinder 1.5 L turbocharged direct-injection (DI) engine by approximately 50% in the LA#4 mode. This was achieved

[1] Innovative Research Excellence
E-mail: hiroki_takeori@jp.honda
[2] Honda Motor Co., Ltd.
[3] Tosoh Corporation

by optimization of zeolites for the exchanged ions to maintain adsorption of the various sizes of HC contained in exhaust gas up to high temperatures[12]. However, large-displacement V6 engines have a higher exhaust flow rate than L4 engines, which increases the treatment amount per unit of time. This lowers the purification rate, and the slower rise in the exhaust gas temperature also lengthens the time until the light-off temperature is reached. As a result, emissions during cold-start increase and can account for 90% of total emissions[13]-[15]. Thus, while HCTC are expected to provide a high reduction effect immediately after cold-start, performance demands on HCTC are also increasing, such as an increased adsorption amount and a higher adsorption rate per unit time due to the greater flow rate, making it necessary to demonstrate the reduction effect of HC traps under such harsh environments. Therefore, this study verified the influence and effect of the HCTC arrangement in the exhaust layout of a V6 engine.

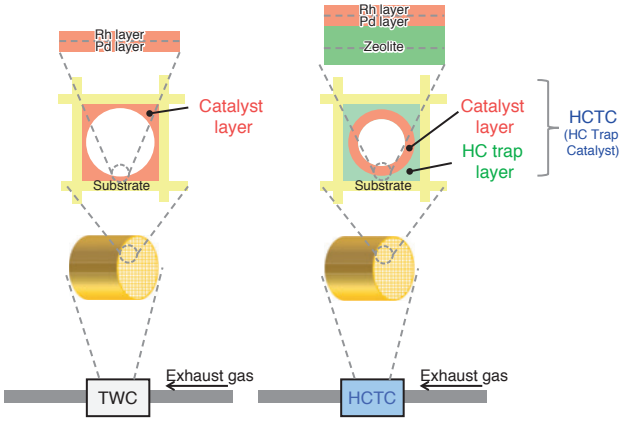


Fig. 1 HC trap catalyst configuration

2. HC Trap Goals

Consider the goals for increasing the emission reduction effect. First is to increase the adsorption amount in order to reduce the large emissions amount immediately after cold-start. Second is to purify the adsorbed HC without desorbing it as much as possible so as not to lessen the reduction effect. Third is to maintain the TWC purification performance other than immediately after cold-start at the same level as the TWC alone. Figure 2 shows the target purification characteristics of an HCTC that achieves these goals, and the following four points can be given as challenges.

(1) The HC contained in exhaust gas include three types with different chemical bonds and structures: paraffins, olefins, and aromatics. In addition, each type has carbon numbers ranging from 1 to 9, in other words a distribution of sizes, which calls for the efficient adsorption of these diverse HC.

(2) If the trapped HC are not retained until the purification start temperature of the co-used HC purification catalyst is reached, they may desorb and be emitted without purification, necessitating the ability to retain adsorbed HC until the purification start temperature or higher.

(3) The adsorbed HC are purified by the HC purification catalyst, but the trap performance or purification performance is greatly influenced by how the HC trap materials and the HC purification catalyst coexist. This makes it necessary to maximize the adsorption and purification performance by optimizing the balance between the arrangement and amounts of both.

(4) At the stoichiometric air-fuel ratio (A/F: stoichiometry), only the amount of oxygen needed for the three-way reaction is supplied, so there is zero oxygen after the TWC. This means that when the trapped HC desorbs, the condition with insufficient oxygen for purification results (A/F: rich), so some method is needed to supplement oxygen in order to maintain desorbed HC purification performance.

To address these issues, this study approached challenge (1) with the concept that the use of zeolites with pore sizes close to the HC molecular sizes is expected to increase the trapping probability at low temperatures and enhance the adsorption rate, making it important to select appropriate zeolite materials. An effective approach for challenge (2) was thought to be to enhance the adsorption force between the HC and the zeolite. Specifically, it was thought that desorption could be effectively suppressed by chemisorption, whereby ion species are introduced to the zeolite to enhance the electrostatic interactions with HC that have high electron densities such as those with double and triple C-C bonds (= aromatics and olefins). Effective approaches for challenges (3) and (4) were thought to be optimization of the HC trap material and HC purification catalyst layer arrangement and the catalyst coating amount to enhance HC purification performance, and the use of oxygen storage capacity (OSC) material with oxygen absorption and release functionality.

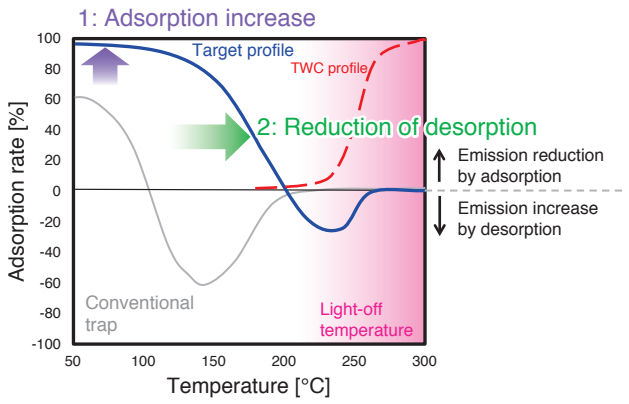


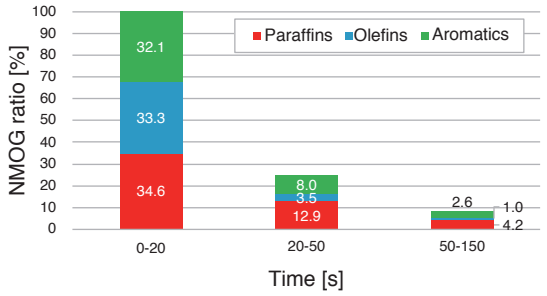
Fig. 2 Direction of enhancement of HC trap material

3. Trap Materials

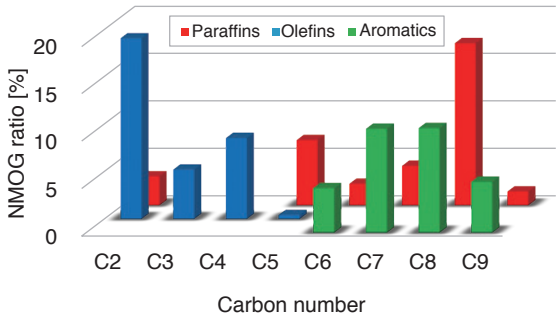
3.1. Analysis of HC Composition

First, in any approach to the trap materials, it is important to understand what kinds of HC are contained in the exhaust gas, so the exhaust gas during engine cold-start was sampled and gas chromatography was used to analyze the HC composition. The detailed method is as follows. A sampling pipe was branched at a position 30 mm from the rear end of the TWC of an inline 4-cylinder 1.5 L turbocharged DI engine, a Tedlar bag was connected a sampling box under depressurization by a pump, and exhaust gas was sampled in each interval of 0-20 seconds, 20-50 seconds, and 50-150 seconds. HC qualitative and quantitative analysis was performed by gas chromatography.

Figure 3(a) shows the transition in the composition of the HC in the exhaust gas after passing through a TWC relative to the time from cold-start, classified into the three types of paraffins, olefins, and aromatics. HC here refers to non-methane organic gases (NMOG), which is the current regulatory criterion in the U.S. state of California. Each value is the relative value of the concentration in each interval with the total NMOG concentration (ppmC) from 0-20 seconds set to 100%. This shows that the emissions for the 0-20 seconds interval are three times the total for the 20-150 seconds interval, which is to say the first 20 seconds accounts for over 70% of the total. Furthermore, Fig. 3(b) shows the emission ratios of each carbon number and HC species relative to the total concentration (ppmC) in the 0-20 seconds interval. Characteristics of a wide size distribution with carbon numbers ranging from 2 to 9, and high ratios of C2-C4 small molecule olefins, C7-C8 aromatics, and C8 paraffins were confirmed.



(a) Relative each group ratio for 0-150 seconds



(b) Relative ratio of carbon number and group for 0-20 seconds

Fig. 3 HC Composition in rear part of TWC

Based on these results, the direction for enhancing the adsorption amount can be classified in terms of molecular size into C2-C4 small molecule HC and C7-C8 large molecule HC. It was decided to efficiently enhance physical adsorption by selecting two zeolites with pore sizes that match the respective molecular sizes. Next, regarding the direction for suppressing HC desorption, there is a high ratio of HC species with double bonds, namely aromatics and olefins, so chemisorption using the aforementioned exchanged ions is considered effective[6), (10)]. However, exchanged ions have a general tendency to thermally degrade, so it is important to consider heat resistance in the selection.

3.2. Selection of Trap Materials (Zeolite)

Zeolites that effectively trap the analyzed HC were investigated. H-type FAU, BEA, MFI, and CHA zeolites made by Tosoh Corporation, as well as these materials ion-exchanged with Ag, Cu, Cr, Co, Ni, and Mo, were used as HC trap materials. Figure 4 shows the structure and pore size of each zeolite used[(16)]. Various zeolites are classified by their framework structure, and each is named with three uppercase letters. Adsorption performance and purification performance were evaluated using a model gas prepared by mixing the gases from various cylinders, etc. at the specified ratio. The model gas for evaluation was mixed by controlling the flow rates of the various cylinders and vaporized liquid HC (toluene, isooctane) to achieve the composition shown in Table 1, and then flowed into the samples as simulated exhaust gas. The gas flow rate was set to 25 L/min (equivalent to SV 50,000 h^{-1}), and the temperature rise rate was set to 40°C/min. The purification rate was calculated using Eq. (1), and the inflow HC concentration was set to 1200 ppmC.

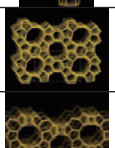
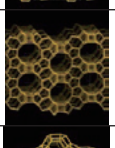
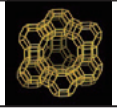
Zeolite	Structure	Pore size [nm]
CHA		0.37
MFI		0.47
BEA		0.60
FAU		0.74

Fig. 4 Structure and pore size of each zeolite (cited from Database of Zeolite Structures)

Table 1 Gas composition of model gas for evaluation

Gas	Concentration
Propylene / ppmC	348
Isopentane / ppmC	108
Toluene / ppmC	408
Isooctane / ppmC	336
NO / ppm	500
H_2 / %	0.17
CO / %	0.5
O_2 / %	0.49
CO_2 / %	14
H_2O / %	10
N_2	Balance

$$\text{Adsorption rate [\%]} = \left(1 - \frac{\textit{Catalyst exhaust HC concentration}}{\textit{Catalyst inflow HC concentration}}\right) \times 100 \quad (1)$$

Figure 5 shows the adsorption rate of H-type zeolites with various pore sizes for paraffins of each carbon number. In order to confirm the relationship between the zeolite pore size and the adsorption rate by HC molecular size, zeolites with pore sizes close to the molecular sizes of small and large HC in the exhaust gas are used, and the pore size in order from the smallest is CHA < MFI < BEA < FAU. The samples used for measurement were prepared by making a slurry of each zeolite powder and coating it onto Φ25.4 mm × 60 mm (30 cc) cordierite honeycombs which were then calcinated for 2 hours at 500°C. Instead of the model gas composition in Table 1, this evaluation used the simple composition of only a paraffin (one type with a carbon number from 3 to 8) at 300 ppmC and nitrogen and was conducted sequentially for carbon numbers from 3 to 8. The results confirmed that HC with a molecular size close to the pore size tend to be more easily adsorbed. Based on this,

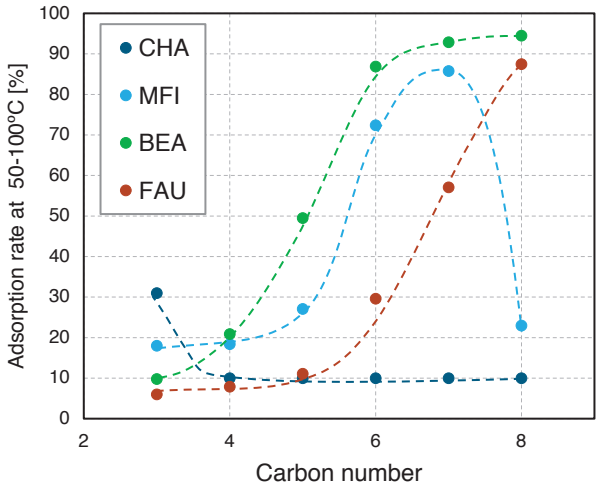


Fig. 5 Influence paraffin carbon number on adsorption rate of various zeolites evaluated with 300 ppmC various paraffin and N_2 gases

Table 2 Propylene adsorption rate on various ion-exchanged FAU

Element	Cu	Ag	Cr	Co	Ni	Mo
Adsorption rate [%]	5.1	13.0	-	-	-	-

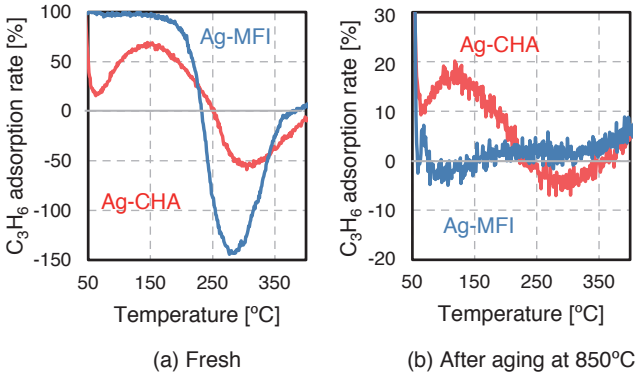
CHA and MFI were selected as candidate zeolites for small-molecule HC, and BEA and FAU for large-molecule HC.

Next, exchanged ions for promoting chemisorption were individually selected for small- and large-molecule HC.

3.3. Selection of Exchanged Ion for Adsorption of Small-Molecule HC

First, in order to select the exchanged ion for small-molecule HC, various exchanged ions were introduced to FAU zeolite and the measured HC adsorption amounts were compared using propylene, which is a small-molecule HC. There have been reports on adsorption with small HC molecules using transition metal ions[17), (18)], but there are few studies on HC trap materials that include heat resistance, so this was actually confirmed. Table 2 shows the results, with only Cu and Ag confirmed to exhibit adsorption performance, with Ag in particular providing the highest adsorption performance, so Ag was selected as the exchanged ion for small-molecule HC.

The performance of Ag zeolite was then evaluated. The heat resistance of exchanged ions is known to vary depending on the combined zeolite type[7], and it is important to evaluate performance after aging. The conditions were set to a temperature (850°C) equivalent to underfloor arrangement, with an atmosphere alternating between flows of propylene-oxygen mixed gas for 80 seconds and air simulating fuel cut for 20 seconds, while simultaneously adding water to maintain a constant 10 vol%. Figure 6 shows the results of evaluating the propylene adsorption performance before aging (Fresh) and after aging of Ag-CHA and Ag-MFI, which are actual candidate zeolites for small molecules with Ag introduced. As shown in Fig. 6(a), fresh Ag-MFI exhibited higher adsorption performance than fresh



(a) Fresh (b) After aging at 850°C

Fig. 6 Propylene adsorption and desorption profile of Ag-CHA and Ag-MFI

Ag-CHA, whereas after aging, the Ag-MFI adsorption function disappeared but Ag-CHA maintained its adsorption function as seen in Fig. 6(b). To analyze these factors, changes in the Ag ion and the zeolite states were analyzed. Figure 7 shows the Ag ion and cluster ratios obtained by ultraviolet-visible absorption spectroscopy (UV-Vis) as a physical property after aging. Figure 7(a) shows that CHA has smaller variation in the ratios before and after aging, whereas Fig. 7(b) shows that Ag aggregation progressed on MFI. Previous reports confirmed that X-ray diffraction (XRD) also showed that CHA exhibited smaller variation in crystallinity fluctuation before and after aging[19]. This suggests that Ag on CHA is more stable, which in turn helps stabilize the zeolite structure, enabling retention of adsorption functions.

3.4. Selection of Exchanged Ion for Adsorption of Large-Molecule HC

Next, focusing on materials for C5 or higher large-molecule HC, particularly aromatics, exchanged ions similar to those in Section 3.3. were introduced to FAU zeolite and the HC adsorption amounts were compared. Figure 8 shows the adsorption rate measurement results, which reveal that Cu alone has unique adsorption characteristics in the temperature range from 150 to 300°C. Reports thus far of adsorption as HC trap materials only concerned adsorption observed with ions other than Cu in the temperature range from 50 to 150°C, and this specificity of Cu is noteworthy. This adsorption characteristic on the high-temperature side is thought to be effective for delaying HC desorption, so Cu was selected as the exchanged ion for large molecules. This adsorption is observed only in the presence of water and disappears in its absence; that is, there is known to be an action of Cu and water[20]. Also, like Ag, Cu is compatible with zeolite regarding to heat resistance, so the HC adsorption performance before and after aging was compared for materials where Cu was ion-exchanged with the candidate zeolites for large molecules. Figure 9 shows the results, and while both materials exhibit a second adsorption peak in the fresh state, only Cu-FAU retains it after aging. Therefore, to clarify the differences in the changes in the ionic state in each zeolite, the focus was placed on Cu degradation after aging, and the total amount of Cu^{+} and Cu^{2+} ions in Cu-FAU and their ratio to CuOx were analyzed by electron spin resonance (ESR). Identification of each ion was performed by making use of the property that the measurable ions differ in the presence and absence of water, and it was assumed that all the Cu ions in the fresh state and the entire ion reduction amount after aging are CuOx. As shown in Fig. 10(a), it was confirmed that after aging, 1/3 of the Cu was oxidized and 2/3 retained the ionic state. Furthermore, the relationship between the Cu ion ratio and the second adsorption rate was investigated using FAU with various silica-alumina ratios (SiO_2/Al_2O_3 of zeolite) and Cu content, and the adsorption rate tended to decrease as the Cu ions decreased (and CuOx increased) as shown in Fig. 10(b). This is also thought to indicate that the ability to maintain adsorption performance after aging is related to the Cu ion retention state.

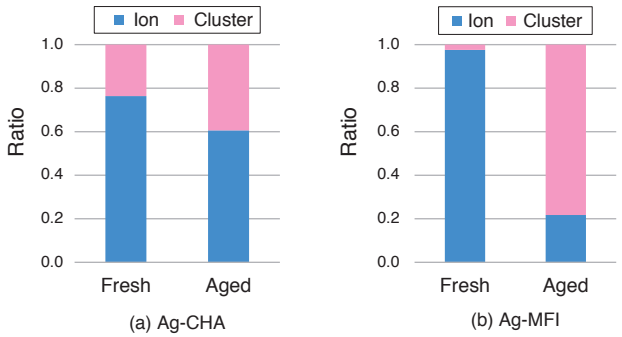


Fig. 7 Variation of Ag ion and cluster ratio calculated from UV-Vis spectra before and after aging

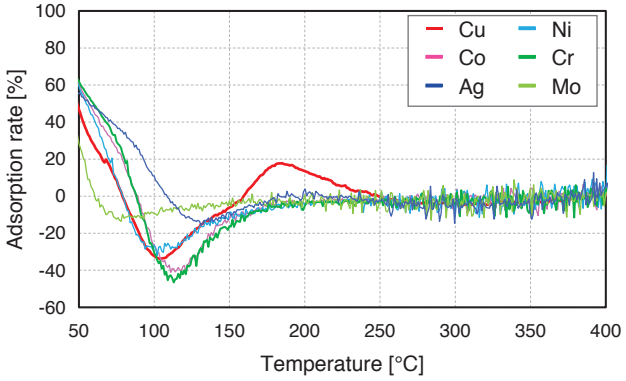


Fig. 8 Profile of HC adsorption on various ion-exchanged FAU after aging

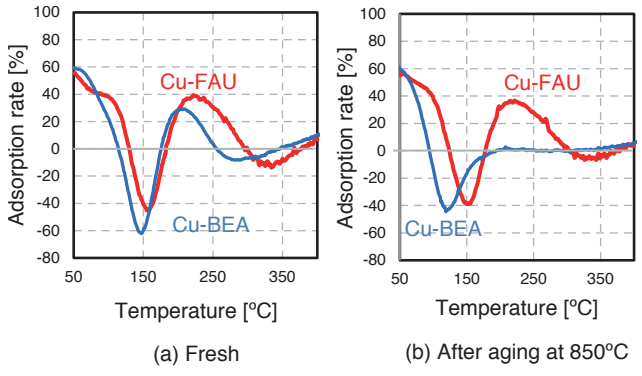
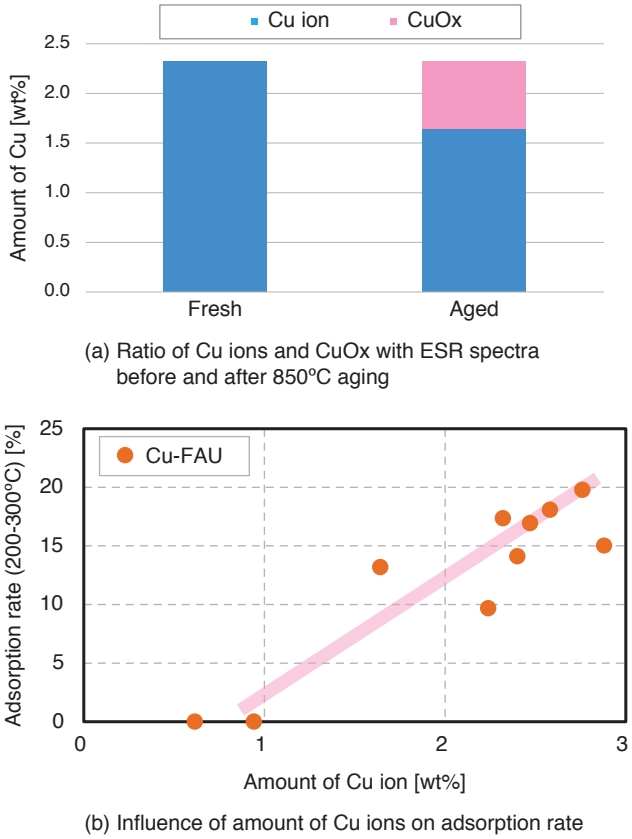


Fig. 9 THC adsorption and desorption profile of Cu-FAU and Cu-BEA



(a) Ratio of Cu ions and CuOx with ESR spectra before and after 850°C aging

(b) Influence of amount of Cu ions on adsorption rate

Fig. 10 Change in Cu ion after aging

3.5. Factors Enhancing the Heat Resistance of Exchanged Ions

It was found that the aforementioned Ag-CHA and Cu-FAU can maintain adsorption functions even after aging, and the main factor for this is thought to be that the zeolite enhanced the stability of the exchanged ion. There are multiple reports on the specific stability of Cu^{2+} ions under hydrothermal aging in CHA[21), (22)], and it has been concluded from simulation analysis that this is due to the more stable interaction between the D6R unit and Cu^{2+} ions in the zeolite framework. In addition, Ag has an electron arrangement similar to that of Cu (one electron in the outermost s orbital and all 10 electrons in the inner d orbitals filled), so the interaction with the D6R unit is also estimated to be similar. Figure 11 shows model drawings of the CHA and FAU frameworks[16], and it can be considered that CHA and FAU both exhibited heat resistance because they both have this D6R that stabilized the exchanged ion.

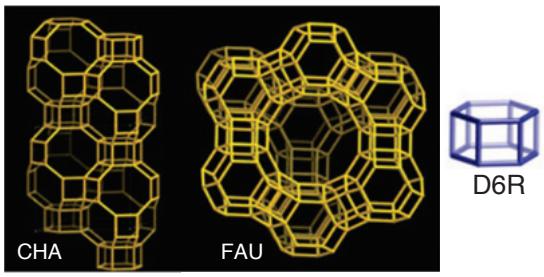


Fig. 11 D6R units in CHA and FAU

4. HC Trap Catalyst

The Ag-CHA and Cu-FAU selected as the HC trap materials were combined with an HC purification catalyst to make an HCTC that completes the process from adsorption to purification. This attempt aimed to achieve an HCTC that makes the most of the adsorption and purification functions.

The catalyst production method was as follows. γ-Al_2O_3 made by Sasol Ltd was used as the Pd support material, and ZrO_2 and CeO_2-ZrO_2 made by DAIICHI KIGENSO KAGAKU KOGYO CO., LTD. were respectively used as the Rh support material and OSC material. Pd and Rh were supported on the support materials by an impregnation method using salt solutions of each, dried for 12 hours at 200°C, and then calcinated for 2 hours at 500°C to obtain Pd/Al_2O_3 and Rh/ZrO_2 powders. These powders were then made into slurries and coated in the specified amounts onto cordierite honeycomb similarly to the HC trap layer.

The methods of addressing the HC purification catalyst challenges described in Chapter 2 (3) and (4) were examined in detail. First, regarding the balance between adsorption and purification, the adsorption and purification performance vary widely depending on the HC trap layer and HC purification catalyst arrangement and amounts. For example, coating the HC purification catalyst on the upper layer side of the trap may result in either lower or higher trap performance. In addition, the coating amount increases by the amount of the trap layer compared to the currently used TWC, so the gas path narrows, leading to a higher gas flow rate (linear velocity) per unit of cross-sectional area of the gas path. This is expected to increase the treatment amount, which in turn may lead to lower adsorption and purification performance. The key is how much these declines in performance can be suppressed and how efficiently the trap effect can be linked to enhancement of the reduction rate.

Next is purification during HC desorption. Here, a lower HC purification rate due to insufficient oxygen leads to an increase in emissions. Use of an OSC material with oxygen absorption and release functionality in the TWC to compensate for this oxygen deficiency was examined.

4.1. Balance between Adsorption and Purification

4.1.1. Layer arrangement of HC trap materials

First, a layer arrangement that can make advantageous use of the HC trap material properties was considered. Based on the material properties, Cu-FAU exhibits interaction with water, so it should be arranged in the upper layer where there is higher probability of contact with water, whereas Ag-CHA adsorption is hindered by water, so arrangement in the lower layer is considered advantageous. Also, considering the ease of diffusion of

large HC molecules within zeolites, if Ag-CHA, which has a pore size smaller than the molecule size, is arranged in the upper layer, large HC molecules will be unable to enter the zeolite pores, thereby lowering the probability of contact with the Cu-FAU arranged in the lower layer. Therefore, a layout with Ag-CHA in the lower layer is expected to be preferable. Figure 12 shows the HC adsorption characteristics as an HC trap layer that combines the materials. Here, (a) shows the case with Cu-FAU arranged in the upper layer and Ag-CHA in the lower layer, and (b) shows the case with the opposite arrangement. As expected, case (a) was confirmed to have a better adsorption rate.

4.1.2. HC trap layer and HC purification catalyst layer arrangement

Next, the effect of the relative arrangement of the HC trap materials and the HC purification catalyst on adsorption and purification was compared. Figure 13 shows the adsorption and purification profiles when equal amounts of HC trap and HC purification catalyst are arranged vertically on the honeycomb. The profiles for the TWC alone and the HC trap layer alone are also shown for comparison.

Here, the HC purification catalyst layer is 150 g/L. When the HC trap layer is in the upper layer, a slight advantage in adsorption can be confirmed, but the purification performance is noticeably lower than the TWC. On the other hand, when the HC purification catalyst is in the upper layer, there is a slight drop in both adsorption and purification, but the impact on purification is less than when the HC trap materials are in the upper layer. This drop in purification performance when the HC purification catalyst is in the lower layer is thought to be largely due to lower gas diffusivity and the delayed temperature rise because of an inability to receive heat directly from the gas. Therefore, in consideration of the total reduction rate, it is important to position the HC purification catalyst in the upper layer.

4.1.3. HC purification catalyst coating amount

The HC trap layer described in section 4.1.1. was set to 200 g/L, which provides the greatest adsorption effect and is equivalent to the coating amount of the currently used TWC. Therefore, if the HC purification catalyst coating amount is equal to that of the TWC, the total will be twice the current amount. Together with the height of the zeolite, the gas flow path will become quite narrow, leading to an increase in the aforementioned linear velocity and adversely affecting adsorption and purification. Therefore, guidelines for the HC purification catalyst are needed.

Figure 14 shows the relationship between the HC purification catalyst layer coating amount and the adsorption rate, purification rate, and reduction rate with the HC trap layer coating amount constant at 200 g/L. Here, the adsorption rate, purification rate, and reduction rate

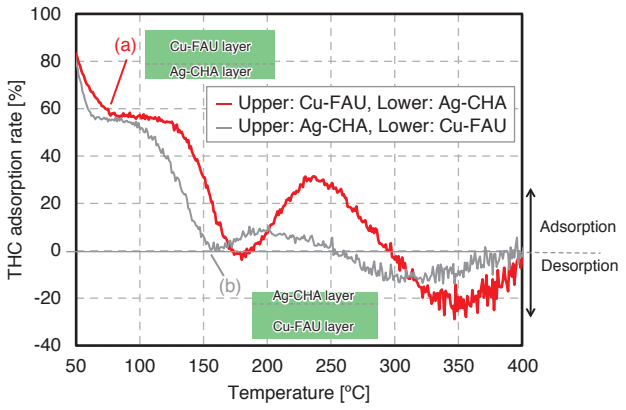


Fig. 12 Relationship between layer composition and adsorption of Cu-FAU and Ag-CHA and TWC

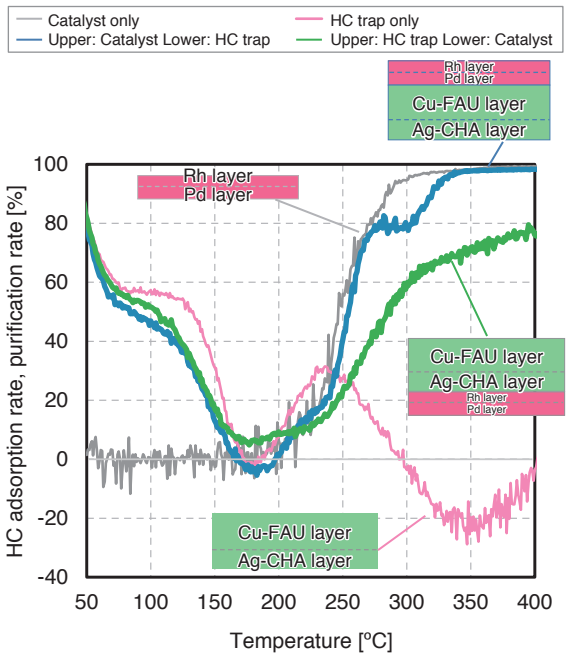


Fig. 13 Effect of zeolite-TWC arrangement on adsorption rate and purification rate

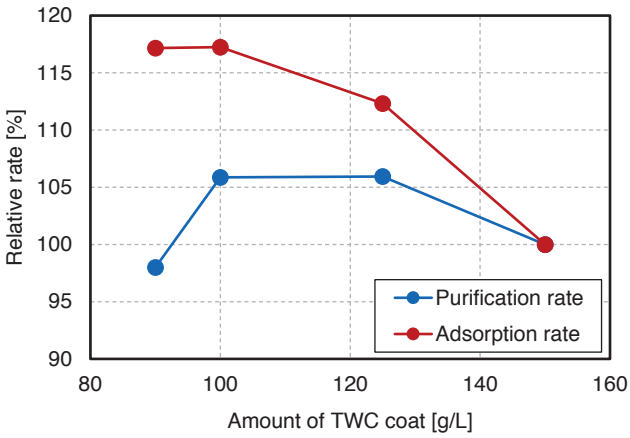


Fig. 14 Relationship between TWC coat and HC absorption rate, purification rate, and reduction rate

are expressed as relative values with the values at 150 g/L set to 100. In addition, the total amount of PGM for each sample was constant at 185 g/ft^3. This figure shows that the adsorption rate exhibits a monotonically increasing trend as the coating amount decreases, indicating that the coating amount of the HC purification catalyst layer in the upper layer has a large influence. Meanwhile, the purification rate shows an initially rising trend as the coating amount decreases, but then turns downward at 100 g/L or less, indicating a tendency to have a maximal value. Figure 15 shows the linear velocity (calculated from the cross-sectional area of the gas path) and the Pd particle size (by the CO adsorption method) relative to the coating amount. Here, the linear velocity is observed to decrease together with the coating amount, indicating that the reaction proceeds more readily. In contrast, the particle size of the Pd, which serves as the active site of the reaction, was confirmed to aggregate with the increase in the Pd density as the coating amount decreases, resulting in lower catalytic activation. This suggests that this characteristic has a maximal value due to the trade-off relationship between these two factors. The above results indicate that it is effective to arrange the HC purification catalyst in the upper layer and the HC trap materials in the lower layer, and to set the HC purification catalyst coating amount to an optimal value to achieve the maximum purification rate value.

