

Teaching a Verification and Validation Course Using Simulations and Experiments With Paper Helicopters

Chanyoung Park

Department of Mechanical and Aerospace Engineering, University of Florida, PO Box 116250, Gainesville, FL 32611-6250
e-mail: cy.park@ufl.edu

Joo-Ho Choi¹

Professor
School of Aerospace and Mechanical Engineering, Korea Aerospace University, Goyang-City 412-791, Gyeonggi-Do, Korea
e-mail: jhchoi@kau.ac.kr

Raphael T. Haftka

Distinguished Professor
Department of Mechanical and Aerospace Engineering, University of Florida, PO Box 116250, Gainesville, FL 32611-6250
e-mail: haftka@ufl.edu

In teaching a course on verification and validation (V&V) in scientific simulations, it is desirable for students to carry out repeated experiments on devices/systems that they can build and that can be easily repeated. This allows them to be exposed to the inherent aleatory uncertainty associated with building and testing when experiments are used to validate the scientific simulation tools. This paper reports on our experience in using paper helicopters for this purpose in a V&V graduate course. Paper helicopters have been used for various courses, from statistics in high school to graduate optimization courses. They are easily made from paper and paper clips and share the feature of autorotation with real helicopters when they are dropped from altitude. For the V&V course, the helicopters permitted comparison of two models of the drag produced by autorotation that slows their fall. A quadratic dependence of the drag on the speed is generally valid for high Reynolds numbers and a linear model appears for low Reynolds numbers. A gratifying result was that some of the helicopters fitted well the linear model and some fitted better the quadratic model, reflecting the fact that the Reynolds number is in an intermediate range. The paper provides details of how the experiments were conducted and analyzed, which would allow them to be used in similar courses. In addition, actual data are provided, which may be useful for teachers who need to cover the subject in a short time that would not allow the physical experiments. The project also allows a verification component of comparing an analytical solution to one obtained by numerical integration.

[DOI: 10.1115/1.4033889]

Keywords: validation, education, experiments, paper helicopter, uncertainty quantification

1 Introduction

Courses on V&V are common for software engineering [1]. Teaching V&V as applied to physical systems has lagged somewhat, but there is a growing interest in developing methodology and courses to teach students and professionals the basics of V&V. There is an interest in courses specific to particular set of physical systems. For example, Sellers et al. [2] described a book chapter intended for a course on V&V of space systems.

In between these two extremes lies the V&V of software codes built to carry out the scientific simulations. Such simulations are now routinely used to predict the crashworthiness of a car as it crashes into a barrier with various scenarios, thus guiding its structural design [3–5]. Similarly, simulations are used to predict the complex flow over aircraft wings or through their engines, guiding the designs of engines and the shape of aircraft wings [6,7]. Consequently, there has been a vigorous activity in the framework development for V&V [8,9] and a rising interest in teaching the subject, along with an excellent textbook [10].

When it comes to scientific simulations, the terminology verification refers to comparing the simulations to known solutions, while validation refers to comparison with physical experiments. Simulation code developers often have the requisite education and skills in numerical analysis to carry out or understand the verification process. However, for validation, they have to collaborate with experimentalists, and their communication is more difficult due to lack of understanding of the experiments.

Therefore, in developing a V&V course for scientific simulations, we felt that it is important that the students will carry out both numerical simulations and experiments. Furthermore, we considered it desirable that the test articles could be manufactured and tested easily enough so that students could repeat the experiments and manufacture and test nominally identical test articles. For that we turned to an old staple, the paper helicopter, which has been used for decades to teach high school and even middle school students' statistics and process control [11]. The paper

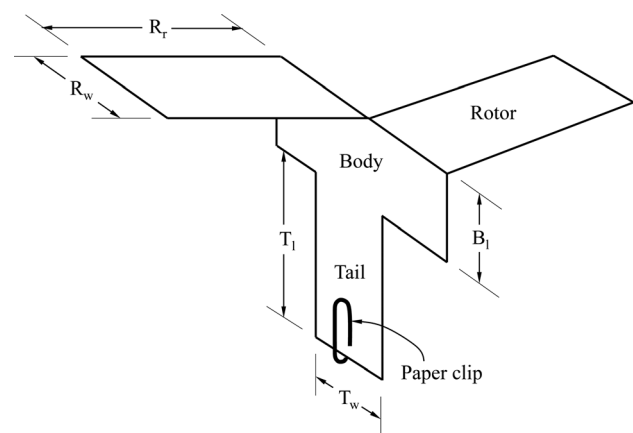


Fig. 1 Paper helicopter dimensions in meters: $R_r = 0.1143$, $R_w = 0.0508$, $B_1 = 0.0381$, $T_1 = 0.0508$, $T_w = 0.0254$ (dimensions in inches: $R_r = 4.5$, $R_w = 2$, $B_1 = 1.5$, $T_1 = 2$, $T_w = 1$)

¹Corresponding Author.

Manuscript received November 25, 2014; final manuscript received May 28, 2016; published online July 15, 2016. Editor: Ashley F. Emery.

helicopter (Fig. 1) can be easily built and tested. Like a real helicopter, its rotor starts rotating as it falls and generates a drag that slows down the fall.

One of us has used the helicopter for teaching experimental optimization with the commonly used model for the drag force that assumes quadratic dependence on the velocity [12]. For the V&V course, we asked the students to validate this quadratic model against a linear model that is known to prevail at low Reynolds numbers [14]. In fact, tests intended to compare the linear and quadratic models for falling and rising spheres [14,15], rotating disks [16], or even irregular shaped objects [17] in fluid have been used in undergraduate laboratory tests.

2 Paper Helicopter and Analytical Model

A paper helicopter is cut from one sheet of paper and folded into the form shown in Fig. 1 with one or two paper clips simulating the weight of the rest of the helicopter. Normally, regular copy-machine paper is used. The paper helicopters are composed of rotor, body, tail, and a paper clip as a flight stabilizer.

The “flight” of the paper helicopter is initiated by dropping it from a high point (typically 10–20 feet) and letting it fall to the ground. The air flow across the rotor causes the helicopter to start rotating, and this process of autorotation slows down the fall. With the assumption that the helicopter is moving down, the net downward force F is

$$F = mg - D \quad (1)$$

where m , g , and D are the mass, the acceleration of gravity, and the drag force, respectively. Using Newton’s second law, the acceleration, which is the first derivative of velocity V with respect to time t , is given as

$$\frac{dV}{dt} = g - \frac{D}{m} \quad (2)$$

The fall time of a helicopter is predicted with an analytical drag model with Eq. (2). Since the drag force increases as the helicopter accelerates, drag is modeled as a function of velocity. It is known that for high Reynolds numbers (i.e., $10^3 < \text{Re} < 2 \times 10^5$), the drag is proportional to the square of the velocity, while for low Reynolds numbers ($\text{Re} < 1$), it is proportional to the velocity (e.g., see Ref. [14]). Since a small helicopter at low speed may fall in an intermediate regime, students were requested to compare the two models.

