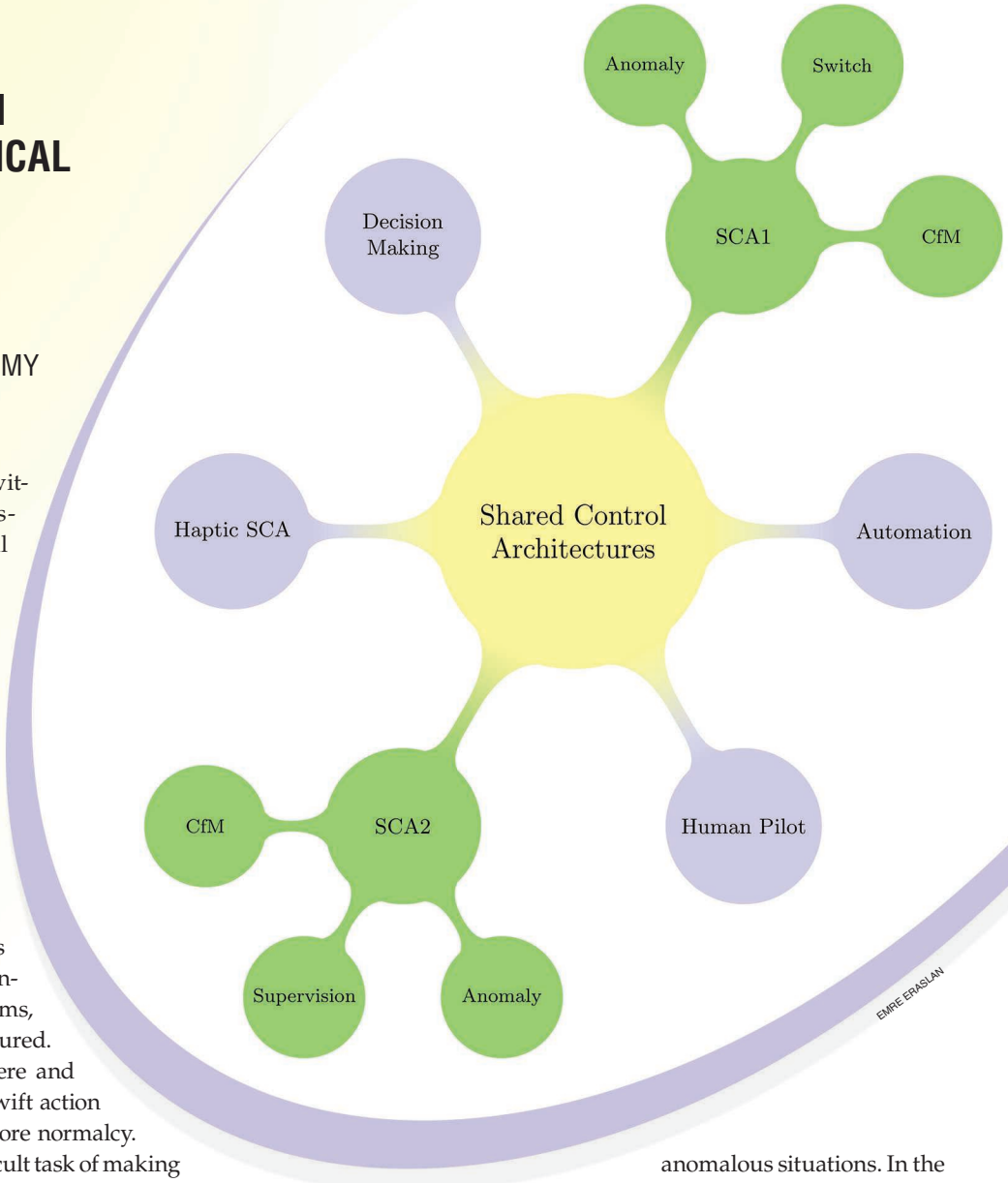
# Shared Control Between Pilots and Autopilots

## AN ILLUSTRATION OF A CYBERPHYSICAL HUMAN SYSTEM

EMRE ERASLAN,
YILDIRAY YILDIZ, and
ANURADHA M. ANNASWAMY




The 21st century is witnessing large transformations in several sectors related to autonomy, including energy, transportation, robotics, and health care. Decision making using real-time information over a large range of operations (as well as the ability to adapt online in the presence of various uncertainties and anomalies) is the hallmark of an autonomous system. To design such a system, a variety of challenges must be addressed. Uncertainties may occur in several forms, both structured and unstructured. Anomalies may often be severe and require rapid detection and swift action to minimize damage and restore normalcy. This article addresses the difficult task of making autonomous decisions in the presence of severe anomalies. While the specific application of focus is flight control, the overall solutions proposed are applicable for general complex dynamic systems.

The domain of decision making is common to both human experts and feedback control systems. Human experts routinely make several decisions when faced with anomalous situations. In the specific context of flight control systems, pilots often make several decisions based on sensory information from the cockpit, situational awareness, and their expert knowledge of the aircraft to ensure a safe performance. All autopilots in a fly-by-wire aircraft are programmed to provide the appropriate corrective input to ensure that the requisite variables accurately follow the specified guidance commands. Advanced autopilots ensure that such a command following occurs, even in the presence of uncertainties and anomalies. However, the process of

*Digital Object Identifier 10.1109/MCS.2020.3019721*
*Date of current version: 13 November 2020*

assembling various pieces of information that may help detect the anomaly may vary between the pilot and autopilot. Once the anomaly is detected, the process of mitigating the impact of the anomaly may also differ between the pilot and autopilot. The nature of perception, speed of response, and intrinsic latencies may all vary significantly between the two decision makers. It may be argued that perception and detection of the anomaly may be carried out efficiently by the human pilot, whereas fast action following a command specification may be best accomplished by an autopilot. The thesis in this article is that a shared control architecture (SCA), where the decision making of the human pilot is judiciously combined with that of an advanced autopilot, is highly attractive in flight control problems where severe anomalies are present (see "Summary").


## Summary

This article considers the problem of control when two distinct decision makers, a human operator and an advanced automation working together, face severe uncertainties and anomalies. We focus on shared control architectures (SCAs) that allow an advantageous combination of their abilities and provide a desired resilient performance. Humans and automation are likely to be interchangeable for routine tasks under normal conditions. However, under severe anomalies, the two entities provide complementary actions. It could be argued that human experts excel at cognitive tasks, such as anomaly recognition and estimation, while fast response with reduced latencies may be better accomplished by automation. This then suggests that architectures that combine their action must be explored. One of the major challenges with two decision makers in the loop is bumpy transfer when control responsibility switches between them. We propose the use of a common metric that enables a smooth, bumpless transition when severe anomalies occur. This common metric is termed capacity for maneuver (CfM), which is a concept rooted in human behavior and can be identified in control systems as the actuator's proximity to its limits of saturation. Two different SCAs are presented, both of which use CfM, and describe how human experts and automation can participate in a shared control action and recover gracefully from anomalous situations. Both of the SCAs are validated using human-in-the-loop experiments. The first architecture consists of a traded control action, where the human expert assumes control from automation when the CfM drops below a certain threshold and ensures a bumpless transfer. The second architecture includes a supervisory control action, where the human expert determines the tradeoff between the CfM and the corresponding degradation in command following and transmits suitable parameters to the automation when an anomaly occurs. The experimental results show that, in the context of flight control, these SCAs result in a bumpless, resilient performance.


Existing frameworks that combine humans and automation generally rely on human experts supplementing flight control automation, that is, the system is semiautonomous, with the automation handling all uncertainties and disturbances, with human control occurring once the environment imposes demands that exceed the automation capabilities [1]–[3]. This approach causes bumpy and late transfer of control from the machine to the human (causing the SCA to fail) as it is unable to keep pace with the cascading demand (which may cause actual accidents [4]–[7]). This suggests that alternate architectures of coordination and interaction between the human expert and automation are needed. Various architectures have been proposed in the literature that can be broadly classified into three forms: a trading action, where humans take over control from automation under emergency conditions [8]; a supervisory action, where the pilot assumes a high-level role and provides the inputs and setpoints for the automation to track [9]; and a combined action, where both the automation and the human expert participate at the same timescale [10]. We label all three forms under a collective name of SCA, with SCA1 denoting the trading action, SCA2 the supervisory action, and SCA3 the combined action (see Figure 1 for a schematic of all three forms of SCAs). This article focuses on SCA1 and SCA2 and their validation using human-in-the-loop simulations.

To have an efficient coordination between humans and automation, two principles from cognitive engineering are utilized that study how humans add resilience to complex systems [4], [5], [7]. The first principle is capacity for maneuver (CfM), which denotes a system's reserve capacity that will remain after the occurrence of an anomaly [5]. It is hypothesized that resiliency is rooted in and achieved via monitoring and regulation of the system CfM [7]. The notion of CfM exists in an engineering context as well, as the reserve capacity of an actuator. Viewing the actuator input power as the system capacity (and noting that a fundamental capacity limit exists in all actuators in the form of magnitude saturation), CfM can be defined for an autopilot-controlled system as the distance between the control input and saturation limits. The need to avoid actuator saturation (and, therefore, increasing CfM) becomes even more urgent in the face of anomalies, which may push the actuators to their limits. This implies that there is a common link between pilot-based decision making with an autopilot-based one in the form of CfM. This commonality is utilized in both SCA1 and SCA2, presented in this article.