4.2. Material Selection Based on OSC

Purification of the desorbed HC is described using the results in Fig. 16. The vertical axis represents the TWC light-off temperature in the model gas performance evaluation test, and the horizontal axis represents the concentration of oxygen in the model gas. Normally, at the stoichiometric A/F, just the precise amount of oxygen needed for the purification reaction of the TWC is supplied, providing conditions for achieving the greatest purification performance. The reference temperature in this figure is set to 285°C, which is the light-off temperature of the reference TWC under stoichiometric conditions. In contrast, the light-off temperature of the HCTC was confirmed to be 322°C, which is more than 30°C higher. The factor for this is that in the HCTC, although HC that desorbs in the light-off temperature range needs to be oxidized together with the HC contained in the exhaust gas, it is thought that the reaction does not readily proceed due to insufficient oxygen supply in actual operation. It was therefore thought that by adding oxygen in the model gas by the amount needed to purify the desorbed HC, the light-off temperature would return to the original level. An equivalent oxygen concentration was added and tested in the same manner, and the light-off temperature was confirmed to return to the same level as the reference TWC. This clearly indicates that the issue is insufficient oxygen, which placed the focus on how to supplement the oxygen needed for the desorbed HC.

Attention was thus turned to the OSC material, which is one of the main components used in TWC. In order to maintain the state in the TWC near A/F stoichiometry that provides the highest purification rate even in A/F fluctuations, the oxygen excess or deficiency at that time is compensated by the oxygen absorption and release function of the OSC material, thereby suppressing purification rate fluctuations[23]. This function was also considered to be an effective means in this study, and helped to enhance the desorbed HC purification performance in terms of increasing oxygen supply with the OSC material. The approach was the CeO_2 ratio in the OSC material. The OSC material mainly comprises a composite oxide of CeO_2 and ZrO_2, which has two aspects: an oxygen adsorption and release function by the valence change of CeO_2, and a function to enhance heat resistance by crystal stabilization of ZrO_2. Therefore, the CeO_2/ZrO_2 ratio was examined that enables retention of heat resistance while helping to increase the oxygen supply during HC desorption. The OSC retention time was used as an indicator of the oxygen supply amount. The OSC retention time was evaluated by maintaining a flow of 0.5% oxygen (nitrogen balance) for 5 minutes in the 600°C steady state to fill the OSC material with oxygen, and then switching to 1% CO (nitrogen balance) to establish and maintain the zero-oxygen state

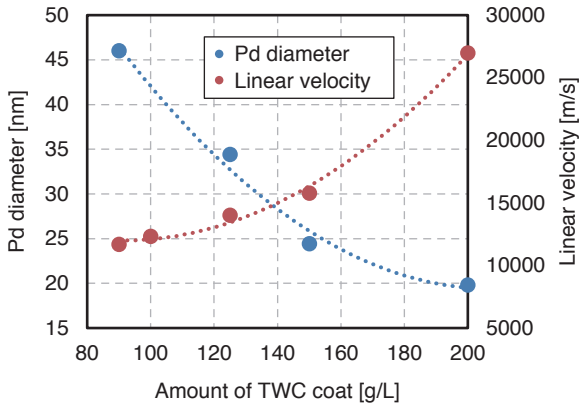


Fig. 15 Influence of coating amount on Pd particle size and linear velocity

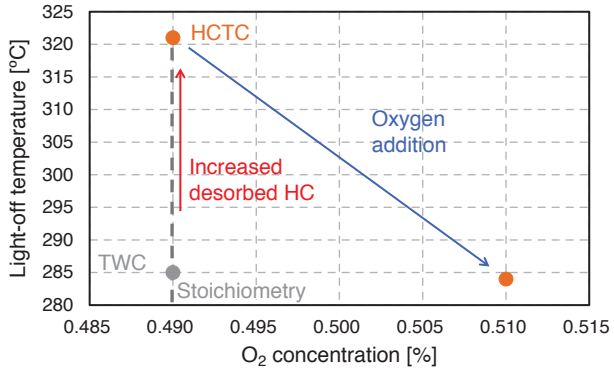


Fig. 16 Effect of oxygen concentration on desorbed HC purification temperature

for 5 minutes. That is, the OSC retention time indicates how many seconds the oxygen in the OSC material can continue to be supplied to oxidize CO and is defined here as the duration until the CO purification rate drops to 50%. The blue plots in Fig. 17 show the results of confirming the relationship between the CeO_2 ratio and the OSC retention time and reveal that the OSC retention time was longest at a CeO_2 ratio of 50%. The reason is thought to be that when the CeO_2 ratio is above this level, the ZrO_2 ratio drops, promoting thermal degradation of the CeO_2. Conversely, when below this level, the lower CeO_2 ratio means that less oxygen can be supplied. Furthermore, a similar tendency was confirmed in the relationship between the CeO_2 ratio and the HC purification rate (Fig. 17, 2nd vertical axis, brown plots, purification rate in the interval from 200 to 400°C). In other words, it was indicated that increasing the amount of oxygen supply extends the OSC retention time and is effective for purification of desorbed HC.

Figure 18 compares the performance before and after

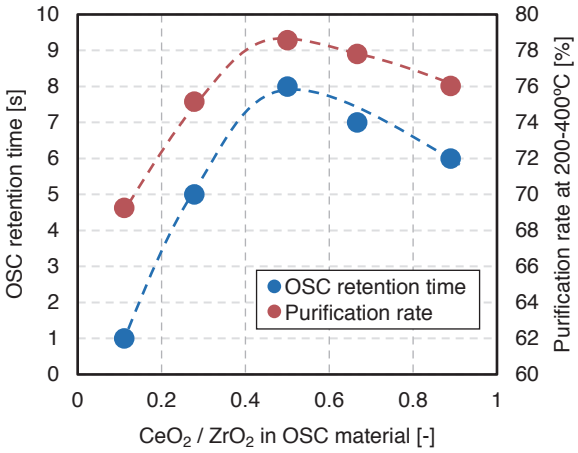


Fig. 17 Relationship between CeO_2 ratio, OSC retention time, and HC reduction rate

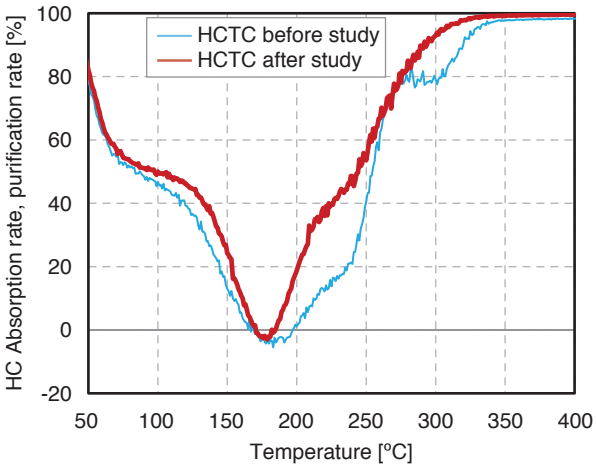


Fig. 18 Effect of HCTC study on temperature characteristics

optimizing the HC purification catalyst coating amount and the CeO_2/ZrO_2 ratio of the OSC material based on the above results. It was confirmed that the decline in adsorption and purification performance can be suppressed by setting the HC purification catalyst coating amount within the range that has both the smallest impact on adsorption and the lowest catalyst degradation, and by setting the CeO_2 ratio in the OSC material to maximize the desorbed HC purification rate.

5. Emission Reduction Effects in a V6 Engine

Figure 19 shows the HC trap system of a V6 engine based on the results of the above HC trap material and HCTC examinations. One TWC is placed close-coupled to the exhaust outlets of the three cylinders on each side, so there are no changes in the close-coupled layout, and the TWC installed underfloor in a conventional V6 engine is

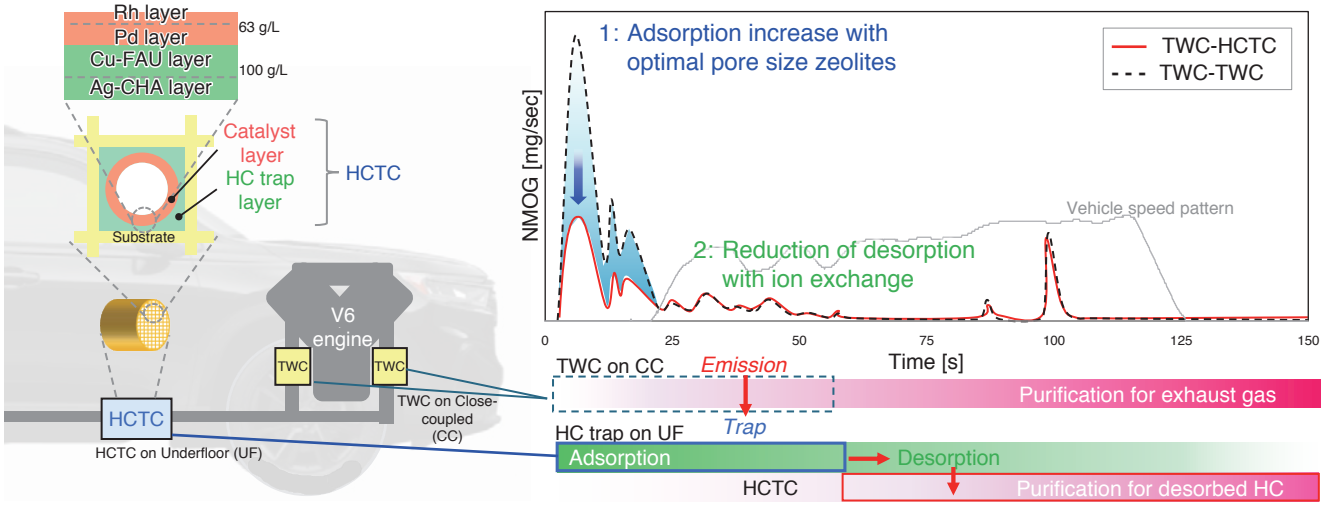


Fig. 19 Emissions reduction strategy using HCTC

replaced with the HCTC. The following purification scheme was devised based on this arrangement. In the 0-20 second interval, the exhaust gas passes through the unlit close-coupled TWCs but is adsorbed by the underfloor HCTC. In the 20-50 second interval, the close-coupled TWCs light-off, but the HC trap layer in the HCTC continues to retain the initially adsorbed HC. Then, in the 50-150 second interval, the gas flowing in from the engine is purified by the mostly activated TWC. Meanwhile, the adsorbed HC begins to desorb from the HC trap layer, but the HC purification catalyst layer of the HCTC is also activated, enabling purification of the desorbed HC.

5.1. Emission Test

The arrangement characteristics and the reduction effects in the final form were verified with an actual engine. The engine was a V6 3.5 L DI engine with LEV III certification gasoline used as the fuel. An identical TWC (PGM 140 g/ft^3, aging equivalent to close-coupled (with only the temperature changed to 950°C in the previously described aging conditions), honeycomb cell density 600 cell/inch^2, cell wall thickness 3.5 mil, and capacity 0.85 L was installed close-coupled to each bank. A TWC (PGM 185 g/ft^3, aging equivalent to underfloor arrangement, honeycomb cell density 400 cell/inch^2, cell wall thickness 3.5 mil, and capacity 1.0 L) was arranged underfloor as a reference, and evaluation was performed by replacement with the HCTC. The exhaust gas test was conducted in the LA#4 (cold-start) mode, and the HC used NMOG.

5.2. Characteristics of HC Trap Material and HC Purification Catalyst Arrangement

First, the effects of the optimal arrangement and volumes of the HC purification catalyst and the HC trap materials, which are the HCTC components, were compared in an actual engine. Figure 20 shows the comparison results (the HC purification catalyst is noted as CAT). Here, the vertical axis represents the reduction rate based on the emission values from 0 to 300 seconds with a TWC as the underfloor unit. Note that all arrangements used the same HC purification catalyst layer and HC trap layer coating amounts. First, at a volume of 1.0 L and an arrangement with the HC trap layer in the lower layer and the HC purification catalyst layer in the upper layer, a 36% reduction in emissions was observed compared to the TWC alone. Furthermore, arrangements with the HC purification catalyst on the front side and the HC trap materials on the rear side (1.0 L volume each) and vice versa were examined with the objective of lowering the impact of each coat on gas diffusion, and showed reductions of 39% and 44%, respectively. In particular, the latter showed an 8% reduction compared to the vertical arrangement, suggesting that a front and rear arrangement can effectively lessen the impact on diffusion. Furthermore, the arrangement with the HC purification catalyst and HC trap materials arranged vertically on a 2.0 L honeycomb provided the highest reduction effect of 48%. This is considered to be the contribution of the effect to lower the linear velocity because the larger volume resulted in thinner coats, and the effect that the greater length in the gas flow direction had on easing the adsorption and purification conditions.

5.3. Demonstration of HC Trap System Effects

An actual engine was used to verify the purification scheme of an exhaust system with an underfloor 2.0 L HCTC, which provided the best purification performance based on the above examinations of trap materials, trap catalysts, and arrangement. As a result, Fig. 21 shows that total NMOG emissions were reduced by 48% compared to the emissions using only TWC, and the profile over time in Fig. 22 exhibits a large early reduction effect as intended. As it was previously reported[12] that adding an underfloor HCTC to a close-coupled 2-bed TWC arrangement in a 1.5 L turbocharged engine achieved reduction effects of 51%, the reduction effects of an underfloor HCTC arrangement have now been demonstrated in multiple engines.

These results showed that an emission reduction rate of around 50% could be achieved during the period from 0 to 300 seconds in either engine, and demonstrated the effectiveness of the HC trap concept in this study from materials to system.

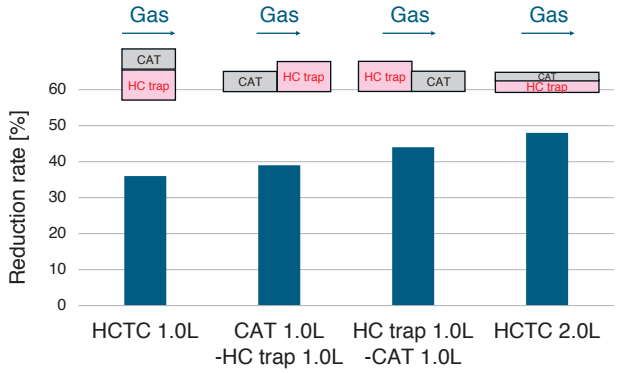


Fig. 20 Relative reduction rate of various arrangement and volume HCTC with V6 engine

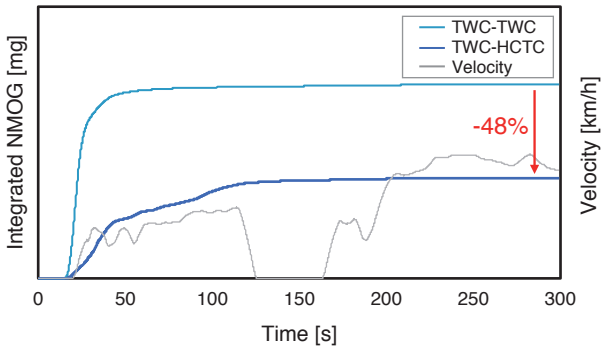


Fig. 21 Comparison of emissions between TWC and final HCTC system with V6 engine

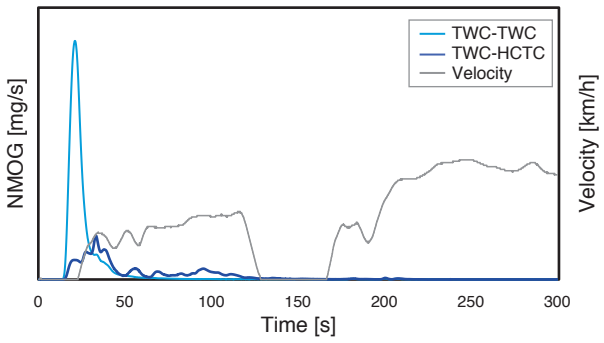


Fig. 22 Changes in HC emission profile over time using final HC trap system in V6 engine

6. Conclusion

Research on an HC trap system with high reduction effects revealed the following.

(1) Increased stability of the exchanged ions that promote chemisorption is key to adsorbing HC up to high temperatures. Use of CHA zeolite with Ag ions introduced for the small-molecule HC trap and FAU zeolite with Cu ions introduced for the large-molecule HC trap enables increasing the adsorption and retention temperature to around 200°C.

(2) Use of OSC material consisting of CeO_2 and ZrO_2 in a 1:1 ratio enables purification of the desorbed HC even under TWC stoichiometric conditions.

(3) The greatest HC purification performance can be achieved by positioning a TWC catalyst containing OSC material in the uppermost layer as a catalyst for purification of trapped HC, Cu-FAU in the middle layer, and Ag-CHA in the lowermost layer.

(4) An underfloor HCTC arrangement is highly effective not only in an L4 engine but also in a V6 engine, and reduced NMOG by 48% in the LA#4 mode test.

Reference

(1) Rood, S., Eslava, S., Manigrasso, A., Bannister, C.: Recent advances in gasoline three-way catalyst formulation: A review, Journal of Automobile Engineering, Vol. 234, Issue 4, p. 936-949 (2020), doi: 10.1177/0954407019859822

(2) Hashimoto, M., Nakanishi, Y., Koyama, H., Inose, S., Takeori, H., Watanabe, T. , Narishige, T., Okayama, T., Suehiro, Y.: Development of Low Temperature Active Material for Three Way Catalyst, SAE 2016 World Congress and Exhibition, SAE Technical Paper 2016-01-0932, (2016), doi: 10.4271/2016-01-0932

(3) Bartley, G. J.: Identifying Limiters to Low Temperature Catalyst Activity, SAE 2015 World Congress & Exhibition, SAE Technical Paper 2015-01-1025, (2015), doi: 10.4271/2015-01-1025

(4) Williamson, W., Dou, D., Robota, H.: Dual-Catalyst Underfloor LEV/ULEV Strategies for Effective Precious Metal Management, SAE Technical Paper 1999-01-0776, (1999), doi: 10.4271/1999-01-0776

(5) Hofstetter, J., Boucharel, P., Atzler, F., Wachtmeister, G.: Fuel Consumption and Emission Reduction for Hybrid Electric Vehicles with Electrically Heated Catalyst, SAE Int. J. Adv. & Curr. Prac. in Mobility, Vol. 3, Issue 1, p. 702-714, (2021), doi: 10.4271/2020-37-0017

(6) Takamitsu, Y., Ariga, K., Yoshida, S., Ogawa, H., Sano, T.: Adsorption of Toluene on Alkali Metal Ion-Exchanged ZSM-5 and β-Zeolites under Humid Conditions, Bulletin of the Chemical Society of Japan, Volume 85, Issue 8, p. 869-876, (2012), doi: 10.1246/bcsj.20120036

(7) Kobatake, Y., Momma, K., Elangovan, S. P., Itabashi, K., Okubo, T., Ogura, M.: "Super Hydrocarbon Reformer Trap" for the Complete Oxidation of Toluene Using Iron-Exchanged β-Zeolite with a Low Silicon/Aluminum Ratio, ChemCatChem, Vol. 8, Issue 15, p. 2516-2524, (2016), doi: 10.1002/cctc.201600550

(8) Endo, Y., Nishikawa, J., Iwakura, H., Inamura, M., Wakabayashi, T., Nakahara, Y., Ogasawara, M., Kato, S.: Development of Highly Durable Zeolite as Hydrocarbon Trap Materials for Automobile Catalyst, SAE Technical Paper 2018-01-0947, (2018), doi: 10.4271/2018-01-0947

(9) Lupescu, J., Xu, L., Jen, H., Harwell, A., Nunan, J., Alltizer, C., Denison, G.: A New Catalyzed HC Trap Technology that Enhances the Conversion of Gasoline Fuel Cold-Start Emissions, SAE Int. J. Fuels Lubr., Vol. 11, Issue 4, p. 411-425, (2018), doi: 10.4271/2018-01-0938

(10) Nakao, K., Mitsuhashi, R., Asakawa, T.: Vehicle Emission Control Technology Utilizing Zeolite Adsorption Function, TOSOH Research & Technology Review, Vol. 63, p. 19-25, (2019) (in Japanese), https://www.tosoh.co.jp/technology/assets/19-2-3.pdf, (accessed 2024/11/5)

(11) Yamamoto, S., Takaya, M., Matsushita, K.: Hydrocarbon-Adsorption Type Three-Way-Catalyst System for Reducing Cold-Start Hydrocarbon Emissions, NIPPON KAGAKU KAISHI, Vol. 2001, Issue 6, p. 333-343, (2001) (in Japanese), doi: 10.1246/nikkashi.2001.333

(12) Takeori, H., Nemoto, K., Oku, Y., Ikutomo, R., Mori, T., Nakao, K., Usui, T., Mitsihashi, R., Nakazawa, N., Matsuo, Y.: Study of HC Trap System to Reduce HC Emissions during Engine Cold-Start, 2024 JSAE Annual Congress (Spring), No. 17-24, 20245072, (2024)

(13) Takahashi, H., Ishizuka, Y., Tomita, M., Nishizawa, K.: Engine-Out and Tail-Pipe Emission Reduction Technologies of V-6 LEVs, SAE Technical Paper 980674, (1998), doi: 10.4271/980674

(14) Mitsuishi, S., Mori, K., Nishizawa, K., Yamamoto, S.: Emission Reduction Technologies for Turbocharged Engines, SAE Technical Paper 1999-01-3629, (1999), doi: 10.4271/1999-01-3629
(15) Miyamoto, K., Tanada, H., Yamashita, M., Tanabe, K.: Development Of Technology for 3.0 L V6 Super Low Emission Vehicle, Transactions of Society of Automotive Engineers of Japan, Vol. 39, Issue 4, p. 4_65-4_71, (2008) (in Japanese), doi: 10.11351/jsaeronbun.39.4_65
(16) Database of Zeolite Structures: http://www.iza-structure.org/databases/, (accessed 2024/11/5)
(17) Liu, X., Lampert, J. K., Arendarskiia, D. A., Farrauto, R.: FT-IR spectroscopic studies of hydrocarbon trapping in Ag^{+}-ZSM-5 for gasoline engines under cold-start conditions, Applied Catalysis B: Environmental, Vol. 35, Issue 2, p. 125-136, (2001), doi: 10.1016/S0926-3373(01)00247-8
(18) Kyriakidou, E. A., Lee, J., Choi, J-S., Lance, M., Toops, T. J.: A comparative study of silver- and palladium- exchanged zeolites in propylene and nitrogen oxide adsorption and desorption for cold-start applications, Catalysis Today, Vol. 360, p. 220-233, (2021), doi: 10.1016/j.cattod.2020.05.019
(19) Takeori, H., Oku, Y., Ikutomo, R., Mori, T. Nemoto, K., Matsuo, Y., Nakao, K., Usui, T., Mitsuhashi, R., Nakazawa, N., et al.: Adsorption effect of Ag-CHA on small olefins and its application to HC Trap, 133th CATSJ Meeting Proceedings, 1P28, (2024) (in Japanese)
(20) Oku, Y., Takeori, H., Ikutomo, R., Mori, T., Nemoto, Y., Matsuo, Y., Nakao, K., Usui, T., Mitsuhashi, R., Nakazawa, N., et al.: High-temperature adsorption effect of Cu-FAU for HC trap, 133th CATSJ Meeting Proceedings, 1P29, (2024) (in Japanese)
(21) Usui, T., Liu, Z., Igarashi, H., Sasaki, Y., Shiramata, Y., Yamada, H., Ohara, K., Kusamoto, T., Wakihara, T.: Identifying the Factors Governing the Early-Stage Degradation of Cu-Chabazite Zeolite for NH_3-SCR, ACS Omega, Vol. 4, Issue 2, p. 3653-3659, (2019), doi: 10.1021/acsomega.8b03409
(22) Xu, R., Wang, Z., Liu, N., Dai, C., Zhang, J., Chen, B.: Understanding Zn Functions on Hydrothermal Stability in a One-Pot-Synthesized Cu&Zn-SSZ-13 Catalyst for NH_3 Selective Catalytic Reduction, ACS Catalysis, Vol. 10, Issue 11, p. 6197-6212, (2020), doi: 10.1021/acscatal.0c01063
(23) Yamada, T., Ashizawa, H., Nagata, M.: Multiple Regression Analysis of OSC Characteristics under Transient TWC Conditions, SAE Technical Paper 2008-01-0713, (2008), doi: 10.4271/2008-01-0713

■ Author ■


Hiroki TAKEORI


Koji NEMOTO


Yuki OKU


Ryohei IKUTOMO


Keita NAKAO


Toyohiro USUI

Integrated Simulation Technology for Design, Manufacturing, and Materials to Reduce Surface Distortion of Automotive Outer Panels

Toru FURUSAWA[1] Yoichi TOYOOKA[1] Eiichiro YOKOI[1]
Shigeru TAKAMI[2] Masao UCHIYAMA[2] Yoshinori TANAKA[2]

ABSTRACT

When producing highly stylized exterior designs for automobiles, conventional manufacturing processes can cause surface distortions on the outer body panels, which require countermeasures in order to meet specifications. To address this issue, we have developed a new development-stage simulation technology. This technology integrates design, manufacturing, and materials factors to predict the potential for a design to be achieved without visible distortions.

We began by proposing a digital development concept that predicts distortions throughout all the plant manufacturing processes (stamp forming, spot welding, and paint curing) based on models made prior to the production of vehicle body design drawings. We then created a simulation technology to analyze the distortions that accumulate throughout all the manufacturing processes, and verified it on actual vehicles.

The technology was then integrated into the development operational rules for new models. This made it possible to reduce reworking due to distortions occurring in plants that had resisted resolution for many years, and thereby enable efficient production of stylized surface designs.

1. Introduction

In recent years, automobile development has increasingly focused on using digital development environments to shorten development times as well as initiatives to increase performance and quality(1). Within this scenario, in body development, design specifications from drawings are being turned into numerical requirements, and management baselines for manufacturing processes are being established. In terms of performance, CAE is being used for rigidity, crashworthiness, strength and durability, and NVH. In manufacturing processes as well, CAE is used for individual processes, including stamp forming, welding, and painting, and in body development, MBD and digital twin simulation analysis technologies are continuing to evolve.

However, the distortions that are detected in the quality control and inspection processes that determine the final style quality of outer body panels are the result of the accumulated deformations and distortions from the manufacturing processes that are based on design specifications. The sensitivity and interactions of the influencing factors in each process are not well understood. In vehicle body design, as models evolve, exterior styling, light weight, safety, and functionality are being enhanced and specifications diversified. In manufacturing, there are advances in electric vehicle production technologies and the reduction of factory environmental impacts, while manufacturing conditions differ from one plant to the next. In terms of vehicle body materials, the use of high-strength, lightweight materials and recycled materials is on the increase, and materials specifications are diversifying. As a result of this growing diversification of requirements, factors impacting outer panel distortions are becoming more numerous and complex. Thus, conventional reporting(2)-(6) and existing technologies are proving insufficient to address the issue.

In order to achieve highly stylized exterior designs for automobiles, therefore, we utilized a digital development environment to link factors relating to the design, manufacture, and materials of outer panels in automobile development, and developed CAE technology to predict surface distortions on outer panels.

[1] Material Research Center
E-mail: toru_furusawa@jp.honda
[2] Honda Motor Co., Ltd.

2. Development Goals

Our goal is to forecast and terminate, at the design stage, the occurrence of outer panel distortions by developing an integrated simulation technology that would encompass vehicle body materials, parts, and all manufacturing processes.

First, by developing CAE technology to analyze accumulated stresses and distortions from manufacturing processes based on design specifications (stamp forming, spot welding, and paint oven curing), it will be possible to predict changes in distortion over time in the actual processes.

Next, by conducting simulations within the baseline conditions and ranges of variation of the vehicle body materials, design specifications, and manufacturing processes, reasons for occurrence will be identified and specifications and conditions established that will not result in visible distortions.

3. Technology

3.1. Developing an Integrated Simulation Technology that Encompasses All the Vehicle Body Materials, Parts, and Manufacturing Processes

3.1.1. Approach to resolving outer panel surface distortions

(1) Influencing factors on outer panel surface distortions

Figure 1 illustrates the process leading to the detection of distortion in vehicle body panels. In vehicle body development, designers create a model layout that meets the required specifications and produce drawings. A prototype is then fabricated using manufacturing processes from stamp forming and spot welding to painting and assembly. From this, the manufacturing conditions that will become the baselines are determined. However, if distortions are visually detected in the exterior inspection process, manufacturing conditions need to be reset, and parts redesigned. In particular, distortions often occur at the surface of the T-intersection of the upper center pillars, on the lower sidesill, and the area around the roof panel arch. These are surfaces that tend to accumulate and concentrate stresses in response to external forces in the manufacturing processes, such as deformation of parts during processing, thermal stresses, or torsion. It is believed that when loads exceed the material's yield point, parts are subject to plastic deformation, which in turn causes outer panel surface distortion.

Although specifications that reflect countermeasures for past occurrences have been integrated into drawings, and verification performed on prototype vehicles, it has not been possible to predict occurrences in advance. The reasons for this are thought to be the application of unprecedented shape specifications in the pursuit of more advanced exterior styling for new models, and differences in production technologies and equipment between factories. There are hundreds of factors that impact panel surface distortion on actual vehicle bodies, including panel gaps at welding points and residual stresses that cannot be measured except with CAE analysis can have a large impact, but are not checklist items in production engineering management for conventional manufacturing processes. These factors were therefore added as analysis items.

