DATA ORIENTED EVALUATION OF FLIGHT CHARACTERISTIC MODELS BY USING BAYES FACTOR

Masaru Naruoka[1] & Tetsujiro Ninomiya[1]

[1]Japan Aerospace Exploration Agency

Abstract

This study describes a model evaluation method using the Bayes factor for application to the modeling of aircraft flight characteristics with flight data. The Bayes factor is the ratio of the marginal likelihoods computed from two postulated models of interest using common data. It can be used as one of the metrics to indicate which model is more plausible in realizing the data. Using this Bayes factor method, the authors re-evaluate our previous findings and results in terms of model evaluation. First, numerical examples confirm that a Student's t-distributed error model, which was proposed to ensure robustness against outliers for modeling, is superior to a normally distributed model. Then, two issues regarding the previous flight characteristics modeling are discussed. The first concerns the inclusion of the pitch rate effect in the lift coefficient modeling. The second is the decomposition of the wave drag effect for the drag coefficient modeling. The first and second responses can be evaluated quantitatively in terms of model evaluation using Bayes factors. The results are consistent with those of the previous studies, which were only qualitatively evaluated.

Keywords: Flight characteristics, Bayesian inference, Markov-chain Monte-Carlo, Bayes factor

1. Introduction

Flight characteristics, i.e., how aircraft in a configuration flies in an environmental condition, are essential to understand aircraft performance and capability. One of the ways to acquire the characteristics is to build a mathematical model by using time history data gathered with flying aircraft. The time history data is so redundant and noisy to be used for other applications that this modeling procedure is performed. Thus, in order to summarize the original data into an accurate and reliable model, the modeling procedure was studied mainly from the aspect of time [1] and frequency [2] domains.

While these conventional studies are still valuable, the series of our previous studies [3, 4] were another approach; a data driven method based on Bayesian inference. The feature of our studies were to treat statistical models, which provide distribution estimation of model parameters instead of their point estimation. The estimation of the parameters was effectively performed with Markov-chain Monte-Carlo (MCMC) sampling, one of the Bayesian inference methods and well summarized in Chapter.11 of Bishop [5]. The robustness of the estimation was also studied by postulating a Student's t-distribution, which is less sensitive to outliers than a normal distribution. To extend our previous studies, a practical guide to evaluate the models will be provided in this study. In parallel to finding the best parameters in a certain model, which was focused in our previous studies, this guide is intended to select the best model among various possible models. For the demonstration purpose, the models consisting of the aerodynamic coefficients and the control surface effectiveness of the conventional fixed-wing aircraft are focused as same as our previous studies.

The reason why the authors have been focused on such established field of the modeling of the flight characteristics is the current digital transformation movement of aircraft design. The aim of this movement is to reduce trial production and physical tests by replacing them to their corresponding digital computation for minimization of the development cost. In terms of the activities to acquire the flight characteristics, calculation with computational fluid dynamics (CFD) is demanded to replace wind tunnel tests (WTT) and flight tests. The improvement of the CFD will result in shortening the conventional procedure of the data acquisition, in which the CFD predicts the characteristics in combination with WTT results, then, the flight tests are performed to verify the prediction.

Using the difference among the CFD, WTT and flight test results is one of the methods to improve the accuracy and reliability of the CFD. Actually, the CFD has been well tuned by using the WTT results. On the other hand, it is much more difficult to use the flight test data as feedback to the CFD compared to the WTT results. There are two major disadvantages for the flight tests in terms of the measurement environment. The first is that the environment cannot be configured arbitrarily. For example, the combinations of the airspeed and angle of attack (AoA), both of which are essential parameters, are limited such that the aircraft continues to fly safely. Secondly, the environment varies unintentionally due to irregularities of airflow such as gusts.

In order to overcome these drawbacks to synthesize the flight test results into reasonable estimates of the flight characteristics, the authors have studied the application of Bayesian inference. As the demonstration of the method, the lift and drag of JAXA's research aircraft "Hisho", a Cessna C680 business jet, were successfully modeled by using its flight data [3]. In addition, it was found that the robustness to outliers included in the flight data was improved with postulation that the residual between the data and the model followed Student's t-distribution instead of a normal distribution [4].

Although the effectiveness of the Bayesian estimation was validated in terms of the model parameter estimates, the discussion on the model structure, i.e., which parameters and explanatory variables should be used to form a mathematical model, has still been open. Therefore, the authors propose to use the Bayes factor for the model evaluation in this study. The Bayes Factor is the ratio of the marginal likelihood of one model to that of another model when a data is given.If the Bayes Factor takes a value greater than one, there is evidence in favor of the former model relative to the latter, in other words, the data are more probable to realize under the former model. This Bayes factor method is more intuitive than the conventional methods of the model evaluation using p-value or Akaike's information criterion (AIC). In addition, sampling based approximation methods of the Bayes factor developed in the recent studies like Gronau et al. [6] is preferable for this study. While these methods are originally intended to overcome the practical problem on the analytical calculation of the Bayes factor, they have the advantage to reuse the MCMC sampling results that the authors used for the parameter estimation.

In the following, the framework to perform the model evaluation of the flight characteristics by using the Bayes factor is explained. Firstly, the Bayes factor is briefly explained with its definition and derivation in Sec. 2. The tools used for the calculation of the Bayes factor are also introduced. Then, application of model evaluation to a simple regression problem is performed with numerical simulation to discuss the effectiveness and calculation setup of the Bayes factor. In Sec. 3, actual applications of the Bayes factor to the model evaluation for the flight characteristics are discussed. There are two example targets; pitch rate effect in the lift modeling, and wave drag contribution in the drag modeling, both of which were treated in the authors previous study. In this section, the flight data used for the model evaluation is also briefly explained. Finally, this study is concluded in Sec. 4.