The foundations for SCA1 and SCA2 come from the results of [11]–[13], with [11] and [12] corresponding to SCA1 and [13] corresponding to SCA2. In [11], the assumption is that the autopilot is designed to accommodate satisfactory operation under nominal conditions, using the requisite tracking performance. An anomaly is assumed to occur in the form of loss of effectiveness of the actuator (which may be due to damage to the control surfaces caused by a sudden change in environmental conditions or a compromised

engine due to bird strikes). The trading action proposed in [11] is for the pilot to assume control from the autopilot, based on the pilot perception that the automation is unable to address the anomaly. This perception is based on the CfM of the actuator, and when it exceeds a certain threshold, the pilot is proposed to assume control. A well-known model of the pilot in [1]–[3] is utilized to propose the specific sequence of control actions that pilots take once they assume control. It is shown through simulation studies that when the pilot carries out this sequence of perception and control tasks, the effect of anomaly is contained, and the tracking performance is maintained at a satisfactory level. A slight variation of this trading role is reported in [12], where, instead of the "autopilot is active → anomaly occurs → pilot control" sequence, the pilot transfers control from one autopilot to a more advanced one following the perception of an anomaly. This article limits attention to the SCA1 architecture in [11] and validates its performance using human-in-the-loop simulations.

The SCA2 architecture utilized in [13] differs from that of SCA1, as mentioned previously, as the trading action from the autopilot to the pilot is replaced by a supervisory action by the pilot. In addition to using CfM, this architecture utilizes a second principle from cognitive engineering denoted as graceful command degradation (GCD). GCD is proposed as an inherent metric adopted by humans [4] that will allow the underlying system to function so as to retain a target CfM. As the name connotes, GCD corresponds to the extent to which the system is allowed to relax its performance goals. When subjected to anomalies, it is reasonable to impose a command degradation; the greater this degradation, the larger the reserves that the system will possess when recovered from the anomaly. A GCD can then be viewed as a control variable tuned to permit a system to reach its targeted CfM. The role of tuning this variable so that the desired CfM is retained by the system is relegated to the pilot in [13]. Specifically, a parameter $\mu$ is transferred to the pilot from the autopilot, which is shown to result in an ideal tradeoff between the CfM and GCD by the overall closed-loop system with SCA2. This article validates SCA2 using a high-fidelity model of an aircraft and human-in-the-loop simulations.

The subjects used in the human-in-the-loop simulations include an airline pilot (who is also a flight instructor) with 2600 h of flight experience as well as human subjects who were trained in a systematic manner by the experiment designer. The details of the training are explained in the "Validation with Human-in-the-Loop Simulations" section. It is shown that SCA1 in [11] and SCA2 in [13] lead to better performance, as conjectured therein, when the human expert completes his or her assigned roles in the respective architectures. The resulting solution is an embodiment of an efficient cyberphysical human system, a topic that is of significant interest of late.

In addition to the aforementioned references, SCAs have been explored in haptic shared control (HSC), which has

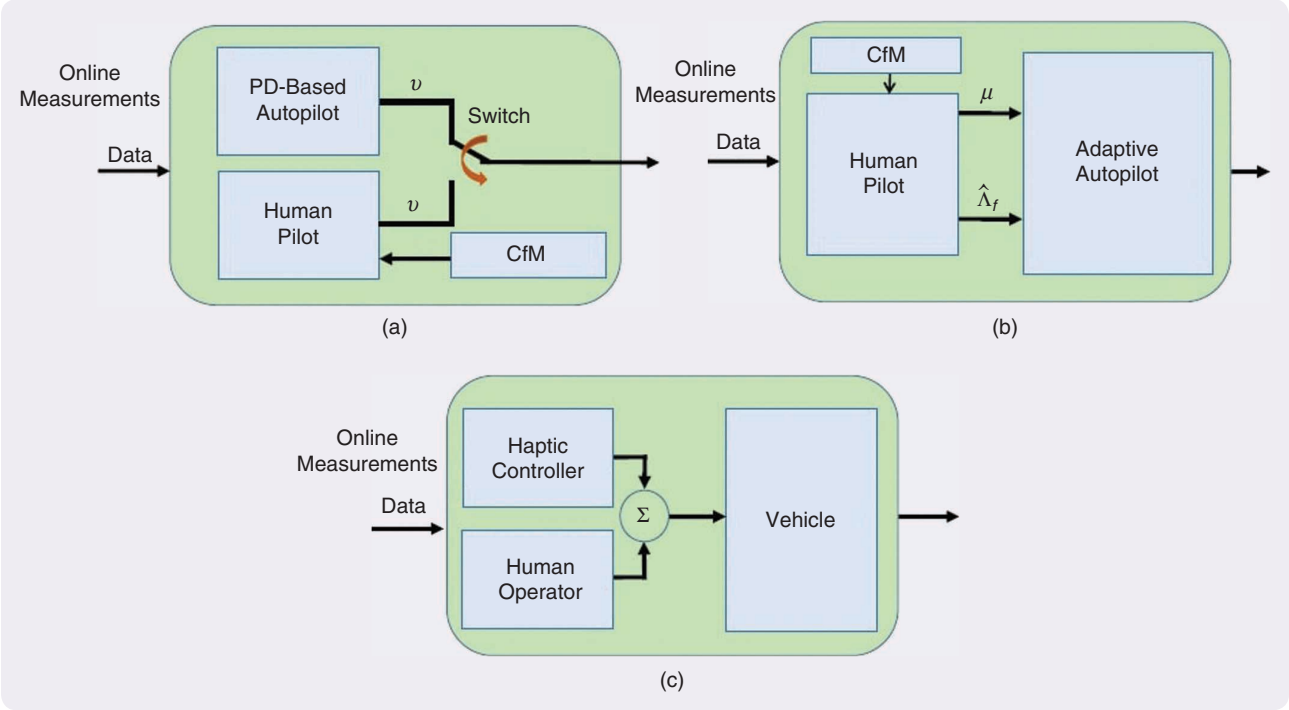



**FIGURE 1** An overview of shared control architectures (SCAs). (a) SCA1, in which the autopilot takes care of the control task until an anomaly, which is then transferred to the pilot based on capacity for maneuver (CfM). (b) SCA2, in which the pilot undertakes a supervisory role. CfM values approaching dangerously small numbers may trigger the pilot to assess the situation and convey necessary auxiliary inputs $\mu$ and $\hat{\Lambda}_{f_p}$ to the autopilot. (c) SCA3, which is based on the combined control effort continuously transmitted by both the haptic controller and human operator, who has the choice of abiding by or dismissing the control support of the automation. PD: proportional derivative.

applications both in the aerospace [14] and automotive [10], [15] domains. In HSC, both the automation and human operator exert forces on a common control interface (such as a steering wheel or a joystick) to achieve their individual goals. In this regard, HSC can be considered to have an SCA3 architecture. In [14], the goal of the haptic feedback is to help an unmanned aerial vehicle operator avoid an obstacle. In [15], one of the goals is lane-keeping, and in [10], the goal is assisting the driver while driving around a curve. All of these approaches provide situational awareness to the human operator, who has the ability to override the haptic feedback by exerting more force to the control interface. An interesting HSC study is presented in [16], where the haptic guidance authority is modified adaptively according to the intensity of the user grip, where challenging scenarios are created by introducing force disturbances or incorrect guidance. Another SCA approach is to enable the automation to affect the control input directly instead of using a common interface [17], which can also be viewed as belonging to SCA3. In this approach, the human operator can override the automation by disengaging it from the control system. Since SCA1 and SCA2 architectures help with anomaly mitigation and SCA3 is not shown to be relevant in this respect, this article focuses on SCA1 and SCA2 and their validations using human-in-the-loop simulation studies.

In the specific area of flight control, SCAs have also been proposed for greater situational awareness. A bumpless transfer between the autopilot and pilot is proposed in [18] using a more informative pilot display. It is argued in [18] that a common problem in pilot–automation interaction is the pilot's unawareness of the automation state and the aircraft during the flight (which makes the automation system opaque to the pilot), prompting the use of such a pilot display. The employment of the display is discussed in [18], in the presence of automation developed in [19] and [20] that provides flight envelope protection with a loss of control logic. A review of recent advances on human–machine interaction can be found in [21] (where the focus is adaptive automation), in which the automation monitors the human pilot to adjust itself accordingly so as to lead to improved situational awareness. No severe anomalies are considered in [18]–[21], which is the focus of this article. The goal of the SCAs discussed in this article is to lead to a bumpless efficient performance in the presence of severe anomalies. Therefore, discussions do not focus on these situational awareness-based research directions.

The following sections introduce the two main components of SCA: the autopilot and the human pilot. The "Autopilot" section provides the details of the employed controllers. The "Human Pilot" section first reviews the main mathematical models of pilots proposed in the literature in the absence of a severe anomaly. The human model used in the development of the proposed SCAs is then proposed. The working principles of these SCAs are later presented in detail in the section "Shared Control." Finally, the "Validation with Human-in-the-Loop Simulations" section shows how the underlying ideas are tested with human subjects.

## AUTOPILOT

This section describes the technical details related to the autopilot discussed in [11] and [13].

### *Dynamic Model of the Aircraft*

Since the autopilot in [13] is assumed to be determined using feedback control, the starting point is the description of the aircraft model, which has the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\Lambda_f u(t) + d + \Phi^T f(x), \\ y(t) &= Cx(t), \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ are deviations around a trim condition in aircraft states and control input, respectively, $d$ represents uncertainties associated with the trim condition, and $\Phi^T f(x)$ represents higher-order effects due to nonlinearities. $A$ is an $(n \times n)$ system matrix and $B$ is an $(n \times m)$ input matrix [both of which are assumed to be known, with $(A, B)$ controllable], and $\Lambda_f$ is a diagonal matrix that reflects a possible actuator anomaly with unknown positive entries $\lambda_{fi}$. $C$ is a known matrix of size $(k \times n)$, chosen so that $y \in \mathbb{R}^k$ corresponds to an output vector of interest. It is assumed that the anomalies occur at time $t_a$, so that $\lambda_{fi} = 1$ for $0 \le t < t_a$, and $\lambda_{fi}$ switches to a value that lies between zero and one for $t > t_a$. It is finally postulated that the higher-order effects are such that $f(x)$ is a known vector that can be determined at each instant of time, while $\Phi$ is an unknown vector parameter. Such a dynamic model is often used in flight control problems [22].