2.1 Quadratic Model. The quadratic model defines drag as

$$D = \frac{1}{2} \rho_{\text{air}} V^2 A C_D \quad (3)$$

where ρ_{air} is the air density, V is the downward velocity of the paper helicopter, A is the disk area spanned by the rotor πR_r^2 , and C_D is the drag coefficient [12]. The purpose of the drag model is to define drag force in terms of downward velocity with two constants; the disk area and the drag coefficient to be calibrated. Note that the area of the two blades can be used as A instead of the disk area, but the use of different measure of area makes no difference to fall time predictions, but only leads to a different value of calibrated C_D . When the drag and helicopter weight are equal, there is no acceleration at a steady-state velocity, which is calculated using Eqs. (2) and (3)

$$V_s = \sqrt{g/c} \text{ and } c = \frac{\rho_{\text{air}} A C_D}{2m} \quad (4)$$

The velocity at t is obtained by solving the differential equation of motion, Eq. (2). The flight distance $d(t)$ at time t is obtained by integrating the velocity over $[0, t]$ analytically. The velocity is obtained as

$$V(t) = V_s \frac{(1 - e^{-2cV_s t})}{(1 + e^{-2cV_s t})} \quad (5)$$

where e^x is the exponential function.

The corresponding falling distance at t is

$$d(t) = V_s t + \frac{\log(e^{-2V_s c t} + 1) - \log 2}{c} \quad (6)$$

Instead of using Eq. (6), it is also possible to obtain the distance at t by numerically integrating the velocity over time. This is more typical of complex problems where it is impossible to integrate the differential equations analytically. Therefore, the students were assigned to solve by Euler’s forward integration and verify numerical solutions of the trajectory of their helicopter against the analytical solutions. Next time when we teach the course, we are likely to ask them to go further and verify that the convergence of the Euler integration with respect to the time-step is linear. Appendix A provides a detailed derivation of Eq. (6).

2.2 Linear Model. The linear model defines drag as

$$D = \frac{1}{2} \rho_{\text{air}} C_D A V_0 V \quad (7)$$

where V_0 is an arbitrary but a typical velocity value that affects the magnitude of the drag coefficient. $V_0 = 3$ (ft/s) was suggested in order to make the calibrated drag coefficient to have the same order of magnitude with that of the quadratic model. Note that the magnitude of V_0 makes no change to fall time predictions but only changes the magnitude of the calibrated drag coefficient. A steady-state velocity is obtained as

$$V_s = g/c \text{ and } c = \frac{\rho_{\text{air}} A C_D V_0}{2m} \quad (8)$$

By solving the differential equation, the velocity is obtained as

$$V(t) = V_s (1 - e^{-ct}) \quad (9)$$

The corresponding falling distance at t is

$$d(t) = V_s \left(t + \frac{e^{-ct}}{c} \right) - \frac{V_s}{c} \quad (10)$$

The simple models of the drag coefficient and the equations of motion contain two major assumptions that may severely limit the accuracy of the model predictions. The first assumption is about the dependence of the drag on velocity: quadratic and linear models. The second assumption is that the unsteady rotation can be modeled by a constant drag coefficient that is independent of speed. In the integration of the equations of motion, we also neglect the time period required to start the autorotation, which for the helicopters built for the project appeared to be of the order of a fraction of a second. Because of these assumptions, there is a need for validation experiments.

3 Paper Helicopter Constructions and Experiments for Validation

The students were directed to conduct the experiments under three conditions for validating the assumptions of constant drag coefficient and linear or quadratic models. Data for the first condition are used to calibrate the parameters of the models, i.e., the mean and standard deviation of the drag coefficient. Data for the second condition are used in order to validate the two possible drag models by changing the number of paper clips and hence the mass. As can be seen from Eqs. (4) and (8), for the quadratic model, the steady-state velocity is proportional to the square root

of the mass, while for the linear model, it is proportional to the mass. Data for the third condition are needed to validate the constant drag coefficient assumption by assessing the predictive capability of the model for height change. In the practice, data for calibration are collected and predictions are made with a calibrated model. The other two sets of data are collected to validate the predictions.

There are deficiencies and uncertainties in the experiments. First, there are construction uncertainties, in that each constructed helicopter is different from other nominally identical helicopters. Second, there are testing uncertainties due to differences in the way the helicopters are dropped, small variations in the height from which they were dropped, and inaccuracies in measuring the fall time. To assess the construction uncertainty, students built three nominally identical helicopters using regular size copy-machine paper of a specific maker and standard small paper clips. To quantify the uncertainties in testing, they repeatedly dropped ten times. We assume that the data for different conditions are independent from each other. However, students collected data for those three conditions through experiments at the same time. So some degree of correlation may have been present.

The first task is to construct three identical helicopters. The students were instructed to individually construct their own helicopters. The students used the same approaches for reducing uncertainty in the construction. In order to eliminate as much uncertainty as possible in the measurement uncertainty in the construction, dimensions of the helicopter were drawn using electrical drawing tools and were printed. Students then cut out the helicopters and folded the rotor blades. One student constructed five helicopters and randomly chose three of them.

For the simulation, the students needed to estimate the mass of the helicopter. A single helicopter is so light that weighing it would require a very accurate scale. Therefore, two options were suggested to students. One option was to measure the mass of several helicopters and use the average mass. The second option was to calculate the mass of helicopter based on the area of paper rotor and body by measuring the average paper mass per unit area and the average mass of a clip. Students obtained the average masses by weighing multiple sheets of paper and a group of clips. Some students measured the mass of individual helicopter using a very accurate scale. The reported mass of students' helicopters had variability because students used clips of various brands. However, little variability was observed for the same brand clips. Please refer to Appendix B for details about measuring weight and the variability of helicopter weights.

In contrast to construction, we allowed student collaboration for collecting fall time data, since helicopter experiments are difficult to do alone. Most student groups performed experiments in a building with an internal wide space, such as an atrium. One student stood at an upper level and dropped the helicopters for every test from both heights and with two and one paper clips. The other student was the timer. Students were asked to videotape at least one of their experiments. Two students used a stand and an electrical releaser for their experiments. However, no significant reduction of uncertainty in fall time was observed with those tools.

Students collected fall time data of three helicopters for those three different conditions. The drop height was required to be greater than 10 ft to reduce the effect of uncertainty in fall time due to the transition to autorotation after release. Students measured the drop height repeatedly, but little variability in the measurements was observed.

The need for managing outliers in the collected data provides students with a more vivid experience as engineers. Unstable flights due to long transition time to autorotation or wind gusts were occasionally a problem. Students who did not remove the outliers for such tests obtained poor data. For example, in one submitted data set, the average fall time for the height of 148.5 in. with two clips was slightly larger than that from the height of 181 in. with one clip. Obviously, this is possible only with some unstable flights from 181 in.

4 Bayesian Calibration and Uncertainty Quantification

In the project, the distribution of the drag coefficient C_D reflecting the test variability is assumed to follow a normal distribution with the mean μ_{C_D} and standard deviation σ_{C_D}

$$C_D \sim N(\mu_{C_D}, \sigma_{C_D}) \quad (11)$$

The assumption of normal distribution reflects the extended coverage of Bayesian calibration for the distribution parameters and previous experience with the paper helicopter when it was used in a course on experimental optimization [12]. The first task is Bayesian calibration of the mean and standard deviation of the drag coefficient distribution. Here, calibration is the process of improving the agreement of a code calculation and reduces the influence of the normal distribution assumption by finding those parameters that minimize the error between the prediction from the model and observed fall time from the experiments [18]. Data for the first condition are used for this.