2. Model evaluation by using Bayes factor

In this section, the method to compare the models by using the Bayes factor is elaborated. Firstly, the Bayes factor, which is the ratio indicating how a model is more plausible to realize a data than another model, is explained in terms of its definition and derivation. The tools to calculate the Bayes factor and its configuration are also described. Then, the numerical examples of the simple linear regression problem are introduced. The examples are intended to show not only the fundamental effectiveness to use the Bayes factor, but also the difference resulted from the setup of the model, tools, and source data. The insight obtained in the examples will be reflected to the actual model evaluation of the flight characteristics in Sec. 3.

2.1 Bayes factor

In the following, the model evaluation with Bayes factor, whose details are well described in Gronau's tutorial [6], is briefly explained. It is assumed that there are two candidates of models described as $\mathfrak{M}_i$ $(i = 1 \text{ or } 2)$. According to ordinal Bayesian inference framework, which the model is superior to the other model is discussed by comparing their posterior probability $p(\mathfrak{M}_i|y)$ in which data y has been observed. In other words, the odds o_{12} of $\mathfrak{M}_1$ against $\mathfrak{M}_2$,

$$o_{12} \equiv \frac{p(\mathfrak{M}_1|y)}{p(\mathfrak{M}_2|y)}, \tag{1}$$

is greater than one, the more plausible $\mathfrak{M}_1$ is. According to Bayes' theorem, the posterior probability of the model $p(\mathfrak{M}_i|y)$ is proportional to the multiplication of its likelihood and prior probability:

$$\underbrace{p(\mathfrak{M}_i|y)}_{\text{posterior}} = \frac{p(y|\mathfrak{M}_i)\,p(\mathfrak{M}_i)}{p(y)} \propto \underbrace{p(y|\mathfrak{M}_i)}_{\text{likelihood}}\,\underbrace{p(\mathfrak{M}_i)}_{\text{prior}} \quad . \tag{2}$$

Thus, Eq. (1) yields

$$\underbrace{o_{12}}_{\text{posterior odds}} = \underbrace{\frac{p(y|\mathfrak{M}_1)}{p(y|\mathfrak{M}_2)}}_{\text{Bayes factor}} \times \underbrace{\frac{p(\mathfrak{M}_1)}{p(\mathfrak{M}_2)}}_{\text{prior odds}}, \tag{3}$$

which shows that the posterior odds can be discussed by using the first term on the right side called Bayes factor, because the second term of the prior odds is user defined (usually one). Therefore, the comparison of the models is accomplished by calculating their Bayes factor, more specifically, the likelihood $p(y|\mathfrak{M}_i)$.

In order to calculate the likelihood $p(y|\mathfrak{M}_i)$, suppose that the model $\mathfrak{M}_i$ consists of parameters $\theta_j\,(j = 1, 2, \cdots)$. The model is probabilistically configured; it is possible that the model $\mathfrak{M}_i$ having different values of the parameters θ_j generates the same data y. Thus, $p(y|\mathfrak{M}_i)$ is calculated by marginalizing over θ_j as

$$p(y|\mathfrak{M}_i) = \iint \cdots\; p(y|\theta_1, \theta_2, \cdots, \mathfrak{M}_i)\; p(\theta_1, \theta_2, \cdots|\mathfrak{M}_i)\; d\theta_1 d\theta_2 \cdots. \tag{4}$$

Eq. (4) has multiple integrals, and the analytical difficulties arise when the model is complex, for example, if it has many parameters or nonlinearities. In order to overcome such problem, the sampling based approximation methods represented by bridge sampling [7], have been developed. These methods use samples of θ_j and their corresponding likelihood $p(y|\theta_j)$, and θ_j are generated base on different distributions depending on the method. For instance, the naive Monte Carlo estimator, the simplest approximation method, uses the prior distribution $p(\theta_j)$ for the generation.

In short, $p(y|\mathfrak{M}_i)$ for the Bayes factor can be approximated by using the samples that the authors generated in the previous studies. This is because the generated samples for the parameter estimation were governed by the relation of the posterior and prior distributions of θ_j as

$$\underbrace{p(\theta_1, \theta_2, \cdots|y)}_{\text{posterior}} \propto \underbrace{p(y|\theta_1, \theta_2, \cdots)}_{\text{likelihood}}\;\underbrace{p(\theta_1, \theta_2, \cdots)}_{\text{prior}}. \tag{5}$$

Although the model symbols $\mathfrak{M}_i$ are dropped in Eq. (5) due to the description for a certain model, the integrand terms of Eq. (4) are the same as the right side of Eq. (5). In addition, there is a tool to calculate the Bayes factor in cooperation with a MCMC sampler. In this study, the authors use the "bridgesampling" [8] and "rstan" [9] R packages for the Bayes factor calculation and the MCMC sampling, respectively. The actual work flow to calculate the Bayes factor after the data acquisition is composed of three steps; firstly, the models for the evaluation are prepared in the format of Stan, a statistical modeling platform [10]. The models are so essntial that they will be shown in the Stan format in the following demonstration. Then, *rstan::sampling* function is used with the data and models in order to generate the samples of the model parameters. The additional arguments explicitly passed to the function are 2000, 3, and 4 for iterations, thinning counts, and chains, respectively. Finally, the

marginal likelihood $p(y|\mathfrak{M}_i)$ are estimated with the samples by using *bridgesampling::bridge_sampler* function. For the estimation method options of this function, in addition to the default "normal", the "warp3", which transforms the posterior distribution for the better estimation, is used. The resultant marginal likelihood is transformed to the Bayes factor by using *bridgesampling::bf* function.
It is noted that the prior distribution for the Bayes factor calculation is known to affect the results more than that for the parameter estimation. This is due to approximation of integration by using samples, and for accurate approximation, the samples must be generated to cover a sufficiently large integral space. The prior $p(\theta_1, \theta_2, \cdots|\mathfrak{M}_i)$ in Eq. (4), which is included into the integrant terms, has an important role to scatter the samples. Thus, the prior for the Bayes factor should be carefully configured. According to Rouder [11], normal and Cauchy distributions having heavily tails recommended as the default priors, which is widely applicable and computationally convenient. In addition, a normal distribution whose variance parameter is governed by an inverse gamma distribution is also recommended for multivariate cases. They differ from the weak informative distribution used in the previous study for the parameter estimation. In order to discuss the effect to the Bayes factor, multiple priors are applied in the following numerical examples.