As the focus of this article is the design of a control architecture in the context of anomalies, actuator constraints are explicitly accommodated. Specifically, assume that $u$ is position/amplitude limited and modeled as

$$\begin{aligned} u_i(t) &= u_{\max_i} \mathrm{sat}\left(\frac{u_{c_i}(t)}{u_{\max_i}}\right) \\ &= \begin{cases} u_{c_i}(t), & |\, u_{c_i}(t)\,| \le u_{\max_i} \\ u_{\max_i}(t)\,\mathrm{sgn}(u_{c_i}(t)), & |\, u_{c_i}(t)\,| > u_{\max_i}, \end{cases} \end{aligned} \tag{2}$$

where $u_{\max_i}$ for $i = 1, \ldots, m$ are the physical amplitude limits of actuator $i$ and $u_{c_i}(t)$ are the control inputs to be determined by the SCA. The functions $\mathrm{sat}(\cdot)$ and $\mathrm{sgn}(\cdot)$ denote saturation and sign functions, respectively.

### *Advanced Autopilot Based on Adaptive Control*

The control input $u_{c_i}(t)$ will be constructed using an adaptive controller. To specify the adaptive controller, a reference model that specifies the commanded behavior from the plant is constructed and of the form [23]

$$\begin{aligned} \dot{x}_m(t) &= A_m x_m(t) + B_m r_0(t), \\ y_m(t) &= C x_m(t), \end{aligned} \tag{3}$$

where $r_0 \in \mathbb{R}^k$ is a reference input, $A_m$ $(n \times n)$ is a Hurwitz matrix, $x_m \in \mathbb{R}^n$ is the state of the reference model, $(A_m, B_m)$ is controllable, and $C$ is defined as in (1) (that is, $y_m \in \mathbb{R}^k$ corresponds to a reference model output). The goal of the adaptive autopilot is then to choose $u_{c_i}(t)$ in (5), so that if an error, $e$, is defined as

$$e(t) = x(t) - x_m(t), \tag{4}$$

then all signals in the adaptive system remain bounded with error $e(t)$ tending to zero asymptotically.

The design of adaptive controllers in the presence of control magnitude constraints is first addressed in [24], with guarantees of closed-loop stability through modification of the error used for the adaptive law. The same problem is addressed in [25] using an approach termed *$\mu$-mod adaptive control*, where the effect of input saturation is accommodated through the addition of another term in the reference model. Another approach based on a closed-loop reference model (CRM) is derived in [26] and [27] to improve the transient performance of the adaptive controller. The proposed autopilot is based on both the $\mu$-mod and CRM approaches. Using the control input in (2), this controller is compactly summarized as

$$u_{c_i}(t) = \begin{cases} u_{\mathrm{ad}_i}(t), & |\, u_{\mathrm{ad}_i}(t)\,| \le u^{\delta}_{\max_i} \\ \dfrac{1}{1+\mu}\left(u_{\mathrm{ad}_i}(t) + \mu \operatorname{sgn}(u_{\mathrm{ad}_i}(t)) u^{\delta}_{\max_i}\right), & |\, u_{c_i}(t)\,| > u^{\delta}_{\max_i} \end{cases}, \tag{5}$$

where

$$u_{\mathrm{ad}_i}(t) = K_x^T(t)x(t) + K_r^T(t)r_0(t) + \hat{d}(t) + \hat{\Phi}^T(t)f(x), \tag{6}$$

$$u^{\delta}_{\max_i} = (1-\delta)u_{\max_i}, \quad 0 \le \delta < 1. \tag{7}$$

A buffer region in the control input domain $[(1-\delta)u_{\max_i}, u_{\max_i}]$ is implied by (5) and (7), and the choice of $\mu$ allows the input to be scaled somewhere in between. The reference model is also modified as

$$\dot{x}_m(t) = A_m x_m(t) + B_m\left(r_0(t) + K_u^T(t)\Delta u_{\mathrm{ad}}(t)\right) - Le(t), \tag{8}$$

$$\Delta u_{\mathrm{ad}_i}(t) = u_{\max_i} \mathrm{sat}\left(\frac{u_{c_i}(t)}{u_{\max_i}}\right) - u_{\mathrm{ad}_i}(t), \tag{9}$$

and $L < 0$ is a constant or a matrix selected such that $(A_m + L)$ is Hurwitz. Finally, the adaptive parameters are adjusted as

$$\begin{aligned} \dot{K}_x(t) &= -\Gamma_x x(t) e^T(t) PB, \\ \dot{K}_r(t) &= -\Gamma_r r_0(t) e^T(t) PB, \\ \dot{\hat{d}}(t) &= -\Gamma_d e^T(t) PB, \\ \dot{\hat{\Phi}}(t) &= -\Gamma_f f(x(t)) e^T(t) PB, \\ \dot{K}_u(t) &= \Gamma_u \Delta u_{\mathrm{ad}} e^T(t) PB_m, \end{aligned} \tag{10}$$

where $P = P^T$ is a solution of the Lyapunov equation (for $Q > 0$)

$$A_m^T P + PA_m = -Q, \tag{11}$$

with $\Gamma_x = \Gamma_x^T > 0, \Gamma_r = \Gamma_r^T > 0, \Gamma_u = \Gamma_u^T > 0$.

The stability of the overall adaptive system specified by (1), (2), and (3)–(11) is established in [25] when $L = 0$. The stability of the adaptive system when no saturation inputs are present is also established in [27]. A very straightforward combination of the two proofs can be easily completed to prove that, when $L < 0$, the adaptive system in this article has globally bounded solutions if the plant in (1) is open-loop stable and bounded solutions for an arbitrary plant if all initial conditions and the control parameters in (10) lie in a compact set. The proof is skipped due to page limitations.

The adaptive autopilot in (5)–(11) provides the required control input $u$ in (1) as a solution to the underlying problem. The autopilot includes several free parameters, including $\mu$ in (5); $\delta$ in (7); the reference model parameters $A_m$, $B_m$, and $L$ in (8); and the control parameters $K_x(0)$, $K_r(0)$, $K_u(0)$, $\hat{d}(0)$, $\hat{\Phi}(0)$ in (10). As shown in the next section, the parameters $\delta$ and $\mu$ are related to CfM and GCD.

## Quantification of Capacity for Maneuver and Graceful Command Degradation and Tradeoffs

The control input $u_{c_i}$ in (5) is shaped by two parameters, $\delta$ and $\mu$, both of which help tune the control input with respect to its specified magnitude limit, $u_{\max_i}$. These two parameters are used to quantify CfM, GCD, and the tradeoffs between them as follows.

### *Capacity for Maneuver*

Qualitatively, CfM corresponds to a system's reserved capacity, which is quantitatively formulated in the current context as the distance between a control input and its saturation limits. Specifically, define CfM as

$$\mathrm{CfM} = \frac{\mathrm{CfM}^R}{\mathrm{CfM}_d}, \tag{12}$$

where

$$\begin{aligned} \mathrm{CfM}^R &= \mathrm{rms}\left(\min_i (c_i(t))\right) \Big|_{t_a}^{T}, \\ c_i(t) &= u_{\max_i} - |\, u_i(t)\,|, \end{aligned} \tag{13}$$

where $c_i(t)$ is the instantaneous available control input of actuator $i$, $\mathrm{CfM}^R$ is the root mean squared CfM variation, and $\mathrm{CfM}_d$ (which denotes the desired CfM) is chosen as

$$\mathrm{CfM}_d = \max_i (\delta u_{\max_i}). \tag{14}$$

In these equations, min and max are the minimum and maximum operators over the $i$th index, rms is the root-mean-square operator, and $t_a$ and $T$ refer to the time of anomaly and final simulation time, respectively. From (13), note that $\mathrm{CfM}^R$ has i) a maximum value $u_{\max}$ for the trivial case when all $u_i(t) = 0$, ii) a value close to $\delta u_{\max}$ if the control inputs approach the buffer region, and iii) a value of

zero if $u_i(t)$ hits the saturation limit $u_{\max}$. Since $\mathrm{CfM}_d = \delta u_{\max}$, it follows that CfM (the corresponding normalized value) is greater than unity when the control inputs are small and far away from saturation, unity as they approach the buffer region, and zero when fully saturated.

### *Graceful Command Degradation*

As mentioned earlier, the reference model represents the commanded behavior from the plant being controlled. To reflect the fact that the actual output may be compromised if the input is constrained, a term was added that depends on $\Delta u_{\mathrm{ad}}(t)$ in (9) to become nonzero whenever the control input saturates (that is, when the control input approaches the saturation limit, $\Delta u_{\mathrm{ad}_i}$ becomes nonzero), thereby suitably allowing a graceful degradation of $x_m$ from its nominal choice, as in (8). This degradation is denoted as $\mathrm{GCD}_i$ and quantified as

$$\mathrm{GCD}_i = \frac{\mathrm{rms}\left(y_{m,i}(t) - r_{0,i}(t)\right)}{\mathrm{rms}\left(r_{0,i}(t)\right)}, \quad t \in T_0, \tag{15}$$

where $T_0$ denotes the interval of interest and $y_{m,i}$ and $r_{0,i}$ indicate the $i$th elements of the reference model output and reference input vectors, respectively. It should be noted that once $\mu$ is specified, the adaptive controller automatically scales the input into the reference model through $\Delta u_{\mathrm{ad}}$ and $K_u$, such that $e(t)$ remains small and the closed-loop system has bounded solutions.

### $\mu$

The intent behind the introduction of the parameter $\mu$ in (5) is to regulate the control input and move it away from saturation when needed. For example, if $|u_{\mathrm{ad}_i}(t)| > u^{\delta}_{\max_i}$, the extreme case of $\mu = 0$ will simply set $u_{c_i} = u_{\mathrm{ad}}$, thereby removing the effect of the virtual limit imposed in (7). As $\mu$ increases, the control input would decrease in magnitude and move toward the virtual saturation limit $u^{\delta}_{\max_i}$, that is, once the buffer $\delta$ is determined, $\mu$ controls $u_i(t)$ within the buffer region $[(1-\delta)u_{\max_i}, u_{\max_i}]$, bringing it closer to the lower limit with increasing $\mu$. In other words, as $\mu$ increases, CfM increases as well in the buffer region.