Bayesian calibration is applied, since it is capable of estimating uncertainty in the calibrated parameters and performs well with a small number of samples. Fall time data were converted to the corresponding drag coefficients using Eqs. (4) and (6) or (8) and (10), which is denoted as $C_{D,\text{test}}^i$, where the superscript i indicates the i th sample. Then, the likelihood of observing a drag coefficient given the parameters is expressed as a conditional probability density function (PDF)

$$p(C_{D,\text{test}}^i | \mu_{C_D}, \sigma_{C_D}) \quad (12)$$

Bayesian calibration technique allows to incorporate the knowledge of parameters to reduce the uncertainty in the calibrated parameters using a prior distribution. However, we used noninformative priors for the mean and standard deviation, which are constant and reciprocal of the variance, respectively [19].

The posterior distribution of the two parameters is proportional to the multiplication of the likelihoods and the priors and is expressed as

$$p(\mu_{C_D}, \sigma_{C_D} | C_{D,\text{test}}^1, \dots, C_{D,\text{test}}^N) \propto \prod_{i=1}^N p(C_{D,\text{test}}^i | \mu_{C_D}, \sigma_{C_D}) \frac{1}{\sigma_{C_D}^2} \quad (13)$$

The posterior distribution represents our belief on the unknown mean and standard deviation of the drag coefficient C_D . The uncertainties due to the finite data are incorporated into the distribution. From the posterior distribution, samples of the mean and standard deviation can be drawn using sampling methods and the generated samples are possible candidates for the true mean and standard deviation under the observed data. In this study, Markov Chain Monte Carlo (MCMC) technique is employed to draw samples from the posterior distribution [20].

5 Predictive Validation of the Model

Validations of the two assumptions, one for the quadratic versus linear drag model and the other for the constant-drag-coefficient, were carried out by comparing the predicted fall time distributions and the data from experiments under the other two conditions, which are the mass change and height change, respectively. The experiments were conducted in the same way as those for the calibration, that is, fall times were measured ten times under the validation conditions for each of the same three helicopters. The predicted fall time distributions are defined by continuous cumulative distribution functions (CDFs) and the actual data are defined as an empirical CDF. To compare those two, the project suggested the students the use of the validation area metric, which is the area between the predicted CDF from the model and the empirical CDF from the experiment [10].

In view of the uncertainty, it is noted that the prediction model includes the epistemic uncertainties in calibrated C_D due to the finite data [21]. Two approaches of modeling the epistemic uncertainties were suggested to the students, with one the predictive distribution and the other the concept of p-box. In the former approach, the effect of epistemic uncertainty is included in the form of marginal distribution. The steps in the predictive approach are outlined as follows:

- (1) Generate N pairs of the mean and the standard deviation for the normally distributed C_D from Eq. (13) using the MCMC method as a result of the calibration against the experimental data.
- (2) Randomly generate a single sample of C_D and the corresponding fall time using Eq. (11) from each pair of the parameters. From the N pairs of the parameters, N fall time samples are collected, from which a CDF, commonly referred as a predictive distribution, is obtained. The value of N should be large enough so that a fairly smooth CDF is obtained.
- (3) Calculate the area between the empirical CDF from the experiments and the predictive distribution, which we call as direct area metric, represents a measure of mismatch between the model and the experiment.

Note that the approach is implemented for the data from each helicopter instead of all three due to the reason addressed subsequently. Hence, three results are obtained from three helicopters independently. The weak point of this approach is that the epistemic uncertainty (due to finite sampling) is not fully accounted for and can create difference between the simulation and experiment even if the simulation was a perfect model of reality. For example, if we calculated an area metric between the empirical CDF from ten data and the true CDF based on those data from a perfect model, the calculated metric can still be a nonzero value.

In order to cover fully the epistemic uncertainty, p-box approach is widely used, which represents the effect of epistemic uncertainty by the confidence bounds of CDFs and calculate area metric with these bounds [10]. Steps 2 and 3 are changed as follows:

- (2) Construct a CDF of C_D from each pair of the mean and standard deviation and compute the corresponding CDF of fall time using Eq. (11). From the N pairs of the parameters, N fall time CDFs are obtained. Calculate the lower and upper percentiles of fall time for different probability values, which constitute the bound curves of a p-box. The percentiles of 2.5 and 97.5 that correspond to the 95% confidence level were suggested to students.
- (3) Calculate the area between the empirical CDF from the experiments and the upper and lower bounds of the p-box.

Figure 2 illustrates the difference of the area metric for the two approaches with student A's data for the first condition of helicopter 1. The data are presented in Table 6 in Appendix B. The direct area metric captures the area due to the finite number of samples and the error in the predicted fall time distribution, while the p-box area metric captures the area beyond the acceptable margin of the epistemic uncertainty with the width of the distribution band.

6 Observations From Data Sets Submitted From Students

In this section, data sets of three students are presented and analyzed. Each data set was collected from the experiments under the three conditions made for the calibration, validation of the drag models, and validation of the constant drag assumption.

The first student's data set was judged to be of high quality, collected from well-controlled experiments. The coefficients of variation (COV) for different helicopters and conditions are smaller than those in the other data sets. Also the data set has few extreme data, commonly referred to as outliers. The student will be

referred to as student A. Ten fall time data for each helicopter were obtained from the height of 149 in. with two clips, which is condition 1. H1, H2, and H3 represent helicopter 1, helicopter 2, and helicopter 3, respectively. Detailed data of all three conditions are presented in Table 6 of Appendix B. Table 1 summarizes the statistics of fall time data for the three conditions. The students measured fall time data with stopwatches with a precision of 0.01 s. Furthermore, the values are an average fall time of ten fall time data with less than 10% (mostly less than 5%), which is likely to improve the accuracy of the centisecond digit.

Figure 3 shows a comparison of the updated joint posterior distributions of the two parameters, namely the mean and the standard deviation, using the MCMC sampling technique with $N = 100,000$ samples. To observe the effect of the number of data, we calculated one posterior distribution based on the first two fall times and another based on all the ten data of helicopter 1. In this case, a quadratic drag model is employed. The distribution with ten data is much narrower than the distribution with only two data, illustrating the effect of repeated experiments on reducing uncertainty. As such, the benefit of employing Bayesian statistics is well recognized, in that the uncertainties of the estimated parameters are quantified in the form of distributions.

Table 2 shows summary statistics of the finally updated posterior distributions. From the table, it is seen that the COV of testing variability in drag coefficient between drops is between 5 and 8% (i.e., COV of helicopter 1 is $0.062/0.778 = 7.9\%$), while the COV of manufacturing variability is about 12% (COVs of mean of the mean). Repeated tests significantly reduce the uncertainties in the estimated mean and STD of drag coefficient, for instance, the STD of the mean, which quantifies the uncertainty in the average of tests, is smaller than the mean of STD, which represents the uncertainty in a single test, by one third for all helicopters. Because the manufacturing variability is large, the helicopter performances may be substantially different; even though they are manufactured based on the same helicopter dimensions, their results are treated as different helicopters in the validation experiments.

The validation metrics corresponding to the data of each helicopter are calculated with the constructed predictive distributions and p-boxes. As stated above, 2.5th and 97.5th percentiles are used for the 95% confidence interval for the p-box. Note that the magnitude of the area metric has no physical meaning [10]. In the calibration, the area metric represents just a degree of disagreement due to the uncertainty from the finite number of data and model inadequacy. In the event of infinite data and perfect model, the area metric may converge to zero. In the validations, the area metric provides a relative measure of prediction accuracy as compared to that of the calibration. If the area value is found to be similar, the prediction at that condition is regarded as validated at the best of our effort.