(2) Evaluating distortions accumulated throughout all processes

Figure 2 illustrates the distortion evaluation process. Depending on the model being developed, the design

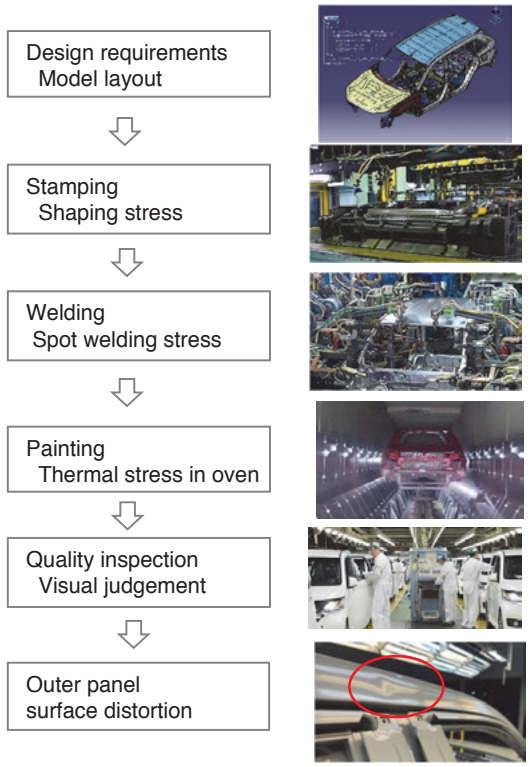


Fig. 1 Surface distortion in automotive body panel manufacturing

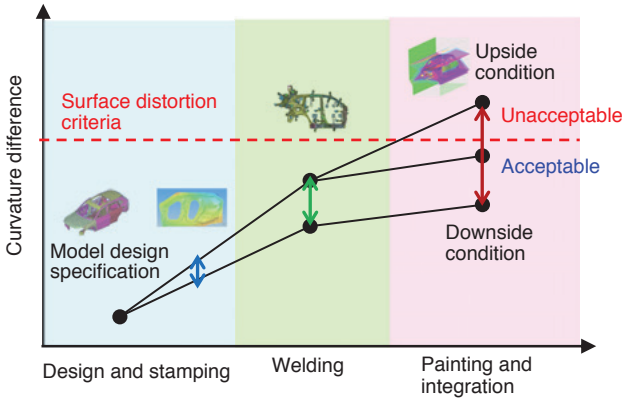


Fig. 2 Evaluation approach of surface distortion criteria for outer panels during development process

shape and size of the outer panels differ, as do the design specifications for the inner panels and the stiffeners between them. In addition, the equipment specifications and processing conditions vary from plant to plant. Within the production process, in the steps from delivered metal sheets to parts and vehicle body processing, the shape and dimensions of the outer panels change, and residual stress and distortions accumulate.

To predict these changes, simulation evaluations were performed based on baseline factors that affect distortion within each process and within ranges of variation. Next, the curvature distribution was determined for outer panels both with and without distortion, and the curvature difference between them was used as an index to evaluate the amount of distortion in the outer panels. The amount of distortion detectable by visual inspection has been accumulated as correlation data relative to the curvature difference[7]. Therefore, it was possible to compare the curvature difference of the distortion amount found in simulations to past data, which in turn enables the setting of specifications to mitigate distortion occurrence.

3.1.2. Developing an end-to-end integrated simulation method

Conventionally, CAE solvers specialized for each manufacturing process (stamp forming, welding, and paint curing) are used, and integrated analysis methods that link all processes have not been available. Here, a method was established to extract and transfer data that could accommodate differences in analysis settings (meshes, node data formats) between processes. Also, factors that impact outer panel surface distortion were reflected as input data for the CAE analysis method. Figure 3 illustrates the manufacturing process steps and digital linking between processes using simulations. For factors against which distortions are especially sensitive, evaluations were performed from the baseline level to the ranges of variation in conditions at factories around the world, and the details are as follows.

(1) Target of analysis

As representative parts, the upper part of the center pillar, sidesill, and roof sections were chosen as targets for outer panel surface distortion analysis.

For the analysis, CAE models were needed that reflected past design know-how and design data from design requirements (baseline dimensions, curvature radius at baseline lines, step dimensions at gaps and offsets, bead surface rigidity, buckling strength, etc.). Here, the mesh sizes and joint modeling used in vehicle performance CAE analysis were repurposed, and analysis performed without the need for additional man-hours to create new models.

(2) Materials specifications

The materials constituting the outer panel parts and joints include steel sheets, aluminum sheets, joint bonding agents, sealers, and so on. The material properties that are entered into the analysis have varied stress-strain charts and temperature dependencies, according to the strength and composition of the base material. The analysis results also vary greatly according to the sheet thickness standard value and the upper and lower limits of tolerances. Therefore, an in-house materials database[8] was used, based on accumulated and detailed measurements of material properties.

(3) Stamping process

In the outer panel stamp forming process, molding

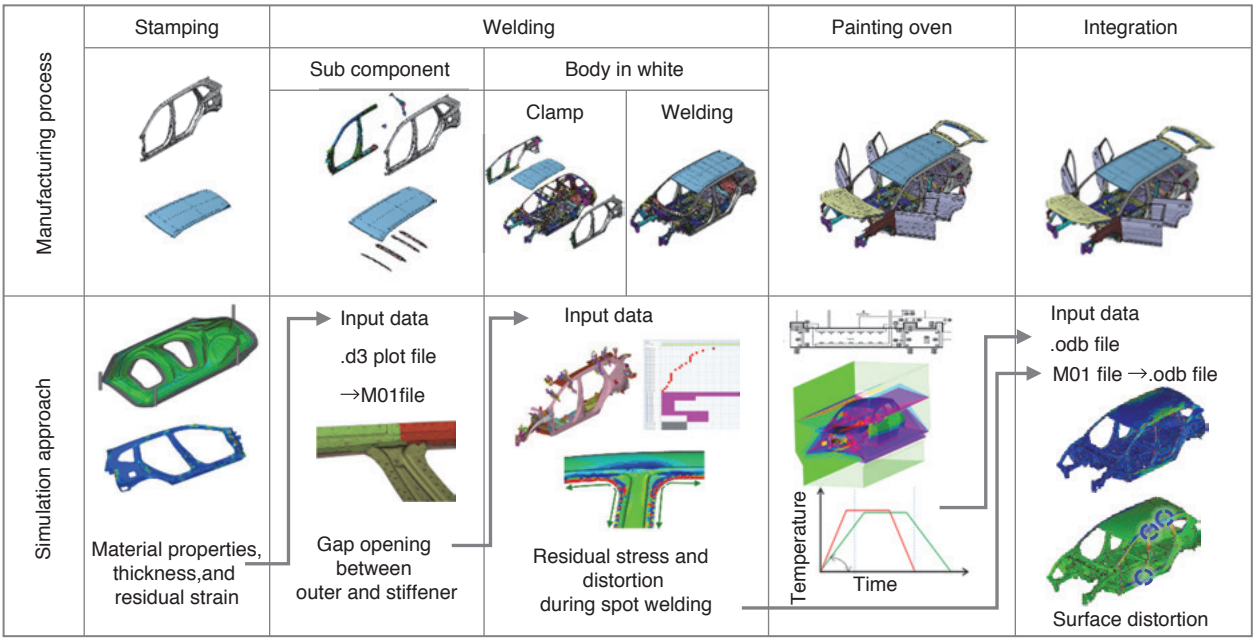


Fig. 3 Integrated simulation approach along manufacturing processes

processes such as draw-and-form molding have a larger impact on distortions than cutting and bending processes such as trimming, restriking, bending, piercing, and cutting. Here, JSOL Corporation's JSTAMP was used. As an explicit solver, the simulator was used to calculate plastic distortions and sheet thickness reduction ratios when drawing and forming flat sheets. Changes in steel sheet yield strength caused by plastic distortion are a key factor in outer panel surface distortions that occur due to compression stresses. Thus, the analysis reflected plastic distortions caused by process hardening. The impact of self-weight was also reflected, and the analysis results carried over to the analysis of the next process of welding.

(4) Welding process

The outer panel welding process specifications include a large number of conditional parameters, such as equipment specifications, welding specifications, clamp specifications, and so on. Factors among these that contribute to distortion are the position of fixing jigs, sheet gaps at spot-welding points, and the effect of residual stress propagation resulting from the spot weld order. These factors are as determined under production engineering requirements. Here, the implicit solver Visual-Assembly, produced by ESI Japan, was used and loaded with the welding clamp locations, spot welding order, and sheet gap parameters. The data for residual stress at the spot-welding points, distortions, and deformation shapes were carried over to the analysis of the next process, that of paint oven curing.

(5) Paint oven curing process

In the outer panel painting process, the vehicle upper body is suspended on a cart and moved over 100 m through a series of curing ovens (electrodeposition oven, intermediate coat oven, topcoat oven, and repair oven), the entire process requiring more than one hour. During this time, the temperature is increased up to about 200°C above the normal temperature to bake and cure the paint, before cooling again. The outer panels that directly face the high-temperature jets from the wall nozzles inside the curing ovens heat up first. The inner parts heat up more slowly as a result of the circulating hot air. The inconsistent thermal expansion caused by the heat difference between the outer and inner panel is a factor in distortion. However, performing thermal fluid analysis with CFD requires large amounts of computational time, making it impractical for the given development schedules. To shorten computation times, the heat transfer analysis solver THESEUS-FE from ARRK Engineering GmbH was used to model thermal fluid heat transfer coefficients and heat radiation. 3D models were first created that reflect the specifications of the different curing ovens at each plant, making it possible to recreate actual paint curing conditions and thereby grasp the spatial distribution and changes over time of the vehicle body temperature. From the analysis results, temperature profile data for each mesh node of the vehicle body model were output and used in the next process as input load data for the final structural analysis of the vehicle body.

(6) Structural analysis reflecting manufacturing history

Since the distortions from the manufacturing processes discussed in (1) to (5) are cumulative, it is necessary for CAE analysis to carry over the history from each process. Therefore, a script was created that brought together the data formats from all the analysis files formats; namely, the .d3 plot file from the stamp process, the M01 file from the welding, and .odb files from the paint curing and final structural analysis. The implicit solver Abaqus, from Dassault Systèmes SE, was used for the stress analysis of the final vehicle body structure.

In this way, the accumulated distortions of the vehicle body design specification DPM, input materials properties, and manufacturing processes (stamp forming, spot welding, and paint curing) were linked as factors to create an integrated simulation that encompasses all development processes.

3.1.3. Verification approach and measurement method construction

On actual manufacturing lines or verification-purpose lines, a prototype vehicle was used to measure factors relating to distortion, and the results compared to the simulation for verification.

(1) Verification environment

The verification experiments were conducted at the Yorii Plant, which is a global baseline standard for plants. An SUV Type 1 model with a large outer panel surface area was used as the prototype vehicle to represent a model progressing through the line.

The Yorii Plant mass-production stamp equipment was used for the stamping process. The welding process was performed on the Yorii Plant verification prototyping line and was equivalent to mass-production welding. For the painting process, the vehicle was passed through the curing ovens on the mass-production line.

(2) Measuring conditions

In order to clarify the amount of distortion within and between the processes, the 3D forms, engineering strain (hereafter, "strain"), sheet gaps, and temperature changes were monitored over time, and the correlations between the timing and location of plastic distortion occurrences verified. Figure 4 shows the verification method. Non-contact 3D shape measurements were taken before and after each manufacturing process to determine the detailed deformation behavior of the actual vehicle. In the welding process, the gaps, and the changes in the gaps, between the outer and inner panels were measured with feeler gauges during clamping, welding, and unclamping. To ascertain the amount of distortion continuously over time between the welding and paint curing processes, three-dimensional deformations were measured from surface-distributed distortions

detected with optical fibers. In the paint curing process, thermocouples were placed on the outer and inner panels, as well as the stiffeners located between them, in order to evaluate the impact on distortion of temperature differences between the inner and outer panels. During these measurements, the optical fibers and thermocouple measuring instruments, and the data loggers were positioned inside specially fabricated heat-resistant boxes, enabling measurement of changes in temperature and distortions over time at approximately 200°C within the vehicle curing ovens.

3.1.4. Verification results by manufacturing process

(1) Stamp process

Side panel deformation during stamp forming was measured using a scribed circle as shown in Fig. 5 (the distortion was calculated from the changes in the long and short axes before and after deformation). The reduction ratio of the sheet thickness was also measured. Figure 6 shows a comparison of the results for the actual measured sheet thickness reduction ratio and the simulation of section 3.1.2.(3). The difference between reduction ratios was found to be less than 6.6%, indicating good correlation.

○ : Measured, — : Not measured

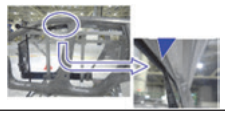
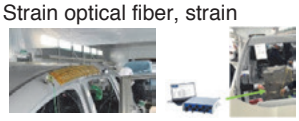
Verification test methods, measurement data	Process		
	Stamping	Welding	Painting
Optical 3D scan, 3D shape 	○	○	○
Feeler gauge and gap width 	—	○	—
Strain optical fiber, strain 	—	○	○
Thermocouple, temperature 	—	—	○

Fig. 4 Verification test methods for vehicles in factory

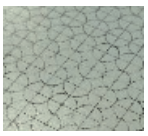


Fig. 5 Measurement method of stamping strain (scribed circle)

(2) Welding process

With regard to changes to strain due to the spot-welding order, which has a large impact on outer panel surface distortion, Fig. 7 shows a comparison of the results between the measurements from the optical fibers and the simulation described in section 3.1.2.(4). The welding spot timing and the changes in strain were found to correlate well at the roof, roof side, and sidesill evaluation locations.

(3) Paint oven curing process

For the painting process simulation described in

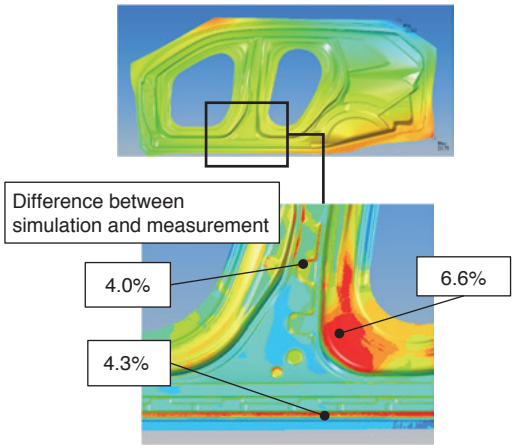


Fig. 6 Simulation of outer panel stamping and comparison of thickness results between simulation and measured reduction in sidesill area

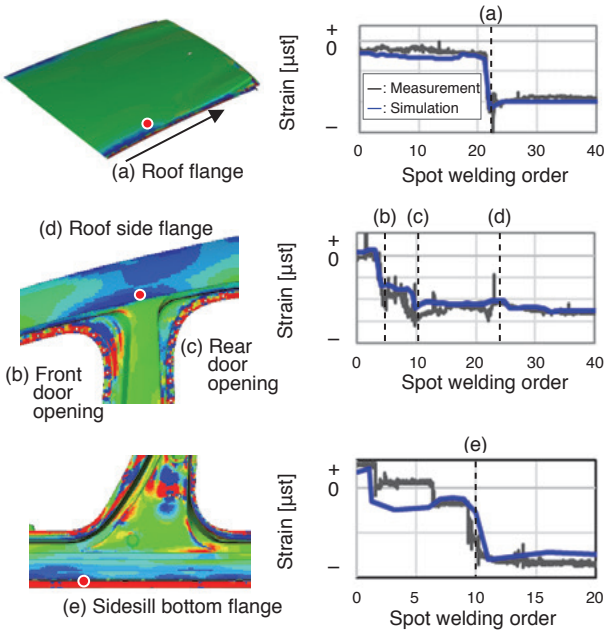


Fig. 7 Comparison of strain results between simulation and measurement along spot welding order

section 3.1.2.(5), Fig. 8 shows a comparison between the calculated and measured sidesill outer panel temperatures, under the temperature conditions and inlet flow rate of the electrodeposition curing process at the Yorii Plant. It was found that the temperature profiles correlated well. Figure 9 shows the difference in temperature between the sidesill outer panel and the stiffeners on the inner side at those locations, as well as the strain measurements made using optical fibers. It was shown that the maximum absolute value for strain variation coincided with the largest temperature difference. From this, it was inferred that plastic deformation occurs when the stress caused by thermal expansion differences during manufacturing exceeds the yield stress.

(4) Structural analysis verification reflecting manufacturing history

To verify the effectiveness of distortion analysis reflecting manufacturing history, the actual measured displacement from the initial sidesill design surface to the post-paint curing state was compared to the simulation results. This comparison is shown in Fig. 10. Compared to the actual measurement (a), simulation (b), which does not consider certain parts of the manufacturing history, showed much less deformation. On the other hand, in simulation (c), which considers the entire manufacturing history, the same deformation in the negative direction (indentation) was visible in the same location in both the measurements and simulation, indicating good correlation.

Next, Fig. 11 shows a comparison of the sidesill distortion amount over time, between the actual measurement taken using optical fibers during paint oven curing and the calculation. The stress value obtained by multiplying the measured strain value and the modulus of elasticity of the outer panel steel sheet material was plotted on the vertical axis, showing the region

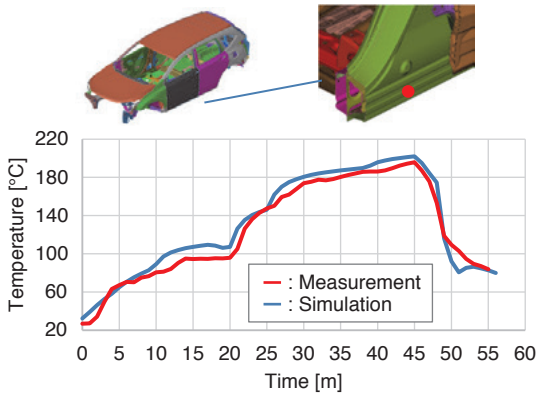


Fig. 8 Comparison of temperature results between simulation and measurement during paint oven process

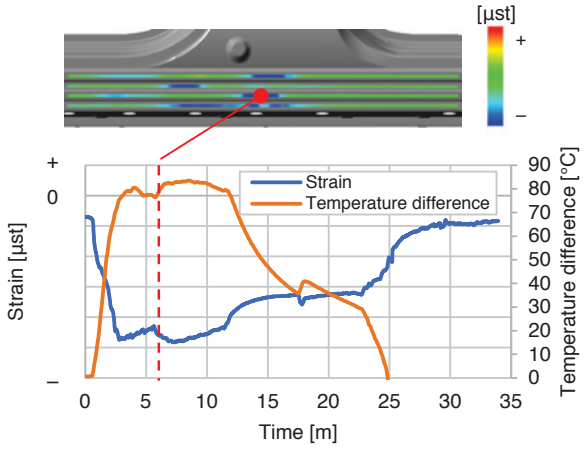


Fig. 9 Measurement results of strain on sidesill outer and temperature difference between sidesill outer and stiffeners

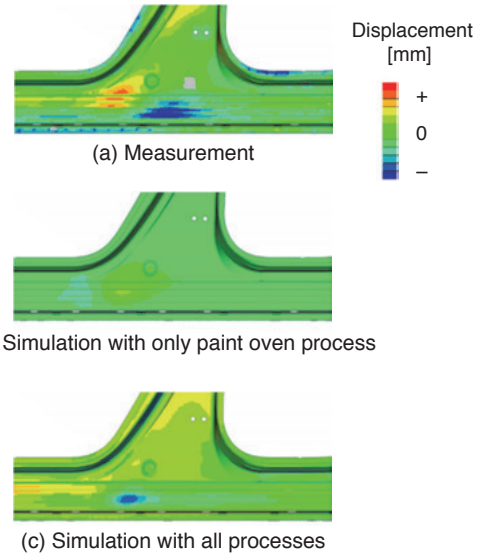


Fig. 10 Comparison of displacement results between measurement and simulation with processes on sidesill

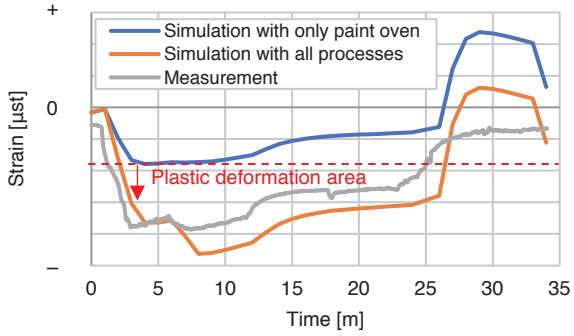


Fig. 11 Comparison of strain results on sidesill between measurement and simulation during paint oven process

where plastic deformation occurred as yield strength was exceeded. Distortion grew as heating began, and measurements showed that plastic deformation occurred after several minutes. In the simulation results, the simulation that did not consider the paint oven curing process showed that stress was within yield strength and no plastic deformation occurred. On the other hand, when the manufacturing history was taken into account, the results showed the same tendency as the measurements, where yield strength was exceeded and plastic deformation occurred.

As the above shows, this simulation method, which integrates the manufacturing processes, was able to predict both the timing of plastic deformation and the amount of distortion that is shown in the measurements, thereby confirming its utility.

3.2. Predicting Feasibility Using Integrated Simulations to Establish Specifications to Suppress Outer Panel Surface Distortion

Simulations were run for specifications aimed at suppressing strain that occurs under various conditions that impact outer panel surface strain distortion of the roof, sidesill, and upper center pillar of a vehicle model. In order to determine the feasibility of specifications, base conditions and variation ranges were compared for panel thickness and outer panel surface curvature in the design specification, differing thickness reduction ratios in the stamp forming process, panel gaps at welding spots during the welding process, and heating rate in the paint curing process. Figure 12 shows that for the sidesill, the impact of sheet thinning during stamp forming and panel gaps during welding was relatively minor. However, when comparing the base condition with the upper-limit condition for heating rate during paint curing, a significant difference was observed. Here, when measurements were made on actual panels, the results showed that under the baseline heating rate for paint curing, the same as used in the simulation, no distortions occurred. On the other hand, when the upper-limit condition was exceeded, the distortion exceeded the criteria, correlating with the simulation results.

At the same time, Fig. 13 shows that a comparison of the simulation results (a) with the measurement results (b) showed a similar distribution trend in the curvature differences of the distortion. From these results, it was possible to reassess the range of variation and apply the optimal heating rate to the manufacturing conditions.

Next, we performed simulations for the roof panel. Figure 14 shows that under the baseline conditions, distortion increased during welding, but the effect of the paint curing process was relatively small. Also, an analysis of the upper-limit of welding condition variation showed that distortions increase, but that even adding paint curing conditions at the upper-limit of the range had only a small impact. This indicated that the welding process had the greater impact on distortion.

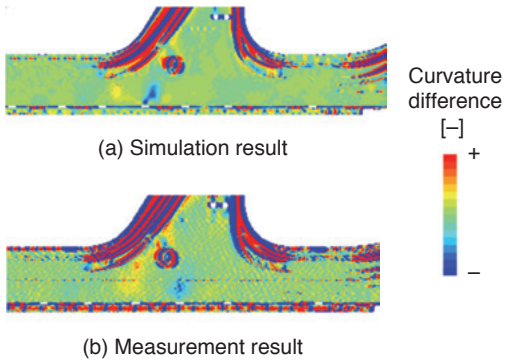


Fig. 13 Simulation and measurement results of curvature difference with all processes on sidesill

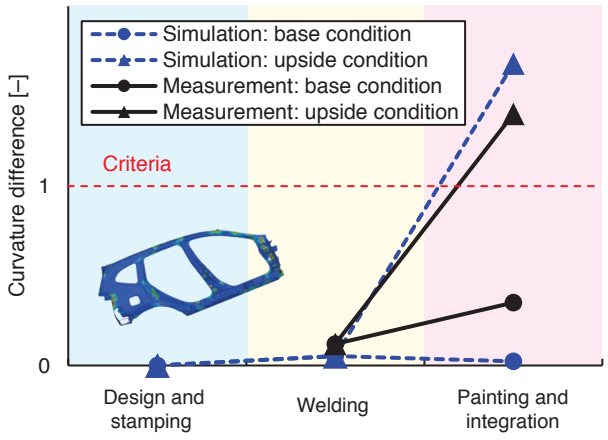


Fig. 12 Comparison of curvature difference results on sidesill

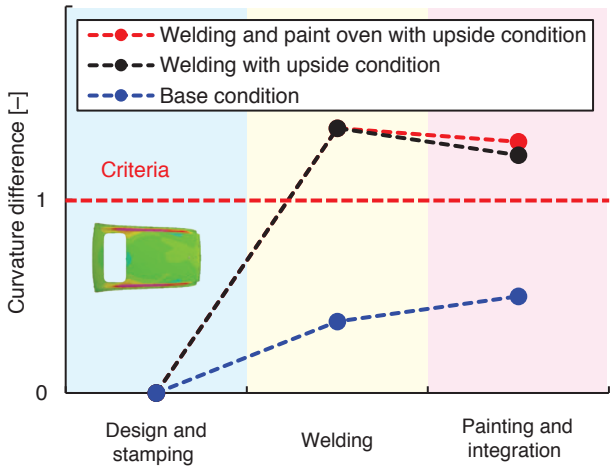


Fig. 14 Simulation comparison results of curvature difference on roof panel

As shown in Fig. 15, the larger the gap between panels at the spot-welding points, the greater the deformation of the rectangular structure. It is thought that this causes excess stress on the roof surface and thus distortion to occur. By modifying the welding conditions and reflecting the optimized manufacturing condition, it was possible to suppress these outer panel surface distortions.

It was also found that the dominant factors that cause outer panel surface distortions vary depending on the parts involved.

From the above, it was possible to predict the feasibility of specifications by referencing the design specifications for the outer panels of a vehicle and taking into account the accumulated distortion from each manufacturing process. It also became clear that the sensitivity to and interactions of influencing factors differ depending on the component or part, and an end-to-end integrated simulation method is effective for suppressing occurrences of distortion among the complex conditions of an entire vehicle.

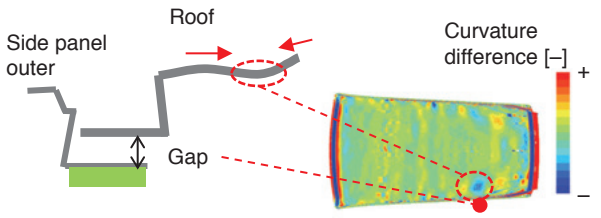


Fig. 15 Simulation result of curvature difference due to gap on roof panel

4. Conclusion

In order to achieve highly stylized exterior designs for automobiles, we developed a simulation to predict, at the design stage, the occurrence criteria for outer panel surface distortion. By incorporating this technology into the development management steps for new models, we were able to end the need for reworking in response to outer panel surface distortion occurring at plants around the world, and thereby increase design development efficiency.

(1) We developed a CAE technology that carries over accumulated distortion and stress data, enabling end-to-end evaluations of vehicle body materials, design specifications, and manufacturing processes (stamp forming, spot welding, and paint curing).
The results of verifications on actual vehicles enabled us to predict the occurrence timing of plastic deformation.

(2) Using the integrated simulation method that we developed, we were able to understand the mechanisms under which different factors have large impacts on distortion, depending on the outer panel part. By optimizing design specifications and manufacturing conditions that constitute these factors, we were able to suppress the occurrence of distortion.

References

(1) https://www.nvidia.com/en-us/learn/learning-path/digital-twins/#source=solr, (accessed 2025/7/1)
(2) Fukumoto, K., Takaki, Y., Iwase, T.: Thermal Deformation Analysis for Automotive Panel Design, Research and development, Kobe Steel engineering reports, Vol. 62, No. 2, p. 29-33, (2012) (in Japanese)
(3) Hamada, K.: FEM Analysis for Thermal Distortion on the Outer Panel, NIPPON STEEL TECHNICAL REPORT, No. 122, p. 88-92, (2019), https://www.nipponsteel.com/en/tech/report/pdf/122-14.pdf, (accessed 2025/7/1)
(4) Kim, M. S., Lee, M. S., Park, S. C., Jeong, J. Y., Shin, Y. C., Shin, S. C., Joo, G. H., Ko, S. Y., Jang, J. H., Lee, I. H., et al.: Analysis of Car Body Performance and Thermal Distortion in Paint Drying Oven when applying a Laser Brazed Aluminum Roof, 2020 JSAE Congress (Autumn) Proceedings, (2020)
(5) Kashiyama, T., Ohori, T., Shiohara, M., Murano, Y., Matsumoto, T.: Measurements of Car Body Deformation in ED Paint Tank and Oven by Water- and Heat-Resistant Displacement Meter and Comparison with CAE, Transactions of Society of Automotive Engineers of Japan, Vol. 53, Issue 4, p. 821-826, (2022) (in Japanese), doi: 10.11351/jsaeronbun.53.821
(6) Cai, W., Bondy, M., Carlson, B., Baylis, M.: Prediction of automotive body-in-white distortion in paint baking process, Manufacturing Letters, Vol. 41, p. 270-280, (2024), doi: 10.1016/j.mfglet.2024.09.032
(7) Toyooka, Y., Furusawa, T., Yokoi, E., Takami, S., Uchiyama, M., Tanaka, Y.: Development of Evaluation Method for Automotive Outer Panel Surface Distortion Using Curvature Difference and Application to Body Design, Honda R&D Technical Review, Vol. 38, p. 117-122
(8) Ito, T., Ishii, S., Honda, H., Sugimoto, N., Toyooka, Y., Furusawa, T.: Material Database for Efficient Development Using Materials Informatics, Honda R&D Technical Review, Vol. 32, No. 2, p. 110-117

■ Author ■


Toru FURUSAWA


Yoichi TOYOOKA


Eiichiro YOKOI


Shigeru TAKAMI


Masao UCHIYAMA


Yoshinori TANAKA

Development of Evaluation Method for Automotive Outer Panel Surface Distortion Using Curvature Difference and Application to Body Design

Yoichi TOYOOKA[1] Toru FURUSAWA[1] Eiichiro YOKOI[1]
Shigeru TAKAMI[2] Masao UCHIYAMA[2] Yoshinori TANAKA[2]

ABSTRACT

A quantitative evaluation method that correlates with human visual inspection is important in helping to predict outer panel surface distortion that may occur in the development phase of the automobile manufacturing process. This study proposes a quantitative evaluation method for distortion using curvature difference, in order to determine specifications through predictive simulations in the development phase. Curvature refers to the rate of change of the slope of a curved surface. The degree of distortion can be quantified by defining the difference in curvature between the standard shape and distorted areas as the curvature difference. The proposed method achieved reliable correlation between outer panel surface distortion and visual evaluation. Therefore, by incorporating curvature difference as one criterion in design phase simulations, it can be utilized as a consistent evaluation index from development to the manufacturing site. This paper reports the details and effectiveness of this evaluation method.

1. Introduction

Initiatives using digital development environments to shorten development periods and enhance quality and performance in automobile development are currently accelerating. In body development, these initiatives include defining specification requirements through design drawings, establishing control standards in the manufacturing process, and conducting simulations for a wide range of performance criteria such as rigidity, crash safety performance, strength, durability, and NVH (noise, vibration, harshness). Simulations have also been introduced to manufacturing processes such as press forming, welding, and painting, while developments in model-based development (MBD) and digital twin technology are further advancing simulation technology for body development. The authors previously developed a simulation technology capable of predicting outer panel surface distortion[(1)] by having design, manufacturing, and materials departments cooperate and handle factors related to outer panels in an integrated manner. However, visual evaluation is a key element of distortion inspection at manufacturing sites, and the remaining issue was that dimension-based prediction results from simulations did not match on-site visual judgments in some cases. This made it necessary to develop a method capable of quantitatively evaluating distortion without relying on visual evaluation, and to incorporate that method into simulations.