2.2 Numerical examples

To demonstrate the evaluation of models by using the Bayes factor, the numerical examples with simulated data by using an univariate linear model will be shown. Especially, the postulation that the residuals are followed by a Student's t-distribution instead of a normal distribution, which was proposed in the previous study [4] to improve the robustness against the outliers, will be discussed from the aspect of the evaluation by using the Bayes factor.
The setup of the examples is explained in the following three parts; input data, postulated models, and prior distributions. The input data consists of 1000 samples per data set, and 50 data sets are prepared. Each sample is composed of x and y according to a linear equation

$$y = x + \varepsilon, \tag{6}$$

where x is randomly drawn from a normal distribution $N(0,1)$, which represents the mean and standard deviation of 0 and 1, respectively. The error ε is randomly drawn from another normal distribution $N(0,0.1)$ or a uniform distribution ranging from -1 to 1. The ε derived from the uniform distribution are intended to act as outliers, although the normal distribution behaves like measurement noise. A predefined mixing ratio of outliers ϕ_u regulates how many of the ε samples are derived from the uniform distribution to generate the 1000 y samples. By changing the ratio ϕ_u, multiple subsets of a data set are formed. More precisely, firstly, two 1000 samples are generated by using the normal and uniform distributions. Then, by concatenating the first $1000 \times (1-\phi_\text{u})$ and the first $1000 \times \phi_\text{u}$ samples of the former and later samples, respectively, the ε samples in a subset are made.
The postulated model is the same linear equation as the sample generator Eq (6), but has an additional slope parameter a to be estimated as

$$y \sim ax + \varepsilon. \tag{7}$$

The a is assumed to follow a normal distribution for the estimation. The assumption for ε is either normal distributed $\sim N(\mu, \sigma)$ or Student's t-distributed $\sim \text{Student\_t}(\nu, \mu, \sigma)$. All the parameters of these distributions are estimated, i.e., the mean μ and standard deviation σ for the normal distributions, and additionally the degrees of freedom ν for the Student's t-distribution. Because of the addition of the degrees of freedom to the estimated parameters, the Student's t-distribution is capable to have thicker tails than the normal distribution. Thus, this setting is intended to show that the error postulation by using the Student's t-distribution will be supported instead of the normal distribution when the outliers are included.
As described in Sec. 2.1, two distributions are examined as the prior of the slope parameter a; normal and Cauchy called as the prior $(\mathfrak{n})$ and $(\mathfrak{c})$, respectively. In addition, the uniform distribution over parameter boundaries, which is used as the default prior in Stan when no specific prior is given, is investigated as the prior $(\mathfrak{u})$. The priors for the error ε is not specified, which means the default prior is used for μ and σ, and also ν if Student's t-distribution is selected.

The above description is implemented as Listing 1 and 2 for the normally distributed and Student's t-distributed error models, respectively. They are the input codes to Stan, and define the postulated models and the priors. The lines starting from the comment-out marks "//" in the model sections will be activated by removing the marks when the normal $(\mathfrak{n})$ or Cauchy $(\mathfrak{c})$ prior is assumed, while these lines are unchanged for the default uniform $(\mathfrak{u})$ prior.

Listing 1: Normally distributed error model in Stan format

```
data {
  int<lower=0> N;
  vector[N] x;
  vector[N] y;
}
parameters {
  real a;
  real mu;
  real<lower=0> sigma;
}
model {
  target += normal_lpdf(y | a * x + mu, sigma);
  // target += normal_lpdf(a | 0, 1); //  Activate when the prior (n)
  // target += cauchy_lpdf(a | 0, 1); //  Activate when the prior (c)
}
```

Listing 2: Student's t-distributed error model in Stan format

```
data {
  int<lower=0> N;
  vector[N] x;
  vector[N] y;
}
parameters {
  real a;
  real mu;
  real<lower=0> sigma;
  real<lower=0> nu;
}
model {
  target += student_t_lpdf(y | nu, a * x + mu, sigma);
  // target += normal_lpdf(a | 0, 1); //  Activate when the prior (n)
  // target += cauchy_lpdf(a | 0, 1); //  Activate when the prior (c)
}
```

The results are summarized in Tables 1-3, and estimated with the warp3 method of the *bridge_sampler* function. In the following, the results with the normal method of the function are omitted because their differences from the warp3 results are sufficiently small such as the four most significant digits are the same. Table 1 shows the logarithm of the estimated marginal likelihood, which are transformed to the logarithm of the Bayes factors indicated in Tables 2-3. All items in the tables except for (0), which will be explained later, consist of the mean and the standard deviation of the values calculated with the 50 data sets.

A Bayes factor is a relative value, and two values of the marginal likelihood are required for its calculation like Eq. (3) shows. The denominators for the Bayes factors listed in Tables 2 are the marginal likelihood values of the normally distributed error model and the $(\mathfrak{u})$ prior for each mixing ratio of the outliers. The denominators for Table 3 are the values of a subset without the outliers for each error model and prior. Although a Bayes factor is originally used to compare two models on the same data set, the Bayes factor obtained by applying two different data sets to the same model as Table 3 are useful for examining the robustness of the model to the data. The (0) symbols in the tables represent those items used as the base likelihood for the calculation. Note that since the Bayes factors in the tables are natural logarithmic, being a positive number rather than greater than one, the original criterion, means that the corresponding model is more plausible than the base model. In addition, according to Kass [12], a Bayes factor greater than 150 means that there is a very strong evidence to select the model corresponding to the numerator instead of the denominator. This criterion is also used by transforming the threshold to $5 \approx \log(150)$.