Table 3 presents the area metrics of the two approaches, i.e., predictive and p-box for the three conditions and the three helicopters. As can be seen in the table, the first condition is the calibration condition, while the conditions 2 and 3 are for validating the quadratic/linear model and the constant drag coefficient assumption, respectively. The area metric obtained in the calibration is compared to others to determine the validity of the assumptions.

For the observed height changes and weight changes, the weight change led to larger change of area metrics. That means the prediction uncertainties due to weight change (from two clips to one clip) are larger than those due to height change (from 149 in. to 173 in.). The observation is somewhat expected in the sense that weight change leads to large changes in the steady-state velocity and the acceleration performance of the helicopter, while height change leads to no change in the flight characteristics. Therefore, there are likely more uncertainties in the prediction for height change than that for weight change. Note, in particular, that the metric increase is remarkable for helicopter 2 under the weight change, that is, the weight change of the helicopter leads to

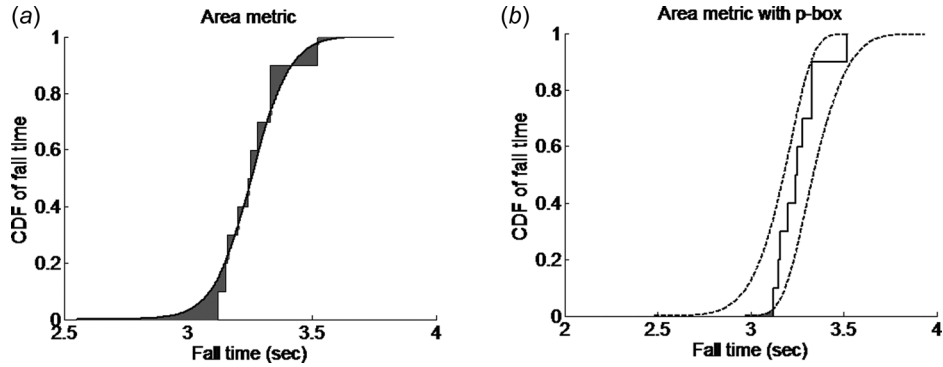


Fig. 2 Area metrics (shaded in gray) of the same data (helicopter 1 and condition 1 of student A in Appendix B) for different epistemic uncertainty models: a predictive distribution and a p-box of 2.5 and 97.5 percentiles (left and right dashed lines) using N of 10,000 samples: (a) validation area metric of a predictive distribution (0.032) and (b) validation area metric of a p-box (0.001)

significantly larger area metric increase than the other helicopters. The sample statistics (Table 1) show that helicopter 2 tends to have shorter average fall time than the others for the first and third conditions, whereas the case is opposite for the second condition, i.e., the fall time is longer than the others. From this observation, we may presume that the data of helicopter 2 with one clip may follow the linear model. This viewpoint agrees with our assumption that the Reynolds number of the helicopter design may lie in the intermediate region between those for the quadratic and linear models. The area metrics with the linear model as described in the following text support this viewpoint. Note that the way described in Sec. 5 was applied to calculate the prediction uncertainties and the corresponding area metrics.

Validity of the linear model is investigated with the same data set. The drag coefficient has to be recalibrated for the selected $V_0 = 3$ ft/s. The area metrics of condition 1 for calibration and 2 for weight change are calculated and compared to examine the validity of the linear mode. The results are given in Table 4. In contrast to the quadratic model, the metrics increase significantly for helicopters 1 and 3 but not so much for helicopter 2. The observation supports that the quadratic model works poorly with helicopter 2 for this condition, and we may conclude that helicopter 2 fits into the linear model much better than the quadratic model.

Note that a simpler method to validate the drag model of a helicopter is available and could be of use in courses where V & V have to be covered in less time. This is possible because the helicopters were found to reach the steady-state speed after only 0.2–0.3 s. Therefore, it is possible to approximate the dependence of the fall time on the mass and the height by assuming that the

helicopter reaches the steady-state speed instantaneously. In case of the quadratic model, fall time is then approximately calculated using Eq. (4) to obtain

$$t \approx h/V_s = h \sqrt{\frac{\rho_{\text{air}} AC_D}{2gm}} \quad (14)$$

The expected fall time of a helicopter is, therefore, proportional to the height and inversely proportional to the square root of the mass. Therefore, the ratio of the expected fall times with two heights h_1 and h_2 , $E(t_{h_1})/E(t_{h_2})$, and with two masses m_1 and m_2 , $E(t_{m_1})/E(t_{m_2})$, are given by

$$\frac{E(t_{h_1})}{E(t_{h_2})} \approx \frac{h_1 E(t_{m_1})}{h_2 E(t_{m_2})} \approx \sqrt{\frac{m_2}{m_1}} \quad (15)$$

where the subscript of t denotes the given condition. For example, t_{h_2} is the fall time for height h_2 and t_{m_2} is the fall time for mass m_2 .

For the linear model, the same assumption of instantaneous transition to steady-state speed provides the approximated fall time using Eq. (8) as

$$t \approx h/V_s = h \frac{\rho_{\text{air}} AC_D V_0}{2gm} \quad (16)$$

So for different heights, the expected time ratio is still the same as the height ratio as Eq. (15), but for different masses we now get

Table 1 Fall time data statistics of students A, B, and C

	H1	H2	H3
Conditions for student A			
Average fall time (s) and COV (%)			
1 Weight: 2 clips; height: 149 in. (for calibration)	3.26 (3.6%)	2.91 (2.3%)	3.15 (3.1%)
2 Weight reduced to 1 clip (drag model validation)	3.70 (1.9%)	3.84 (3.8%)	3.71 (1.6%)
3 Height increased to 173 in. (constant drag validation)	3.75 (2.8%)	3.42 (2.5%)	3.70 (1.6%)
Ratio of fall times for the height change	1.15	1.18	1.17
Conditions for student B			
1 Weight: 2 clips; height: 148.5 in. (for calibration)	3.78 (9.8%)	3.67 (9.0%)	3.56 (5.1%)
2 Weight reduced to 1 clip (drag model validation)	4.16 (8.3%)	4.34 (3.4%)	4.29 (2.6%)
3 Height increased to 181 in. (constant drag validation)	4.52 (4.5%)	4.46 (6.9%)	4.34 (4.0%)
Ratio of fall times for the height change	1.20	1.22	1.22
Conditions for student C			
1 Weight: 2 clips; height: 124 in. (for calibration)	2.84 (4.9%)	3.12 (4.7%)	2.82 (4.7%)
2 Weight reduced to 1 clip (drag model validation)	2.13 (3.5%)	2.39 (3.6%)	2.65 (9.6%)
3 Height increased to 139 in. (constant drag validation)	3.46 (1.9%)	3.22 (3.1%)	3.35 (1.8%)
Ratio of fall times between two heights	1.218	1.032	1.188

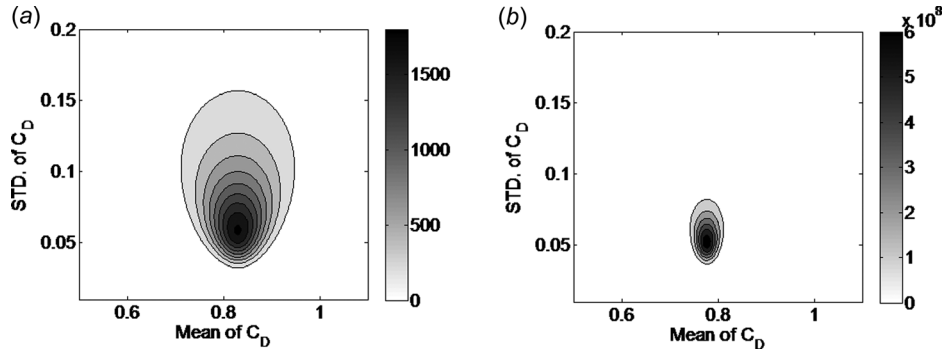


Fig. 3 Process of updating posterior distributions of the mean and standard deviation of C_D of helicopter 1, for student A data (see Table 1): (a) after 2nd update and (b) after 10th update