Human vision is sensitive to disturbances in surface smoothness and continuity, as well as discontinuous and sudden changes in curvature. For example, the sense of incongruity when the reflection of light changes in an unnatural manner, or when a continuous image bends, is due to the sudden change in curvature that negatively impacts the smoothness perceived by humans[(2), (3)]. Especially on glossy surfaces such as automotive outer panels, even slight unevenness can change the reflection of light, greatly impacting visual evaluation. Past methods for quantification and evaluation of distortion are listed below, but each faces challenges, and they have not been put to practical use in Honda body development.

[1] Material Research Center
E-mail: yoichi_toyooka@jp.honda
[2] Honda Motor Co., Ltd.

(1) Evaluation based on dimensional differences through non-contact 3D measurement[4)-6)]

While this method is able to capture fine surface waviness with high precision, correlation between visual evaluation of light waviness and bending by the human eye and evaluation by non-contact 3D measurement has not been demonstrated, making it a challenge to set distortion amounts that match visual evaluation.

(2) Evaluation using curvature[7), 8)]

Similar to the evaluation method in (1) above, the distribution of surface distortion is quantitatively evaluated by measuring the 3D shape of the outer panel surface and calculating the surface curvature and changes in the curvature. However, due to their design, the design surfaces representing the final product shapes described in the product drawings addressed in this paper inherently have large curvatures or changes in curvature. Therefore, small changes in the curvature due to outer panel surface distortion become buried in the curvature of the original design surface, making it a challenge to extract only the curvature of distorted areas as extreme values.

(3) Line-scan measurement using a profiler

Differences in the height of the surface over a certain measurement length are continuously measured using contact or non-contact surface profilers, and utilized to quantify the surface distortion. However, measurement as a surface is not possible, making it a challenge to evaluate wide areas such as actual panels.

Therefore, the goal of this study was to develop a new quantitative evaluation method in which the prediction results of outer panel surface distortion simulations match on-site visual evaluation.

2. Development Goals

To evaluate outer panel surface distortion, it is necessary to establish criteria to serve as judgment standards, and this calls for quantification of outer panel surface distortion. This paper focuses on a quantification method that examines the difference in curvature between the design surface and the outer panel surface after deformation in the body manufacturing process. The reason that humans are able to visually perceive distortion is that the reflection changes unnaturally, the result of a sudden change in curvature partway along a smooth transition in curvature. Therefore, it was thought that areas with distortion could be effectively extracted by identifying the difference between the original curvature and the curvature that contains distortion. First, human visual evaluation was correlated with the calculated curvature difference values for actual vehicle outer panels. Next, the design specifications were investigated in simulations, using the quantitative index established based on this correlation as criteria. Lastly, validity was verified by comparison with actual vehicles manufactured using design specifications that are effective against distortion.

3. Method for Quantification of Outer Panel Surface Distortion

The principle of using curvature difference to quantify outer panel surface distortion is described below. When a certain curve on the xy coordinate plane is represented as $y = f(x)$, the x direction curvature κ is defined by Eq. (1). Labelling the curvature of the design surface as κ_1 and the curvature of the surface that contains distortion as κ_2, the difference between the two given by Eq. (2) is defined as the curvature difference $\Delta\kappa$.

$$\kappa = \frac{f''(x)}{(1+f'(x)^2)^{3/2}} \tag{1}$$

$$\Delta\kappa = \kappa_1 - \kappa_2 \tag{2}$$

Figure 1 shows an example using a simple curve. Figure 1(a) is the curve intended by the design, and Fig. 1(b) is a curve distorted in one area compared to the design curve. The two curves are nearly indistinguishable on the graphs. Next, Fig. 1(c) and (d) respectively show the curvature distribution of both curves, and here too, the presence of distortion cannot be determined from the curvature distributions. However, when the difference in curvature distribution is obtained and plotted as shown in Fig. 1(e), the extreme value of the difference in curvature distribution appears at the distortion location. If it can be demonstrated that the calculated curvature difference results and human visual evaluation are correlated, this would enable quantitative evaluation.

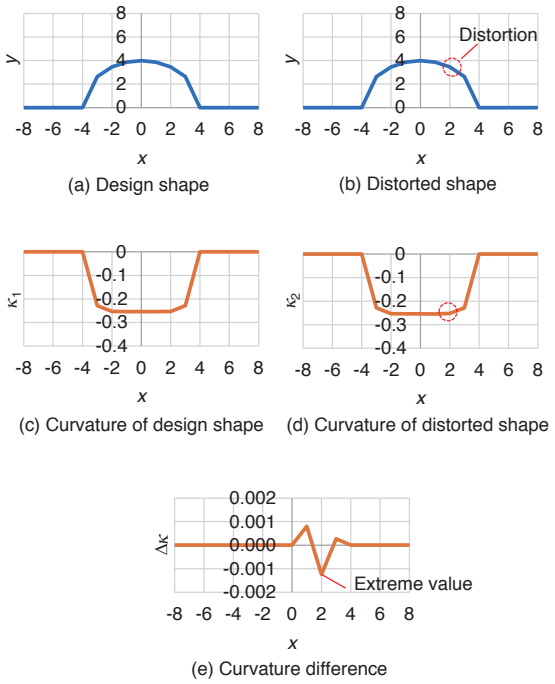


Fig. 1 Example of distortion extraction using curvature difference

4. Proof of Principle Using Actual Vehicles

4.1. Method for Calculating Curvature Difference Using Actual Vehicle Measurement Data

The curvature difference of the outer panels of actual vehicles was calculated to investigate the correlation between quantitative values based on the curvature difference of outer panel surface distortion and visual evaluation. This paper used optical scanning by GOM GmbH (measurement device: ATOS 5X) to measure the 3D shape data of actual vehicles. The data was output as random mesh data with an average mesh size of 1 mm. Figure 2 shows the curvature difference calculation flow. To calculate curvature difference from the outer panel shape data, the overall process was divided into four steps.

・Step 1: Acquisition of 3D mesh data

The panel surface to evaluate distortion was arranged parallel to the *xy*-plane as data on an *xyz*-axis 3D coordinate system, and the design data was converted from CAD data to mesh data.

Optical scanning measurement data was used as the 3D data of the actual vehicles.

・Step 2: Conversion to equal-interval mesh data

x and *y* coordinates were spaced at equal intervals on the random mesh data, and linear interpolation was performed for the *z* coordinates corresponding to those *xy* coordinates.

・Step 3: Calculation of curvature difference

The first and second derivatives in the *x*-direction of the design and measurement data were calculated using the central difference method, and the curvature κ_1 of the design data in the *x*-direction and the curvature κ_2 of the measurement data in the *x* direction were calculated. Finally, the difference between curvature κ_1 and curvature κ_2 was obtained and used as the curvature difference $\Delta\kappa$.

As described hereafter in Section 4.2., this paper focused on the roof panel, roof side, and side sill to confirm the correlation between visual evaluation and curvature difference. Therefore, the longitudinal direction of each area was aligned with the *x*-axis, and in consideration of past situations where distortion appeared, evaluation was performed only for the *x*-direction.

・Step 4: Mapping of curvature difference

The calculated curvature difference values were visualized by mapping to the mesh data coordinates set in step 2. In order to intuitively grasp the tendencies of the

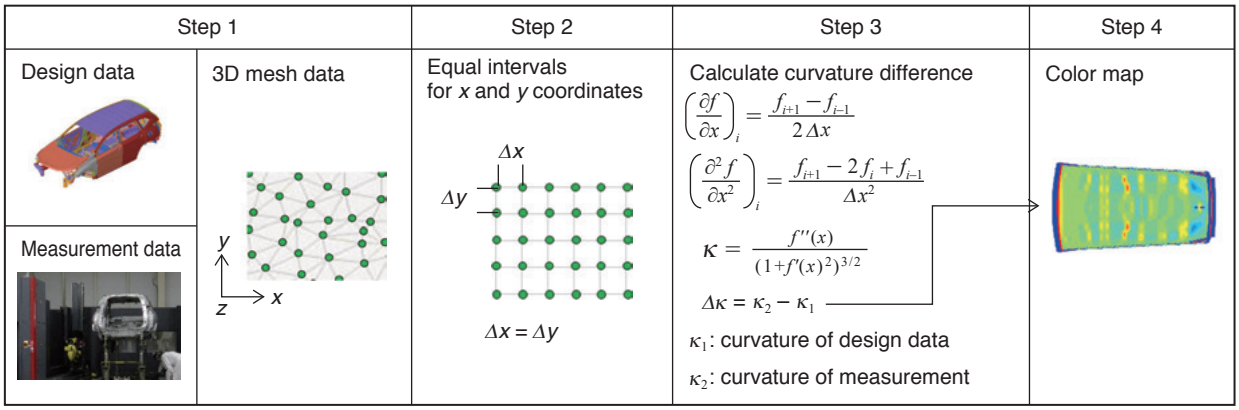
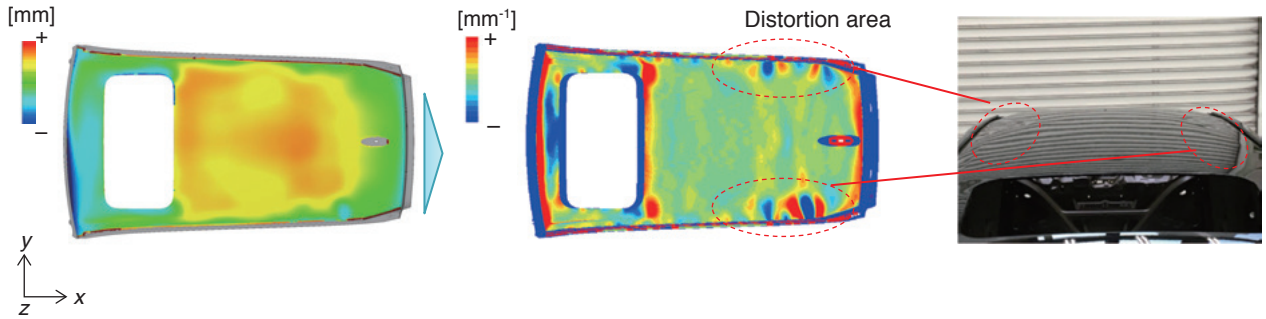


Fig. 2 Curvature difference calculation flow



(a) Dimensional difference (*z*-axis)

(b) Curvature difference (*x*-direction)

(c) Actual roof panel

Fig. 3 Example of dimensional difference and curvature difference in measurement data

deformation, the signs of the curvature difference were defined as positive in the convex direction and negative in the concave direction relative to human vision, and displayed using a color map.

Figure 3 shows the results of calculating and color mapping the dimensional difference and curvature difference of an actual roof panel. The displacement in the *z*-axis positive direction at the center of the roof panel can be confirmed in (a), but there are cases where areas with large dimensional difference may not be visually recognized as distortion. This is because while the dimensional difference indicates the displacement between the design and measurement data, local deformation and waviness resulting in visual unnaturalness also exist as small displacement, and that small displacement becomes buried, making it a challenge to capture. In contrast, the waviness at the roof edges is clearly extracted in the curvature difference (b), and was confirmed to be consistent with the distortion areas in the photograph (c). This showed that curvature difference is highly correlated with visual evaluation, and is an effective evaluation index close to visual inspection.

4.2. Verification of Correlation between Visual Evaluation and Curvature Difference

To confirm the correlation between visual evaluation and curvature difference, verification was conducted using five sample vehicles prepared with varying degrees of distortion. The prepared vehicle types included two sedans and three SUVs with large outer panel areas prone to distortion. For visual evaluation, the points where the curvature difference values are extreme values were extracted in advance for the three areas of the roof panel, roof side, and side sill, as shown in the examples in Fig. 4. Those points were visually scored using the 10-point scale shown in Table 1, and a score of 6 or more was defined as borderline acceptable for a mass-produced product. Evaluation was conducted in an inspection environment with fluorescent lights arranged at equal intervals to facilitate detection of distortion by changes in light reflection. 55 expert inspectors scored the extracted points, and the average value was used as the score. These scores were compared with the curvature difference values at each point to confirm the correlation.

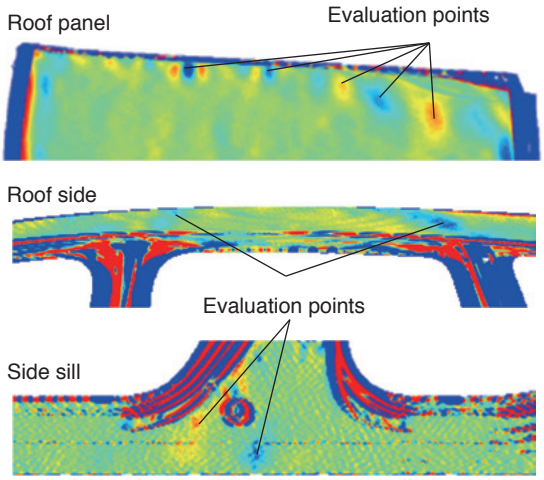


Fig. 4 Example images of visual evaluation points

Table 1 Visual evaluation scores

Score	Judgement	Reactions
10	EXCELLENT	No distortion
9	VERY GOOD	Undetectable even by experts
8	GOOD	Detectable by experts
7	ACCEPTABLE	Recognized but acceptable by experts
6	BORDERLINE	Acceptable but noticeable if customers pay attention
5	UNACCEPTABLE	Detectable by customers
4	UNACCEPTABLE	Complaint
3	POOR	Strong complaint
2	VERY POOR	Repair request
1	N/A	Rejection

Figures 5, 6, and 7 show the respective confirmation results for the roof panel, roof side, and side sill. The vertical axis of the graphs represents the visual evaluation, and the horizontal axis represents the curvature difference. The relationship between the average value of the visual evaluation scores and the curvature difference at the points shown in Fig. 4 is plotted for each vehicle type. Regarding the side sill in Fig. 7, the sedan models use resin garnish parts, so these were not evaluated, and only the SUV plots are shown.

The results clearly show a tendency in which the lower the visual evaluation score, the greater the absolute value of the curvature difference for any outer panel area, confirming the expected correlation. The acceptable criteria range differed for each of the roof panel, roof side, and side sill, which is thought to be due to differences in the curvature of the design for each area. When the curvature of the design surface is small, that is to say a smoother surface, it is easier to visually detect distortion even if the curvature difference is small. Conversely, when the curvature of the design surface is large, it may be a challenge to visually detect distortion even in case of a large curvature difference. Based on this result, human vision and a quantitative index of distortion in actual vehicles can be correlated by setting the range of curvature difference resulting in a visual evaluation score of 6 points or more as the criteria for distortion. These criteria enable comparison using simulations, as well as investigation of potential factors in the event of issues, and accurate predictions when making design changes.

These criteria were established to cover the outer panels of a wide range of existing models, and are considered capable of accommodating even the development of new models. However, efforts are needed to further enhance accuracy by accumulating more detailed data in preparation for the appearance of greatly differing designs as bodies evolve going forward.

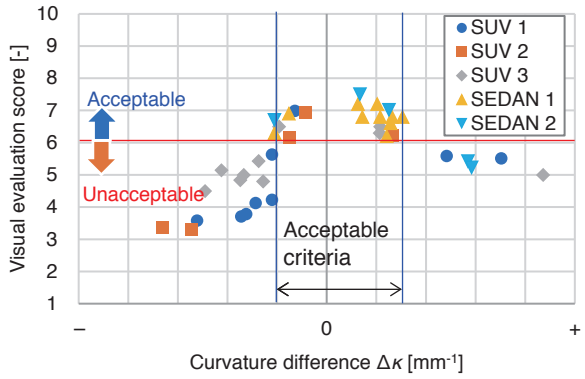


Fig. 5 Correlation between visual evaluation scores and curvature difference of roof panel

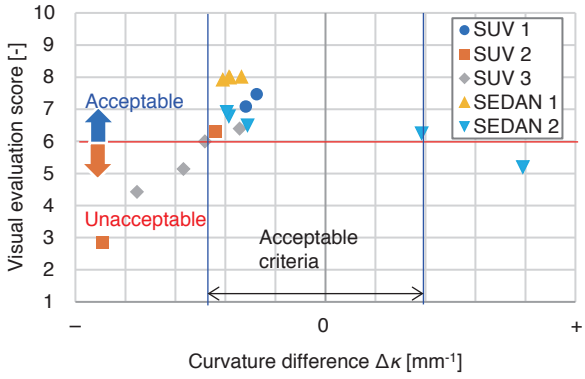


Fig. 6 Correlation between visual evaluation scores and curvature difference of roof side

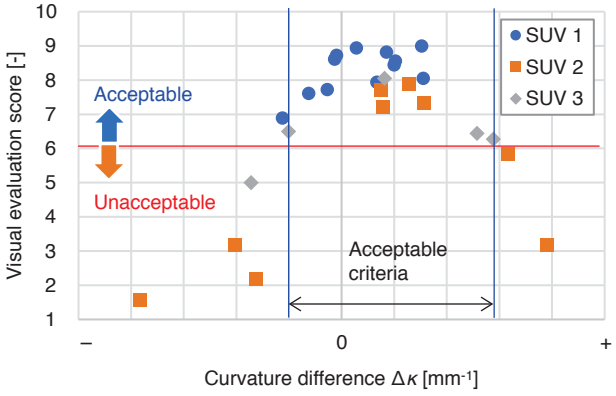


Fig. 7 Correlation between visual evaluation scores and curvature difference of side sill

5. Application to Simulation

Figure 8 shows an example of the design process for reducing distortion in the roof of the 2022 model year CR-V. The roof edges have spot welds between the body frame and the roof panel, making these locations relatively susceptible to distortion. Figure 8(a) and (b) compare the results of curvature difference calculated by predictive simulations of outer panel surface distortion based on two design proposals: a first design and a resistance-enhanced design. In the first design (a), welding of the weld flanges resulted in displacement and distortion in the visible area of the roof, and the curvature difference in the red areas did not meet the acceptable criteria. As the design for enhancing resistance (b), the dimensions of the two cross-sectional shapes of the visible surface and the vertical face were considered within the range that does not negatively impact other design requirements such as press formability, functionality, and aesthetics. For the visible surface, the curvature was reduced and the rigidity was increased. For the vertical face, a two-step shape was established and the dimensions were set so that deformation in welding does not easily extend to the design surface. As a result, the curvature difference values were confirmed to be within the acceptable criteria.

Figure 9 shows the curvature difference measurement results for the actual vehicle roof. The design with enhanced distortion resistance (b) was confirmed to fall within the acceptable criteria range.

The above results show that by quantifying distortion using curvature difference, it is possible to demonstrate congruity between predictive simulations of outer panel surface distortion in the design phase and the actual body.

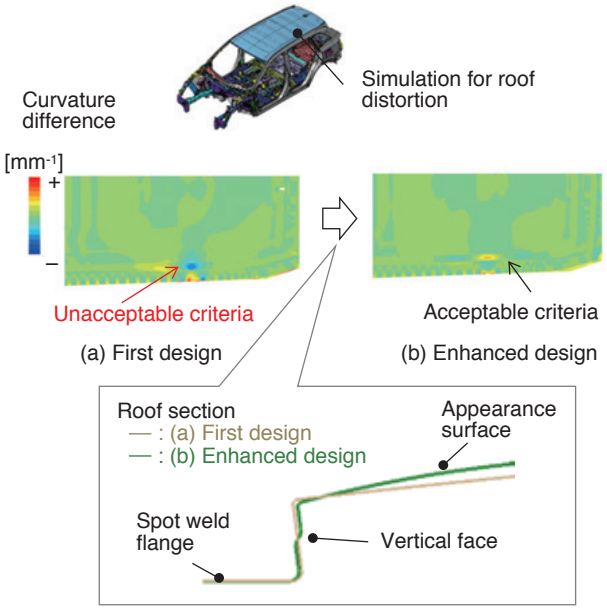


Fig. 8 Design process to reduce distortion in roof

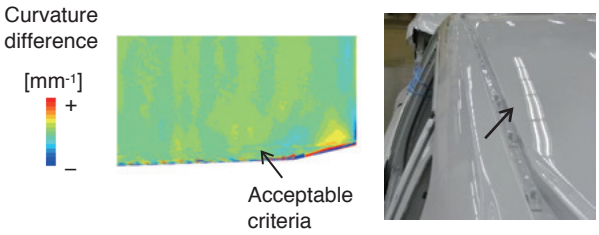


Fig. 9 Measurement of actual roof curvature difference

6. Conclusion

Curvature difference was used to establish a quantitative index correlated with human vision in order to achieve a quantitative evaluation method for distortion, which is needed to develop simulation technology for body outer panel surface distortion. Establishment of a calculation process that enables the calculation of differences in curvature between actual measurements and simulation mesh data made it possible to verify specifications using simulation data. This calculation process can be used together with predictive simulations of outer panel surface distortion in actual development sites, and will enable the development of technology for enhancing the efficiency of vehicle body development.

References

(1) Furusawa, T., Toyooka, Y., Yokoi, E., Takami, S., Uchiyama, M., Tanaka, Y.: Integrated Simulation Technology for Design, Manufacturing, and Materials to Reduce Surface Distortion of Automotive Outer Panels, Honda R&D Technical Review, Vol. 38, p. 109-116
(2) Sande, G., Yoshida, S., Sakurai, T., Igarashi, M.: Jiyu Kyokusen, Kyokumen no Namamekasa no Hyoka Hoho nitsuite, Journal of Graphics Science of Japan, Vol. 38, p. 129-134, (2004) (in Japanese), doi: 10.5989/jsgs.38.Supplement1_129
(3) Ohsaki, M.: Fairness and Mechanics of Curves, traverse: kyoto university architectural journal, Vol. 4, p. 47-54, (2003) (in Japanese), doi: 10.14989/traverse_4_47
(4) Fukumoto, K., Takaki, Y., Iwase, T.: Thermal Deformation Analysis for Automotive Panel Design, Research and development, Kobe Steel Engineering reports, Vol. 62, No. 2, p. 29-33, (2012) (in Japanese)
(5) Hamada, K.: FEM Analysis for Thermal Distortion on the Outer Panel, NIPPON STEEL TECHNICAL REPORT, No. 122, p. 88-92, (2019), https://www.nipponsteel.com/en/tech/report/pdf/122-14.pdf, (accessed 2025/7/22)
(6) Kashiyama, T., Ohori, T., Shiohara, M., Murano, Y., Matsumoto, T.: Measurements of Car Body Deformation in ED Paint Tank and Oven by Water- and Heat-Resistant Displacement Meter and Comparison with CAE, Transactions of Society of Automotive Engineers of Japan, Vol. 53, Issue 4, p. 821-826, (2022) (in Japanese), doi: 10.11351/jsaeronbun.53.821
(7) Yonemura, S., Yoshida, T., Uenishi, A., Hiwatashi, S., Kosugi, S., Nakagawa, J.: Advanced Technologies Related to Simulation and Evaluation for Performance of Automotive Exterior Panels, Nippon Steel Technical Report, Vol. 393, p. 11-17, (2012) (in Japanese), https://www.nipponsteel.com/tech/report/nsc/pdf/393-03.pdf, (accessed 2025/7/22)
(8) Ichijo, N., Iwata, N., Iwata, T., Mita, T., Tsutamori, H.: Development of Surface Evaluation Method and Stamping Simulation for Surface Deflection of Automotive Outer Panels, Journal of the Japan Society for Technology of Plasticity, Vol. 54, Issue 628, p. 421-425, (2013), doi: 10.9773/sosei.54.421

■ Author ■


Yoichi TOYOOKA


Toru FURUSAWA


Eiichiro YOKOI


Shigeru TAKAMI


Masao UCHIYAMA


Yoshinori TANAKA

Short-time Prediction Tool for Flow Fields around Side Mirrors Using Pix2Pix GAN

Takuro HASHIMOTO[1] Koji OKABE[1]
Takeshi YAMATODA[2] Takafusa KIDO[2]

ABSTRACT

To examine the design direction of side mirrors in the early stages of development, a Pix2Pix GAN deep learning model was used to develop an easy-to-operate tool that enables prediction of flow fields with sufficient accuracy in a short period of time and output of the results. In addition to a shorter prediction time with a deep learning model, further time reduction was achieved by omitting the signed distance function for shape recognition that has a high computational load. The tool was also developed to maintain prediction accuracy even with a small amount of training data. Furthermore, output of fluid physical quantities required for verification on the tool operation screen was minimized, allowing performance of the process from input to output solely by button operations, making it easy to use for designers.

1. Introduction

Automobile side mirrors are parts used to understand the situation to the rear of the vehicle. As they are mounted on the exterior of the body, their part shape can disrupt the surrounding airflow at high speeds, potentially causing the mirror to vibrate and reduce rear visibility. Therefore, computational fluid dynamics (CFD) is used to predict performance and examine shapes to help minimize the potential impact of this flow. While the evolution of graphics processing units (GPU) continues to shorten CFD calculation times, it still requires around three days to calculate the flow velocity (three velocity components) around the side mirror, the surface streamlines, and the surface pressure distribution.

The side mirror shape affects the vibration of the mirror surface, and also impacts the styling. To achieve a good balance between functionality and styling, designers and styling designers should ideally be able to discuss matters on the spot. CFD calculations currently take time, however, making it necessary to take designs back and conduct performance predictions. Due to the inability to reconcile the optimal shape determined by designers as well as the design desired by styling designers during that period, inconsistencies in direction often lead to reworking, which becomes a bottleneck for shortening the development period.

To address this issue, it was thought that if there was a tool capable of performing prediction in a short time of around five minutes, designers and styling designers would be able to repeat trial and error while viewing the prediction results, enabling determination of a shape with good balance during meetings.

Active research is underway on surrogate models based on deep learning as a method for performing prediction in a short time, and many reports have indicated its effectiveness in shortening prediction times[(1)].

Surrogate models using deep learning generally adopt an architecture based on CNN autoencoder (CNN-AE) that combines a CNN encoder and CNN decoder that use convolutional neural networks (CNN) as shown in Fig. 1[(2)]. However, CNN-AE is known to have the disadvantage that detailed information is lost as noise in the encoding (compression) process for feature extraction, resulting in blurred shape outlines. This is addressed by converting shape data to a signed distance function (SDF) for input to the CNN encoder. SDF is a function that provides geometric information about a shape and includes codes for calculating the distance from a certain point or determining which points are inside and which are outside the model

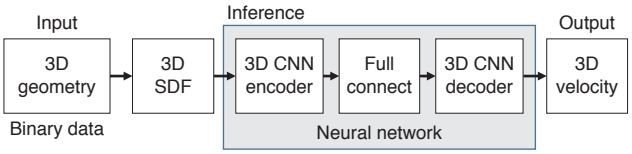


Fig. 1 Basic architecture of previous study

[1] Honda Motor Co., Ltd.
E-mail: takuro_hashimoto@jp.honda
[2] IDAJ Co., LTD.

surface. While this enhances accuracy, there is also the disadvantage that time is needed to calculate the distances for all points every time in both training and inference.

It is also generally known that data on the order of 10^3 is needed to provide sufficient accuracy in development of deep learning models, and Akasaka et al. conducted research with a similar condition[1]. This presented a challenge for this study, as using CFD results to prepare this order of training data would take some five months.

An additional challenge is that the display method and operability need to be considered when designers use prediction tools. Since discussion time is limited, it is more important to be able to swiftly and easily display only the information needed for consideration, rather than confirming events from multiple perspectives.

To address these challenges, this study examined the following three points. First, to shorten the prediction time, the use of deep learning and the feasibility of omitting SDF were examined. Second, a deep learning model capable of providing prediction accuracy with a small amount of training data was investigated. Third, minimization of the outputs required for verification and an operation screen that enables operation solely by button operations were investigated to help prevent designer operations from becoming complicated.

2. Selection of Deep Learning Model and Development of Prediction Tool

To address the above challenges, the optimal deep learning model was selected and pre-checked, and a display screen that enables easy output by designers was investigated.

2.1. Shorter Prediction Time

To shorten the prediction time, this study also applied a surrogate model based on deep learning. Pix2Pix Generative Adversarial Networks (Pix2Pix GAN)[3] was adopted as that deep learning model, and further speeding up by omitting SDF was also examined. Figure 2 shows the GAN that forms the basis of that algorithm. GAN is a generator training method proposed by Goodfellow et al.[4], and consists of generator and discriminator neural networks. The generator generates data using random noise as the input, and the discriminator judges whether that data is real or fake. Training proceeds by having the generator and discriminator compete to enhance the quality of the generated data (adversarial training). As a result, GAN are characteristic in that training continues until sharp outputs are achieved, as opposed to CNN-AE, which may produce blurred shape contours.

An extension of GAN is the conditional generative adversarial network (conditional GAN: cGAN) shown in Fig. 3 proposed by Mirza et al.[5] This enables specification of conditional inputs to the generator, making it possible to enhance the prediction accuracy by training a generator with outputs corresponding to arbitrary inputs.

Pix2Pix GAN is a further extension of cGAN and is widely used as an image generation algorithm. By learning the relationship between a pair of images, it can generate another image that reflects that relationship based on a single image. 2D images are commonly used as conditional inputs, but to predict the spatial flow velocity (three velocity components) and pressure around a side mirror, the architecture is extended to three dimensions and STL-format 3D data is used instead of conditional images. While there is a case of using 3D expansion for brain tumor segmentation by Cirillo et al.[6], there are no known cases of application to fluid prediction. Thus, a generator that inputs voxel data of shapes and outputs fluid physical quantities was used. Furthermore, a major feature of Pix2Pix GAN is the use of U-Net proposed by Ronneberger et al.[7] as the generator. Figure 4 shows the architecture used in this study, which is an extension of the U-Net CNN layer to three dimensions. U-Net is an architecture consisting of an encoder and a decoder. Features extracted by the encoder

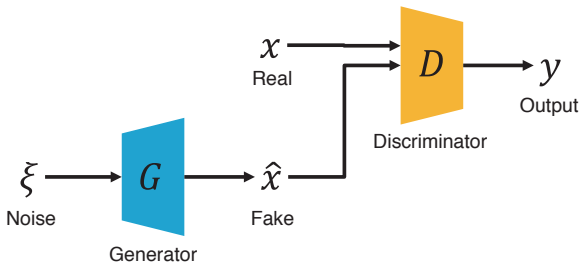


Fig. 2 GAN

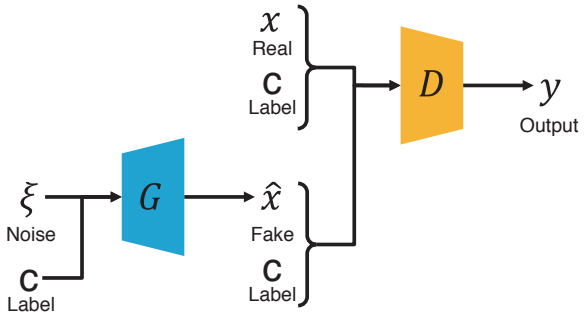


Fig. 3 Conditional GAN

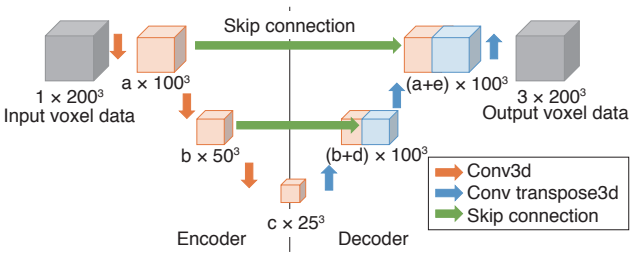


Fig. 4 Extension of U-Net to 3D

are restored to the original-size image by the decoder, and the skip connection between the encoder and decoder enables transmission of detailed information to the decoder layers of the same resolution without the loss of any data. With this architecture, spatial information and resolution details that tend to be lost by the encoder can be transmitted directly to the decoder, enabling retention of even fine features that might be lost in the CNN-AE encoding process. It was examined whether use of these GAN and U-Net would enable shortening of prediction times while maintaining prediction accuracy, even without use of SDF.