Table 1 – Logarithm of marginal likelihood

Error model		Normal			Student's t		
Prior		$(\mathfrak{u})$ (Default)	$(\mathfrak{n})$ Normal	$(\mathfrak{c})$ Cuachy	$(\mathfrak{u})$ (Default)	$(\mathfrak{c})$ Normal	$(\mathfrak{n})$ Cauchy
Mixing ratio of outliers (ϕ_u)	0%	834.3 ± 1.7	832.9 ± 1.7	832.4 ± 1.7	829.4 ± 1.8	828.0 ± 1.7	827.6 ± 1.8
	1%	826.6 ± 6.4	825.1 ± 6.4	824.8 ± 6.4	814.5 ± 4.1	813.0 ± 4.2	812.6 ± 4.2
	2%	802.1 ± 12.0	800.7 ± 12.0	800.3 ± 12.1	804.6 ± 4.4	803.2 ± 4.4	802.7 ± 4.5
	3%	770.3 ± 15.4	768.8 ± 15.3	768.4 ± 15.4	795.9 ± 4.2	794.4 ± 4.2	794.0 ± 4.2
	4%	734.9 ± 17.2	733.4 ± 17.2	732.9 ± 17.1	787.5 ± 4.1	786.0 ± 4.1	785.6 ± 4.0
	5%	700.5 ± 18.4	698.9 ± 18.4	698.6 ± 18.4	779.9 ± 4.3	778.6 ± 4.4	778.1 ± 4.4
	10%	535.0 ± 22.8	533.5 ± 22.9	533.0 ± 22.9	743.8 ± 4.6	742.5 ± 4.6	742.1 ± 4.5

Table 2 – Logarithm of Bayes factor per each mixing ratio of outliers

Error model		Normal			Student's t		
Prior		$(\mathfrak{u})$ (Default)	$(\mathfrak{n})$ Normal	$(\mathfrak{c})$ Cuachy	$(\mathfrak{u})$ (Default)	$(\mathfrak{c})$ Normal	$(\mathfrak{n})$ Cauchy
Mixing ratio of outliers (ϕ_u)	0%	(0)	-1.4 ± 0.2	-1.8 ± 0.2	-4.8 ± 0.6	-6.2 ± 0.6	-6.7 ± 0.7
	1%	(0)	-1.5 ± 0.3	-1.8 ± 0.3	-12.1 ± 5.0	-13.6 ± 5.0	-14.0 ± 5.0
	2%	(0)	-1.5 ± 0.6	-1.9 ± 0.6	2.4 ± 9.3	1.1 ± 9.2	0.6 ± 9.2
	3%	(0)	-1.5 ± 0.8	-1.9 ± 0.6	25.6 ± 12.4	24.1 ± 12.4	23.7 ± 12.3
	4%	(0)	-1.5 ± 1.0	-1.9 ± 0.9	52.7 ± 14.1	51.2 ± 14.0	50.8 ± 14.1
	5%	(0)	-1.6 ± 1.0	-1.9 ± 1.0	79.5 ± 14.8	78.1 ± 14.8	77.7 ± 14.8
	10%	(0)	-1.6 ± 1.1	-2.1 ± 1.1	208.8 ± 18.8	207.4 ± 18.8	207.0 ± 18.9

Table 3 – Logarithm of Bayes factor per each model and prior

Error model		Normal			Student's t		
Prior		$(\mathfrak{u})$ (Default)	$(\mathfrak{n})$ Normal	$(\mathfrak{c})$ Cuachy	$(\mathfrak{u})$(Default)	$(\mathfrak{c})$ Normal	$(\mathfrak{n})$ Cauchy
Mixing ratio of outliers (ϕ_u)	0%	(0)	(0)	(0)	(0)	(0)	(0)
	1%	-7.7 ± 6.9	-7.7 ± 6.8	-7.6 ± 6.9	-15.0 ± 3.6	-15.0 ± 3.6	-15.0 ± 3.7
	2%	-32.1 ± 12.3	-32.2 ± 12.2	-32.2 ± 12.3	-24.9 ± 4.0	-24.8 ± 4.0	-24.9 ± 4.1
	3%	-64.0 ± 15.5	-64.1 ± 15.4	-64.0 ± 15.5	-33.6 ± 3.8	-33.7 ± 3.8	-33.6 ± 3.9
	4%	-99.4 ± 17.2	-99.5 ± 17.2	-99.5 ± 17.1	-41.9 ± 3.8	-42.0 ± 3.8	-42.0 ± 3.7
	5%	-133.8 ± 18.4	-134.0 ± 18.4	-133.9 ± 18.5	-49.5 ± 4.2	-49.5 ± 4.2	-49.5 ± 4.2
	10%	-299.2 ± 23.2	-299.4 ± 23.3	-299.4 ± 23.3	-85.6 ± 5.2	-85.6 ± 5.1	-85.5 ± 5.0

In short, Table 2 shows which combination of model and prior is superior in data sets having the same number of the outliers, while Table 3 indicates how a setup of a model and a prior is affected by the outliers. They are easy to understand, and the results will be discussed based on Tables 1-3 instead of Table 1.

Firstly, the error models will be compared by using the Bayes factors in Table 2. The normally distributed error model is better than the Student's t-distributed error model, whose Bayes factors are significantly negative values, at the mixing ratio of the outliers are 0% and 1%. On the other hands, neither error model is superior at the 2% mixing ratio, and after 3%, the Student's t-distributed error model is clearly superior because their Bayes factors are sufficiently greater than the threshold value, 5. The fact that the Student's t-distributed error model is superior under the input data having more outliers is natural, because the Student's t-distribution can have a thicker tail than the normal distribution. However, it is interesting that Student's t-distributed error model, which can work well with the input data having less outliers by adjusting its parameter of degrees of freedom, is not always superior. It is concluded that the error model should be configured based on the nature of the input data.