Table 2 Summary statistics of the posterior distributions of C_D for helicopters of students A and B

Student A	H1	H2	H3
Mean of the mean	0.778	0.607	0.721
STD of the mean	0.021	0.010	0.016
Mean of the STD	0.062	0.031	0.048
STD of the STD	0.017	0.008	0.013
Student B			
Mean of the mean	0.896	0.842	0.783
STD of the mean	0.069	0.060	0.034
Mean of the STD	0.190	0.154	0.084
STD of the STD	0.059	0.049	0.028

$$\frac{E(t_{m_1})}{E(t_{m_2})} \approx \frac{m_2}{m_1} \quad (17)$$

By comparing the ratio of expected fall times with the ratio of measured average fall times, we can determine the valid drag model for each helicopter. Upper part of Table 5 presents mass ratios, calculated expected fall time ratios by the quadratic and the linear models, and ratio of measured fall time ratios that are

calculated from the average fall times in Table 1. Even with the constant speed assumption and the uncertainty in average fall times due to the finite number of data, this approach draws the same conclusion as that from the area metric that the comparison of ratios can be a simple alternative validation approach for this problem. For helicopter 2, the average fall time ratio is very close to the expected fall time ratio for the linear model, which implies that the linear drag model is a good fit for helicopter 2. In the same manner, the quadratic models are good for helicopter 1 and 3. The same results have been found for the data of other students. Also, as seen in the lower part in Table 5, for all three helicopters, the assumption of constant drag coefficients is supported by the observation that the ratio of measured fall times is very close to the ratio of the height for constant drag validation to the original height, which is $173/149 = 1.161$.

The next results are from a different student referred to as student B. The complete data set is given in Table 7 of Appendix B. The data have larger variability in fall time data than all the other data sets. Table 1 shows the statistics of the fall time data. We can see that the COVs of student B's data set are larger than those of student A. Particularly, the COVs of the data for calibration are two to three times larger than student A's. That implies that student B calibrated the drag coefficient parameters with poor data that includes large uncertainty. Despite the substantial variation in

Table 3 Validation area metrics of three helicopters of students A, B, and C for three conditions (quadratic model). The one-clip results support the validity of the quadratic model.

	H1	H2	H3
Conditions for student A			
Validation area metric with predictive distribution			
1 Weight: 2 clips; height: 149 in. (for calibration)	0.032	0.017	0.028
2 Weight reduced to 1 clip (drag model validation)	0.091	0.475	0.091
3 Height increased to 173 in. (constant drag validation)	0.043	0.060	0.078
Validation area metric with p-box			
1 Weight: 2 clips; height: 149 in. (for calibration)	0.001	0.000	0.000
2 Weight reduced to 1 clip (drag model validation)	0.017	0.411	0.026
3 Height increased to 173 in. (constant drag validation)	0.003	0.013	0.016
Conditions for student B			
Validation area metric with predictive distribution			
1 Weight: 2 clips; height: 148.5 in. (for calibration)	0.127	0.088	0.060
2 Weight reduced to 1 clip (drag model validation)	0.146	0.321	0.372
3 Height increased to 181 in. (constant drag validation)	0.286	0.151	0.068
Validation area metric with p-box			
1 Weight: 2 clips; height: 148.5 in. (for calibration)	0.004	0.001	0.001
2 Weight reduced to 1 clip (drag model validation)	0.012	0.106	0.227
3 Height increased to 181 in. (constant drag validation)	0.044	0.004	0.005
Conditions for student C			
Validation area metric with predictive distribution			
1 Weight: 2 clips; height: 124 in. (for calibration)	0.043	0.047	0.044
2 Weight reduced to 1 clip (drag model validation)	0.438	0.418	0.163
Validation area metric with p-box			
1 Weight: 2 clips; height: 124 in. (for calibration)	0.001	0.000	0.001
2 Weight reduced to 1 clip (drag model validation)	0.344	0.317	0.083

Table 4 Validation area metrics of three helicopters of students A, B, and C for validating the linear model

	H1	H2	H3
Conditions for student A			
Validation area metric with predictive distribution			
1 Weight: 2 clips; height: 149 in. (for calibration)	0.032	0.018	0.028
2 Weight reduced to 1 clip (drag model validation)	0.642	0.040	0.464
Validation area metric with p-box			
1 Weight: 2 clips; height: 149 in. (for calibration)	0.001	0.000	0.000
2 Weight reduced to 1 clip (drag model validation)	0.513	0.008	0.360
Conditions for student B			
Validation area metric with predictive distribution			
1 Weight: 2 clips; height: 148.5 in. (for calibration)	0.108	0.094	0.056
2 Weight reduced to 1 clip (drag model validation)	0.453	0.276	0.114
Validation area metric with p-box			
1 Weight: 2 clips; height: 148.5 in. (for calibration)	0.002	0.001	0.001
2 Weight reduced to 1 clip (drag model validation)	0.116	0.048	0.013
Conditions student C			
Validation area metric with predictive distribution			
1 Weight: 2 clips; height: 124 in. (for calibration)	0.046	0.044	0.044
2 Weight reduced to 1 clip (drag model validation)	0.204	0.171	0.348
Validation area metric with p-box			
1 Weight: 2 clips; height: 124 in. (for calibration)	0.001	0.001	0.01
2 Weight reduced to 1 clip (drag model validation)	0.124	0.085	0.267

fall times, however, the ratios of average fall times between the two heights are surprisingly close to the ratio of the height for constant drag validation to the original height ($181/148.5 = 1.219$).

The large uncertainty in data is transferred into the large epistemic uncertainty in calibration. Table 2 shows the statistical characteristics of posterior distributions of three helicopters of student B. We can observe that the magnitudes of STD of the mean are larger than those of student A by two- to sixfolds.

Large epistemic uncertainty consequently leads to large area metrics. Table 3 shows area metrics with predictive distribution of the condition 1 are considerably larger than those of the student A's data. For area metrics with the p-box, however, the area metrics are not so different from those of student A. That is because the effect of the large epistemic uncertainty is compensated by the correspondingly wide confidence interval of the p-box. This is one of good features of using the p-box, in that the magnitude of area metrics is comparable even for very different degree of the epistemic uncertainty.

For this data set, helicopter 1 shows a different area metric behavior from the other helicopters, in that the quadratic model predicts the fall time distribution for helicopter 1 with much smaller error than the other helicopters. Since we observed similar behavior in student A's data set, assuming that helicopter 1 follows the quadratic model while the others follow the linear model is not a leap of faith.

Table 4 shows area metrics with the linear model. The behavior of the area metrics confirms that the assumption is reasonable. The difference between area metrics of helicopter 1 is almost two times larger than those of helicopters 2 and 3. Thus, we may

conclude that the quadratic model is suggested for helicopter 1, while so is the linear model for helicopters 2 and 3.