It is easy to see from (8) and (9) that, similar to CfM, as $\mu$ increases, GCD also increases. This is because an increase in $\mu$ increases $\Delta u_{\mathrm{ad}_i}(t)$, which, in turn, increases the GCD. While a larger CfM improves the responsiveness of the system to future anomalies, a lower bound on the reference command is necessary to finish the mission within practical constraints. In other words, $\mu$ must be chosen so that GCD remains above a lower limit while maintaining a large CfM. The task of selecting the appropriate $\mu$ is relegated to the human pilot.

### Choice of the Reference Model Parameters

In addition to $\mu$ and $\delta$, the adaptive controller in (2) and (3)–(11) requires the reference model parameters $A_m$, $B_m$, and $L$ and the control parameters $K_x(0)$, $K_r(0)$, and $K_u(0)$ at time $t = 0$. If no anomalies are present, then $\Lambda_{\mathrm{nom}} = \Lambda_f = I$, which implies that $A_m$ and $B_m$ (as well as the control parameters) can be chosen as

$$\begin{aligned} A_m &= A + BK_x^T(0), \\ K_r^T(0) &= -(A_m^{-1}B)^{-1}, \\ B_m &= BK_r^T(0), \\ K_u^T(0) &= -A_m^{-1}B, \end{aligned} \tag{16}$$

where $K_x(0)$ is computed using a linear-quadratic regulator method and the nominal plant parameters $(A, B)$ [28] and $K_r(0)$ in (16) are selected to provide unity low-frequency dc gain for the closed-loop system. When anomalies occur $(\Lambda_f \neq I)$ at time $t = t_a$ and assuming an estimate $\hat{\Lambda}_f$ is available, a choice similar to that in (16) can be completed using the plant parameters $(A, B\hat{\Lambda}_f)$ and the relations

$$\begin{aligned} A_m &= A + B\hat{\Lambda}_f K_x^T(t_a), \\ K_r^T(t_a) &= -(A_m^{-1}B\hat{\Lambda}_f(t_a))^{-1}, \\ B_m &= B\hat{\Lambda}_f K_r^T(t_a), \\ K_u^T(t_a) &= -A_m^{-1}B\hat{\Lambda}_f(t_a), \end{aligned} \tag{17}$$

with the adaptive controller specified using (2)–(11) for all $t \geq t_a$. Finally, $L$ is chosen as in [27], and lower parameters $\hat{d}(0)$, $\hat{\Phi}(0)$ are chosen arbitrarily. Similar to $\mu$, the task of assessing the estimate $\hat{\Lambda}_{f_p}$ is also relegated to the human pilot.

## *Autopilot Based on Proportional-Derivative Control*

To investigate SCAs, another autopilot employed in the closed-loop system is the proportional-derivative (PD) controller. Assuming a single control input [11], the goal is to control the dynamics [1]

$$Y_p(s) = \frac{1}{s(s+a)}, \tag{18}$$

which represents the aircraft transfer function between the input $u$ and an output $M(t)$, which is assumed to be a simplified version of the dynamics in (1). The input $u$ is subjected to the same magnitude and rate constraints as those in (2). Considering the transfer function (18) between the input $u$ and output $M(t)$, the PD controller can be chosen as

$$u(t) = K_p(M(t) - M_{\mathrm{cmd}}(t)) + K_r(\dot{M}(t)), \tag{19}$$

where $M_{\mathrm{cmd}}$ is the desired command signal that $M$ is required to follow. Given the second-order structure of the dynamics, it can be shown that suitable gains $K_p$ and $K_r$ can be determined so that the closed-loop system is stable, and for command inputs at low frequencies, a satisfactory tracking performance can be obtained.

It is noted that, during the experimental validation studies, certain anomalies are introduced to (18) in the form of unmodeled dynamics and time delays. Therefore, according to the crossover model [29], the human pilots must adapt themselves to demonstrate different compensation

characteristics (such as pure gain, lead, or lag) based on the type of the anomaly. This creates a challenging scenario for the SCA.

## HUMAN PILOT

This section discusses mathematical models of human pilot decision making on the basis of the absence and presence of flight anomalies. A great deal of research was conducted on mathematical human pilot modeling, assuming that no failure in the aircraft or severe disturbances in the environment are present. Since decision making will differ significantly whether the aircraft is under nominal operation or subjected to severe anomalies, the corresponding models are entirely different as well and discussed separately in what follows.

### *Pilot Models in the Absence of Anomaly*

In the absence of an anomalous event(s), mathematical models of human pilot control behavior can be classified according to control-theoretic, physiological, and, more recently, machine learning methods [30], [31]. One of the most well-known control-theoretic methods in the modeling of human pilots (namely, the *crossover model*) is presented in [29] as an assembly of the pilot and the controlled vehicle for single-loop control systems. The open-loop transfer function for the crossover model is

$$Y_h(j\omega)Y_p(j\omega) = \frac{\omega_c e^{-\tau_e j\omega}}{j\omega}, \tag{20}$$

where $Y_h(j\omega)$ is a transfer function of the human pilot, $Y_p(j\omega)$ is a transfer function of the aircraft, $\omega_c$ is the crossover frequency, and $\tau_e$ is the effective time delay pertinent to the system delays and human pilot lags. The crossover model is applicable for a range of frequencies around the crossover frequency $\omega_c$. When a "remnant" signal is introduced to $Y_h(j\omega)$ to account for the nonlinear effects of the pilot-vehicle system, the model is called a quasi-linear model [32].

Other sophisticated quasi-linear models can be found in the literature as the *extended crossover model* [32] [which works especially for conditionally stable systems, that is, when a pilot attempts to stabilize an unstable transfer function of the controlled element, $Y_p(j\omega)$] and *the precision model* [32] (which treats a wider frequency region than the crossover model). The single-loop control tasks are covered by quasi-linear models. They can be extended to multiloop control tasks by the introduction of the *optimal control model* [33], [34].

Another approach in modeling the human pilot is employing the information of sensory dynamics relevant to humans to extract the effect of motion, proprioceptive, vestibular, and visual cues on the control effort. An example of this is the *descriptive model* [35], in which a series of experiments are conducted to distinguish the influence of vestibular and visual stimuli from the control behavior. Another example is the *revised structural model* [36], where the human pilot is modeled as the unification of proprioceptive, vestibular, and visual feedback paths. It is hypothesized that such cues help alleviate the compensatory control action taken by the human pilot.

It is noted that, due to their physical limitations, models such as that in (20) are valid for operation in a predefined boundary or envelope where the environmental factors are steady and stable. In the case of an anomaly, they may not perform as expected [37].

### *Pilot Models in the Presence of Anomaly*

When extreme events and failures occur, human pilots are known to adapt themselves to changing environmental conditions, which overstep the boundaries of automated systems. Since the hallmark of any autonomous system is its ability to self-govern (even under emergency conditions), modeling of the pilot decision making upon the occurrence of an anomaly is indispensable, and examples are present in the literature [1]–[3], [11]–[13], [38], [39]. This article assumes that the anomalies can be modeled either as an abrupt change in the vehicle dynamics [1]–[3], [11], [12], [39] or a loss of control effectiveness in the control input [13], [38]. In either case, the objective is the modeling of the decision making of the pilot so as to elicit a resilient performance from the aircraft and recover rapidly from the impact of the anomaly. Since the focus of this article is shared control, among the pilot models that are developed for anomaly response, we exploit those that explain the pilot behavior in relation to the autopilot. These models are developed using the CfM concept, the details of which are explained in the following sections.

#### Pilot Models Based on Capacity for Maneuver

A recent method in modeling the human pilot under anomalous events utilizes the CfM concept [11], [13], [38]. As discussed in the "Autopilot" section, CfM refers to the remaining range of the actuators before saturation, which quantifies the available maneuvering capacity of the vehicle. It is hypothesized that surveillance and regulation of a system's available capacity to respond to all events help maintain the resiliency of a system, which is a necessary merit to recover from unexpected and abrupt failures or disturbances [7]. Two different types of pilot models are proposed, both of which use CfM, but in different ways.

##### *Perception Trigger*

The pilot model is assumed to assess the CfM and implicitly compute a perception gain based on it. The quantification of this gain $K_t$ is predicated on $\text{CfM}^R$ with the definition in (13). The perception trigger is associated to the gain $K_t$, which is implicitly computed in [11]. The perception algorithm for the pilot is

$$K_t = \begin{cases} 0, & |F_0| < 1 \\ 1, & |F_0| \geq 1, \end{cases} \tag{21}$$

where

$$F_0 = G_1(s)\,[F(t)],$$
$$F(t) = \frac{\frac{d}{dt}(\mathrm{CfM}^{R_m}) - \mu_p}{3\sigma_p}$$
$$\mathrm{CfM}^{R_m} = u_{\max} - \mathrm{rms}\,(u(t)) \tag{22}$$

$G_1(s)$ is a second-order filter introduced as a smoothing and lagging operator into the human perception algorithm, $F(t)$ is the perception variable, $\mathrm{CfM}^{R_m}$ is a slightly modified version of $\mathrm{CfM}^{R}$ in (13), $\mu_p$ is the average of $(d/dt)(\mathrm{CfM}^{R_m})$, and $\sigma_p$ is the standard deviation of $(d/dt)(\mathrm{CfM}^{R_m})$, both of which are measured over a nominal flight simulation. The computation of these statistical parameters is further elaborated in the "Validation of Shared Control Architecture 1" section. The hypothesis is that the human pilot has such a perception trigger $K_t$, and when $K_t = 1$, the pilot assumes control from the autopilot. The "Validation of Shared Control Architecture 1" section validates this perception model.