Then, how the error models are affected by the outliers will be investigated based on the Bayes factors in Table 3. The mixing ratio increasing, the Bayes factors of both the normally and Student's t-distributed error models decrease. In addition, the normally distributed error model is more rapidly deteriorated than the Student's t-distributed error model. Thus, the Student's t-distributed error model is more robust against the data. Furthermore, this property is confirmed with the standard deviations, which is nearly the same for all mixing ratios in the Student's t-distribution error model, while the standard deviation increases with mixing ratio in the normally distributed error model.

Based on the above discussions from the aspect of the error model comparisons, the authors propose to use the Student's t-distributed error model for the modeling of flight characteristics. This is because, unlike the numerical examples, it is much more difficult to know the quantity and distribution of error in actual flight data. This proposal is consistent to the authors previous study [4], which shows the effectiveness of the Student's t-distribution in the parameter estimation context.

Finally, the priors will be examined with the comparison in the same error models. According to both Tables 2 and 3, there are differences in terms of the priors, but they are not sufficiently large. Every prior seems to work similarly. However, it is possible that this fact is specific to these numerical examples. Therefore, to be conservative, the priors recommended by Rouder [11] will be used for the following actual problems.

3. Application to model evaluation of flight characteristics

In this section, the applications of the model evaluation of the flight characteristics by using the Bayes factor will be demonstrated. The used data and target models were treated in the authors previous study. Firstly, the source data is briefly explained. Then, two targets of the model evaluation are described; the pitch rate component of the lift modeling and the wave drag component in the drag modeling.

3.1 Flight data

The source flight data is gathered with “Hisho” depicted in Fig. 1, one of research aircraft operated by JAXA. This aircraft was modified to install basic research equipment represented by a data acquisition system, and its original type was Cessna 680, which is twin-powered fixed-wing business jet aircraft. The acquisition system mainly consists of sensors, a network switch to control data flow, monitor consoles, and a data recorder. The time stamps obtained with a GPS receiver embedded in the network switch are attached to all the data acquired with the system, which is useful for general research. In this study, the airflow data such as airspeed and AoA, the attitude like the pitch angle, the control surface deflection, and the thrust setting are used as the flight data. The weight and thrust in flight are additionally estimated by using the acquired data. In order to indicated that the used data sufficiently covers the flight envelope, the plot of the airspeed and altitude of the data is depicted as colored lines in Fig. 2. The data is collected in 2016, 2018 and 2020, and its length is sufficient long, 31.9 hours. The details of the data is referred in [4], especially in Fig. 6.

Figure 1 – Target aircraft “Hisho”

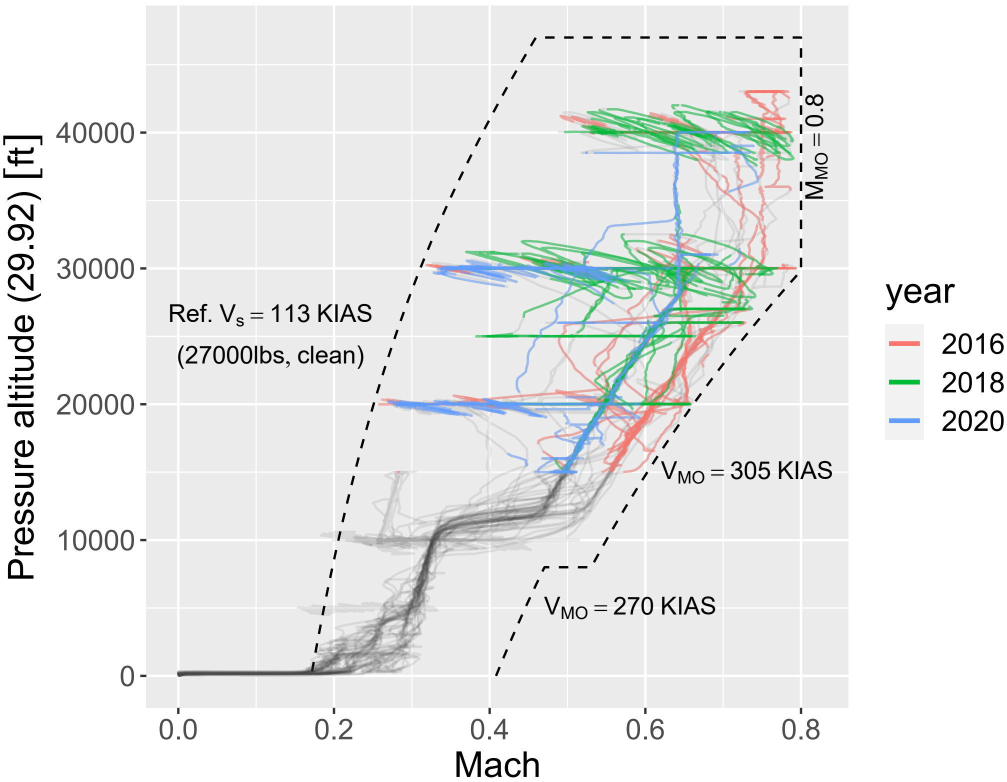

Figure 2 – Flight data

3.2 Application to pitch rate effect in lift modeling

In the previous study [4], the lift coefficient C_L of "Hisho" was modeled with a Student's t-distributed error model as the model $\mathfrak{M}_{\text{base}}$

$$\mathfrak{M}_{\text{base}} : C_L \sim \text{Student\_t}\left(\nu, C_{L0} + C_{L\alpha}\alpha + C_{L\delta_e}\delta_e, \sigma\right) , \tag{8}$$

where α and δ_e were angle of attack and elevator deflection, respectively. C_{L0}, $C_{L\alpha}$ and $C_{L\delta_e}$ were the estimated parameters. Note that the stabilizer deflection is dropped from the model because it is strongly correlated to the angle of attack and deteriorate the estimation. As the red dots in Fig. 3 shows, the residuals calculated by subtracting the model outputs from the flight data were correlated with the pitch rates q that were not included into the model, so the modeling was margin to improve. Therefore, the model was updated to the model $\mathfrak{M}_{\text{base}+q}$ by adding another parameter C_{Lq} of the contribution derived form the pitch rate change to

$$\mathfrak{M}_{\text{base}+q} : C_L \sim \text{Student\_t}\left(\nu, C_{L0} + C_{L\alpha}\alpha + C_{L\delta_e}\delta_e + C_{Lq}q, \sigma\right) , \tag{9}$$

and the difference of the flight data and the model outputs were improved to the blue dots in Fig. 3.