Another student C's fall time data statistics are presented in Table 1 (the complete data set is given in Table 8). This student collected data under different conditions from the other students, in which one clip is used for the first and third conditions to examine height change, and weight is increased by adding one clip in the second condition. In the data, it was found that the uncertainty in the data of Helicopter 3 is much higher than the other helicopters in terms of COV for the second condition. In addition, the ratios of fall times between different heights are different by themselves and also different from the ratio of the height for constant drag validation to the original height ($139/124 = 1.121$). This result seems to indicate that all three helicopters do not seem to validate the assumption of constant drag coefficient, and this may reflect some problems in the tests.

In Tables 3 and 4, the area metrics of helicopter 3 confirms the same observation from the other two students, that the helicopters can have different drag models even if they are manufactured to be nominally identical by the same student. As seen in the comparison of first and second conditions, for Helicopter 3, the quadratic model predicts the fall time better with much smaller area metric, while so is the linear model for the other helicopters.

Note that checking the validity of experimental data is a very important step. One student drew a very strange conclusion because the student used a bad data set in which the average fall time of higher height is slightly less than that of lower height. Paper helicopters with the given design sometime become unstable during flight. For example, a helicopter may descend in a helical flight pass with asymmetric rotor rotation, which is a sign of

Table 5 Weight and height ratios, the corresponding average fall time ratios for drag models and actual average fall time ratios for student A

		H1	H2	H3
Mass change	Mass ratio (1 clip/2 clips)	0.737	0.735	0.740
	Calculated ratio of expected fall times by quadratic model	1.164	1.166	1.162
	Calculated ratio of expected fall times by linear model	1.356	1.359	1.351
	Ratio of measured average fall times	1.135	1.318	1.178
Height change	Height ratio	1.161	1.161	1.161
	Ratios of measured average fall times	1.150	1.175	1.175

unstable flight. The student did not exclude data with unstable flight and it led to draw the strange conclusion. For the next course, we selected a smaller design, which reduces the chance of being unstable, and required to videotape all helicopter flights to prevent such mistakes in data collection.

7 Instruction and Project Administration

Before distributing the 4-week project, lecturers introduced the students to the required knowledge for validating the paper helicopter. The way of estimating the mean and standard deviation of the drag coefficient from experiments using the Bayesian calibration approach was presented early in the course. MCMC technique was also taught and practiced to generate samples from posterior distributions. Later in the course, validation procedures were covered along with the area metric and the p-box.

A one-hour lecture was allocated for introducing students to the paper helicopter, its mathematical model based on Newton's second law, and the drag model. In order to introduce them to the concepts of verification, a numerical solution using Euler explicit integration was also covered. Homework was then assigned to students for analysis code implementation and code verification as a preparation of the project. Students reported that they spent significant amount of time for manufacturing the helicopters and executing the experiments. All the lectures were video-recorded to accommodate remote learning.

The class had 24 students. The students were encouraged to conduct the experiments and submitted fall time data of three identically built helicopters for three conditions. For each condition, ten fall time data were collected. A maximum of 20% bonus points were given to students who submitted data early, and collected data were shared with other students.

8 Remarks

It is important to teach a V&V course with a project including a physical experiment that allows students to have an opportunity to apply methods on a simplified real problem and to experience unexpected surprises from experiments.

The paper described teaching experience of a V&V course using simulations and experiments with paper helicopters. The benefits of including experiments with paper helicopters are that they can be easily built from paper and paper clips, and the experiments provide a real experience of validating simulations based on the observed data. The experiments are also easy to perform and monitor. The validation process includes executions of experiment and collecting data so that students can experience the perspectives of both analysts and experimentalists.

Students experienced difficulty in interpreting validation results from the experimental data with unexpected surprises. All the data sets collected by students contain fall time data of three identically built helicopters for three conditions. Very different drag characteristics were observed between helicopters even though they were nominally identical. The possibility of bad data was first suspected as the reason, but careful analysis of the data indicated that helicopters may not follow the same drag models even if they were identically constructed by the same person. This may reflect that the helicopters operate in an intermediate Reynolds number regime. While this phenomenon was not expected when the project was conceived, we consider it a happy finding for exposing students to this phenomenon.

Based on the experience gained in the course, we may be able to probe further the differences between the helicopter following the linear model and the ones following the quadratic model in the next course. In particular, students will be given that challenge and also be asked to videotape the two types so as to look for additional clues.

For evidence on the effectiveness of the project as an educational tool, we had only a comment from a single student in the course evaluation that "the last portion of the course really

brought everything together for me. I know that was part of the point of the project, but I wonder if that could not be accomplished early on by emphasizing more of the terminology in homework and tests earlier. The last test probably could be omitted. The project was just as effective in ensuring the topics were learned." We planned to ask every student to give a comment on the project next time in order to solicit ideas for improving it as well as for assessing its impact.

Acknowledgment

Part of this work was supported by the U.S. Department of Energy, National Nuclear Security Administration, Advanced Simulation and Computing Program, as a Cooperative Agreement under the Predictive Science Academic Alliance Program (Grant No. DE-NA0002378), and the grants from the International Collaborative R&D Programs funded by the Agency for Defense Development (ADD, Grant No. UD130062GD).

Nomenclature

A	= area of rotor rotation
C_D	= drag coefficient
$C_{D,\text{test}}$	= converted drag coefficient from a fall time datum
CDF	= cumulative density function
COV	= coefficient of variation
D	= drag
g	= acceleration of gravity
m	= helicopter mass
MCMC	= Markov Chain Monte Carlo
PDF	= probability density function
Re	= Reynolds number and similar abbreviations do not use italics
STD	= standard deviation
V_s	= steady-state speed
V_0	= velocity constant for linear model
μ_{C_D}	= mean of drag coefficient
ρ_{air}	= density of air
σ_{C_D}	= standard deviation of drag coefficient

Appendix A: Derivation of the Falling Distance Function

The gravity force accelerates a paper helicopter. The drag force of a helicopter is defined with Eq. (1) based on the quadratic drag force model; the force that accelerates the helicopter is the difference between the gravity force and the drag force

$$m \frac{dV}{dt} = mg - \frac{1}{2} \rho_{\text{air}} V^2 A C_D \quad (\text{A1})$$

Equation (A1) is rewritten in terms of a constant c as

$$\frac{dV}{dt} = g - cV^2 \text{ and } c = \frac{\rho_{\text{air}} A C_D}{2m} \quad (\text{A2})$$

Since the drag force increases as the velocity increases, the paper helicopter reaches the steady-state speed. When the drag force becomes equivalent to the gravity force, the acceleration becomes zero. According to the equation, the steady-state speed is a function of the dimension of the helicopter

$$V_s = \sqrt{g/c} \quad (\text{A3})$$

The acceleration can be expressed in terms of the steady-state speed and the velocity. To derive the time-dependent velocity of the helicopter, the differential equation is solved with the boundary condition of $V = 0$ for $t = 0$; the relation between the velocity and time is explicitly expressed as