### *Capacity for Maneuver-Graceful Command Degradation Tradeoff*

The pilot is assumed to implicitly assess the available (normalized) CfM when an anomaly occurs and decide on the amount of GCD that is allowable, to let the CfM become comparable to the $\mathrm{CfM}_d$. In other words, assume that the pilot is capable of assessing the parameter $\mu$ and input this value to the autopilot following the occurrence of an anomaly. That is, the pilot model is assumed to take the available CfM as the input and deliver $\mu$ as the output. The "Validation of Shared Control Architecture 2" section validates this assessment.

## SHARED CONTROL

The focus of this article is on an SCA that combines the decision making of a pilot and autopilot in flight control. The architecture is invoked under alert conditions, with triggers in place that specify when the decision making is transferred from one authority to another. The specific alert conditions of focus in this article correspond to physical anomalies that compromise the actuator effectiveness. The typical roles of the autopilot and pilot in a flight control problem were

## Different Forms of Shared Control Architectures

Human–machine interactions can be varied in nature, and a vast amount of literature of human–machine interactions is available in a number of sectors, including energy, transportation, health care, robotics, and manufacturing. Broadly speaking, the resulting systems that emerge from these interactions come under the rubric of the cyberphysical and human system (CPHS).

The adjective *shared* in the phrase "shared control architectures" (SCAs) is used in a broad sense, where a human expert shares the decision making with automation in some form or the other in the CPHS [S1]. These SCAs can be grouped into three categories: SCA1 as traded [S1], SCA2 as supervised [S2], and SCA3 as combined [S3].

In SCA1, the authority shifts from the automation to the human expert as an emergency override when dictated by the anomaly. For example, in a teleoperation task, if a robot autonomously travels on the remote environment and at certain time intervals, a human operator assumes control and provides real-time speed and direction commands, and then this type of interaction can be considered traded control.

In SCA2, the human expert is on the loop and provides a supervisory, high-level input to the automation. One such example may occur in the teleoperation scenario mentioned in the previous paragraph, where the human operator can provide reference points for the mobile robot and the control system on the robot can autonomously navigate the vehicle, make it avoid obstacles, and reach the reference points.

In SCA3, different from SCA1 and SCA2, the human and the automation are both simultaneously active. For example, in the discussed teleoperation task, if the human controls the robot via a joystick in real time and, at the same time, the robot controller continuously sends force inputs to the joystick to keep the robot collision-free on the terrain, then this can be considered combined control.

The main distinction between different classifications originates from the timing of the human and automation control inputs. The commonality between the classes, however, is that both the human and the automation share the load of control, albeit with different timescales. Therefore, for ease of exposition, shared control is used as the collective rubric for both of the interaction types investigated in this article.

It is noted that the use of the adjective *shared* in this article is different than what has been proposed in [S1] and [S2], where the term *shared* is reserved exclusively for actions that both humans and automation simultaneously exert at the same timescale. We denote this as a combined action (SCA3) and the collective human–automation control architectures as *shared*. As we better understand the ramifications of these different types of CPHS, these definitions may also evolve.

### REFERENCES


[S1] D. A. Abbink et al., "A topology of shared control systems? Finding common ground in diversity," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 5, pp. 509–525, Oct. 2018. doi: 10.1109/THMS.2018.2791570.

[S2] T. B. Sheridan, Ed. *Monitoring Behavior and Supervisory Control*, vol. 1. Berlin: Springer-Verlag, 2013.

[S3] M. Mulder, D. A. Abbink, and E. R. Boer, "Sharing control with haptics: Seamless driver support from manual to automatic control," *Hum. Factors J. Hum. Factors Ergonom. Soc.*, vol. 54, no. 5, pp. 786–798, May 2012. doi: 10.1177/0018720812443984.

described in the "Autopilot" and "Human Pilot" sections, respectively. In the "Autopilot" section, autopilots based on PD control and adaptive control were described, with the former designed to ensure satisfactory command following under nominal conditions and the latter to accommodate parametric uncertainties (including the loss of control effectiveness in the actuators). In the "Pilot Models in the Presence of Anomaly" section, two different models of decision making in pilots were proposed, both based on the monitoring of CfM of the actuators. This section proposes two different SCAs using the models of the autopilots and pilots described in the previous sections (see "Different Forms of Shared Control Architectures").

### *Shared Control Architecture 1: A Pilot With a Capacity for Maneuver-Based Perception and a Fixed-Gain Autopilot*

The first SCA can be summarized as a sequence {autopilot runs, anomaly occurs, pilot takes over}. That is, it is assumed that an autopilot based on PD control [as in (19)] is in place, ensuring satisfactory command tracking under nominal conditions. The human pilot is assumed to consist of a perception component and an adaptation component. The perception component consists of monitoring $\mathrm{CfM}^{R_m}$, through which a perception trigger $F_0$ is calculated using (22). The adaptation component consists of monitoring the control gain in (21) and assuming control of the aircraft when $K_t = 1$. The details of this shared controller and its evaluation using a numerical simulation study can be found in [11]. Figure 2 illustrates the schematic of SCA1.

### *Shared Control Architecture 2: A Pilot With Capacity for Maneuver-Based Decision Making and an Advanced Adaptive Autopilot*

In this SCA, the role of the pilot is a supervisory one, while the autopilot takes on an increased and more complex role. The pilot is assumed to monitor CfM of the resident actuators in the aircraft following an anomaly. In an effort to allow CfM to stay close to $\mathrm{CfM}_d$ in (14), the command is allowed to be degraded; the pilot then determines a parameter $\mu$ that directly scales the control effort through (5) and indirectly scales the command signal through (8) and (9). Once $\mu$ is specified by the pilot, then the adaptive autopilot continues to supply the control input using (5)–(11). If the pilot has high situational awareness, he or she provides $\hat{\Lambda}_{f_p}$ as well, which is an estimate of the severity of the anomaly. The details of this shared controller and its evaluation using a numerical simulation study can be found in [13] and [38]. Figure 3 displays the schematic of SCA2.

## VALIDATION WITH HUMAN-IN-THE-LOOP SIMULATIONS

The goal in this section is to validate two hypotheses of the human pilot actions, namely, SCA1 and SCA2. Despite the

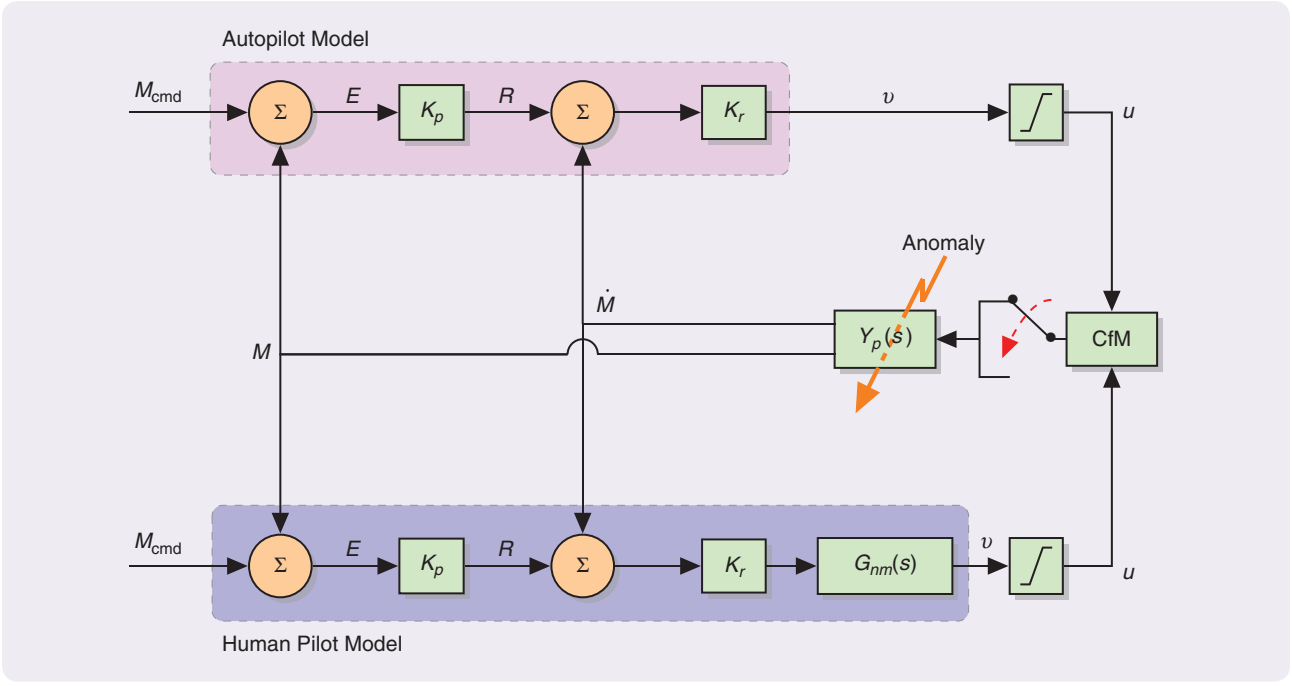



**FIGURE 2** A block diagram of shared control architecture 1 (adapted from [11]). The autopilot model consists of a fixed gain controller, whereas the human pilot model comprehends a perception part based on the capacity for maneuver (CfM) concept and an adaptation part governed by empirical adaptive laws [3]. The neuromuscular transfer function [42] is $G_{nm}(s)$, which corresponds to control input formed by an arm or leg, $G_{nm}(s) = (100/s^2 + 14.14s + 100)$. When an anomaly occurs, the plant dynamics $Y_p(s)$ undergo an abrupt change by rendering the autopilot insufficient for the rest of the control. At this stage, the occurrence of an anomaly is captured by the CfM such that the control is transferred to the human pilot model for resilient flight control.