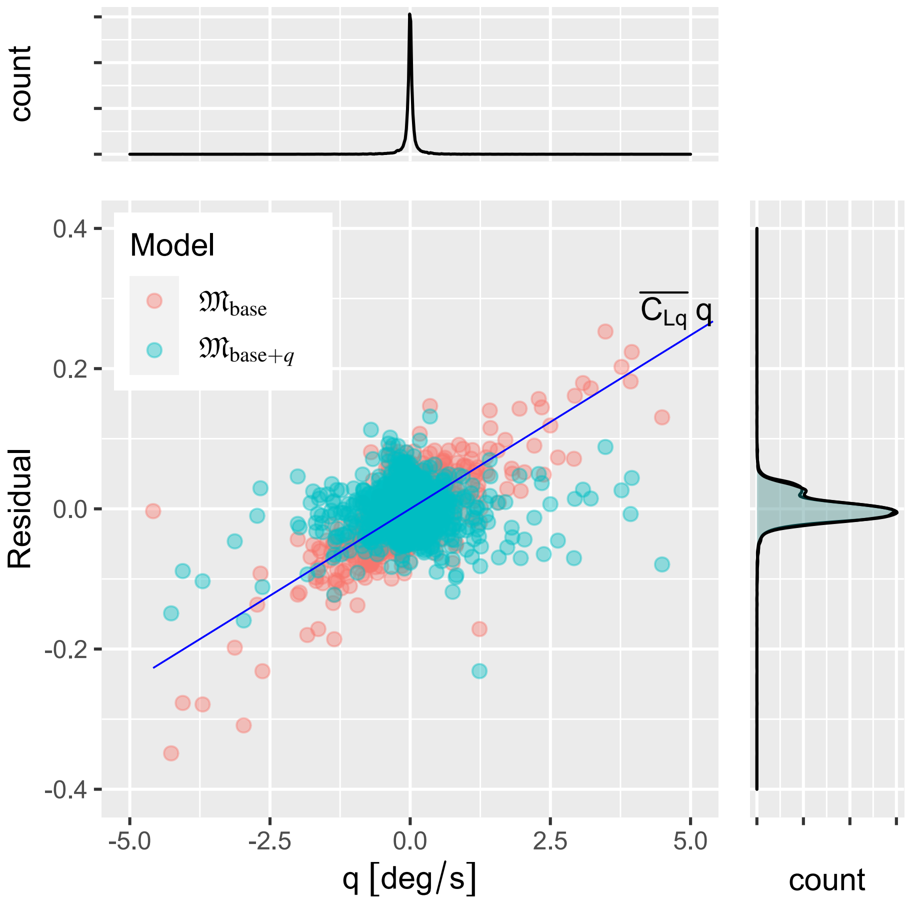

Figure 3 – Residual in models with and without pitch rate contribution

In this study, the improvement is investigated again in terms of the model evaluation. Equations (8) and (9) are transformed to the Stan code in Listing 3. The priors for the parameters are the Cauchy distributions based on the discussion in the numerical examples.

Listing 3: Lift coefficient model in Stan format

```
data {
```

```
  int<lower=1> N;
  vector[N] alpha_deg;
  vector[N] de_deg;
  vector[N] q_dps;
  vector[N] cL;
}
parameters {
  real cL0;
  real cLa;
  real cLde;
  //real cLq; //  Activate when 𝔐base+q, i.e., Eq. (9)
  real<lower=0> sigma_cL;
  real<lower=0> nu_cL;
}
model {
  target += student_t_lpdf(cL |
      nu_cL,
      cL0 + (cLa * alpha_deg) + (cLde * de_deg)
        // + (cLq * q_dps) //  Activate when 𝔐base+q, i.e., Eq. (9)
        ,
      sigma_cL);
  target += cauchy_lpdf(cL0 | 0, 1);
  target += cauchy_lpdf(cLa | 0, 1);
  target += cauchy_lpdf(cLde | 0, 1);
  // target += cauchy_lpdf(cLq | 0, 1) //  Activate when 𝔐base+q, i.e., Eq. (9)
}
```

The Bayes factor of the two models is

$$\log BF\left(\frac{\mathfrak{M}_{\text{base}+q}}{\mathfrak{M}_{\text{base}}}\right) = 777.7\,. \tag{10}$$

This means that the model $\mathfrak{M}_{\text{base}+q}$ is supported again from the aspect of the model evaluation by using the Bayes factor. In other words, it is concluded that the pitch rate effect to the lift coefficient must be considered when using this flight data.

3.3 Application to wave drag component in drag modeling

The next application is the drag modeling. It is well knwon that the drag is decomposed into two parts; components due to lift or not. The former is knwon as the induced drag, which increases the drag coefficient C_D by ${C_L}^2$, the square of the lift coefficient. On the other hand, the latter is the parasite drag, which is often discussed in a more broken-down manner. The textbook written by Raymer [13] says that the parasite drag is composed as Fig. 4 shows. The wave drag depicted by the blue area was focused on in the previous study [3]. It quickly increases when a Mach number M exceeds a critical Mach number Mc, and is sufficiently large part of the parasite drag when M is bigger than a drag divergent Mach number M_{DD}. There is a relationship that $M_c < M_{\text{DD}} < 1$, and a well-known definition is that the difference of the parasite drag coefficients at M_c and M_{DD} is 0.0020. The maximum cruise speed of the target aircraft “Hisho” is 0.8, so it might be possible to model the wave drag by using the flight data.