Figure 5 shows the architecture examined in this study. The inference part used a Pix2Pix GAN. A method that omits the SDF that was required before inference in CNN-AE and inputs the voxel representation of the shape data directly to the generator was examined while verifying the prediction accuracy.

2.2. Provision of Prediction Accuracy with a Small Amount of Training Data

Typically used CNN-AEs are trained with small data units called mini-batches, but as they are unable to consider the entire data distribution, a large amount of training data is needed to obtain high accuracy.

In contrast, Pix2Pix GAN is an effective deep learning model that can achieve high prediction accuracy even with a small amount of training data. The reasons given for this are that adversarial training enables the generative model to grasp the entire data distribution[4], and the U-Net employed in the generator enables high-accuracy prediction. The combination of these factors enables training of the generator to achieve high prediction accuracy using less data than for CNN-AE.

This study set target values for prediction accuracy and the number of training data in order to verify the prediction accuracy with a small amount of training data. As the accuracy of CFD using Reynolds-averaged Navier-Stokes (RANS) is about ±10%, the target prediction accuracy was set to a discrepancy of ±10% or less. The pressure coefficient C_p of the side mirror is in the range of ±0.3, so this target was set to ±0.03 or less, and the target relative to the mainstream velocity of 38.8 m/s was set to ±3.8 m/s (±13.6 km/h). The target amount of training data was set to 100 or less.

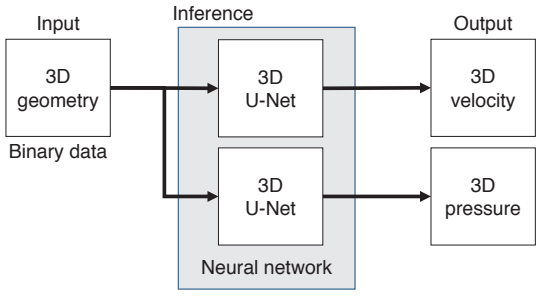


Fig. 5 Proposed architecture

Table 1 Deep learning conditions

	CNN-AE	Pix2Pix GAN
Number of training data	32,500	2,000
Using SDF	Yes	No

Accuracy was verified using the mean squared error (MSE) given by Eq. (1).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \quad (1)$$

Here, Y_i and $\hat{Y}_i$ respectively indicate the CFD results and the results estimated by the prediction tool, and n indicates the number of voxel points.

From this equation, the accuracy targets are an MSE of 750 or less for pressure, and 14.4 or less for velocity.

2.3. Verification of Accuracy Using 2D Images

In developing this tool, the validity of the Pix2Pix GAN prediction accuracy was verified in advance by comparing Pix2Pix GAN with the CNN-AE developed by Guo et al. 2D velocity distribution images calculated for various shapes using the general-purpose CFD software ANSYS Fluent by Synopsys Inc. were used as the training data. Table 1 shows the learning conditions. CNN-AE used 32,500 images as training data, and also converted them to SDF. Pix2Pix GAN used 2,000 images selected randomly from among those 32,500 images as training data, and omitted SDF. Figure 6 shows scooter-shaped binary data as an example of input in prediction, and the SDF calculated from that data. Figure 7 compares the prediction results of the flow velocity U_x component (horizontal direction in the figure) under these conditions. Both CNN-AE and Pix2Pix GAN produced images of the scooter shape and the velocity distribution

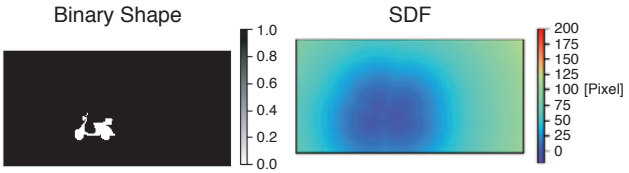


Fig. 6 2D input image

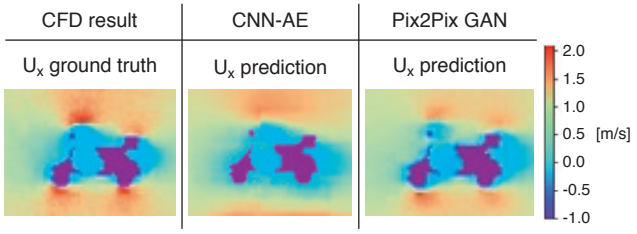


Fig. 7 Comparison between CNN-AE and Pix2Pix GAN

around the scooter with little blur similar to CFD. Both were also able to qualitatively reproduce the relative velocities and corresponding areas. This result indicated that by adopting Pix2Pix GAN instead of CNN-AE, the amount of training data can be reduced and the flow velocity near the body within the flow field can be predicted even without SDF.

2.4. Achievement of an Operation Screen Enabling Easy Output

Since the operation screen needs to be easy to operate for designers, it was designed so that all operations can be performed using only buttons. The buttons needed to perform all operations from loading the input STL data to output of the prediction results are arranged in order, enabling performance of all operations on a single screen. The outputs were narrowed down to the cross-sectional velocity distribution and the fluid information (surface pressure distribution, flow velocity distribution, and streamlines) for the side mirror surface shown in Fig. 8, making it easy to select and display the desired data using only button operations.

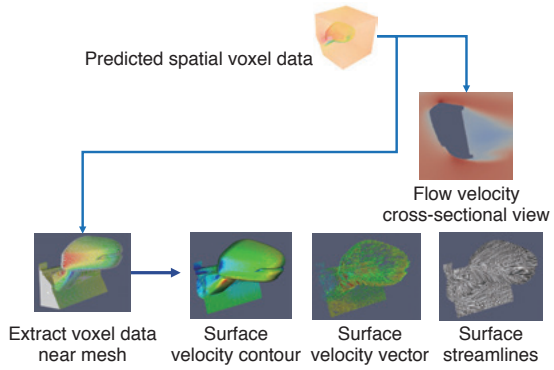


Fig. 8 Mapping from spatial distributions to surfaces

3. Training for Side Mirror Prediction

Training to predict the pressure and flow velocity distributions around the side mirror was conducted using the tool described in Chapter 2.

3.1. Training Process

Figure 9 shows the training process for spatial pressure distribution in the proposed Pix2Pix GAN model used in this study. First, the voxel representation of the side mirror shape was applied to the generator as conditional input data. Here, a pair consisting of the spatial pressure distribution predicted by that generator and the side mirror shape, and a pair consisting of the CFD result (true values) and the side mirror shape were generated. Next, the discriminator performed real/fake judgment using those two pairs as input, and then adversarial training was conducted based on that judgment. Training was also performed similarly for the 3D spatial velocity distribution (X, Y, and Z directions). For reference, this model is capable of predicting spatial pressure and spatial velocity using only the generator when given the STL data of an unknown side mirror shape.

3.2. Generation of Training Data by CFD

The training data used CFD results for a vehicle equipped with side mirrors. Since application to various mirror sizes is assumed, the body with the largest mirror size was selected in order to encompass all mirror sizes. Eleven shape parameters representing the shape, position, and size features of the side mirror were set as shown in Fig. 10. Three levels were set for each shape parameter, and 101 cases of combinations were generated based on randomized design of experiments. Parametric data generation used the commercially available morphing tool DEP MeshWorks by Detroit Engineered Products, and CFD used ANSYS Fluent.

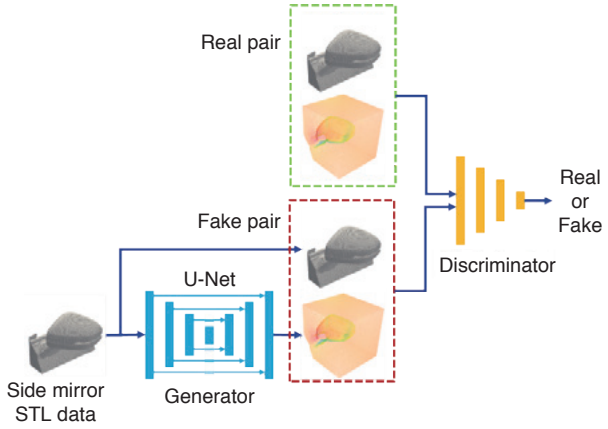
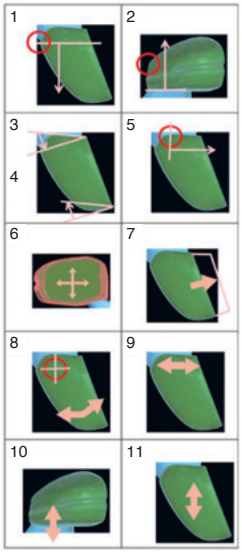


Fig. 9 Deep learning model for pressure distribution



No.	Parameters
1	Peak position (Horizontal direction)
2	Peak position (Vertical direction)
3	Inner plane angle
4	Outer plane angle
5	Leading edge angle
6	Mirror size
7	Length in mainstream direction
8	Mounting position (Rotation)
9	Mounting position (Mainstream direction)
10	Mounting position (Vertical direction)
11	Mounting position (Horizontal direction)

Fig. 10 CFD model creation for training data

Table 2 Mesh conditions

Number of mesh	50 million
Vehicle surface mesh size	4 mm
Circumference	8-16 mm
Boundary layer mesh	2 mm x 1 layer

Table 3 Analysis conditions

Turbulence model	Realizable k-ε
Speed	140 km/h (38.8 m/s)
Density	1.205 kg/m^3
Viscosity coefficient	1.8075e-5 Pa-s

Table 2 shows the CFD mesh conditions, and Fig. 11 shows the mesh placement. A hexahedral mesh was used as the overall base, and a single-layer boundary-layer mesh was positioned to reproduce only the surface shape. Table 3 shows the CFD analysis conditions. The turbulence model used RANS. Figure 12 shows an example of the CFD results.

3.3. Extraction of Training Data

Compared to the CFD results for the overall vehicle, the data volume was reduced, and speedup was achieved while enhancing the shape resolution of the side mirror model by narrowing down the prediction space that uses deep learning to the area around the side mirror. The specific prediction space that uses deep learning was a 400 mm cube encompassing the side mirror as shown in Fig. 13, and

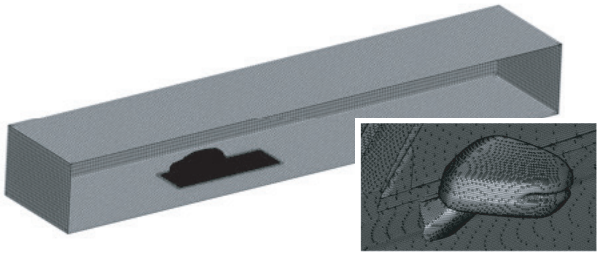


Fig. 11 Mesh placement

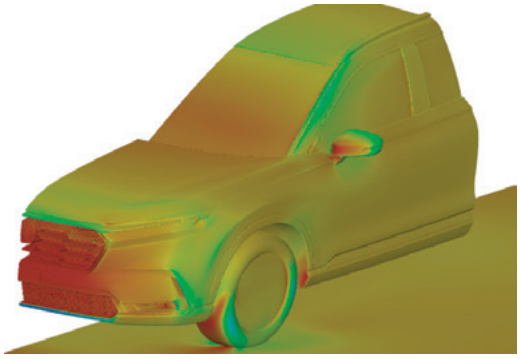


Fig. 12 Example of pressure distribution results using CFD

Table 4 Machine learning conditions

	Velocity	Pressure
Optimizer	Adams	Adams
Epoch	1000	1000
Batch size	1	1
Learning rate	2×10^{-4}	2×10^{-5}

voxels were arranged at 2-mm intervals within that cube. Conversion to the voxel data for the mirror shape shown in Fig. 14 used the vtkSelectEnclosedPoints algorithm from the open-source The Visualization Toolkit (VTK) library.

Figure 15 shows an example of a spatial pressure field used as training data. Conversion from CFD mesh to voxels was performed by substituting the average value of the 10 closest CFD mesh points to the center of each voxel.

3.4. Learning Conditions and Learning Time

Table 4 shows the learning conditions, with 80 of the prepared 101 cases used as training data. A single NVIDIA

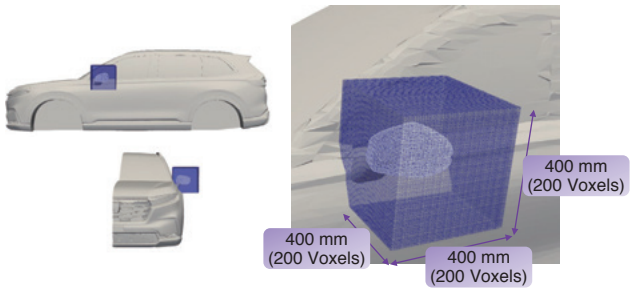


Fig. 13 Training data space for deep learning

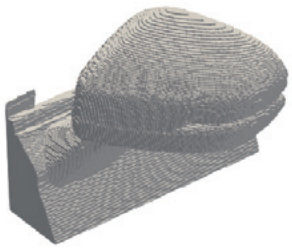


Fig. 14 Side mirror voxel model as noise and condition label

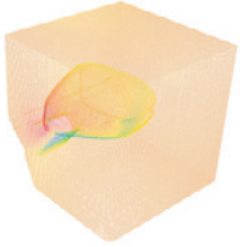


Fig. 15 Example of pressure distribution as training data

RTX A6000 GPU was used for training, and training took approximately 190 hours.

4. Verification of Prediction Tool Accuracy and Effect of Prediction Tool

Verification was conducted for the predicted results of spatial pressure distribution and spatial velocity distribution obtained using this tool, and the time-shortening effect when executed on the created operation screen was also confirmed.

4.1. Spatial Pressure Distribution Prediction Results

Figure 16 compares the CFD results and the surface pressure distribution obtained by projecting the inferred results of spatial pressure distribution onto the surfaces. Although the resolution decreases and unevenness appears in the pressure distribution due to the effects of voxelization, the locations where the pressure increases and decreases qualitatively matched the CFD results.

Figure 17 shows the MSE of the predicted spatial pressure distribution. The horizontal axis represents the ID number assigned to each mirror shape, and the MSE is plotted on the vertical axis. The white circles indicate the MSE for 80 cases of training data in the developed deep learning tool, and the black circles indicate the MSE for 21 cases of verification data. The MSE was about 160 for the overall training data, and 330 for the overall verification data, both of which were within the accuracy target of 10% or less. However, the MSE for the most severe case was 841, which slightly exceeded 10%.

4.2. Spatial Flow Velocity Distribution Prediction Results

Figure 18 compares the CFD results and the surface flow velocity distribution and streamlines obtained by projecting the inferred results of spatial flow velocity distribution onto the surfaces. Mirror vibrations originate from the surface velocity distribution and the direction of the streamlines, which are impacted by factors such as flow

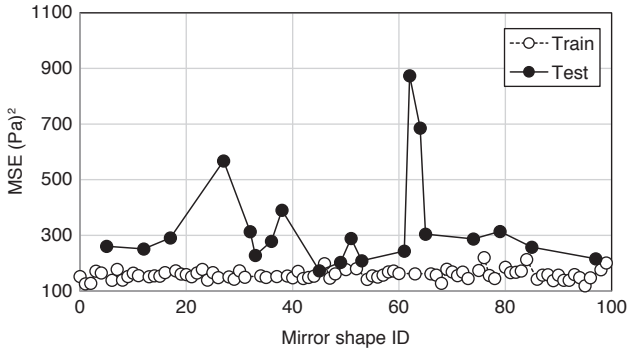


Fig. 17 MSE of inference of pressure distribution

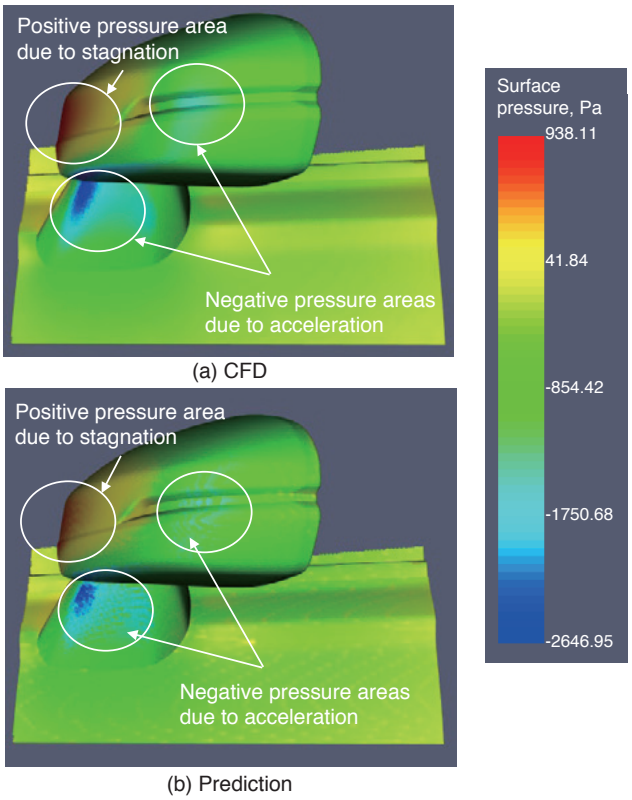


Fig. 16 Comparison of predicted and CFD pressure distribution

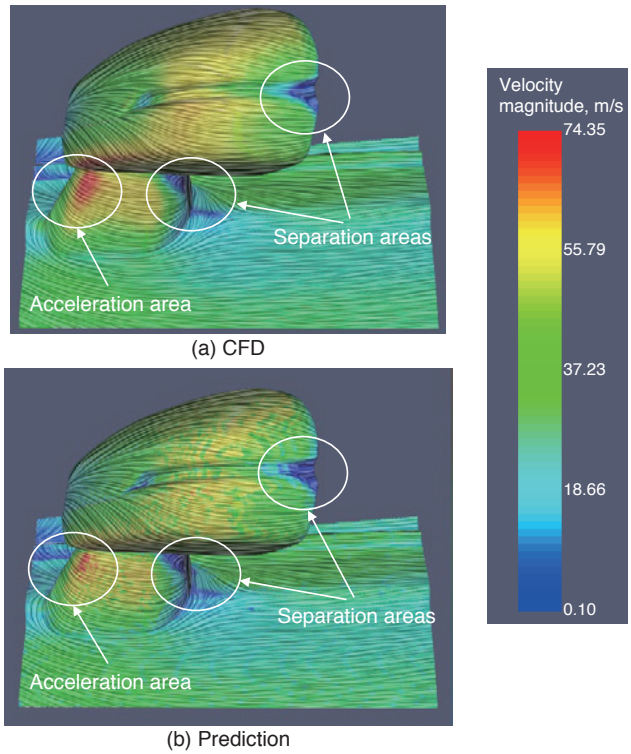


Fig. 18 Comparison of predicted and CFD velocity distribution

separation. From this perspective, the prediction results are generally consistent with the CFD results.

The above results show that the developed prediction tool provides sufficient accuracy even with 80 training data.

Figure 19 shows the MSE for the predicted spatial flow velocity distribution of each mirror shape. The MSE was about 0.1 for the overall training data (80 cases) and about 1.0 for the overall verification data (21 cases). Converted to velocity, the MSE for the overall verification data was about ±1 m/s, which was within the accuracy target of 10% or less. The MSE for even the most severe case was 4.9 (velocity ±1.1 m/s), which was also within 10% or less. These results show that sufficient accuracy can be provided when utilizing this tool, even with a small amount of training data.

4.3. Operation Screen

Figure 20 shows the operation screen created in this study. The buttons for the steps from loading the input STL to outputting the results are arranged in order, enabling operation on the same screen. The required output displays have been narrowed down, and all operations can be easily performed using only buttons.

Table 5 Computer specifications

	CFD	Deep learning
CPU	AMS EPYC7702	Intel Xeon w9-3475X 2.21GHz
Number of CPUs	128 core/node	1
GPU	-	NVIDIA RTX6000 Ada
Memory	512 GB/node	1 TB
OS	Red Hat Enterprise Linux release 8.3	Windows Server 2022

4.4. Calculation Time

Table 5 shows the specifications of the equipment used for each calculation when evaluating the prediction tool time. Additionally, Fig. 21 shows the overall workflow of this prediction tool. Prediction work previously that took 52 hours with CFD could be performed in approximately 100 s.

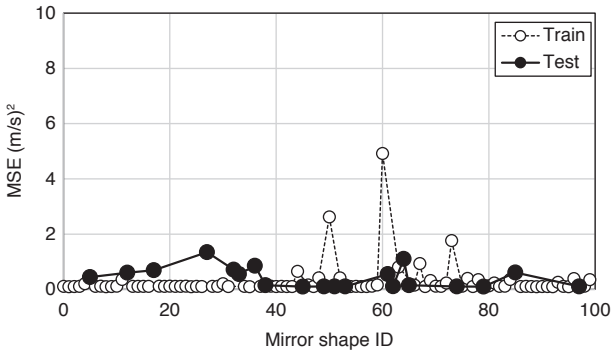


Fig. 19 MSE of inference of velocity distribution

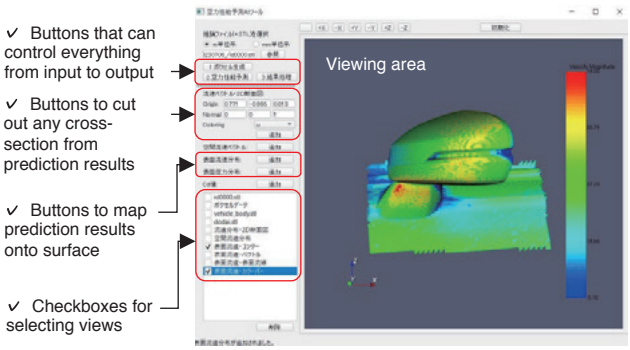


Fig. 20 Graphics of user interface

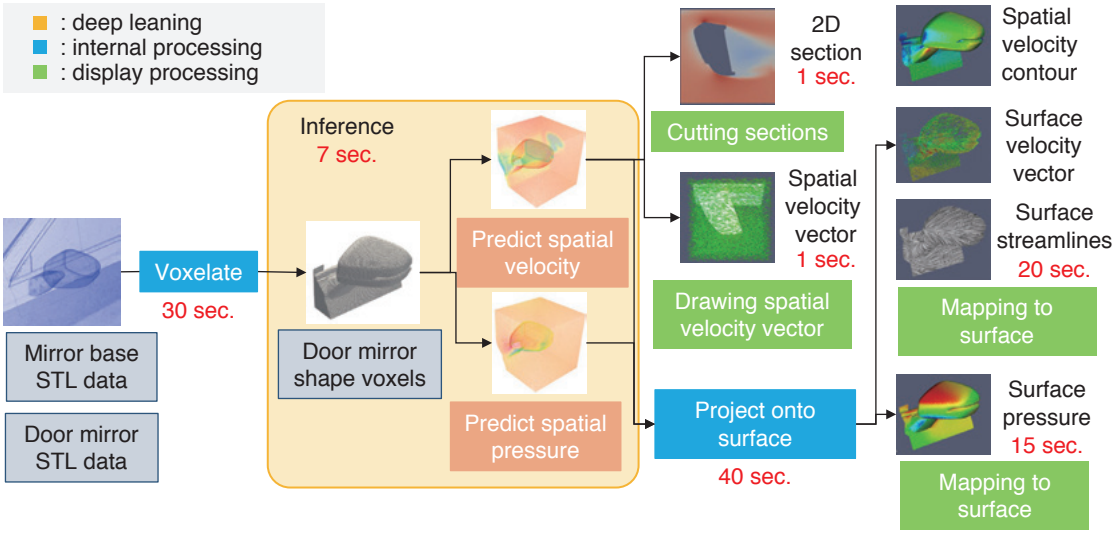


Fig. 21 Workflow of proposed system and total time

4.5. Summary

From the above, use of the developed prediction tool enables designers to make design decisions while exchanging opinions in real-time with styling designers.

Furthermore, continuing discussions will promote deeper exploration and refinement of ideas and stimulate both parties with differing perspectives, which can be expected to lead to the creation of various ideas.

5. Conclusion

This study used Pix2Pix GAN to develop a tool with the following characteristics for predicting flow fields around side mirrors.

(1) In addition to shorter prediction times using a deep learning model, the contributions of the GAN and U-Net architectures, which are characteristics of Pix2Pix GAN, achieved further time reduction by omitting SDF, which has a high computational load. This enabled completion of the entire process from loading STL data to results display in approximately 100 s.

(2) Generative models that use adversarial training are able to grasp the entire data distribution, enabling predictions with a smaller amount of data compared to conventional methods that are trained with mini-batches. As a result, predictions of pressure and flow velocity were within ±10% of the CFD results, even with the small amount of 80 training data.

(3) By minimizing the display of fluid physical quantities required for verification and indicating the functions on the buttons, a display screen was created that enables designers to easily select all operations using only mouse controls.

This method is useful for users who find it challenging to collect training data, and for users who are about to begin introduction and development of such systems. It is anticipated that, going forward, this method will also be widely utilized in other prediction tools.

References

(1) Akasaka, K., Chen, F., Teraguchi, T.: Surrogate Model Development for Prediction of Car Aerodynamics Using Machine Learning, Transactions of Society of Automotive Engineers of Japan, Vol. 52, Issue 3, p. 621-626, (2021) (in Japanese), doi: 10.11351/jsaeronbun.52.621

(2) Guo, X., Li, W., Iorio, F.: Convolutional Neural Networks for Steady Flow Approximation, KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p. 481-490, (2016), doi: 10.1145/2939672.2939738

(3) Isola, P., Zhu, J-Y., Zhou, T., Efros, A. A.: Image-to-Image Translation with Conditional Adversarial Networks, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 5967-5976, (2017), doi: 10.1109/CVPR.2017.632

(4) Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks, Communications of the ACM, Vol. 63, Issue 11, p. 139-144, (2020), doi: 10.1145/3422622

(5) Mirza, M., Osindero, S.: Conditional Generative Adversarial Nets, arXiv, (2014), doi: 10.48550/arXiv.1411.1784

(6) Cirillo, M. D., Abramian, D., Eklund, A.: Vox2Vox: 3D-GAN for Brain Tumour Segmentation, Lecture Notes in Computer Science, Springer, Cham, Vol. 12658, p. 274-284, (2021), doi: 10.1007/978-3-030-72084-1_25

(7) Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, Lecture Notes in Computer Science, Vol. 9351, p. 234-241, (2015), doi: 10.1007/978-3-319-24574-4_28

■ Author ■


Takuro HASHIMOTO


Koji OKABE


Takeshi YAMATODA


Takafusa KIDO

Quantitative Measurement Scale for Effectiveness of Adapted Vehicle

Arisa ENDO[1] Toshihiro KOYAMA[1]
Keiichi SUNAMI[1] Tetsuya GOMI, Ph.D.[1]

ABSTRACT

With a goal of contributing to the enrichment of customers' social lives through adapted vehicles, the authors studied a method for quantitatively expressing the effectiveness of the vehicles in order to understand their effect and communicate their value. Effectiveness was defined as the degree of each customer's engagement in their individual social activities after continued use of the adapted vehicle, with engagement determined using a scale based on customer responses to 17 5-point questions. The authors then used this scale to statistically verify a predictive value for the degree of involvement in social activities. The questionnaire was given to physically disabled drivers and drivers involved in caregiving services. The results indicated that highly accurate predictions could be made for both groups, with accuracy rates of 0.825 and 0.846, respectively. This confirmed that the scale developed for this study can quantitatively express the effectiveness of adapted vehicles.

1. Introduction

Under the slogan "Fun For Everyone. Honda - The Joy of Mobility for Everyone," Honda continues to provide adapted vehicles in order to help physically disabled drivers and drivers involved in caregiving services enjoy mobility and social activities(1).

Adapted vehicles can be broadly classified into driving aids vehicles and wheelchair accessible vehicles. Driving aids vehicles have driver assistance devices that enable drivers with lower limb disabilities to operate the gas pedal and brake with their hands, or drivers with upper limb disabilities to operate the steering wheel with their feet. Figure 1 shows an example of a driving aids vehicle. Wheelchair accessible vehicles are equipped with elevating seats and wheelchair ramps to reduce the labor required to assist wheelchair users in and out of the vehicle. An example is shown in Fig. 2.

Honda introduced their first driving aids vehicle in 1976, a sedan equipped with devices to assist people with upper limb disabilities in driving. It introduced a wheelchair accessible kei car in 1995.

Figure 3 shows sales volumes of adapted vehicles in Japan. The market for adapted vehicles grew more than tenfold between fiscal years 1994 and 2004, peaked in fiscal year 2014, and then began to shrink with each subsequent year. The market is currently less than 30,000 units. Although the adapted vehicle market is small compared to the approximately 4 million passenger cars sold in Japan, automakers have brought many different models to market. This suggests that automakers value the social responsibility side of their businesses adapted vehicles represent and that they strive to emphasize the value of adapted vehicles, even though such value cannot be measured by sales volume or



Fig. 1 Adapted vehicle with hand control device



Fig. 2 Adapted vehicle with rear ramp

[1] Honda Motor Co., Ltd.
E-mail: arisa_endo@jp.honda

sales revenue alone.

Driving aids vehicles are considered to provide value to customers with upper or lower limb disabilities by enabling them to act autonomously when they use these vehicles. Wheelchair accessible vehicles are also considered to give value to customers who provide care and transport to those they care for, and value to those receiving care by easing their mental and physical burden, providing increased freedom to carry out their daily activities and expanding opportunity in their lives.

Additional product enhancement and development is required for adapted vehicles to continue providing customers with the value. This requires information on the effect that adapted vehicles provide, but the scarcity of such information makes it challenging to understand the actual situation. For that reason, a suitable survey instrument needed to be developed.

To understand the effect adapted vehicles provide, qualitative perceptions of that effect need to be expressed in concrete terms so that there is no gap with the customer's perceptions. This is also desirable in that it could help to understand customer needs, enhance products and services, and communicate the value of adapted vehicles. This is due to growing interest in ESG management and the increasing importance of quantitatively disclosing the effectiveness of corporate activities for society[2].

Therefore, the authors attempted to define and quantify the effectiveness of adapted vehicles and to verify their usefulness in conjunction with previously proposed scales for measuring life satisfaction.