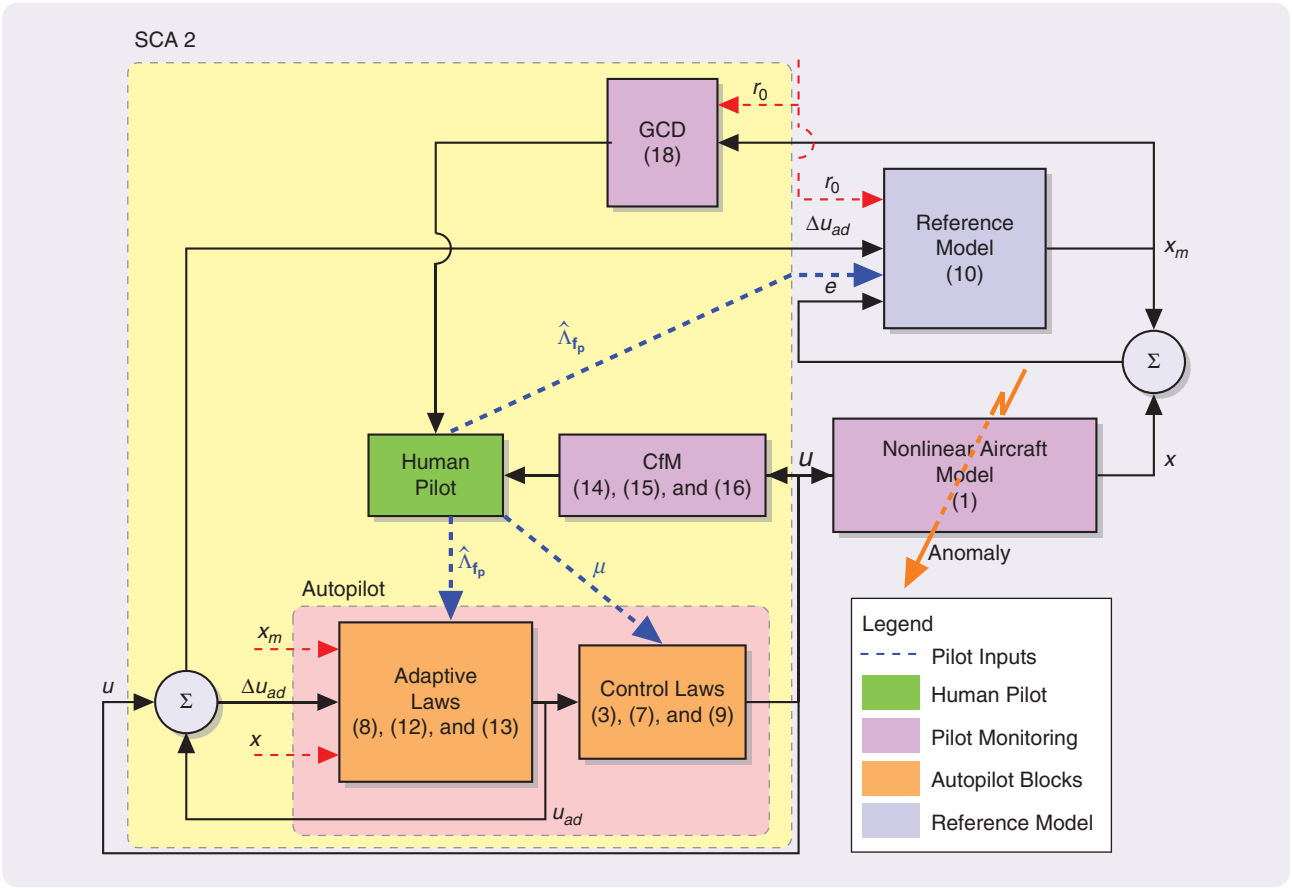



**FIGURE 3** A block diagram of the proposed shared control architecture 2 (SCA2). The human pilot undertakes a supervisory role by providing the key parameters $\mu$ and $\hat{\Lambda}_{f_p}$ to the adaptive autopilot. The blocks are expressed in different colors based on their functions in the proposed SCA2. The numbers in parentheses in each block correspond to the related equations. GCD: graceful command degradation; CfM: capacity for maneuver.

presence of obvious common elements to the two SCAs of a human pilot, an autopilot, and a shared controller that combines their decision making, the details differ significantly. We therefore describe the validation of these two hypotheses separately in what follows. In each case, the validation is presented in the following order: the experimental setup, the type of anomaly, the experimental procedure, details of the human subjects, the pilot model parameters, results, and observations.

Although the experimental procedures used for SCA1 validation and SCA2 validation differ (and are therefore explained in separate sections), they use similar principles. To minimize repetition, main tasks can be summarized as follows. The procedure comprises three main phases: *Pilot Briefing*, *Preparation Tests*, and *Performance Tests*. In *Pilot Briefing*, the aim of the experiment and the experimental setup are introduced. In *Preparation Tests*, the subjects are encouraged to gain practice with the joystick controls. In the last phase, *Performance Tests*, the subjects are expected to conduct the experiments only once to not affect the reliability due to learning. Anomaly introduction times are randomized to prevent predictability. As nominal (anomaly-free) plant dynamics, the simpler model introduced in (18) is used for SCA1 validation, and a nonlinear F-16 dynamics [40], [41] is used for SCA2 validation. The experiment was approved by the Bilkent University Ethics Committee, and informed consent was taken from each subject before conducting the experiment.

It is noted that during the validation studies, SCA1 and SCA2 are not compared. They are separately validated with different human-in-the-loop simulation settings. However, a comparison of SCA1 and SCA2 can be pursued as a future research direction.

## *Validation of Shared Control Architecture 1*

### Experimental Setup

The experimental setup consists of a pilot screen and the commercially available pilot joystick Logitech Extreme 3D Pro (see Figure 4), with the goal of performing a desktop, human-in-the-loop simulation. The flight screen interface for SCA1 can be observed in Figure 4(a) and is separately illustrated in Figure 5. The orange line with a circle in the middle shows the reference to be followed by the pilot. This line is moved up and down according to the desired reference command $M_{\text{cmd}}$ (see Figure 2). The blue sphere in Figure 5 represents the nose tip of the aircraft and is driven by the joystick inputs. When the subject moves the joystick,

(a) (b)

**FIGURE 4** The experimental setups for (a) shared control architecture 1 (SCA1) and (b) SCA2. The SCA1 experiment consists of a pilot screen (see Figure 5) and a commercially available pilot joystick, whose pitch input is used. The SCA2 experiment consists of a different pilot screen (see Figure 6) and the joystick. In SCA2, subjects use the joystick lever to provide input.

he or she provides the control input, $v$, which goes through the aircraft model and produces the movement of the blue sphere. The control objective is to keep the blue sphere inside the orange circle, which translates into tracking the reference command. Similarly, the flight screen for SCA2 is seen in Figure 4(b) and separately illustrated in Figure 6. The details on this screen are provided in the "Validation of Shared Control Architecture 2" section.

## Anomaly

The anomaly is modeled by a sudden change in vehicle dynamics, from $Y_p^{\text{before}}(s)$ to $Y_p^{\text{after}}(s)$ (as in [3]), and it is illustrated in Figure 2. Two different flight scenarios, $S_{\text{harsh}}$ and $S_{\text{mild}}$, are investigated, which correspond to a harsh and a mild anomaly, respectively. In the harsh anomaly, it is assumed that

$$Y_{p,h}^{\text{before}}(s) = \frac{1}{s(s+10)}, \quad Y_{p,h}^{\text{after}}(s) = \frac{e^{-0.2s}}{s(s+5)(s+10)}, \tag{23}$$

whereas in the mild anomaly, it is assumed that

$$Y_{p,m}^{\text{before}}(s) = \frac{1}{s(s+7)}, \quad Y_{p,m}^{\text{after}}(s) = \frac{e^{-0.18s}}{s(s+7)(s+9)}. \tag{24}$$

The specific numerical values of the parameters in (23) and (24) are chosen so that the pilot action has a distinct effect in the two cases based on their response times. More details of these choices are provided in the following section.

The anomaly is introduced at a certain instant of time $t_a$ in the experiment. The anomaly alert is conveyed as a sound signal at $t_s$, following which the pilots take over control at time $t_{\text{TRT}}$, after a certain reaction time $t_{\text{RT}}$. Denoting $\Delta T = t_s - t_a$ as the alert time, the total elapsed time from the onset of anomaly to the instant of hitting the joystick button is defined as

$$t_{\text{TRT}} = t_{\text{RT}} + \Delta T. \tag{25}$$

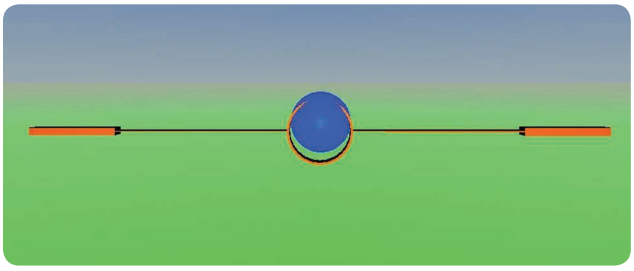


**FIGURE 5** The flight screen interface for shared control architecture 1. The blue sphere represents the aircraft nose and is controlled in the vertical direction via joystick movements. The orange line with a circle in the middle (which also moves in the vertical direction) represents the reference command to be followed. The control objective is to keep the blue sphere inside the orange circle. At the beginning of the experiments, the aircraft is controlled by the autopilot. When an anomaly occurs, the subjects are warned to assume control via a sound signal. The time of initiating this signal is determined using different alert times.

## Experimental Procedure

The experimental procedure consists of three main parts: *Pilot Briefing*, *Preparation Tests*, and *Performance Tests* (see Figure 7). The first part of the procedure is the *Pilot Briefing*, where the subjects are required to read a pilot briefing to have a clear understanding of the experiment. The briefing consists of six main sections: *Overall Purpose*, *Autopilot*, *Anomaly*, *Experimental Setup*, *Flight Screen*, and *Instructions*. In these sections, the main concepts and experimental

hardware (such as the pilot screen and the joystick lever) are introduced to the subject.