In the previous study, the postulated model was

$$C_D \sim \text{Normal}\,(\underbrace{C_{Di}{C_L}^2}_{\text{induced}} + C_{D0} + \underbrace{20\,(\max\,(M - M_c, 0))^4}_{\text{wave drag}},\ \sigma\ )\,. \tag{11}$$

which utilized the wave drag model described in Filippone [14]. As shown in Fig. 5, the comparisons of the estimated wave drag derived from the data and the model output in the black dots and blue line, respectively, concluded that the modeling of the wave drag component was not unambiguously successful.

In this study, the model is updated as

$$\mathfrak{M}_w : C_D \sim \text{Student\_t}\left(\nu, C_{Di}{C_L}^2 + C_{D0} + C_{Dw}\,(\max\,(M - M_c, 0))^4, \sigma\right)\,, \tag{12}$$

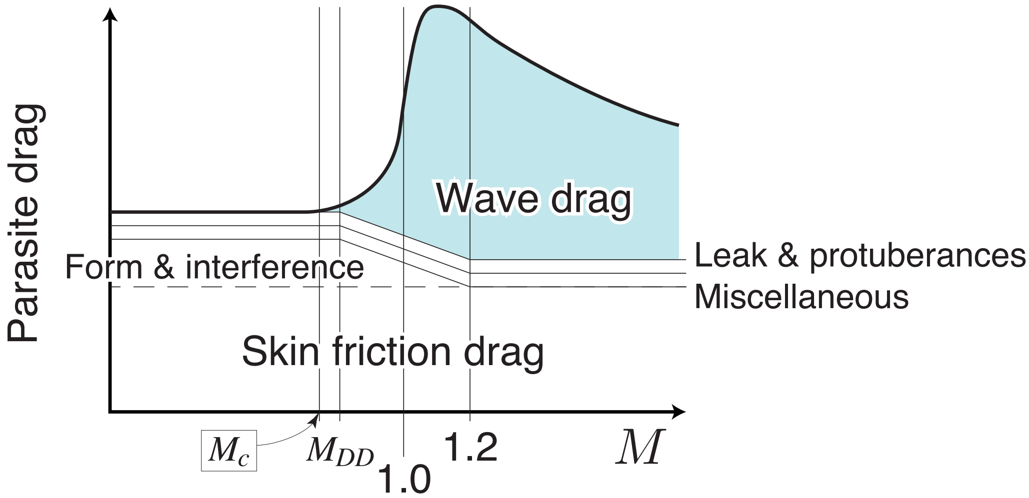

Figure 4 – Parasite drag

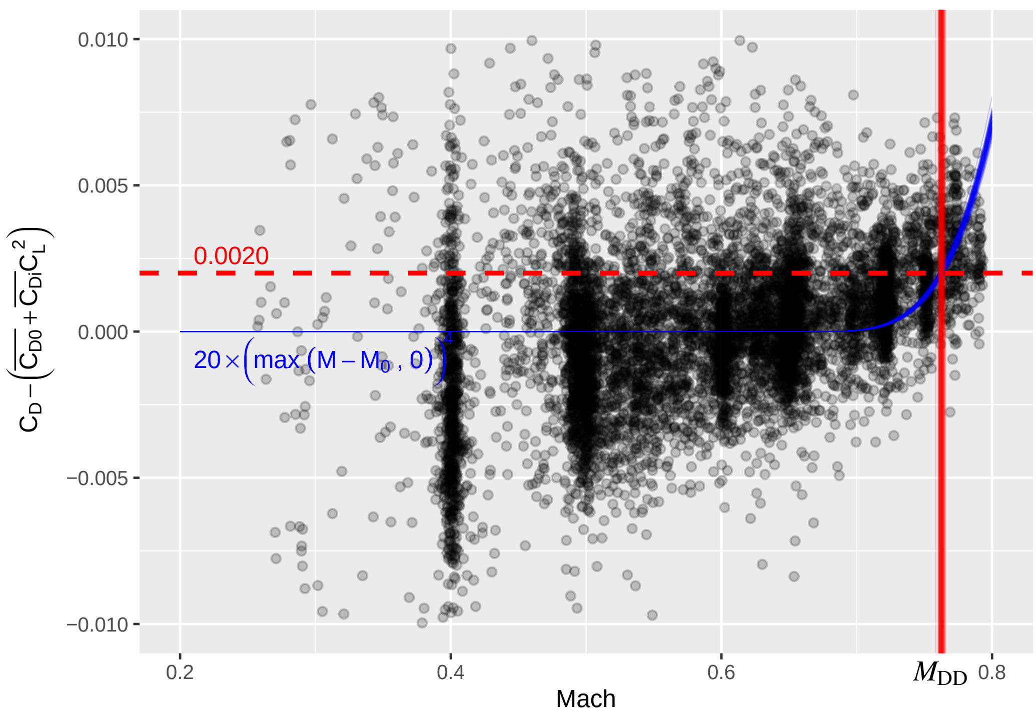

Figure 5 – Estimated wave drag

where the Student's t-distributed error model is used due to its robustness to outliers discussed in the previous section. In addition, the wave drag coefficient C_{Dw} is additionally introduced for flexibility instead of fixed 20 in Eq. (11). To compare with this model, the other model $\mathfrak{M}_0$ as

$$\mathfrak{M}_0 : C_D \sim \text{Student\_t}\left(\nu, C_{Di}{C_L}^2 + C_{D0}, \sigma\right) \tag{13}$$

which does not include the wave drag component, is prepared. These models are implemented as the Stan code in Listing 4. The priors for C_{D0} and C_{Dw} are the Cauchy distributions.