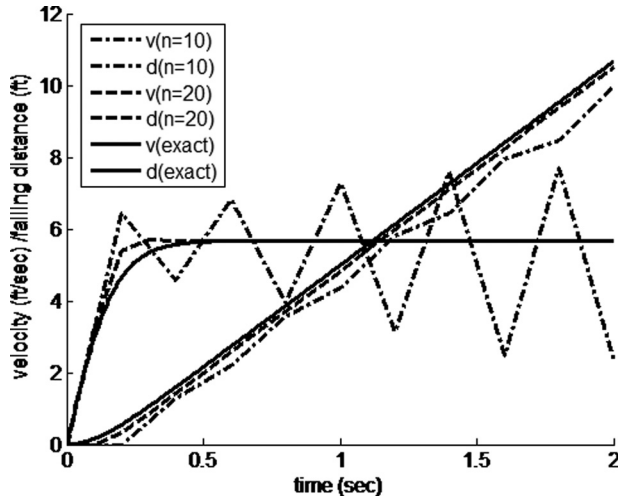
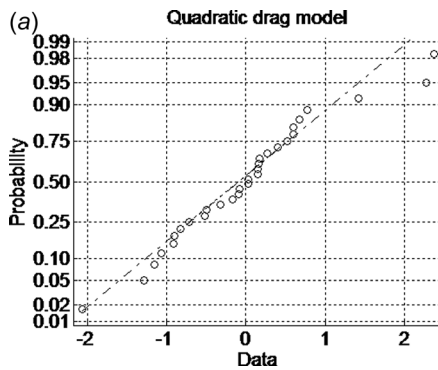


Fig. 4 Velocity (v) and falling distance (d) using numerical schemes with time-steps $n = 10$ and $n = 20$ and exact solutions

$$V = V_s \frac{(1 - e^{-2cV_s t})}{(1 + e^{-2cV_s t})} \quad (\text{A4})$$



The falling height is obtained through integrating the velocity V for the time range 0 to the fall time t . The falling distance $d(t)$ at time t is expressed as

$$d(t) = V_s t + \frac{\log(e^{-2cV_s t} + 1) - \log 2}{c} \quad (\text{A5})$$

We can calculate the fall time t_{fall} using Eq. (A5) using the condition that the falling distance at the fall time should be equal to the falling height h_{fall} as

$$t_{\text{fall}} = d^{-1}(h_{\text{fall}}) \quad (\text{A6})$$

Figure 4 presents velocity and falling distances at time t using the forward Euler method and the trapezoidal integration with 10 and 20 time-steps and the exact solutions using Eq. (A5). There is a little difference between the falling distance curves from the numerical solutions with 20 time-steps and the exact solution. The difference is caused by the error in the velocity calculation in the transition regime, in which the helicopter decelerates and eventually reaches to the steady-state velocity.

Appendix B: Fall Time Data of Three Students

Each student prepared three nominally identical helicopters and measured fall time ten times in repetition for each of them. Conditions 1, 2, and 3 are different in terms of falling height and weight

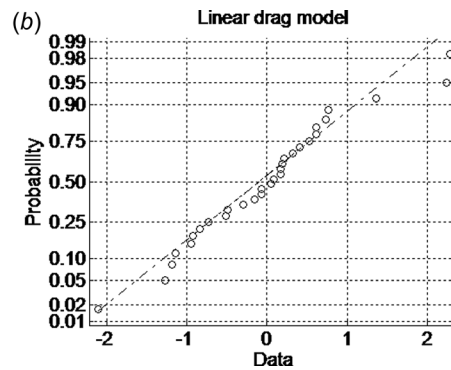


Fig. 5 Normal probability plots of pooled drag coefficient data for the quadratic and linear drag models

Table 6 Fall time data of student A

Condition 1			Condition 2			Condition 3					
Height: 149 in. and clips: 2			Height: 149 in. and clips: 1			Height: 173 in. and clips: 2					
	H1	H2	H3		H1	H2	H3		H1	H2	H3
1	3.52	2.82	3.13	1	3.68	3.98	3.81	1	3.59	3.47	3.65
2	3.2	2.86	2.94	2	3.64	3.87	3.72	2	3.76	3.22	3.75
3	3.12	2.84	3.24	3	3.82	4.04	3.69	3	3.8	3.45	3.67
4	3.24	2.93	3.09	4	3.68	3.95	3.76	4	3.6	3.53	3.76
5	3.33	2.85	3.19	5	3.6	3.82	3.69	5	3.66	3.39	3.7
6	3.16	2.95	3.25	6	3.64	3.62	3.68	6	3.84	3.47	3.66
7	3.25	2.98	3.13	7	3.74	3.58	3.62	7	3.81	3.42	3.76
8	3.28	2.95	3.27	8	3.74	3.8	3.79	8	3.89	3.45	3.68
9	3.15	2.92	3.15	9	3.78	3.87	3.66	9	3.74	3.43	3.61
10	3.33	3.02	3.12	10	3.67	3.87	3.71	10	3.85	3.36	3.8
Average	3.258	2.912	3.151	Average	3.699	3.840	3.713	Average	3.754	3.419	3.704
STD	0.117	0.066	0.096	STD	0.069	0.146	0.059	STD	0.105	0.084	0.061
COV	0.036	0.023	0.031	COV	0.019	0.038	0.016	COV	0.028	0.025	0.016

Table 7 Fall time data of student B

Condition 1				Condition 2				Condition 3			
Height: 148.5 in. and clips: 2				Height: 148.5 in. and clips: 1				Height: 181 in. and clips: 2			
	H1	H2	H3		H1	H2	H3		H1	H2	H3
1	4.63	3.46	3.86	1	4.48	4.59	4.34	1	4.2	4.39	4.4
2	3.43	3.54	3.47	2	3.53	4.18	4.03	2	4.39	4.28	4.14
3	4.05	3.97	3.86	3	4.23	4.28	4.38	3	4.5	4.87	4.12
4	3.81	3.97	3.61	4	4.4	4.37	4.34	4	4.5	5.02	4.34
5	3.36	3.03	3.42	5	4.34	4.15	4.2	5	4.49	4.51	4.37
6	3.67	4.16	3.58	6	3.64	4.3	4.34	6	4.9	4.36	4.43
7	3.9	3.58	3.33	7	3.93	4.3	4.29	7	4.54	4.37	4.28
8	3.86	3.51	3.54	8	4.51	4.21	4.28	8	4.54	3.91	4.12
9	3.59	3.89	3.39	9	4.15	4.55	4.42	9	4.34	4.4	4.56
10	3.51	3.63	3.49	10	4.36	4.43	4.23	10	4.78	4.5	4.61
Average	3.781	3.674	3.555	Average	4.157	4.336	4.285	Average	4.518	4.461	4.337
STD	0.371	0.329	0.182	STD	0.346	0.149	0.111	STD	0.202	0.307	0.175
COV	0.098	0.090	0.051	COV	0.083	0.034	0.026	COV	0.045	0.069	0.040

Table 8 Fall time data of student C

Condition 1				Condition 2				Condition 3			
Height: 124 in. and clips: 1				Height: 124 in. and clips: 2				Height: 139 in. and clips: 1			
	H1	H2	H3		H1	H2	H3		H1	H2	H3
1	2.917	3.185	3.017	1	2.076	2.229	2.733	1	3.364	3.047	3.348
2	2.900	3.169	2.933	2	2.193	2.381	1.995	2	3.415	3.097	3.331
3	3.034	3.135	2.750	3	2.109	2.431	2.833	3	3.364	3.281	3.298
4	2.666	3.437	2.750	4	2.243	2.381	2.632	4	3.482	3.164	3.364
5	2.951	3.136	2.682	5	2.143	2.414	2.800	5	3.381	3.197	3.431
6	2.850	2.984	2.716	6	2.126	2.397	2.599	6	3.549	3.364	3.364
7	2.867	3.001	2.817	7	1.975	2.532	2.817	7	3.498	3.247	3.281
8	2.833	3.152	3.051	8	2.142	2.313	2.498	8	3.465	3.331	3.415
9	2.867	3.102	2.817	9	2.076	2.314	2.833	9	3.415	3.230	3.264
10	2.548	2.901	2.698	10	2.193	2.464	2.733	10	3.515	3.230	3.431
Average	2.843	3.120	2.823	Average	2.127	2.386	2.647	Average	3.445	3.219	3.353
STD	0.140	0.145	0.133	STD	0.075	0.085	0.255	STD	0.066	0.098	0.061
COV	0.049	0.047	0.047	COV	0.035	0.036	0.096	COV	0.019	0.031	0.018

(number of clips). Condition 1 is used for calibrating the parameters, which are the mean and standard deviation of the unknown drag coefficient. Conditions 2 and 3 are used for validating the predicted variability of fall time for different falling height or weight.