The second part is *Preparation Tests*, in which the subjects are introduced to a demonstration test conducted by the experiment designer to familiarize the subjects with the setup. In this test, the subjects observe the experiment designer following a reference command using the joystick [see Figure 4(a)]. They also watch the designer respond to control switching alert sounds by taking over control via the joystick. Following these demonstration tests, the subjects are requested to perform the experiment themselves. To complete this part, three preparation tests (each with a duration of 90 s) are conducted. At the end of each test, the root mean square error (RMSE) $e_{\text{rms}}$ of the subjects is calculated as

$$e_{\text{rms}} = \sqrt{\frac{1}{T_p}\int_{t_a}^{T_p} e(\tau)^2 d\tau}, \tag{26}$$

where

$$e(t) = M_{\text{cmd}}(t) - M(t) \tag{27}$$

and $T_p = 90$ s. It is expected that the $e_{\text{rms}}$ in each trial decreases as a sign of learning.

The third is the *Performance Tests* (each with a duration of 180 s), which aim at testing the performance of the proposed SCA in terms of tracking error $e_{\text{rms}}$, $\text{CfM}^{R_m}$, and a bumpless transfer metric $\rho$ (which is calculated using the difference between $e_{\text{rms}}$ values that are obtained using the 10-s intervals before and after the anomaly). This calculation is performed as

$$\rho = \sqrt{\frac{1}{t_a+10}\int_{t_a}^{t_a+10} e(\tau)^2 d\tau} - \sqrt{\frac{1}{t_a}\int_{t_a-10}^{t_a} e(\tau)^2 d\tau}. \tag{28}$$

## Details of the Human Subjects

The experiment with the harsh anomaly was conducted by 15 subjects (including one flight pilot), whereas the one with the mild anomaly was conducted by three subjects (including one flight pilot). All subjects were more than 18 years old, and four of the subjects were left-handed. However, this did not bring about any problems since an ambidextrous joystick was utilized. Some statistical data pertaining to the subjects are given in Table 1.

## Pilot Model Parameters

The pilot model in (21)–(22) includes statistical parameters $\mu_p$ and $\sigma_p$, the mean and the standard deviation

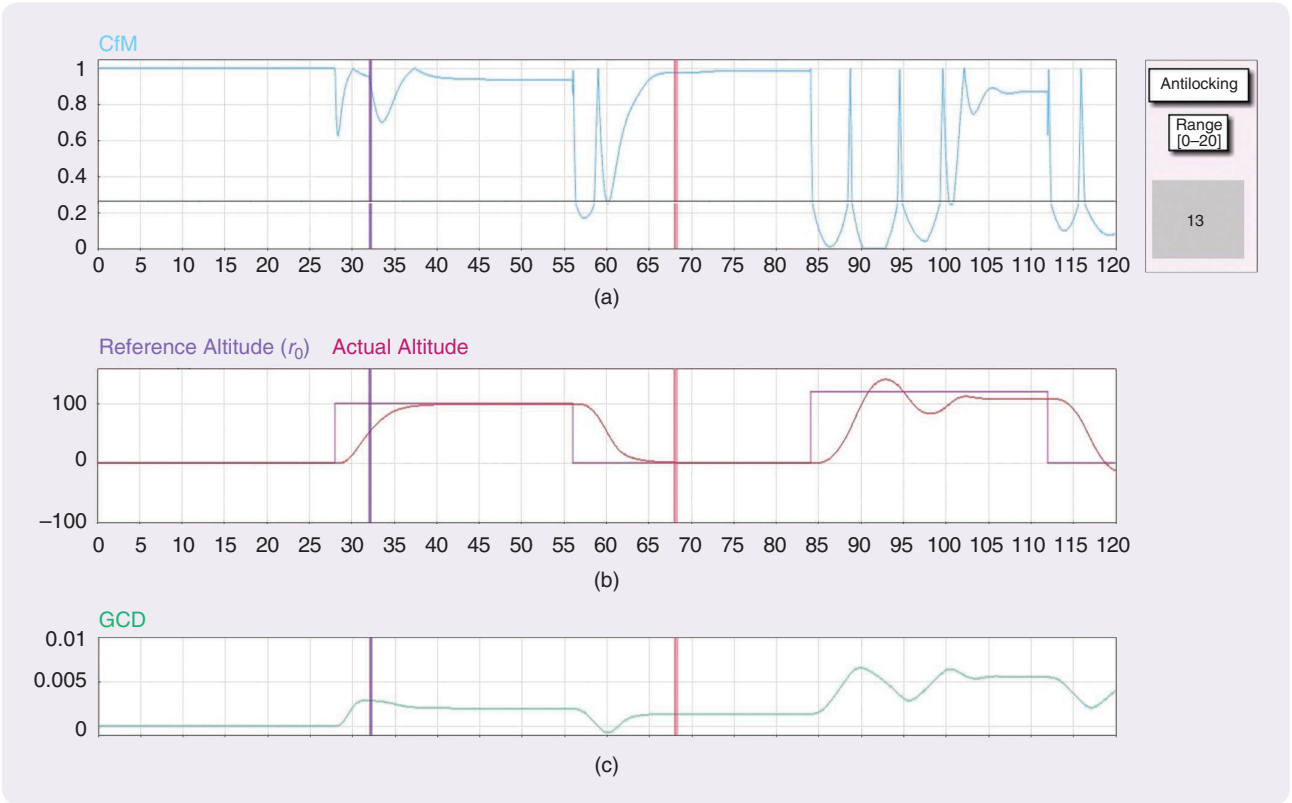



**FIGURE 6** The flight screen interface for shared control architecture 2. Part (a) shows the normalized $c_i(t)$ in (13), which can be considered as the instantaneous capacity for maneuver (CfM). The purple- and maroon-colored vertical lines at $t_{a_1} = 32$ s and $t_{a_2} = 68$ s, respectively, show the instants of anomaly introduction. The horizontal black line is the antilocking border, below which the $\mu$ input becomes effective. It is explained to the subjects as well as demonstrated during training that setting $\mu$ to high values where CfM variation is over this border has no influence on CfM [which is apparent from (5)]. Part (b) shows reference tracking, and (c) is the time variation of the graceful command degradation during flight. GCD: graceful command degradation.

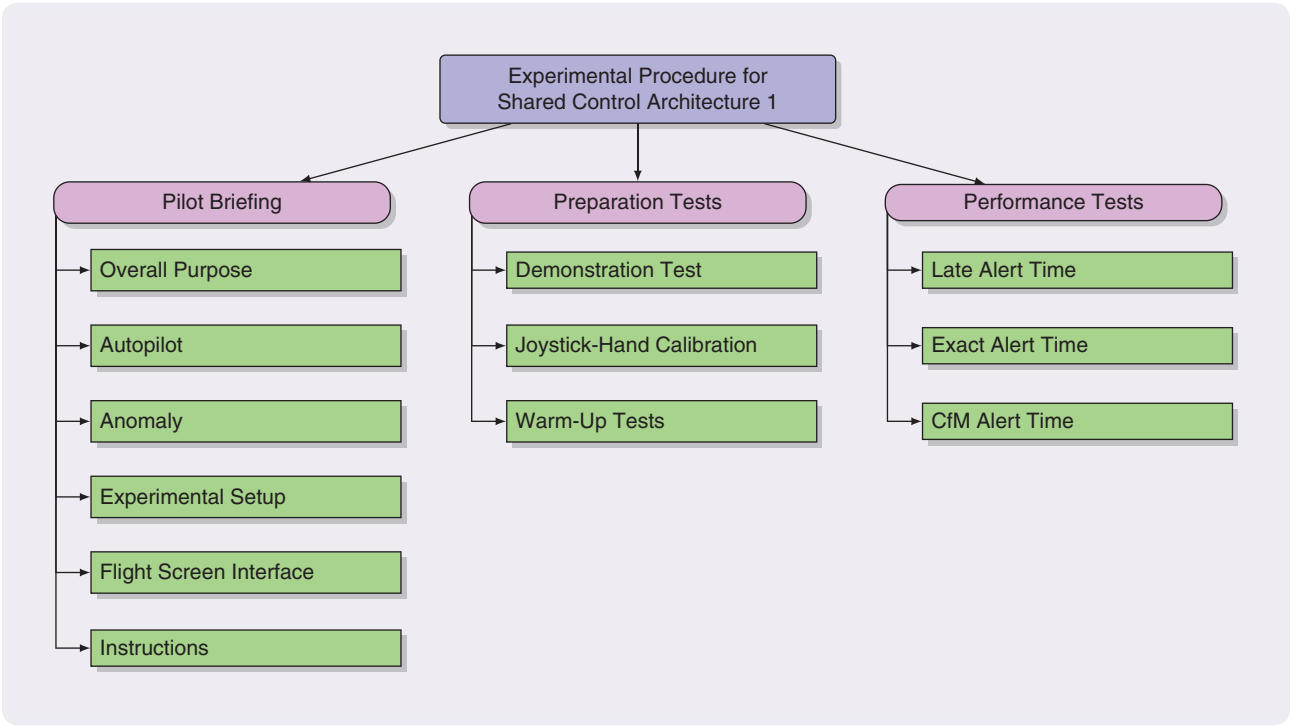



**FIGURE 7** The experimental procedure breakdown for shared control architecture 1. Three main tasks constituting the procedure are shown. In *Pilot Briefing*, the subjects read the pilot briefing, review it with the experiment designer, and have a question and answer session. In *Preparation Tests*, the subjects become familiar with the test via demonstration runs and warm-up tests. Finally, in *Performance Tests*, the real tests are conducted using three different alert times. CfM: capacity for maneuver.

of the time derivative of $\mathrm{CfM}^{R_m}$, respectively, and the parameters of the filter $G_1(s)$. The filter is chosen as $G_1(s) = (2.25/s^2 + 1.5s + 2.25)$ to reflect the bandwidth of the pilot stick motion. To obtain the other statistical parameters, several flight simulations were run with the PD control-based autopilot in closed loop, and the resulting $\mathrm{CfM}^{R_m}$ values were calculated for 180 s, both for the harsh and mild anomalies. The time-averaged statistics of the resulting profiles were used to calculate the statistical parameters as $\mu_p = 0.028, \sigma_p = 0.038$ for the harsh anomaly and $\mu_p = 0.091, \sigma_p = 0.077$ for the mild anomaly.