Listing 4: Drag coefficient model in Stan format

```
data {
  int<lower=1> N;
  vector[N] cL;
  vector[N] mach;
  vector[N] cD;
}
parameters {
  real cD0;
  real<lower=1.0/(pi()*7.731)> cDi; // 1/(eπAR)
  real<lower=0> cDw;
  real<lower=0.55, upper=0.9> machC;
  real<lower=0> sigma_cD;
  real<lower=0> nu_cD;
}
model {
  vector[N] mach_delta;
  for(i in 1:N){
    mach_delta[i] = (mach[i] >= mach0) ? pow(mach[i] - machC, 4) : 0;
  }
  target += student_t_lpdf(cD |
      nu_cD,
      (cDi * (cL .* cL)) + cD0
        // + (cDw * mach_delta) // Activate when M_w, i.e., Eq. (12)
        ,
      sigma_cD);
  target += cauchy_lpdf(cD0 | 0, 1);
  target += cauchy_lpdf(cDw | 0, 1);
}
```

The Bayes factor of these two models is

$$\log BF\left(\frac{\mathfrak{M}_w}{\mathfrak{M}_0}\right) = -80.7\,, \tag{14}$$

which means the model $\mathfrak{M}_w$ is not supported. Note that the update of the error models does not significantly affect the results, because the two Bayes factors of the original $\mathfrak{M}_w$ and $\mathfrak{M}_0$ using the normally distributed error models against $\mathfrak{M}_w$ and $\mathfrak{M}_0$ do not exceed the threshold (± 5). Therefore, it is concluded that the modeling the wave drag component using this flight data is not appropriate from model evaluation perspective. This conclusion is consistent with the previous results.

4. Conclusion

In this study, the application of the Bayes factor to the evaluation for modeling the flight characteristics by using the flight data was proposed. This proposal was the complementary extension of the previous study to use Bayesian inference for the parameter estimation of an already postulated model. The objective of the series of studies was to obtain an accurate model by using flight data. However, the flight data were collected under conditions that were unintentionally variable compared to ground tests such as wind tunnel tests. Consequently, one of the key techniques was to assess the compatibility between the postulated model and the flight data. This was achieved by calculating the Bayes factor. Furthermore, it was also demonstrated that the use of the Bayes factor had the advantage of directly utilizing the samples generated by the MCMC sampler, which was employed in the previous study for the parameter estimation.

In this paper, the derivation of the Bayes factor was first briefly described, and the tools for its calculation were introduced. Then, its application was demonstrated using the numerical examples. By comparing the Bayes factors, it was quantitatively reconfirmed that the robustness against outliers was improved by constructing the error model with the Student's t-distribution, which was proposed in the previous studies. Then, the two modeling issues of the flight characteristics of JAXA's research aircraft "Hisho" were discussed. The first target was the pitch rate effect for the modeling of the lift coefficient. The Bayes factor of the model with the effect against the model without the effect was significantly large, and it was concluded that the effect should be modeled. The second application was the wave drag decomposition for the drag modeling. The calculated Bayes factor did not support the model with the explicit wave drag term. The results of these applications were consistent with the previous results, and the quantitative decisions were available for the first time in this study.

Acknowledgment

This work was supported by JSPS KAKENHI Grant JP21381031.

Contact Author Email Address

naruoka.masaru@jaxa.jp

Copyright Statement

References

[1] Ravindra V. Jategaonkar. *Flight Vehicle System Identification: A Time Domain Methodology (Progress in Astronautics and Aeronautics)*. AIAA, 2006.

[2] Morelli Eugene. A. Real-time parameter estimation in the frequency domain. Technical report, 1999. NASA-aiaa-99-4043.

[3] M. Naruoka. Bayesian approach to simultaneous static and dynamic flight characteristics modeling. In *ICAS2021, 32th International Congress of the Aeronautical Sciences, Shanghai, China*, Sep 2021.

[4] M. Naruoka and T. Ninomiya. Statistical data selection for better flight characteristic modeling. In *ICAS2022, 33th International Congress of the Aeronautical Sciences, Stockholm, Sweden*, Sep 2022.

[5] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[6] Quentin F. Gronau, Alexandra Sarafoglou, Dora Matzke, Alexander Ly, Udo Boehm, Maarten Marsman, David S. Leslie, Jonathan J. Forster, Eric-Jan Wagenmakers, and Helen Steingroever. A tutorial on bridge sampling. *Journal of Mathematical Psychology*, Vol. 81, pp. 80–97, December 2017.

[7] Antony M. Overstall and Jonathan J. Forster. Default bayesian model determination methods for generalised linear mixed models. *Computational Statistics & Data Analysis*, Vol. 54, No. 12, pp. 3269–3288, December 2010.

[8] Quentin F. Gronau, Henrik Singmann, and Eric-Jan Wagenmakers. bridgesampling: An R package for estimating normalizing constants. *Journal of Statistical Software*, Vol. 92, No. 10, pp. 1–29, 2020.

[9] Stan Development Team. RStan: the R interface to Stan, 2023. R package version 2.32.3.

[10] Stan Development Team. Stan modeling language users guide and reference manual, version 2.34.

[11] Jeffrey N. Rouder and Richard D. Morey. Default bayes factors for model selection in regression. *Multivariate Behavioral Research*, Vol. 47, No. 6, p. 877?903, November 2012.

[12] Robert E. Kass and Adrian E. Raftery. Bayes factors. *Journal of the American Statistical Association*, Vol. 90, No. 430, pp. 773–795, June 1995.

[13] D.P. Raymer. *Aircraft Design: A Conceptual Approach*. AIAA education series. American Institute of Aeronautics and Astronautics, 2012.

[14] Antonio Filippone. Comprehensive analysis of transport aircraft flight performance. *Progress in Aerospace Sciences*, Vol. 44, No. 3, pp. 192–236, Apr 2008.