Table 6 shows the fall time data collected by student A. The student measured the masses of helicopters using a very accurate scale with an error of 0.001 g to measure the mass of each helicopter. The student measured helicopters repeatedly and get the average mass of each helicopter. Table 7 shows the fall time data collected by student B. The student constructed five helicopters and chooses three of those five helicopters. Table 8 shows the fall time data collected by student C.

Table 9 Mass of helicopters of students A, B, and C

Student A (g)	H1	H2	H3
With two clips	2.78	2.76	2.77
With one clip	2.05	2.03	2.05
Student B	H1	H2	H3
With two clips	2.31	2.31	2.31
With one clip	1.89	1.89	1.89
Student C	H1	H2	H3
With two clips	2.37	2.40	2.41
With one clip	1.90	1.93	1.94

Table 9 shows measured mass from students A, B, and C. Student A measured the average masses of the helicopters with different number of clips. Student B measured the masses of five helicopters and used the average mass as the representative mass of the helicopters. Student C measured the masses of three helicopters with an accurate scale as student A. Note that the helicopter weights were measured within the limit of measurement accuracy of the scales.

Appendix C: Testing Normality of Drag Coefficient

In this paper, the drag coefficient is assumed to follow a normal distribution. Since the quadratic and linear drag models are considered in this paper, the drag coefficients corresponding to the fall time data for calibration are obtained for each drag model. As described in Sec. 4, ten drag coefficients are converted from the fall time data to obtain the posterior distribution of the drag coefficients for each drag model. To pool the drag coefficients from the students, we standardize the drag coefficients of each student using the mean and standard deviation and pool all the standardized coefficients for each drag model.

The normal probability plot is a common tool to visually evaluate the normality of the pooled drag coefficient data. The MATLAB function “normplot” is used to obtain the normal probability plots. Figure 5 shows that the data are close to the diagonal lines for both the quadratic and linear drag models. That means the normal distribution assumption is very likely reasonable.

References

- [1] Causevic, A., Sundmark, D., and Punnekkat, S., 2010, "An Industrial Survey on Contemporary Aspects of Software Testing," *Third International Conference on Software Testing, Verification and Validation (ICST)*, pp. 393–401.
- [2] Sellers, J. J., Larson, W. J., Kirkpatrick, D., and White, J., 2009, "Redefining Space System Verification and Validation," US Air Force T&E Days, U.S. Air Force T&E Days Conferences, Paper No. *AIAA* 2009-1752.
- [3] Kamal, M. M., 1970, "Analysis and Simulation of Vehicle to Barrier Impact," SAE Technical Paper, Paper No. 700414.
- [4] Happer, A., Araszewski, M., Toor, A., Overgaard, R., and Johal, R., 2000, "Comprehensive Analysis Method for Vehicle/Pedestrian Collisions," *SAE Technical Paper*, Paper No. 2000-01-0846.
- [5] Borovinšek, M., Vesenjāk, M., Ulbin, M., and Ren, Z., 2007, "Simulation of Crash Tests for High Containment Levels of Road Safety Barriers," *Eng. Failure Anal.*, **14**(8), pp. 1711–1718.
- [6] Kroll, N., Rossow, C. C., Schwamborn, D., Becker, K., and Heller, G., 2002, "MEGAFLOW—A Numerical Flow Simulation Tool for Transport Aircraft Design," ICAS Congress, Paper No. 2002-1.10.5.
- [7] Jameson, A., Martinelli, L., and Pierce, N. A., 1998, "Optimum Aerodynamic Design Using the Navier–Stokes Equations," *Theor. Comput. Fluid Dyn.*, **10**(1–4), pp. 213–237.
- [8] Roy, C. J., and Oberkampf, W. L., 2011, "A Comprehensive Framework for Verification, Validation, and Uncertainty Quantification in Scientific Computing," *Comput. Methods Appl. Mech. Eng.*, **200**(25), pp. 2131–2144.
- [9] Ling, Y., and Mahadevan, S., 2013, "Quantitative Model Validation Techniques: New Insights," *Reliab. Eng. Syst. Saf.*, **111**, pp. 217–231.
- [10] Oberkampf, W. L., and Roy, C. J., 2010, *Verification and Validation in Scientific Computing*, Cambridge University Press, New York.
- [11] Tobin, K. G., and Capie, W., 1981, "The Development and Validation of a Group Test of Logical Thinking," *Educ. Psychol. Meas.*, **41**(2), pp. 413–423.
- [12] Siorek, T. L., and Hafitka, R. T., 1998, "Paper Helicopter—Experimental Optimum Engineering Design Classroom Problem," The American Institute of Aeronautics and Astronautics, Paper No. *AIAA*-98-4963.
- [13] Hafitka, R. T., and Jenkins, D. A., 1998, "Classroom Project in Analytical and Experimental Optimization," *Struct. Optim.*, **15**(1), pp. 63–67.
- [14] Owen, J. P., and Ryu, W. S., 2005, "The Effects of Linear and Quadratic Drag on Falling Spheres: An Undergraduate Laboratory," *Eur. J. Phys.*, **26**(6), pp. 1085–1091.
- [15] Timmerman, P., and van der Weele, J. P., 1999, "On the Rise and Fall of a Ball With Linear or Quadratic Drag," *Am. J. Phys.*, **67**(6), pp. 538–546.
- [16] Thompson, B. G., and Smith, P. A., 2004, "An Experiment in Rotational Motion With Linear and Quadratic Drag," *Am. J. Phys.*, **72**(6), pp. 839–842.
- [17] Mohazzabi, P., 2010, "Falling and Rising in a Fluid With Both Linear and Quadratic Drag," *Can. J. Phys.*, **88**(9), pp. 623–626.
- [18] Trucano, T. G., Swiler, L. P., Igusa, T., Oberkampf, W. L., and Pilch, M., 2006, "Calibration, Validation, and Sensitivity Analysis: What's What," *Reliab. Eng. Syst. Saf.*, **91**(10), pp. 1331–1357.
- [19] Murphy, K. P., 2007, "Conjugate Bayesian Analysis of the Gaussian Distribution," *University of British Columbia preprint*, Oct. 3.
- [20] Andrieu, C., De Freitas, N., Doucet, A., and Jordan, M. I., 2003, "An Introduction to MCMC for Machine Learning," *Mach. Learn.*, **50**(1–2), pp. 5–43.
- [21] Oberkampf, W. L., DeLand, S. M., Rutherford, B. M., Diegert, K. V., and Alvin, K. F., 2002, "Error and Uncertainty in Modeling and Simulation," *Reliab. Eng. Syst. Saf.*, **75**(3), pp. 333–357.