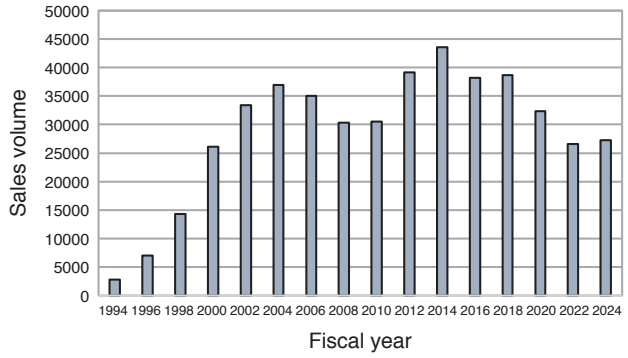


Fig. 3 Domestic sales volume of adapted vehicles

2. Study of Measuring Methods

2.1. Definition of Effectiveness of Adapted Vehicles

The effects of a company's business activities can be expressed over time using a logic model[3]. Additionally, individuals' life satisfaction and happiness have been found to be positively correlated with participation in social activities and a person's willingness to participate in such activities, and with social capital at the individual level[4]-[9]. For that reason, the authors viewed the effectiveness of adapted vehicles as the impact on society and customers resulting from business activities. The authors then used the introduction of adapted vehicles to the market as a substitute for this business activity and laid out the impacts of adapted vehicles over time. The results are shown in Table 1 as a logic model.

The columns in Table 1 represent the initial, mid-term, and long-term outcomes over time after market launch. The rows represent the local community and beneficiaries impacted by adapted vehicles.

Beneficiary 1 represents drivers with upper or lower limb disabilities who use driving aids vehicles. The impact of the driving aids vehicle was assumed to be that using the vehicle would enable unassisted excursions, social participation, and independence.

Beneficiary 2 represents drivers of wheelchair accessible vehicles who transport individuals in their care, such as relatives. The impact of the wheelchair accessible vehicle was assumed to be that using the vehicle would reduce mental and physical burdens and enable both caregiving and social participation.

Beneficiary 3 represents individuals receiving care who are passengers in wheelchair accessible vehicles. The impact of the wheelchair accessible vehicle was assumed to be that its use would reduce mental and physical burdens and increase opportunities to interact with society.

The local community represents the area in which the beneficiary lives. The impact of providing adapted vehicles was assumed to be that it would foster social capital by increasing the number of people who find it easier to move around and get in and out of vehicles and by revitalizing individual action and society. It was considered that, beyond the positive changes in human behavior, the result would be a virtuous cycle of social revitalization and individual well-being.

After examining Table 1, the authors defined the effectiveness of adapted vehicles as "the degree of activation of the customer's individual social activities before and after the use of an adapted vehicle." This would enable surveys to be conducted at individual companies, with the potential for quantification through customers who purchased adapted vehicles.

2.2. Quantification of Effectiveness

The authors proposed the Vitality scale to quantify the degree of activation of the above-mentioned social activities of individuals.

Vitality can be measured using a questionnaires on changes in the behavior and awareness of individuals directly affected by adapted vehicles. Table 2 lists the items covered in the questionnaires.

Respondents answered each question on a 5-point scale. The meaning of each numbered response was as follows: 1.

Table 1 Outcome stages of logic model

Beneficiary classification	Initial outcomes	Mid-term outcomes	Long-term outcomes
Local community	(1) More people can move by car. (2) More people find getting in and out easier.	More people participate in social activities. People are revitalized.	Social capital is fostered. Society is revitalized.
Beneficiary 1	(1) Able to go out by oneself.	(1) Actions become voluntary	(1) Achieve social participation and independent living.
Beneficiary 2	(2) Reduced mental and physical burden.	(2) Balance personal activities and caregiving.	(2) Continue social participation to utilize individual abilities.
Beneficiary 3		(2) Able to move around for daily needs.	(2) Increased opportunities to interact with society.

Virtuous cycle of well-being and social revitalization

Table 2 Vitality questionnaire

	Questions: Which of the following did you experience?
1	Increased shopping frequency.
2	Increased participation in lessons and clubs.
3	Increased participation in local festivals and events.
4	Increased participation in entertainment such as theater and sports events.
5	Increased opportunities to go out for a change of pace.
6	Increased frequency of meetings with friends and acquaintances.
7	Increased frequency of meetings with relatives and family.
8	Increased contact with neighbors.
9	Increased number of new acquaintances.
10	Income has become more stable.
11	More options to earn income.
12	More ability to be considerate of others.
13	Increased interest in new things.
14	Increased opportunities to feel moved with others (or to want to feel moved).
15	Increased motivation to take action.
16	Increased emotional ease and peace of mind.
17	Increased moments of joy (or wanting to feel joy).

Strongly agree; 2. Somewhat agree; 3. Neutral; 4. Somewhat disagree; and 5. Strongly disagree. A response of 1 or 2 was counted as an increase in opportunities to participate in social activities and was given a score of 1 point, while a response of 3 or more was counted as no change and was given a score of 0 points. The degree of activation towards social activities was then quantitatively evaluated based on the total score (out of 17).

2.3. Verification Method

2.3.1. Scale for comparative verification

Three scales were used for comparison against Vitality: the Philadelphia Geriatric Center Morale Scale (PGC)[10]-[12], the Satisfaction With Life Scale (SWLS)[13], [14], and the Organisation for Economic Co-operation and Development Subjective Well-Being - Subjective Assessment of Life in General (OECD)[5].

These are well-known scales to quantify subjective life satisfaction and happiness.

Table 3 Interpretation of correlation coefficient R

Correlation coefficient R	Level
$1.0 \geqq R \geqq 0.7$	Very high
$0.7 \geqq R \geqq 0.5$	High
$0.5 \geqq R \geqq 0.4$	Moderate
$0.4 \geqq R \geqq 0.3$	Fair
$0.3 \geqq R \geqq 0.2$	Low
$0.2 \geqq R \geqq 0.00$	Very low

2.3.2. Correlation analysis

Four online questionnaires (PGC, SWLS, OECD, and Vitality) were administered to the Beneficiary 1 and Beneficiary 2 groups, who use adapted vehicles.

Since the PGC is a questionnaire for people 65 years of age and older, the PGC responses were used with Beneficiary 1 and 2 individuals aged 65 or older.

A total of 118 individuals responded. Of these, 23 of the Beneficiary 1 group were 65 or older and 23 of the Beneficiary 2 group were 65 years or older, while 36 of the Beneficiary 1 group were 64 or younger and 36 of the Beneficiary 2 group were 64 or younger.

The correlations between PGC, SWLS, and OECD against Vitality were determined from the score of each scale as found from the questionnaire results. The evaluation criteria for social surveys[15] shown in Table 3 were used to determine the correlations.

2.3.3. Predictive value validation method

A total of five online questionnaires were administered: the PGC, SWLS, OECD, and Vitality questionnaires about changes in individual behavior and awareness, and a questionnaire about increased involvement in social activity. To ask about the facts concerning whether there was increased social activity, as a direct measure, the authors used a representative question on social activities, namely whether "participation in volunteering, NPOs, and civic activities has increased"[4].

The target survey respondents were the same as in the correlation analysis. However, the questionnaire results described in the previous section were used for PGC, SWLS, OECD, and Vitality. The authors used logistic regression analysis to verify whether the scores from each scale derived from the online questionnaire results could predict responses regarding "whether there was increased participation in volunteering, NPOs, and civic activities." based on the scores obtained by each scale.

The software R version 4.5.0[16] was used to calculate and analyze scores from the scales, and a prediction model was developed using "whether they had opportunities to participate in volunteering, NPOs, and civic activities" as the dependent variable and scores for Vitality, etc. as explanatory variables. The logistic curve representing the prediction model is the standard sigmoid function shown in Eq. (1) and a typical curve is shown in Fig. 4.

$$q_i = \frac{1}{1 + e^{-z_i}} \tag{1}$$

Here, q_i is the incidence rate and e is the natural logarithm. The variable z_i is a linear predictor consisting of the sum of n explanatory variables x_n, partial regression coefficients β_n, and constant term β_0, as shown in Eq. (2).

$$z_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_n x_n \tag{2}$$

A prediction model that forms a monomial in which each score value is substituted into x_1 in Eq. (2) was verified.

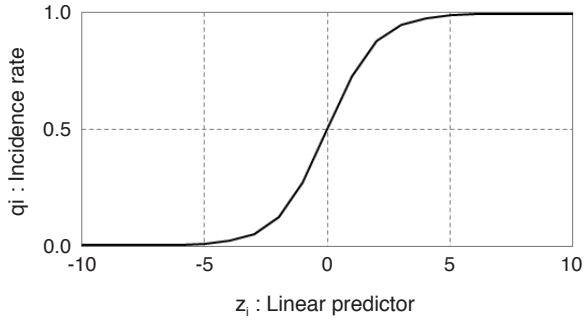


Fig. 4 Diagram of standard sigmoid function

2.3.4. Method of verifying proposed scale accuracy

To verify the accuracy of the Vitality model, 196 individuals from the Beneficiary 1 group and 319 individuals from the Beneficiary 2 group, for a total of 515 individuals, completed two online questionnaires on Vitality and "whether they had opportunities to participate in volunteering, NPOs, and civic activities." Of the 515 individuals, the responses of 59 individuals in Beneficiary 1 group and 59 individuals in Beneficiary 2 group were based on the results of the online questionnaire described in the previous section.

First, between-group differences between those who had increased opportunities to participate in social activities and those who experienced no change were compared to see whether Vitality could be used to determine whether social activities had increased or not based on the results of this online questionnaire. Next, the authors verified the accuracy of the prediction model using the additional 397 individuals as training data.

3. Results and Discussion

3.1. Correlation Analysis

The correlation coefficients between Vitality and the three existing scales of PGC, SWLS, and OECD are shown in Table 4 for each Beneficiary group. The results show a weak correlation between Vitality and OECD, but little correlation with the other scale indicators. Therefore, the authors considered that Vitality can be treated as a scale independent of the existing scales.

3.2. Predictive Value Validation

Table 5 shows the data obtained from logistic regression analysis. p-values of less than 0.05 are considered statistically significant. No individuals in the Beneficiary 2 group aged 65 or older who responded to PGC had increased opportunities to participate in social activities, so these were excluded from the analysis.

The PGC p-value for the Beneficiary 1 group and the SWLS and OECD p-values for the Beneficiary 2 group exceeded 0.05 and thus were not statistically significant. This suggests that these scales are inadequate for distinguishing increased opportunities for social participation for the Beneficiary 1 and Beneficiary 2 groups overall.

On the other hand, Vitality was statistically significant for both the Beneficiary 1 and Beneficiary 2 groups, indicating that this scale can distinguish "whether there was an increase in opportunities to participate in volunteering, NPOs, and civic activities" and can quantify the degree of activation of an individuals' social activities.

Figure 5 shows the logistic curves for each scale and the Beneficiary 1 group. The horizontal axis of the graph shows the score for each scale, and the vertical axis shows the results from the questionnaire on "whether there was an increase in opportunities to participate in volunteering,

Table 5 Logistic regression statistics

Items			Estimate (β_0, β_1)	Standard error	z value	p value
Beneficiary 1	PGC	(Intercept)	-7.924	5.810	-1.36	0.173
		PGC score	0.5718	0.5773	0.990	0.322
	SWLS	(Intercept)	-2.999	1.024	-2.93	0.003
		SWLS score	0.09952	0.05028	1.98	0.048
	OECD	(Intercept)	-3.290	1.020	-3.33	0.001
		Subjective happiness value	0.3208	0.1368	2.35	0.019
	Vitality	(Intercept)	-4.116	1.021	-4.03	$p < 0.001$
		Vitality	0.3430	0.09196	3.73	$p < 0.001$
Beneficiary 2	SWLS	(Intercept)	-1.337	1.114	-1.20	0.230
		SWLS score	-0.01302	0.05514	-0.236	0.813
	OECD	(Intercept)	-1.261	0.9454	-1.33	0.182
		Subjective happiness value	-0.05803	0.1383	-0.368	0.713
	Vitality	(Intercept)	-4.721	1.262	-3.74	$p < 0.001$
		Vitality	0.3591	0.1069	3.36	$p < 0.001$

Table 4 Correlation coefficient R between Vitality and existing indicators

Indicators	R for Vitality	
	Beneficiary 1	Beneficiary 2
PGC	0.11	-0.058
SWLS	0.16	-0.12
OECD	0.25	-0.098

NPOs, and civic activities." An increase is scored as 1 and no change is scored as 0, with responses plotted using a circle mark. Equation (1), the logistic curve, is shown as a solid line, with the constant terms from Table 5 denoted as β_0 and the partial regression coefficient as β_1 for Eq. (2), substituting the score from each scale for x_1. Examining Fig. 5(a) with Vitality as the explanatory variable, the logistic curve shows an S-shape, and the incidence rate q_i exceeds 0.5 for scores between 10 and 15. Logistic regression analysis rounds the incidence rate to a binary classification, enabling the user to distinguish whether there was an increase in opportunities to participate. Such a distinction could be made in Fig. 5(c) SWLS and (d) OECD, but not in (b) PGC.

Next, Figure. 6 shows the results for the Beneficiary 2 group. The logistic curve shows that only Vitality in Fig. 6(a) could distinguish whether there was an increase in opportunities to participate, indicating that it is possible to quantify the degree of activation of individuals' social activities.

3.3. Verifying Proposed Scale Accuracy

3.3.1. Beneficiary 1 group

The authors compared Cohen's d with Cliff's delta, which can handle non-parametric data, to see if there is a between-group difference between those who had increased opportunities to participate in volunteering, NPOs, and civic activities because they had obtained an adapted vehicle and those who experienced no change in the Beneficiary 1 group. Table 6 shows that the between-group difference is large according to both indicators, enabling clear determinations of whether an increase in participation opportunities occurred.

Figure 7 shows a histogram of Beneficiary 1's Vitality and whether there was an increase in opportunities for participation. A comparison of the two groups shows a large between-group difference, as indicated by the fact that the modes are on opposite sides, left and right.

Next, Table 7 uses a confusion matrix to show the results of verification of Vitality prediction accuracy. The table shows actual data and predictions based on Vitality, treating an increase in participation opportunities as positive and no change as negative. The threshold to judge Vitality as positive is when the incidence rate q_i is at least 0.5 on the logistic curve shown in Fig. 5(a). The number of accurate answers for 137 questionnaire respondents was 113, so the model had an accuracy rate of 0.825 and a recall rate of 0.868.

The accuracy of Vitality predictions was also verified using the Area Under the Curve (AUC) indicator, which is used to evaluate the performance of classification models created by machine learning. The results are shown in Fig. 8. The AUC is the area bounded by the Receiver Operating Characteristic (ROC) curve consisting of the false positive and true positive rates, with a value between a minimum of 0.5 and a maximum of 1.0. Larger numbers indicate greater accuracy, and a value of 0.8 or higher is considered indicative of a good model. The fact that the AUC for Vitality is 0.848 means that the model was able to

accurately quantify the effectiveness of adapted vehicles for Beneficiary 1 group.

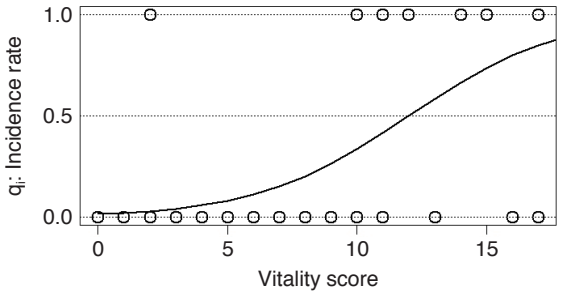
3.3.2. Beneficiary 2 group

The accuracy of the Vitality model for the Beneficiary 2 group was verified in the same way as the Beneficiary 1 group in the previous section. The model's accuracy rate was 0.846, the recall rate was 0.856, and the AUC was 0.925.

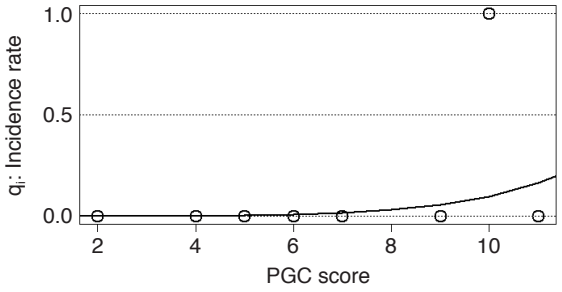
Therefore, the effectiveness of adapted vehicles for the Beneficiary 1 and Beneficiary 2 groups could be accurately quantified by Vitality.

Table 6 Calculation results of Cohen's d and Cliff's delta

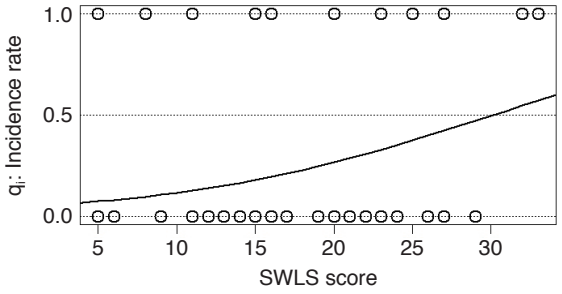
		Increase or no increase in participation opportunities in volunteering, NPOs, and civic activities	
Index		Cohen's d	Cliff's delta
Effect size		-1.90	-0.81
Between-group difference		Large	Large
95% confidence interval	Lower limit	-2.60	-0.92
	Upper limit	-1.19	-0.55



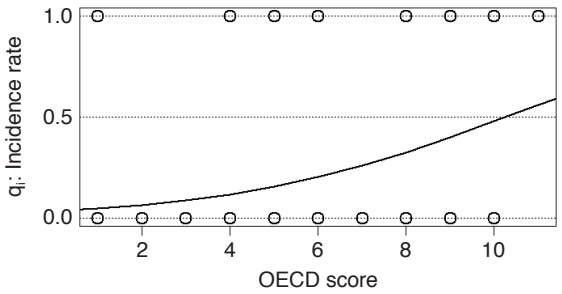
(a) Vitality logistic curve



(b) PGC logistic curve

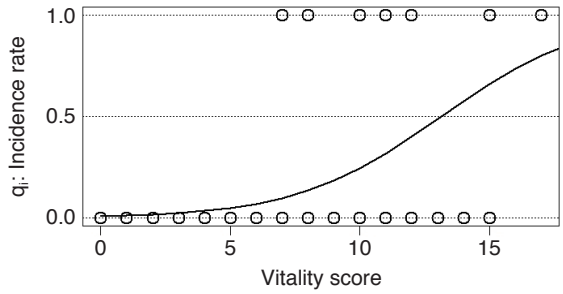


(c) SWLS logistic curve

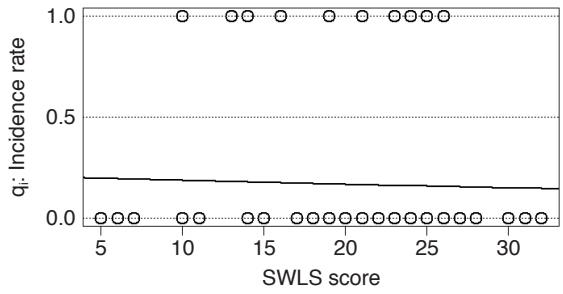


(d) OECD logistic curve

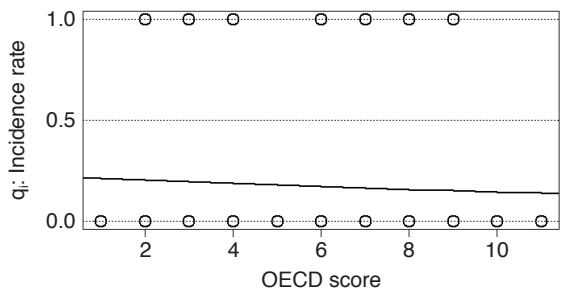
Fig. 5 Logistic curve for Beneficiary 1



(a) Vitality logistic curve



(b) SWLS logistic curve



(c) OECD logistic curve

Fig. 6 Logistic curve for Beneficiary 2

Table 7 Confusion matrix of Vitality in Beneficiary 1 with increased participation in volunteering, NPOs, and civic activities

		Prediction	
		Nonparticipation	Participation
Actual	Nonparticipation	92	10
	Participation	14	21

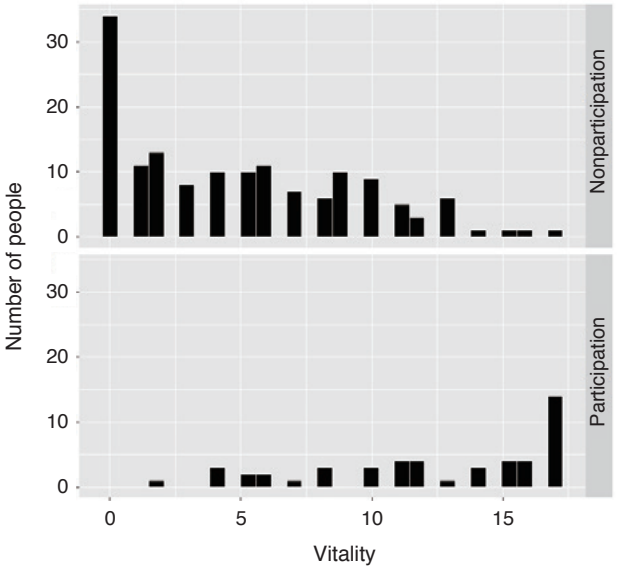


Fig. 7 Histogram of Vitality and increased participation opportunities in volunteer, NPOs, and civic activities in Beneficiary 1

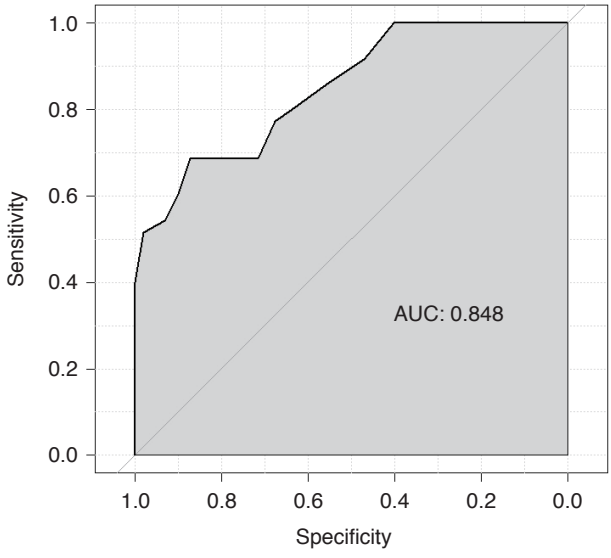


Fig. 8 ROC curve and AUC for Beneficiary 1

4. Conclusion

This paper has proposed a definition and a scale for the effectiveness of adapted vehicles in continually providing customers with the joy of mobility, and reached the following conclusions.

(1) The authors defined the effectiveness of adapted vehicles as “the degree of activation of the customer's individual social activities before and after the use of an adapted vehicle.”

(2) The degree of activation of social activities can be quantified by “Vitality,” which is obtained from the results of a 17-item questionnaire.

(3) The proposed Vitality is a scale independent of existing scales.

(4) This paper statistically verified that Vitality is an effective scale that can distinguish increases in opportunities to participate in social activities, and that it makes this distinction more easily than existing scales.

Going forward, Vitality will be used to understand effect and develop products.

References

(1) https://www.honda.co.jp/welfare/thought, (accessed 2025/6/30) (in Japanese)

(2) Ikuta, T., Fujii, H.: Research Perspectives on Non-Financial Information Disclosure and ESG Management, Review of Environmental Economics and Policy Studies, Vol. 13, Issue 2, p. 44-56, (2020) (in Japanese), doi: 10.14927/reeps.13.2_44

(3) https://www.npo-homepage.go.jp/uploads/h28-social-impact-sokushin-chousa-02.pdf, (accessed 2025/6/30) (in Japanese)

(4) Utsunomiya, K.: Chiiki Koukyou Koutsu to Socail Capital no Kanrensei, Koutsugakukenkyu, Vol. 59, p. 77-84, (2016) (in Japanese), doi: 10.32238/koutsugakkai.59.0_77

(5) Murakami, Y., Takahashi, S.: Beyond GDP - Development of Measures for Understanding Well-Being, Serviceology, Vol. 6, Issue 4, p. 8-15, (2020) (in Japanese), doi: 10.24464/serviceology.6.4_8

(6) https://rilac.or.jp/wordpress/wp-content/uploads/2017/03/%E5%9C%B0%E5%9F%9F%E5%8A%9B%E5%A0%B1%E5%91%8A%E6%9B%B8.pdf, (accessed 2025/2/10) (in Japanese)

(7) Sakurai, M.: Analysis of the Structure of Volunteer Participative Motivation based on the Multiple Motivations Approach: Volunteers of Kyoto City Area, The Nonprofit Preview, Vol. 2, Issue 2, p. 111-122, (2002) (in Japanese), doi: 10.11433/janpora.2.111

(8) Clary, G. E., Snyder, M., Ridge, D. R., Copeland, J., Stukas, A. A., Haugen, J., Miene, P.: Understanding and assessing the motivations of volunteers: A functional approach, Journal of Personality and Social Psychology, Vol. 74, Issue 6, p. 1516-1530, (1998),

doi: 10.1037/0022-3514.74.6.1516
(9) https://www.esri.cao.go.jp/jp/esri/prj/hou/hou075/hou75.pdf, (accessed 2025/2/10) (in Japanese)
(10) Lawton, M. P.: The Philadelphia Geriatric Center Morale Scale: A Revision, Journal of Gerontology, Vol. 30, Issue 1, p. 85-89, (1975), doi: 10.1093/geronj/30.1.85
(11) Kudo, Y.: Koureisya no Well-being ni Kannsuru Shihyou to sono Katuyou, Journal of Japan Academy of Community Health Nursing, Vol. 22, Issue 1, p. 59-64, (2019) (in Japanese)
(12) Feng, T., Kagawa, K.: Koureisya no Taijnnteki Shakai Katsudou Shakudo no Kaihatsu, Graduate School of Humanities and Social Sciences, Okayama University, Vol. 20, Issue 1, p. 119-130, (2005) (in Japanese), doi: 10.18926/1429
(13) Diener, E., Emmons, R. A., Larsen, R. J., Griffin, S.: The Satisfaction With Life Scale, Journal of Personality Assessment, Vol. 49, Issue 1, p. 71-75, (1985), doi: 10.1207/s15327752jpa4901_13
(14) Ohishi, S.: Shiawase wo Kagaku suru –Shinrigaku kara Wakatta Koto–, Shinyosha, p. 47-49, (2009) (in Japanese)
(15) Iwanaga, M., Otsuka, Y., Takahashi, K.: Shakai Chousa no Kiso, Foundation for the Promotion of the Open University of Japan, p. 125, (2003) (in Japanese)
(16) The R Core Team: "R: A Language and Environment for Statistical Computing", Version 4.5.1, (2025), https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf, (accessed 2025/7/22)

■ Author ■


Arisa ENDO


Toshihiro KOYAMA


Keiichi SUNAMI


Tetsuya GOMI

Retraining-free Customizable End-to-end Automatic Speech Recognition

Yui SUDO, Ph.D.[1] Shakeel MUHAMMAD, Ph.D.[1] Naoaki SUMIDA[1]

ABSTRACT

In this paper, the authors propose a method of customizing automatic speech recognition models using editable term lists to increase recognition accuracy for phrases registered at the discretion of the user. The proposed method adds a text encoder and cross-attention layer to the conventional Connectionist Temporal Classification/Attention-based automatic speech recognition model to detect registered terms in speech signals. To increase registered term recognition accuracy, the authors propose an inference algorithm that integrates the probability of detected registered words and Connectionist Temporal Classification. Results of their evaluation using the in-house Japanese-language dataset and the English-language Librispeech corpus showed that using the proposed method lowered the character error rate and the word error rate for registered terms.

1. Introduction

The development of deep learning has helped to advance automatic speech recognition technology. In particular, in recent years, end-to-end automatic speech recognition (E2E-ASR) models, a method that estimates text sequences directly from speech input, have achieved great increases in performance[1]. These dramatic increases in performance have led to the incorporation of ASR into many applications, including smart speakers, car navigation systems, meeting minutes transcription systems, and the Honda Communication Assistance System[2, 3].

E2E-ASR uses a single neural network to convert input speech signals directly into a text sequence, unlike conventional systems that include multiple components such as an acoustic model, lexicon, and language model. Neural networks are optimized using deep learning to mitigate error accumulation between components in conventional systems. A typical ASR model comprises an encoder that encodes the input speech and a decoder that converts the encoded speech features into a text sequence. Previously, various decoders have been proposed, such as Connectionist Temporal Classification (CTC)[4, 5], Recurrent Neural Network Transducer[6, 7], and attention mechanism[8, 9]. Hybrid systems combining these methods have also been proposed[10, 11], among which CTC/Attention-based E2E-ASR is one of widely used methods due to its high performance[10].

The performance of ASR using deep learning, however, tends to be highly dependent on the training data, making it a challenge to correctly recognize words and phrases that rarely appear in the training data. For example, technical terms and personal names are often incorrectly recognized because they rarely appear in training data. However, such words may be important keywords for certain users, and their misrecognition may greatly reduce user satisfaction. It is possible to enhance recognition performance by collecting training data that includes such words and retraining the neural network. However, frequent retraining is not possible because human resources and high-performance GPU servers are required for collecting training data, annotating, and training neural networks. Therefore, a retraining-free ASR model customization method is needed.

Other research has approached this need by proposing customization methods that can register any word[12, 13]. These methods add a cross-attention layer[14] to the decoder, enabling detection of registered terms with any parts of speech without retraining. Furthermore, to increase registered term detection accuracy, methods to introduce auxiliary branches[13, 15] and auxiliary loss functions in the cross-attention layer have been proposed[16]. These methods, however, do not explicitly use auxiliary branches or auxiliary losses during inference, limiting how much their performance with registered terms can be increased.

2. Aims of the Research

This paper proposes a method to customize E2E-ASR models utilizing user-editable term lists without retraining. The proposed method detects registered terms in speech

[1] Honda Research Institute Japan Co., Ltd.

signals by adding a text encoder and cross-attention layer to a CTC/Attention-based E2E-ASR model (Fig. 1)[10], similar to the conventional customization method[16] (Fig. 2). Unlike the conventional customization method[16], however, to further increase registered term recognition accuracy, this paper proposes an inference algorithm that increases the probability of detected registered terms and integrates the output probabilities with a CTC decoder. The aims of this research are as follows.

- Enhance registered term recognition performance by increasing the probability of registered terms detected during inference.
- Enhance overall recognition performance by integrating output probabilities using the CTC decoder during inference.
- Conduct a detailed analysis of performance with a Japanese-language dataset.

This paper extends earlier research[16]. Specifically, in respect to the preceding research[17], the researchers integrated the output probabilities using a CTC decoder and conducted a new detailed performance analysis with a Japanese-language dataset.

3. Conventional Model (CTC/Attention Model)

The CTC/Attention model[10], a conventional ASR model, is shown in Fig. 1. The CTC/Attention model has an audio encoder, an attention decoder, and a CTC decoder. The attention decoder and the CTC decoder share an audio encoder. Given an audio feature sequence $\boldsymbol{X} = [\boldsymbol{x}_1, ..., \boldsymbol{x}_L]$ of length L, posterior probability $P_{\text{att}}(y \mid \boldsymbol{X})$ and $P_{\text{ctc}}(y \mid \boldsymbol{X})$ of text sequence $y = [y_1, ..., y_S]$ of length S are respectively estimated. The decoders and audio encoder are trained jointly, and the output probabilities of the two decoders are integrated during inference. This approach could stabilize learning and enhance performance during inference. The following section describes each component in detail.

3.1. Audio Encoder

A conformer is used for the audio encoder[7], and converts the audio feature sequence $\boldsymbol{X}$ to a hidden state vector $\boldsymbol{H} = [\boldsymbol{h}_1, ..., \boldsymbol{h}_T] \in \mathrm{R}^{T \times d}$ of length T as follows.

$$\boldsymbol{H} = \text{AudioEnc}(\boldsymbol{X}) \tag{1}$$

where d is the dimensions of the hidden state vector.

3.2. Attention Decoder

A transformer decoder is used as the attention decoder[14]. When given the hidden state vector $\boldsymbol{H}$ extracted by the audio encoder shown in Eq. (1) and the text sequence $y_{0:s-1} = [y_0, ..., y_{s-1}]$ up to one frame previously, the attention decoder estimates the posterior probability $P(y_s \mid y_{0:s-1}, \boldsymbol{X})$ of the sth token y_s as follows.

$$P_{\text{att}}(y_s \mid y_{0:s-1}, \boldsymbol{X}) = \text{AttDecoder}(y_{0:s-1}, \boldsymbol{H}) \tag{2}$$

By repeating Eq. (2) for $s = 1, ..., S$, the posterior probability of text sequence y is formulated as follows.

$$P_{\text{att}}(y \mid \boldsymbol{X}) = \prod_{s=1}^{S} P_{\text{att}}(y_s \mid y_{0:s-1}, \boldsymbol{X}) \tag{3}$$

Cross entropy loss is introduced to the above.

$$L_{\text{att}} = -\log P_{\text{att}}(y \mid \boldsymbol{X}) \tag{4}$$

3.3. CTC Decoder

The CTC decoder comprises one linear layer[4, 5]. As Fig. 1 shows, the CTC decoder uses the hidden state vector $\boldsymbol{H}$ output by the audio encoder to estimate the posterior probability $P_{\text{ctc}}(y \mid \boldsymbol{X})$ of text sequence y, as shown below.

$$P_{\text{ctc}}(y \mid \boldsymbol{X}) = \text{CTC}(\boldsymbol{H}) \tag{5}$$

The CTC loss shown below is introduced to the above output.

$$L_{\text{ctc}} = -\log P_{\text{ctc}}(y \mid \boldsymbol{X}) \tag{6}$$

3.4. Joint Training and Joint Decoding

The CTC/Attention model is optimized to minimize loss by joint training using the weight λ shown below for the loss function defined in Eqs. (4) and (6).

$$L = (1 - \lambda) L_{\text{att}} + \lambda L_{\text{ctc}} \tag{7}$$

Additionally, to enhance performance during inference, the output probabilities of the attention decoder and CTC decoder are integrated using the weight μ shown below.

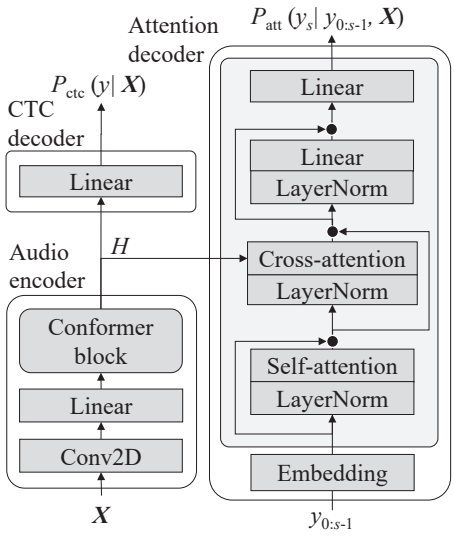


Fig. 1 Overall architecture of CTC/Attention model

$$\alpha = (1 - \mu)\, \alpha_{\text{att}} + \mu \alpha_{\text{ctc}}, \tag{8}$$

$$\alpha_{\text{att}} = \log P_{\text{att}}(y \mid \boldsymbol{X}), \tag{9}$$

$$\alpha_{\text{ctc}} = \log P_{\text{ctc}}(y \mid \boldsymbol{X}), \tag{10}$$

where α, α_{att}, and α_{ctc} are the respective output probabilities of the integration probability, attention decoder, and CTC decoder.

Conventional CTC/Attention models do not assume external input such as term lists, so these models cannot be customized after training, and it is a challenge to accurately recognize words that do not appear in the training data.

4. Network Structure of the Proposed Model

The overall structure of the proposed model is shown in Fig. 2. In addition to the conventional CTC/Attention model, the proposed method has a text encoder for inputting term lists and adds cross-attention to the attention decoder (green parts in Fig. 2). The text encoder encodes a list of terms that can be edited at the discretion of the user, and the cross-attention layer detects registered terms in the input speech.

When given an audio feature sequence $\boldsymbol{X}$ and a term list $\boldsymbol{B} = \{b_0, b_1, ..., b_N\}$, the extended attention decoder estimates the posterior probability $P_{\text{ext}}(y \mid \boldsymbol{X}, \boldsymbol{B})$ of a text sequence y of length S and the posterior probability $P_{\text{idx}}(n \mid \boldsymbol{X}, \boldsymbol{B})$ of index $n = [n_1, ..., n_S]$ of the detected registered terms (Fig. 2). The following section discusses each additional component.

4.1. Text Encoder

A transformer encoder is used as the text encoder[12]. The transformer encoder consists of an embedding layer, a transformer block, a mean pooling layer, and an editable term list $\boldsymbol{B} = \{b_0, b_1, ..., b_N\}$. b_N denotes the text sequence of the Nth registered term (e.g. [morning, assembly]). b_0 is a dummy word used when there is no registered term in the speech signal. The term list $\boldsymbol{B}$ is input to the embedding layer and transformer block and converted into a token-by-token feature sequence $\boldsymbol{G} \in \mathrm{R}^{(N+1) \times I\max \times d}$ as shown below.

$$\boldsymbol{G} = \text{Transformer}(\text{Embedding}(\boldsymbol{B})) \tag{11}$$

$I_{\max}$ is the maximum length of a word in the term list (e.g. $I_{\max} = 2$ if [morning, assembly] is the maximum length). It is then converted into a feature $\boldsymbol{V} = [v_0, \cdots, v_N] \in \mathrm{R}^{(N+1) \times d}$ representing the features of the whole word using the mean pooling layer.

$$\boldsymbol{V} = \text{MeanPool}(\boldsymbol{G}) \tag{12}$$

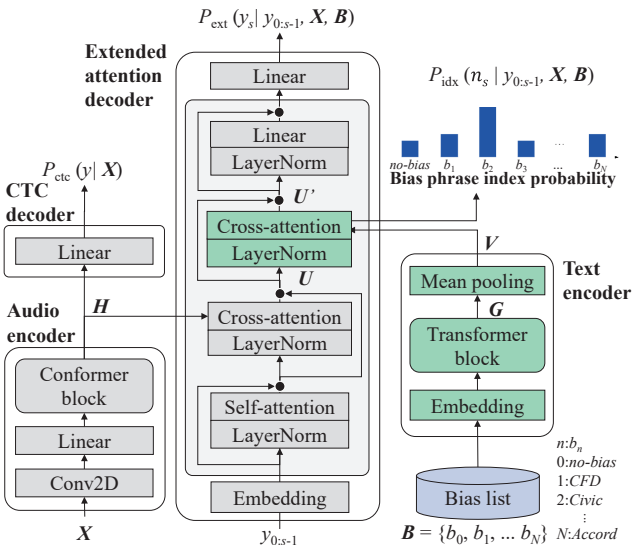


Fig. 2 Overall architecture of proposed method

4.2. Extended Attention Decoder

The extended attention decoder uses the hidden state vectors $\boldsymbol{H}$ and $\boldsymbol{V}$ extracted by the audio encoder and text encoder shown in Eqs. (1) and (12) to estimate the posterior probability $P_{\text{ext}}(y \mid \boldsymbol{X}, \boldsymbol{B})$ of the text sequence and the index probability $P_{\text{idx}}(n \mid \boldsymbol{X}, \boldsymbol{B})$ of the detected registered term[12].

4.2.1. Text sequence probability

When given the hidden state vectors $\boldsymbol{H}$ and $\boldsymbol{V}$ extracted by the audio encoder and text encoder shown in Eqs. (1) and (12), and the text sequence $y_{0:s-1}$ up to one frame previously, the decoder estimates the sth token y_s as follows.

$$P_{\text{ext}}(y_s \mid y_{0:s-1}, \boldsymbol{X}, \boldsymbol{B}) = \text{ExtDecoder}(y_{0:s-1}, \boldsymbol{H}, \boldsymbol{V}) \tag{13}$$

By repeating Eq. (13) for $s = 1, \ldots, S$, the posterior probability of text sequence y is formulated as follows.

$$P_{\text{ext}}(y \mid \boldsymbol{X}, \boldsymbol{B}) = \prod_{s=1}^{S} P_{\text{ext}}(y_s \mid y_{0:s-1}, \boldsymbol{X}, \boldsymbol{B}) \tag{14}$$

Cross entropy loss is introduced to the above.

$$L_{\text{ext}} = -\log P_{\text{ext}}(y \mid \boldsymbol{X}, \boldsymbol{B}) \tag{15}$$

4.2.2. Registered term index probability

Writing the input and output of the cross-attention layer of Fig. 2 as $\boldsymbol{U}$ and $\boldsymbol{U}'$, respectively, the cross-attention layer is defined as follows.

$$\boldsymbol{U}' = \text{Softmax}\left(\frac{\text{LN}(\boldsymbol{U})\boldsymbol{V}^T}{\sqrt{d}}\right)\boldsymbol{V} + \boldsymbol{U} \tag{16}$$

The term Softmax in Eq. (16) is the inner product of the input feature sequence $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_s]$ and the term list feature sequence $\boldsymbol{V} = [\boldsymbol{v}_0, \ldots, \boldsymbol{v}_N]$, so it indicates their degree

of similarity; that is, it can be interpreted as the probability of the index sequence $n = [n_1, \ldots, n_s]$ of registered terms in the input speech, as shown below.

$$P_{\text{idx}}(n \mid \boldsymbol{X}, \boldsymbol{B}) = \text{Softmax}\left(\frac{\text{LN}(\boldsymbol{U})\boldsymbol{V}^T}{\sqrt{d}}\right) \quad (17)$$

The index with the highest probability is a registered term contained in the speech. For example, if two words (1: morning assembly, 2: PR) are registered in the term list and the utterance y = [do PR at morning assembly] is spoken, the index sequence of registered terms (n = [0, 2, 0, 1, 1]. n = 0 [no-phrase]) is used for tokens that are not registered terms. Registered term index loss using cross entropy is introduced to the above output.

$$L_{\text{idx}} = -\log P_{\text{idx}}(n \mid \boldsymbol{X}, \boldsymbol{B}) \quad (18)$$

5. Model Training and Inference Methods

5.1. Training Method

The proposed method is optimized to minimize the loss by multi-task learning using the weights λ and γ shown below for the loss function defined in Eqs. (6), (15), and (17).

$$L = (1 - \lambda) L_{\text{ext}} + \lambda L_{\text{ctc}} + \gamma L_{\text{idx}} \quad (19)$$

During training, any word can be registered to the term list during inference, whereas a term list is randomly created from ground truth for each batch during training. This random selection is intended to allow the model to accommodate a wide range of words and phrases.

5.2. Inference Algorithm

To further enhance recognition performance, this paper proposes a new inference algorithm. Specifically, for the posterior probability $P_{\text{ext}}(y \mid \boldsymbol{X}, \boldsymbol{B})$ of text sequence y estimated by the extended attention decoder, (1) the index probabilities $P_{\text{idx}}(n \mid \boldsymbol{X}, \boldsymbol{B})$ of registered terms are integrated (red part of Fig. 3), and (2) the output probabilities by the CTC decoder are integrated (blue part of Fig. 3). The proposed algorithm is shown in Fig. 3.

(1) First, the extended attention decoder calculates the probability α_{ext} of each sth token based on Eq. (4) (line 5). Next, the sth registered term index n_s is estimated using Eq. (8) and the argmax function (line 6). If a registered term is detected ($n_s \neq 0$), the output probability of the tokens in the registered term is increased based on the weight α_{bonus} (line 9). For example, if the detected registered term is b_n = [morning, assembly], the output probability of the tokens [morning] and [assembly] is increased.

(2) Then, the output probability α_{ctc} of the CTC decoder is calculated using CTC prefix scoring[10]. However, since term list $\boldsymbol{B}$ is not input to the CTC decoder (Fig. 2), the probability of registered terms is not calculated. In other words, only if no registered terms are detected (n_s = 0), output probabilities are integrated using the weight μ (line 16).

6. Evaluation Experiment

6.1. Experimental Conditions

The input audio feature $\boldsymbol{X}$ is an 80-dimensional mel filter bank with a window length of 512 samples. A mel filter bank represents a time-frequency characteristic using a mel scale that is modeled on human auditory characteristics. The audio encoder consists of two convolutional layers, a 256-dimensional linear projection layer, and 12 conformer blocks. The text encoder includes an embedding layer and six transformer blocks and has a mean pooling layer in the final layer. The extended attention decoder comprises an embedding layer, six transformer blocks, and a linear layer. The CTC decoder consists of one linear layer. The weights λ and γ during training as shown in Eq. (19) were respectively set to 0.3 and 1.0. The parameters α_{bonus} and μ during inference, as described in Section 5.2, were respectively set to 1.0 and 0.3.

An in-house dataset (Japanese) and the Librispeech-100 corpus (English)[18] were used for evaluation. The in-house dataset consists of 93 hours of in-house speech data including meeting and morning assembly scenarios,

Algorithm 1 Decoding algorithm

1: *hyps* = {<*sos*>:1.0}; *end_hyps* = {};
2: **for** s = 1 to maxlen **do**
3: *ext hyps* = {}
4: **for** $\boldsymbol{l} \in$ *hyps* **do**
5: $\boldsymbol{\alpha}_{\text{ext}} = \log P_{\text{ext}}(y_s \mid y_{1:s-1}, \boldsymbol{X}, \boldsymbol{B})$
6: $n_s \leftarrow$ Argmax($P(n_s \mid y_{1:s-1}, \boldsymbol{X}, \boldsymbol{B})$)
7: **if** $n_s \neq 0$ **then**
8: **for** token $\in b_n$ **do**
9: $\boldsymbol{\alpha}_{\text{ext}}$[token] $\leftarrow \boldsymbol{\alpha}_{\text{ext}}$[token] + α_{bonus}
10: **end for**
11: **end if**
12: **for** $y_s \in$ Top-$k_{\text{beam}}(\boldsymbol{\alpha}_{\text{ext}})$ **do**
13: Add $\tilde{l} = l \oplus y_s$ and the score α_{ext} to *ext hyps*
14: $\alpha_{\text{ctc}} = \text{CTCPrefixScore}(\tilde{l}, \boldsymbol{H})$
15: **if** $\hat{n}_s = 0$ **then** ("*no-bias*")
16: $ext\_hyps[\tilde{l}] = \mu\alpha_{\text{ctc}} + (1 - \mu)\ \alpha_{\text{ext}}$
17: **else**
18: $ext\_hyps[\tilde{l}] = \alpha_{\text{ext}}$
19: **end if**
20: **if** y_s = <eos> or len($\tilde{l}$) = maxlen **then**
21: Add $\tilde{l}$ to *end_hyps*
22: **end if**
23: **end for**
24: **end for**
25: *hyps* = Top-k(*ext hyps*)
26: **end for**
27: **return** Top-1(*end_hyps*)

Fig. 3 Joint decoding algorithm

the Corpus of Spontaneous Japanese[19], and a database developed by the Advanced Telecommunications Research Institute International[20]. The proposed method was evaluated based on character error rate (CER), CER for registered terms (B-CER), and CER for other (U-CER). The objective of the proposed method is to increase B-CER while minimizing the increase in U-CER and overall CER. With Librispeech-100, word error rate (WER), U-WER, and B-WER were used.

6.2. Verification with Japanese Dataset

Table 1 shows the evaluation results when $N = 203$ words or phrases are registered in the term list. The baseline was a conventional non-customizable CTC/Attention model[10] (A1). Results were also compared with conventional customization methods[15), (21] (B1, 2). Unlike the proposed method, these methods do not explicitly use the probability of registered terms detected during inference. For the proposed method, the researchers also tested the impact of excluding the inference algorithms (registered term index probability integration and probability integration by CTC decoder) described in Section 5.2. (C2, 3).

First, the proposed method was compared with the baseline and conventional customization methods (A1, B1, and B2 vs. C1). C1 decreased B-CER (CER for registered terms) by 11.71 points compared to baseline, although U-CER (CER for other than registered terms) increased by 0.06 points. As a result, overall CER decreased by 1.40 points. The proposed method also showed a decrease in both overall CER and B-CER over conventional customization methods[15), (21].

Next, this paper discusses the impact of excluding registered term index probability integration and probability integration by CTC decoder. As shown in C2 and C3, excluding registered term index probability integration increased the B-CER by 1.96 points, and excluding probability integration by CTC decoder increased the U-CER by 0.97 points. This confirmed that their exclusion is effective for enhancing recognition performance for respective registered terms and other words.

6.2.1. Performance analysis

Registered terms were then classified into words written in katakana, the alphabet (such as acronyms written with the alphabet), kanji, and English, and the recognition accuracy was analyzed. Words written with multiple character types were classified according to the type of character they used the most. For example, “CFD 計算 ” was categorized as written with the alphabet, because it uses mostly alphabetical characters. Words in the alphabet category are those pronounced as readings of the alphabetic characters, such as “PR,” whereas words in the English category are those pronounced otherwise, such as “sketch.”

Table 2 and Fig. 4 show the recognition accuracy for each category of registered terms, as well as examples of typical inference results. In inference results examples 1 to 3, the registered terms were correctly recognized by the proposed method, while in example 4, the registered term was not correctly recognized. Table 2 shows that for words where each character and its pronunciation are uniquely determined, such as in the katakana and alphabet categories, the recognition accuracy was over 80%, whereas the recognition accuracy for the English category was clearly low, just 29%. In particular, the recognition accuracy of English words like “sketch” and “cascade,” which are commonly written in katakana, was far lower.

This can be attributed to the fact that the words registered during training were selected at random. The cross-attention layer added to the text encoder and the extended attention decoder in the proposed method (C1) is trained to detect the pronunciation of registered words from the input speech. This may have resulted in insufficient training for words with uncommon combinations of spelling and pronunciation.

Table 1 Experimental results on in-house dataset

ID	Model	CER [%]	U-CER [%]	B-CER [%]
A1	Baseline	9.85	8.17	21.86
B1	NEA-ASR [21]	9.75	8.11	21.90
B2	Intermediate DB [15]	9.28	8.23	16.93
C1	Proposed method	8.45	8.23	10.15
C2	C1 w/o index probability	8.65	8.19	12.11
C3	C2 w/o CTC integration	9.78	9.16	14.54

Table 2 Performance analysis on in-house dataset

Class	Accuracy [%]	Examples
Katakana	86	ジャイロ，リファイン
Alphabet	84	CFD, PR
Kanji	64	接種，風洞
English	29	SKETCH, CASCADE

1. Reference: 一番のネックは風洞だって事だ
   Baseline: 一番のネッカーフードだって事だ
   Proposed (C1) : 一番のネックは風洞だって事だ
2. Reference: それとは別の**職域接種**
   Baseline: それとは別の職域摂取
   Proposed (C1) : それとは別の職域接種
3. Reference: **被引例情報**のときに**第一引例**って
   Baseline: 非陰例情報のときに第一因例って
   Proposed (C1) : 非引例情報のときに第一引例って
4. Reference: **SKETCH** モードを開いて
   Baseline: スケッチモードを開いて
   Proposed (C1) : スケッチモードを開いて

Boldface: Registered phrase
Red: Incorrectly recognized characters
Blue: Correctly recognized characters

Fig. 4 Typical inference results

6.2.2. Scalability verification

In actual use cases, the number of words registered to a term list tends to increase unless users explicitly delete registered terms. Figure 5 shows the effect of term list size on B-CER (CER for registered words). In this experiment, we compared B-CERs when no words were registered to the term list ($N = 0$), when words appropriate to the use case were registered ($N = 203$), and when unnecessary words were added ($N = 500$ to 2000).

When no words were registered to the term list ($N = 0$), there was no decrease in B-CER, whereas when appropriate words were registered ($N = 203$), B-CER decreased. As the number of words unnecessary for the use case in question increased, the B-CER gradually increased ($N = 500$ to 2000). It was confirmed, however, that the proposed method is robust to increases in term list size, since it maintained a sufficient decrease effect in B-CER compared to the case where no words are registered to the term list ($N = 0$).

6.3. Verification with English Dataset

Next, the researchers evaluated a case of $N = 100$ words registered in a term list as an English dataset (Librispeech-100), since English has different characteristics from Japanese. Since Japanese is an ideographic language and English is a phonetic language, it is important to evaluate languages like these that have different characteristics. Table 3 shows the effectiveness of the proposed method, registered term index probability integration, and probability integration by CTC. A similar trend was observed for the English evaluation results for Japanese. The proposed method decreased the B-CER and thus decreased the overall WER compared to the baseline (A1 vs. C1). The B-WER and U-WER decreased when registered term index probability integration and probability integration by CTC decoder were excluded, thus confirming the effectiveness of these methods in enhancing the recognition performance in each case (C1 vs. C2, 3).

Additionally, Fig. 6 shows the registered term index probabilities estimated by the cross-attention layer based on Eq. (8). The horizontal axis shows the input speech utterances, and the vertical axis shows the registered words.

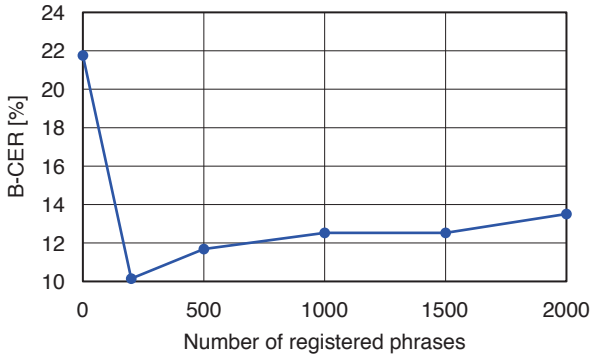


Fig. 5 Robustness to term list size

Table 3 Experimental results on Librispeech-100 h

ID	Model	WER [%]	U-WER [%]	B-WER [%]
A1	Baseline	8.59	5.87	30.71
C1	Proposed method	5.76	4.10	19.30
C2	C1 w/o index probability	6.49	4.83	19.98
C3	C2 w/o CTC integration	7.53	5.27	25.92

These utterances included the registered words “legality” and “legislature.” Figure 6 confirms that the proposed method can accurately detect registered words from the input speech.

The above confirmed that the proposed method is effective for languages with different characteristics (e.g. Japanese and English).

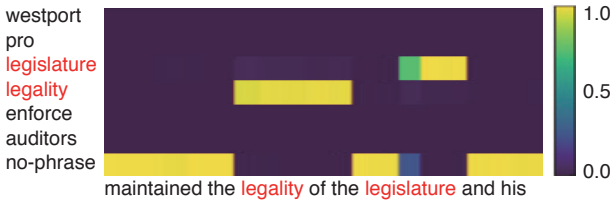


Fig. 6 Registered phrase detection

7. Conclusion

In this paper, the authors propose a method of customizing E2E-ASR models using editable term lists to increase recognition accuracy for phrases registered at the discretion of the user. In the proposed method, a text encoder and cross-attention layer are added to a conventional CTC/Attention-based E2E-ASR model to detect registered terms in speech signals. Additionally, to further increase registered term recognition accuracy, this paper proposes an inference algorithm that increases the probability of detected registered terms and uses probability integration with a CTC decoder. Results of an evaluation using an in-house dataset (Japanese) and the Librispeech corpus (English) demonstrated that the proposed method lowered the character error rate and the word error rate for registered terms.

References

(1) Prabhavalkar, R., Hori, T., Sainath, T. N., Schlüter, R., Watanabe, S.: End-to-End Speech Recognition: A Survey, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 32, p. 325-351, (2023), doi: 10.1109/TASLP.2023.3328283

(2) https://global.honda/jp/stories/003.html, (accessed 2024/9/23) (in Japanese)

(3) https://www.jp.honda-ri.com/en/activity/honda-ca-system/, (accessed 2024/9/23)
(4) Graves, A., Fernández, S., Gomez, F., Schmidhuber, J.: Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks, ICML '06: Proceedings of the 23rd international conference on Machine learning, p. 369-376, (2006), doi: 10.1145/1143844.1143891
(5) Peng, Y., Sudo, Y., Shakeel, M., Watanabe, S.: OWSM-CTC: An Open Encoder-Only Speech Foundation Model for Speech Recognition, Translation, and Language Identification, Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, Vol. 1, p. 10192–10209, (2024), doi: 10.18653/v1/2024.acl-long.549
(6) Graves, A.: Sequence Transduction with Recurrent Neural Networks, International Conference of Machine Learning (ICML) 2012 Workshop on Representation Learning, arXiv preprint, (2012), doi: 10.48550/arXiv.1211.3711
(7) Gulati, A., Qin, J., Chiu, C.-C., Parmar, N., Zhang, Y., Yu, J., Han, W., Wang, S., Zhang, Z., Wu, Y., Pang, R.: Conformer: Convolution-augmented Transformer for Speech Recognition, Proc. Interspeech 2020, p. 5036-5040, (2020), doi: 10.21437/Interspeech.2020-3015
(8) Chorowski, J. K., Bahdanau, D., Serdyuk, D., Cho, K., Bengio, Y.: Attention-Based Models for Speech Recognition, Advances in Neural Information Processing Systems, Vol. 28, (2015)
(9) Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., Sutskever, I.: Robust Speech Recognition via Large-Scale Weak Supervision, Proceedings of the 40th International Conference on Machine Learning, p. 28492-28518, (2023)
(10) Watanabe, S., Hori, T., Kim, S., Hershey, J. R., Hayashi, T.: Hybrid CTC/Attention Architecture for End-to-End Speech Recognition, IEEE Journal of Selected Topics in Signal Processing, Vol. 11, Issue 8, p. 1240-1253, (2017), doi: 10.1109/JSTSP.2017.2763455
(11) Sudo, Y., Shakeel, M., Yan, B., Shi, J., Watanabe, S.: 4D ASR: Joint modeling of CTC, Attention, Transducer, and Mask-Predict decoders, Proc. INTERSPEECH 2023, p. 3312-3316, (2023), doi: 10.21437/Interspeech.2023-1303
(12) Pundak, G., Sainath, T. N., Prabhavalkar, R., Kannan, A., Zhao, D.: Deep Context: End-to-end Contextual Speech Recognition, 2018 IEEE Spoken Language Technology Workshop (SLT), (2018), doi: 10.1109/SLT.2018.8639034
(13) Huang, K., Zhang, A., Yang, Z., Guo, P., Mu, B., Xu, T., Xie, L.: Contextualized End-to-End Speech Recognition with Contextual Phrase Prediction Network, Proc. INTERSPEECH 2023, p. 4933-4937, (2023), doi: 10.21437/Interspeech.2023-767
(14) Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I.: Attention is all you need, NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, p. 6000-6010, (2017)
(15) Shakeel, M., Sudo, Y., Peng, Y., Watanabe, S.: Contextualized End-to-end Automatic Speech Recognition with Intermediate Biasing Loss, Proc. Interspeech 2024, p. 3909-3913, (2024), doi: 10.21437/Interspeech.2024-1257
(16) Huber, C., Hussain, J., Stüker, S., Waibel, A.: Instant One-Shot Word-Learning for Context-Specific Neural Sequence-to-Sequence Speech Recognition, 2021 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), (2021), doi: 10.1109/ASRU51503.2021.9687898
(17) Sudo, Y., Shakeel, M., Fukumoto, Y., Peng, Y., Watanabe, S.: Contextualized Automatic Speech Recognition With Attention-Based Bias Phrase Boosted Beam Search, ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), p. 10896-10900, (2024), doi: 10.1109/ICASSP48485.2024.10447782
(18) Panayotov, V., Chen, G., Povey, D., Khudanpur, S.: Librispeech: An ASR corpus based on public domain audio books, 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), p. 5206-5210, (2015), doi: 10.1109/ICASSP.2015.7178964
(19) Maekawa, K.: Corpus of Spontaneous Japanese: Its design and evaluation, Proc. ISCA/IEEE Workshop on Spontaneous Speech Processing and Recognition, (2003), https://www.isca-archive.org/sspr_2003/maekawa03_sspr.pdf, (accessed 2024/9/23)
(20) Kurematsu, A., Takeda, K., Sagisaka, Y., Katagiri, S., Kuwabara, H., Shikano, K.: ATR Japanese speech database as a tool of speech recognition and synthesis, Speech Communication, Vol. 9, Issue 4, p. 357-363, (1990), doi: 10.1016/0167-6393(90)90011-W
(21) Sudo, Y., Hata, K., Nakadai, K.: Retraining-free Customized ASR for Enharmonic Words Based on a Named-Entity-Aware Model and Phoneme Similarity Estimation, Proc. INTERSPEECH 2023, p. 491-495, (2023), doi: 10.21437/Interspeech.2023-1320

■ Author ■


Yui SUDO


Shakeel MUHAMMAD


Naoaki SUMIDA

Enhancement of Water Resistance by Methanation Catalysts with LDHs Structure

Hideaki YONEDA[1] Kazuki YANAGISAWA[1]
Yohei IIJIMA[1] Junpei YOSHIDA[1]

ABSTRACT

In order to enhance the efficiency of the synthetic methane production process, the authors developed an innovative catalyst that converts synthesis gas (syngas) obtained from the co-electrolysis reaction of water vapor and CO_2 into methane, and studied the allowable proportion of water vapor in the source gas, i.e., water resistance. Evaluation of the activity using the continuous-flow method showed that NiAl-LDHs utilizing layered double hydroxides exhibited higher catalytic reactivity for the generation of methane from syngas consisting of CO and H_2 at 0.1 MPa and 300°C. In addition, the authors experimented with the methane conversion reaction of syngas containing water vapor, which is an issue in the synthetic methane production process, when combined with a co-electrolysis reactor. The results showed that when up to 20% water vapor was present, CO conversion and CH_4 selectivity were equivalent to cases without water vapor, and the catalyst exhibited enhanced performance compared to the conventional metal-supported catalysts.

1. Introduction

The conversion of carbon dioxide (CO_2) into valuable chemicals continues to attract much attention in terms of supporting the separation, capture, and effective utilization of carbon dioxide[(1)]. Among the approaches tried, the production of synthetic methane as a gas fuel is attracting attention in Japan and abroad. In Japan, the Ministry of Economy, Trade and Industry is leading a methanation project, while gas companies and automobile-related companies have also undertaken efforts[(2)].

There are three main methods of synthetic methane generation: one method produces methane from biogas-derived CO_2, and the other two are Power-to-Gas (P2G) methane production (methanation) technologies. The first P2G process is the Sabatier reaction, shown in Fig. 1(a), which uses as feedstock CO_2 recovered from the atmosphere or factory exhaust gases and hydrogen (H_2) obtained by electrolyzing room-temperature water to produce methane in a catalytic thermochemical reaction. The second process, called innovative methanation, is shown in Fig. 1(b). Here, a mixture of synthesis gas (syngas), i.e., carbon monoxide (CO) and H_2, is obtained by co-electrolysis, which is a simultaneous electrolytic reaction of CO_2 and higher-temperature water vapor (H_2O). Then, using this syngas as feedstock, methane is produced by a catalytic thermochemical reaction[(3)]. In this study, the authors investigate innovative methanation, which has the potential to produce methane very efficiently while consuming little energy outside of the reaction pathway[(4)].

As a concrete example of innovative methanation, the system shown within the dashed lines of Fig. 2 has been investigated[(4)]. To generate the end product methane, the system consists of a co-electrolysis reactor with stacked solid oxide electrolysis cells (SOEC), which are expected to provide higher electrolysis efficiency, upstream and a

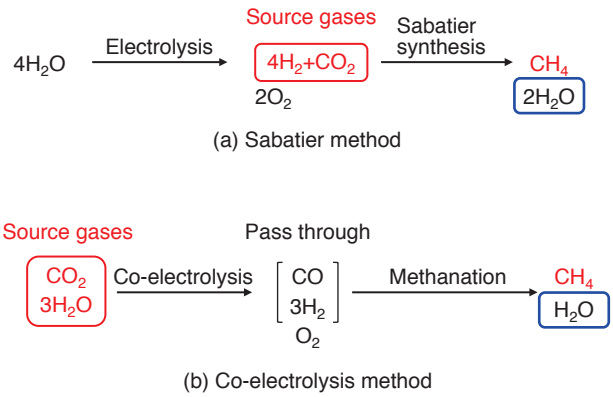


Fig. 1 Comparison of methane synthesis processes

[1] Innovative Research Excellence, Power Unit & Energy
E-mail: hideaki_yoneda@jp.honda

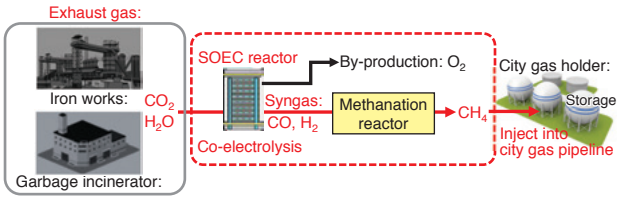


Fig. 2 Co-electrolysis methanation system via syngas

thermochemical reactor (methanation reactor) downstream to effectively utilize the heat generated by the methanation process, thereby converting the syngas obtained from the co-electrolysis reactor to methane.

2. Achieving Higher Methane Production Efficiency

2.1. Synthetic Methane Production Process

In a co-electrolysis reactor, the reactions in Eqs. (1) and (2) occur simultaneously at the fuel electrode to generate CO, H_2, and oxygen ions (O^{2-}) from the source gas and electrons (e^-).

$$CO_2 + 2e^- \rightarrow CO + O^{2-} \quad (1)$$

$$H_2O + 2e^- \rightarrow H_2 + O^{2-} \quad (2)$$

In addition, the water gas shift reaction shown in Eq. (3) occurs in parallel as a side reaction.

$$CO + H_2O \Leftrightarrow CO_2 + H_2 \quad (3)$$

At the oxygen electrode, oxygen (O_2) is generated from O^{2-} as shown in Eq. (4).

$$O^{2-} \rightarrow 1/2O_2 + 2e^- \quad (4)$$

During methanation using syngas as feedstock, 3 moles of H_2 are needed for 1 mole of CO, as shown in Eq. (5). The CO:H_2 ratio adjustment should ideally be achieved by controlling the electrolytic reaction in the co-electrolysis reactor and the flow rates of CO_2 and H_2O at the inlet.

$$CO + 3H_2 \rightarrow CH_4 + H_2O \quad (5)$$

2.2. Methanation Operating Challenges

The co-electrolysis reactor consists of multiple stacked solid oxide electrolysis cells, and when the source gas passes through the stacked cells, gas that comes in contact with the cells is converted to CO and H_2 through electrolytic reaction. In other words, source gas that does not come in contact with cells does not react and is discharged out of the reactor as unreacted gas. To help prevent cell damage due to over-voltage occurring, for example, when the source gas is depleted, 5% to 20% of the source gases (CO_2 and H_2O) in use are often passed through during the syngas generation process. Therefore, gas that has not been passed through and converted also remains in the syngas as an unreacted residual component[5]-[7]. Water vapor, the H_2O in the above residual component, is usually separated and removed before it gets to the methanation reactor because it is a factor that reduces catalytic activity. In the innovative methanation process as studied in the past, the proportion of water vapor in the generated gas after the co-electrolysis reactor was assumed to be 16% to 19%, and a drain was installed before the methanation reactor to cool, liquefy, and separate the water vapor, thereby reducing the proportion to 5% to 6%, after which the gas was fed into the methanation reactor[8]. The energy consumption associated with separation was estimated to be up to 58 kW, accounting for several percentage points of system energy efficiency. This study sought to develop a reliable water-resistant catalyst suitable for innovative methanation systems, thereby reducing the energy needed to separate the water vapor in the residual component, as well as reducing the overall process size and cost.

3. Catalyst Development Goals

There are three performance requirements for methanation catalysts, and two of the three concern the reaction shown in Eq. (5). The first is conversion, which indicates the proportion of the source gas CO that is converted to methane and other substances. The second is selectivity, which indicates the proportion of the converted material that is converted to the target product, methane. The CO conversion and CH_4 selectivity should be higher because the generated synthetic methane is to be used again as gas fuel. In addition, in methanation, H_2O is present around the catalyst as a byproduct of methanation, as shown in Eq. (5), as well as in the unreacted portion from the co-electrolysis reactor at the earlier step, and the CO obtained from the co-electrolysis reactor is converted back to CO_2 by the water gas shift reaction in Eq. (3). As a result, CO_2 entering the process is emitted from the process without being converted to valuable chemicals. Therefore, the third performance requirement for the catalyst in this study was that the catalyst should have a higher tolerance to water vapor contamination (water resistance), so that the water gas shift reaction does not occur even if water vapor is contained.

In summary, the following three performance requirements should be met.

(1) Higher CO conversion
(2) Higher CH_4 selectivity
(3) Higher tolerance to water vapor contamination (water resistance)

In conventional methanation catalysts, Ni-based catalysts prepared by the coprecipitation method or impregnation method are widely used because of

their lower cost, higher activity, and higher methane selectivity[9]. The activity of catalysts prepared by the impregnation method usually depends on the basic properties of the carrier and on the preparation method. Various carriers such as SiO_2, ZrO_2, γAl_2O_3, and CeO_2 have been studied[10]. However, the porous structure of these carriers makes them susceptible to H_2O adsorption, which covers the active sites of the catalyst and reduces its activity. Therefore, H_2O needs to be separated and removed upstream of the methanation reactor by a drain or other means. In this study, the authors focused on the structure of layered double hydroxides (LDHs), usually called hydrotalcites, which already contain H_2O. LDHs have a two-dimensional layered structure, as shown in Fig. 3[11], and it is well known that there is a great diversity of structures, including the metal cation ratios that can be prepared, various interactions of the layers composed between host and guest, and higher specific surface area[12]. Furthermore, since each layer of LDHs contains H_2O molecules and the H_2O adsorption sites are already buried, the catalyst has water repellency; that is, there is no adsorption when new water vapor is introduced, which is considered to be an advantage over conventional supported catalysts. In addition, because the structure of the LDHs results in much more highly dispersed Ni nanoparticles containing more reducible Ni species (Ni^{2+} to Ni^0) than conventional Ni-γAl_2O_3 catalysts (where a small amount of nickel ions support the formation of mixed oxides with alumina in the form of Ni^{3+})[13], the authors believed this would increase the number of reaction sites and increase CO_2 conversion and CH_4 selectivity. Therefore, LDHs are expected to be applied to hydrogen generation reactions and photoelectrochemical water oxidation processes and as functional catalysts[14), (15].

In this study, the LDHs catalyst developed by the authors, which uses Ni as the active metal, is compared to catalysts prepared by conventional coprecipitation and impregnation methods, and the performance and characteristics of the LDHs catalyst and its effect on the system are described.

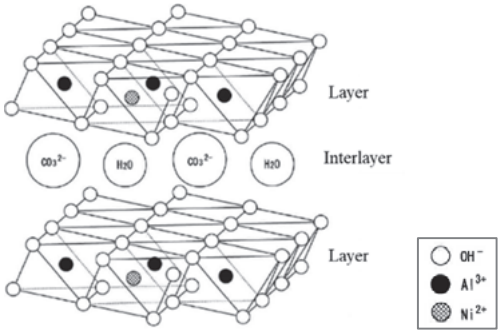


Fig. 3 LDHs structure model

4. Experimental Method

4.1. Catalyst Preparation

4.1.1. Preparation of Ni coprecipitation catalyst

A Ni-based coprecipitation catalyst was prepared using the following procedure. $Ni(NO_3)_2 \cdot 6H_2O$ was dissolved in distilled water, and the resulting solution was dropped into a beaker together with Na_2CO_3 solution (1 mol/L, room temperature) at a slow rate of 2 mL/min. At this time, the pH value of the mixture was maintained at 8.5 at 80°C. The precipitate was then allowed to stand at room temperature for 24 hours, followed by filtration and washing to completely remove residual Na_2CO_3 from the precipitate. The precipitate was then dried at 60°C in air for 12 hours. Finally, it was heated at 2°C/min and maintained at 550°C for 4 hours for calcination.

4.1.2. Preparation of Ni supported catalyst

The catalyst in which Ni was supported by carriers with the impregnation method was prepared using the following procedure. Commercially available carriers (SiO_2, γAl_2O_3, ZrO_2, CeO_2) were calcinated at 550°C for 4 hours to remove surface impurities, and Ni was introduced into the prepared carriers by the wet impregnation method. $Ni(NO_3)_2 \cdot 6H_2O$ was dissolved in deionized water and impregnated uniformly onto the carrier. Each catalyst was prepared so that the weight of Ni was 20 wt% of the total. The resulting mixture was then exposed to a vacuum at room temperature for 1 hour and then dried at 60°C in air for 12 hours. Finally, it was heated at 2°C/min and maintained at 550°C for 4 hours for calcination.

4.1.3. Preparation of NiAl-LDHs

To achieve higher dispersion of Ni nanoparticles and increase the reaction sites, NiAl-LDHs were synthesized by a hydrothermal method using urea. First, 2.18 g of $Ni(NO_3)_2 \cdot 6H_2O$ and 0.94 g of $Al(NO_3)_3 \cdot 9H_2O$ were dissolved together in 60 ml of deionized water. Separately, 6 g of urea was dissolved in 60 ml of deionized water. Each solution was stirred at room temperature for 30 minutes at a stirring speed of 300 rpm, and then the two solutions were mixed and stirred for another 5 minutes. The mixture was then sealed in two 100 ml Teflon-coated stainless-steel autoclaves, heated at 3°C/min, and maintained at 120°C for 12 hours. The mixture was allowed to stand without stirring during this period. Then, the prepared precipitates were washed with 200 ml of deionized water each time by centrifugation until the pH of the solution was neutral. The resulting product was dried at 60°C in air for 12 hours. Finally, to help prevent the dried catalyst from losing its structure due to bonding with oxygen, it was heated at 2°C/min in a nitrogen atmosphere and calcined at 500°C for 4 hours.

4.2. Catalyst Characterization

4.2.1. Specific surface area

The specific surface area of the prepared catalysts

was measured by N_2 physisorption using a Micromeritics 3Flex specific surface area and microporous gas adsorption analyzer. Prior to measuring, all samples were pretreated at 250°C for 10 hours under vacuum to remove surface contaminants.

4.2.2. Structural analysis by X-ray diffraction

To confirm the crystal structure of the prepared catalysts, X-ray diffraction (XRD) measurements were taken using a Rigaku RINT 2400 system X-ray diffractometer.

4.3. Evaluation of Catalyst Activity

Catalytic activity was evaluated by the continuous-flow method in a fixed-bed reactor with an inner diameter of 8 mm, as shown in Fig. 4. The catalyst was loaded into the reaction tube according to the following procedure. First, an appropriate amount of quartz wool was put in, then a mixture of 2 g of quartz sand and 0.25 g of catalyst prepared to a particle size of 400 to 850 μm was poured in. Then, 2 g of glass beads were added to preheat the gas during the reaction, and finally quartz wool was put on the top surface.

The catalyst was reduced with H_2 at 40 ml/min for 8 hours at 400°C prior to the reaction. After maintaining at 300°C for 8 hours at 0.1 MPa, 20 mL/min, $H_2/CO = 3.0$, the activity of each catalyst was measured.

The gases generated were detected by three on-line gas chromatographs equipped with a Shimadzu Corporation GC-2014AT thermal conductivity detector (TCD) and GC-2014AF flame ionization detector (FID). The TCD measures CO_2 upstream and H_2 downstream, respectively, while the FID measures CO and hydrocarbons.

CO conversion was calculated based on Eq. (6), CH_4 and CO_2 selectivity on Eqs. (7) and (8), and CH_4 yield on Eq. (9). Ar in Eq. (6) is used as an internal standard substance. Here, Eq. (6) is the proportion of CO converted to valuable products, Eq. (7) is the proportion of CH_4 to total converted hydrocarbons and CO_2, and Eq. (8) is the proportion of CO_2. Equation (9) shows the proportion of the reactants from which the target product CH_4 was obtained.

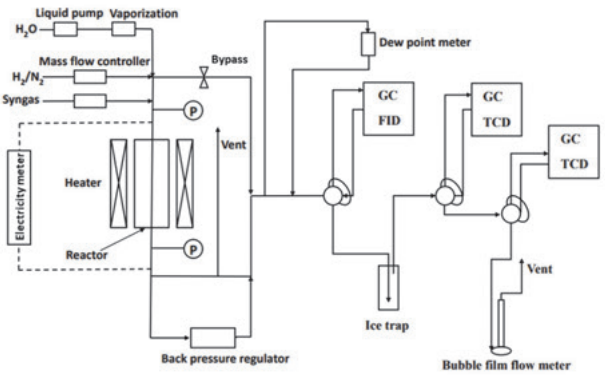


Fig. 4 Reaction devices diagram of fix-bed reactor

$$CO\ conversion\ [\%] = \frac{Area\left(\frac{CO}{Ar}\right)^{in} - Area\left(\frac{CO}{Ar}\right)^{out}}{Area\left(\frac{CO}{Ar}\right)^{in}} \times 100 \quad (6)$$

$$CH_4\ selectivity\ [\%] = \frac{n\,(CH_4)}{n\,(Total\ hydrocarbon + CO_2)} \times 100 \quad (7)$$

$$CO_2\ selectivity\ [\%] = \frac{n\,(CO_2)}{n\,(Total\ hydrocarbon + CO_2)} \times 100 \quad (8)$$

$$CH_4\ yield\ [\%] = CO\ conversion \times CH_4\ selectivity \times 100 \quad (9)$$

n: [C-mol/h]

5. Experiment Results and Discussion

5.1. Specific Surface Area, XRD Measurement Results

Results for the specific surface area of the prepared catalysts are shown in Fig. 5. The specific surface area of NiAl-LDHs was lower than that of 20 wt% Ni/SiO_2, but was comparable or higher than that of the other carriers.

XRD results for the prepared NiAl-LDHs are shown in Fig. 6. The positions of the peaks at 37°, 43°, and 62° are attributed to the Ni oxide NiO on the LDHs, and the clear presence of the peaks indicates that Ni-based LDHs have been prepared.

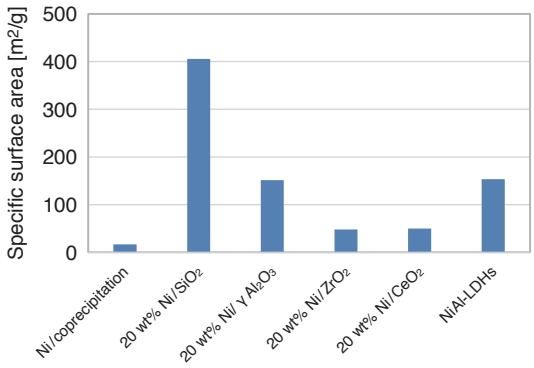


Fig. 5 Comparison of specific surface area

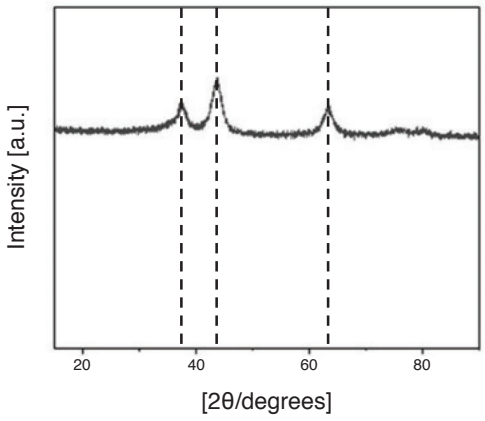


Fig. 6 XRD measurement result of NiAl-LDHs

Table 1 Results of methane synthesis processes

Catalysts	CO conversion [%]	Selectivity [%]						CH_4 yield [%]
		CO_2	CH_4	C_2	C_3	C_4	C_{5+}	
Ni-coprecipitation	55.9	0	97	2.8	0.3	0	0	54.2
20 wt% Ni/SiO_2	99.5	2.9	93.7	2.9	0.4	0	0	93.2
20 wt% $Ni/\gamma Al_2O_3$	92	0.4	87.9	7.4	3.3	0.8	0.2	80.9
20 wt% Ni/ZrO_2	82.2	0.1	89.7	7.1	2.5	0.4	0.1	73.7
20 wt% Ni/CeO_2	47	7	50.6	15.1	14.1	5.7	7.5	23.8
NiAl-LDHs	100	0.2	99.8	0	0	0	0	99.8

5.2. Catalytic Activity

A comparison of CO conversion, selectivity, and yield of NiAl-LDHs with catalysts with different carriers using Ni is shown in Table 1. First, other than NiAl-LDHs, 20 wt% Ni/SiO_2 performed more effectively than other catalysts, with 99.5% CO conversion, 93.7% CH_4 selectivity, and 93.2% CH_4 yield. However, CO_2 was also generated in addition to CH_4. On the other hand, NiAl-LDHs achieved 100% CO conversion, 99.8% CH_4 selectivity, and the highest yield of 99.8%. NiAl-LDHs showed greater activity than 20 wt% Ni/SiO_2 despite having lower specific surface area, indicating that it is a reliable catalyst for methanation.

5.3. Reaction Characteristics

Syngas methanation characteristics were evaluated for NiAl-LDHs at different temperatures. The reaction conditions other than temperature were the same as in the catalytic activity experiment. The results for different temperatures are shown in Fig. 7. Both CO conversion and CH_4 selectivity were lower at temperatures below 250°C, while CO conversion reached 100% at 275°C, and CH_4 selectivity reached 100% at 300°C. In general, methanation reactions often show higher reactivity at temperatures from 250°C to 500°C[10), 12)] and pressures from 0.1 to several MPa for Ni-based catalysts, but lower temperatures and higher pressures are advantageous for exothermic reactions. With the NiAl-LDHs developed in this study, the reaction temperature is comparable to those values, while the above results were obtained at a reaction pressure of 0.1 MPa. This confirms the advantageous properties of higher activity at lower temperatures and lower pressures and the ability to reduce pressurization energy.

5.4. Impact When There Is Water Contamination of the Source Gas

Under the same conditions as in the catalytic activity experiment, the proportion of water vapor supplied in the syngas was calculated based on Eq. (10).

$$Water\ vapor\ proportion\ [\%] = \frac{F_{H_2O}\,(gas)}{(F_{H_2O}\,(gas) + F_{syngas})} \quad (10)$$

As shown in Fig. 8, the CO conversion of the benchmark 20 wt% Ni/SiO_2 decreased as the proportion of water vapor increased, and the CO conversion decreased to 50% when the water vapor content was 20%. On the other hand, the CO conversion of NiAl-LDHs remained close to 100% even when the water vapor proportion increased from 20% to 65%. As shown in Fig. 9, at water vapor proportions up to 20%, the CH_4 selectivity of LDHs was equivalent to that when no water vapor contamination was present, and started to gradually decrease after a water vapor proportion

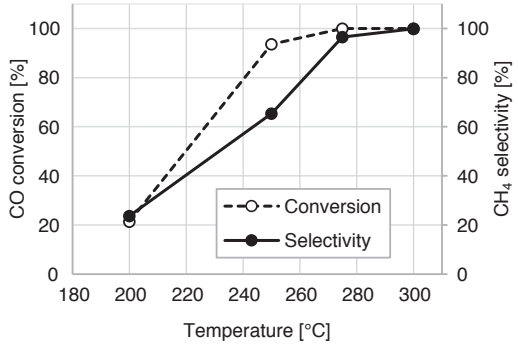


Fig. 7 Reaction characteristics of NiAl-LDHs

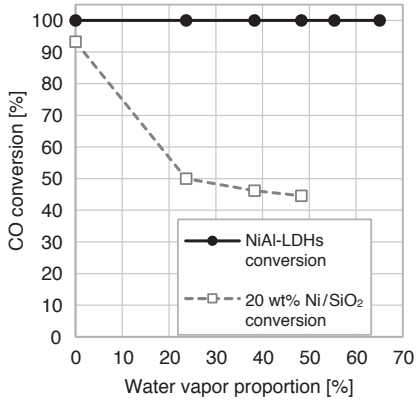


Fig. 8 Conversion of each catalyst with water vapor

of 20%. When the water vapor contamination reached 65%, CH_4 selectivity was 55%. CO_2 selectivity increased as CH_4 selectivity decreased, suggesting that a water gas shift occurred when the NiAl-LDHs exceeded the allowable moisture proportion. These results indicate that the catalyst is more water resistant than conventional supported catalysts and that the allowable proportion of water vapor in the source gas is 20%.

The use of the NiAl-LDHs developed in this study could help to eliminate the need to remove residual water vapor from the source gas, potentially reducing energy consumption by up to 58 kW.

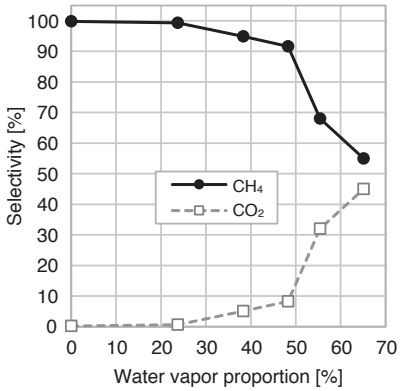


Fig. 9 Selectivity of NiAl-LDHs when mixed with water vapor

6. Conclusion

NiAl-LDHs were developed as a thermo-chemical catalyst to generate methane from syngas to increase the efficiency of the synthetic methane production process, and the following results were obtained.

(1) NiAl-LDHs exhibited 100% CO conversion and almost 100% CH_4 selectivity at 0.1 MPa and 300°C. Among the catalysts used for comparison, 20 wt% Ni/SiO_2 had about the same CO conversion as NiAl-LDHs, but the NiAl-LDHs had higher CH_4 selectivity and CH_4 yield than even those of 20 wt% Ni/SiO_2.

(2) Comparing NiAl-LDHs and 20 wt% Ni/SiO_2, a Ni based catalyst used for comparison, in terms of the allowable proportion of water vapor in the source gas showed that the CO conversion and CH_4 selectivity of 20 wt% Ni/SiO_2 decreased immediately when there was water vapor contamination, whereas NiAl-LDHs maintained its activity without decrease at up to 20% water vapor contamination, showing reliable water resistance.

Based on the above, the use of the NiAl-LDHs developed in this study is expected to contribute to the building of more efficient methane production processes in the future by reducing energy consumption in the dehydration process and downsizing it upstream of the methanation reactor, thereby increasing process efficiency and reducing process costs.

References

(1) Agency for Natural Resources and Energy: Annual Report on Energy for FY2023 (Energy White Paper 2024), p. 252-254, (2024), https://www.enecho.meti.go.jp/about/whitepaper/2024/pdf/whitepaper2024_all.pdf, (accessed 2025/01/13) (in Japanese)

(2) Agency for Natural Resources and Energy: Document 4-2, Methanation Torikumi Map 2023, The 9th Methanation Suishin Kanmin Kyougikai, https://www.meti.go.jp/shingikai/energy_environment/methanation_suishin/pdf/009_04_02.pdf, (accessed 2025/01/13) (in Japanese)

(3) Agency for Natural Resources and Energy: Document 3-1, The 27th Advisory Committee for Natural Resources and Energy, Electricity and Gas Industry Committee, Electricity and Gas for Small Enterprise, WG for Reviewing the Electricity and Gas industry system, https://www.meti.go.jp/shingikai/enecho/denryoku_gas/denryoku_gas/gas_jigyo_wg/pdf/027_03_01.pdf, (accessed 2025/01/13) (in Japanese)

(4) Mouri, M., Yanagisawa, K., Maki, M., Yoshida, J.: Investigation of CO_2 and Steam Co-electrolysis Conditions for Synthesizing Carbon Recycled Methane, Honda R&D Technical Review, Vol. 34, No. 1, p. 74-82

(5) Shimada, H., Yamaguchi, T., Araki, T., Mizusawa, T., Mori, M., Fujishiro, Y.: 2. Kotai Sankabutsugata Denkai Seru wo mochiita Kyoudenkai Gijyutsu no Kaihatsu, Electrochemistry, The Electrochemical Society of Japan, Vol. 84, Issue 8, p. 611-615, (2016) (in Japanese), doi: 10.5796/electrochemistry.84.611

(6) Araki, T., Mizusawa, T.: 3. SOEC wo mochiita Suijouki Denkai Shisutemu no Teijyou Saikuru Keisan, Electrochemistry, The Electrochemical Society of Japan, Vol. 84, Issue 8, p. 616-619, (2016) (in Japanese), doi: 10.5796/electrochemistry.84.616

(7) Matsunaga, K., Yoshino, M., Watanabe, H.: Hydrogen Production and Hydrogen Power Storage System Using Solid Oxide Electrolysis Cells, TOSHIBA REVIEW, Vol. 71, No. 5, p. 41-45, (2016) (in Japanese), https://www.global.toshiba/content/dam/toshiba/migration/corp/techReviewAssets/tech/review/2016/05/71_05pdf/b04.pdf, (accessed 2025/01/13)

(8) Suito, F., Nagata, Y., Araki, T., Yanagisawa, K., Mouri, M., Maki, M., Yoshida, J.: Effect of the Air Contamination in Infrastructure Emission CO_2 Source Gas of Co-Electrolytic SOEC System on the Energy Efficiency, Transactions of Society of Automotive Engineers of Japan, Vol. 53, Issue 5, p. 973-978, (2022) (in Japanese), doi: 10.11351/jsaeronbun.53.973

(9) Meng, F., Li, X., Li, M., Cui, X., Li, Z.: Catalytic performance of CO methanation over La-promoted Ni/Al_2O_3 catalyst in a slurry-bed reactor, Chemical Engineering Journal, Vol. 313, p. 1548-1555, (2017), doi: 10.1016/j.cej.2016.11.038
(10) Yang, X., Wang, X., Gao, G., Wendurima, Liu, E., Shi, Q., Zhang, J., Han, C., Wang, J., Lu, H., et al.: Nickel on a macro-mesoporous Al_2O_3@ZrO_2 core/shell nanocomposite as a novel catalyst for CO methanation, International Journal of Hydrogen Energy, Vol. 38, Issue 32, p. 13926-13937, (2013), doi: 10.1016/j.ijhydene.2013.08.083
(11) JP, 2024-063277, A
(12) He, F., Zhuang, J., Lu, B., Liu, X., Zhang, J., Gu, F., Zhu, M., Xu, J., Zhong, Z., Xu, G., et al.: Ni-based catalysts derived from Ni-Zr-Al ternary hydrotalcites show outstanding catalytic properties for low-temperature CO_2 methanation, Applied Catalysis B: Environmental, Vol. 293, Article 120218, (2021), doi: 10.1016/j.apcatb.2021.120218
(13) He, L., Lin, Q., Liu, Y., Huang, Y.: Unique catalysis of Ni-Al hydrotalcite derived catalyst in CO_2 methanation: cooperative effect between Ni nanoparticles and a basic support, Journal of Energy Chemistry, Vol. 23, Issued 5, p. 587-592, (2014), doi: 10.1016/S2095-4956(14)60144-3
(14) He, H., Xiao, J., Liu, Z., Yang, B., Wang, D., Peng, X., Zeng, L., Li, Z., Lei, L., Qiu, M., et al.: Boosting the hydrogen evolution of layered double hydroxide by optimizing the electronic structure and accelerating the water dissociation kinetics, Chemical Engineering Journal, Vol. 453, Part 1, Article 139751, (2023), doi: 10.1016/j.cej.2022.139751
(15) Morimoto, K.: Syntheses of layered double hydroxides and their application to functional materials, Japanese Magazine of Mineralogical and Petrological Sciences, Vol. 48, Issue 1, p. 46-50, (2019) (in Japanese), doi: 10.2465/gkk.190130

■ Author ■


Hideaki YONEDA


Kazuki YANAGISAWA


Yohei IIJIMA


Junpei YOSHIDA

Honda R&D Co., Ltd.
Honda R&D Technical Review Editorial Committee

Material Research Center, Honda R&D Co., Ltd.
Innovative Research Excellence, Honda R&D Co., Ltd.
Innovative Research Excellence, Power Unit & Energy, Honda R&D Co., Ltd.
Motorcycle and Power Product Operations, Honda Motor Co., Ltd.
Automobile Development Operations, Honda Motor Co., Ltd.

Editorial Afterword

Thank you for reading the Honda R&D Technical Review Vol. 38 to the end. I would like to express my heartfelt gratitude to everyone that made this issue possible, including the contributing authors, the people who supported the publication process, and everyone involved in the many peer review meetings. We can all be proud of the hard work we've done.

This is my second year participating in the peer reviews for this publication. Many of the papers and technology descriptions were outside of my own field of technical expertise, and the level of technology was extremely high, so it took time for me to fully understand each of the papers. In gaining this understanding, I have striven to make the articles easier and more pleasurable to read. Of course, members of the review committee and area specialists were also involved in the process, and the result is a Honda R&D Technical Review that offers depth to specialists while presenting content that people outside of the field can also understand. I believe that many people read the papers with great interest.

I myself have been involved in research and development at Honda since joining the company more than 30 years ago, and I have always been mindful of trends or principles that can be found in test results. In product development and R&D, there is always some kind of pattern or principle waiting to be discovered, regardless of the scale or breadth of the work. Such discoveries can lead to all manner of improvements, such as to performance, quality, or efficiency. When you find principles in your work that you think can be applied more broadly, I urge you to submit a paper to the Honda R&D Technical Review. While it may take some time, the resulting sense of achievement will surely more than make up for the effort.

I look forward to receiving more research results in the future, and dream of the day that people will say, "that Nobel Prize-winning technology first appeared in the pages of the Honda R&D Technical Review."

Toru KITAMURA

©2026 Honda R&D Co., Ltd. (All rights reserved)

ISSN 2187-381X

Published by:
Honda R&D Co., Ltd.
1-4-1, Chuo Wako-shi, Saitama-ken,
351-0193, Japan

Edited by:
Honda R&D Co., Ltd.
Honda R&D Technical Review Editorial Committee

Printed:
Mar. 1, 2026

Honda R&D Co., Ltd.