The pilot model in (21)–(22) implies that the pilot perceives the presence of the anomaly at the time instant when $K_t$ becomes unity, following which the control action switches from the autopilot to the human pilot. The action of the pilot, based on this trigger, is introduced in the experiment by choosing $t_s$ (the instant of the sound signal) to coincide with the perception trigger. The corresponding $\Delta T = t_s - t_a$, where $t_a$ is the instant when the anomaly is introduced, is denoted as CfM based. To benchmark this CfM-based switching action, two other switching mechanisms are introduced, one defined to be "exact" (where $t_s = t_a$, so that $\Delta T = 0$) and another to be "late" (where $\Delta T$ is chosen to be significantly larger than the CfM-based one). These choices are summarized in Table 2.

**TABLE 1** The statistical data of the subjects in the shared control architecture 1 experiment. $\mu()$ and $\sigma()$ represent the average and the standard deviation operators, respectively.

| Scenario | Number of Participants | Female | $\mu$(Age) | $\sigma$(Age) | Left-Handed |
|---|---|---|---|---|---|
| $S_{\text{harsh}}$ | 15 | 1 | 22.9 | 3.6 | 3 |
| $S_{\text{mild}}$ | 3 | 0 | 26.0 | 5.2 | 0 |

**TABLE 2** The timeline of anomalies. Based on the switching mechanism, the anomaly is reported to the subject with a sound signal at $t_s$. $\Delta T$, the alert time, is defined as the time elapsed between the sound signal and the occurrence of the anomaly. CfM: capacity for maneuver.

| Switch | $t_a(S)$ | $t_s(S)$ | $\Delta T(S)$ |
|---|---|---|---|
| Scenario with the harsh anomaly, $S_{\text{harsh}}$. | | | |
| Late | 50 | 55.5 | 5.5 |
| Exact | 50 | 50 | 0 |
| CfM based | 50 | 51.1 | 1.1 |
| Scenario with the mild anomaly, $S_{\text{mild}}$. | | | |
| Late | 64 | 74 | 10 |
| Exact | 64 | 64 | 0 |
| CfM based | 64 | 70.2 | 6.2 |

**TABLE 3** The averaged $e_{\text{rms}}$ and capacity for maneuver ($\text{CfM}^{R_M}$) values for $S_{\text{harsh}}$. "Sim" refers to simulation, and "Exp" refers to experiment. The autopilot shows the worst tracking error performance. The reason for a high $\text{CfM}^{R_M}$ amount for the autopilot is the inability to effectively use the actuators to accommodate the anomaly.

| $S_{\text{harsh}}$ | Auto | $\text{Sim}_{\text{late}}$ | $\text{Exp}_{\text{late}}$ | $\text{Sim}_{\text{exact}}$ | $\text{Exp}_{\text{exact}}$ | $\text{Sim}_{\text{CfM-based}}$ | $\text{Exp}_{\text{CfM-based}}$ |
|---|---|---|---|---|---|---|---|
| $e_{\text{rms}}$ | 478 | 363 | 383 | 319 | 354 | 318 | 348 |
| $\text{CfM}^{R_m}$ | 8.92 | 7.93 | 6.84 | 7.83 | 6.80 | 7.83 | 7.06 |

**TABLE 4** Mean, $\mu$, and standard error, $\sigma_M = \sigma/\sqrt{n}$, where $\sigma$ is standard deviation and $n$ is the subject size of $e_{\text{rms}}$ (on the left) and capacity for maneuver ($\text{CfM}^{R_M}$) (on the right) for a harsh anomaly.

| Experiment | $\mu$ | $\sigma_M$ | Experiment | $\mu$ | $\sigma_M$ |
|---|---|---|---|---|---|
| $\text{Exp}_{\text{late}}$ | 383 | 16 | $\text{Exp}_{\text{late}}$ | 6.84 | 0.10 |
| $\text{Exp}_{\text{exact}}$ | 354 | 15 | $\text{Exp}_{\text{exact}}$ | 6.80 | 0.12 |
| $\text{Exp}_{\text{CfM-based}}$ | 348 | 14 | $\text{Exp}_{\text{CfM-based}}$ | 7.06 | 0.11 |

**TABLE 5** The averaged bumpless transfer metric $\rho$ and standard error $\sigma_M(\rho) = \sigma(\rho)/\sqrt{n}$, where $\sigma$ is the standard deviation of $\rho$, $n$ is the subject size, averaged reaction times is $t_{\text{RT}}$, and total reaction times is $t_{\text{TRT}}$ for $S_{\text{harsh}}$. The least amount of bumpless transfer of control is in the case of capacity for maneuver (CfM)-based shared control architecture.

| Switch | $\rho$ | $\rho_M(\rho)$ | $t_{\text{RT}}$(s) | $t_{\text{TRT}}$(s) |
|---|---|---|---|---|
| Late | 216 | 4.27 | 1.07 | 6.64 |
| Exact | 82 | 1.37 | 0.98 | 0.98 |
| CfM based | 26 | 0.72 | 0.99 | 2.12 |

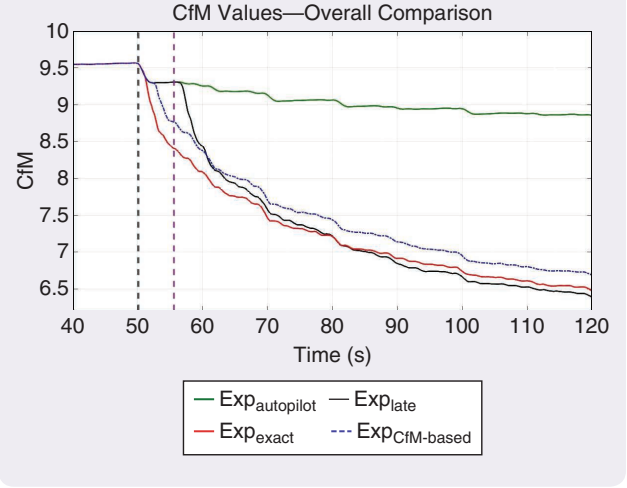



**FIGURE 8** The capacity for maneuver ($\text{CfM}^{R_m}$) variation for the autopilot and all of the alert timing mechanisms. $\text{CfM}^{R_m}$s of $\text{Exp}_{\text{late}}$ and $\text{Exp}_{\text{exact}}$ show changing trends during the simulation. However, $\text{CfM}^{R_m}$ of $\text{Exp}_{\text{CfM-based}}$ prominently stays at the top, especially after the transient effects of the anomaly ($t > t_a + 10$).

### *Results and Observations*

#### Scenario 1: Harsh Anomaly

The results related to the harsh anomaly defined in (23) are presented for various alert times transmitted to the subjects. To compare the results obtained from the subjects, numerical simulation results are completed, where the participants are replaced with the pilot model (21)–(22) using the same alert times. The results obtained are summarized in Table 3 using both the tracking error $e_{\text{rms}}$ and the corresponding $\text{CfM}^{R_m}$.

All numbers reported in Table 3 are averaged over all 15 subjects. $e_{\text{rms}}$ (scaled by $10^4$) was calculated using (26) with $T_p = 180$ s, while $\text{CfM}^{R_m}$ was calculated using (22) with $u_{\max} = 10$. The statistical variations of both $e_{\text{rms}}$ and $\text{CfM}^{R_m}$ over the 15 subjects are quantified for all three SCA experiments. The late, exact, and CfM-based alert times are summarized in Table 4. The average bumpless transfer metric $\rho$ and its standard error $\sigma_M(\rho)$ are also determined for these three cases in Table 5. Table 5 also provides the average reaction times $t_{\text{RT}}$ and average total reaction times $t_{\text{TRT}}$ of the subjects.

#### Observations

The first observation from Table 3 is that the $e_{\text{rms}}$ for the CfM-based case is at least 25% smaller than that for the case with the autopilot alone. The second observation is that, among the SCA experiments, the one with the CfM-based alert time has the smallest tracking error, while the one with the late alert time has the largest tracking error. However, it is noted that the difference between the exact and CfM-based cases is not significant.

The third observation from Table 3 is that CfM-based control switching from the autopilot to the pilot provides not only the smallest tracking error but also the largest $\text{CfM}^{R_m}$ value compared to other switching strategies. This can also be observed in Figure 8, where $\text{CfM}^{R_m}$ values (averaged over all subjects) are provided for different alert times. As noted earlier, the large $\text{CfM}^{R_m}$ value of the autopilot results from inefficient use of the actuators, which manifests itself with a large tracking error. The final observation (which comes from Table 5) is that among different alert times, the CfM-based alert time provides the smoothest transfer of control (with a bumpless transfer metric of $\rho = 26$), while the late alert provides $\rho = 216$ and the exact alert gives $\rho = 82$.

These observations imply the following. The pilot reengagement after an anomaly should not be delayed until it

**TABLE 6** The averaged $e_{rms}$ and capacity for maneuver ($CfM^{R_M}$) values for $S_{mild}$. The autopilot still shows the worst performance. However, the error introduced is less than that of the harsh anomaly.

| $S_{mild}$ | Auto | $Sim_{late}$ | $Exp_{late}$ | $Sim_{exact}$ | $Exp_{exact}$ | $Sim_{CfM\text{-}based}$ | $Exp_{CfM\text{-}based}$ |
|---|---|---|---|---|---|---|---|
| $e_{rms}$ | 408 | 281 | 259 | 297 | 236 | 243 | 217 |
| $CfM^{R_m}$ | 8.78 | 7.90 | 7.75 | 7.64 | 7.55 | 8.13 | 7.79 |

becomes too late, which corresponds to a late-alert switching strategy. At the same time, immediate pilot action right after the anomaly detection might not be necessary either, which is indicated by the fact that the bumpless transfer metric and the $CfM^{R_m}$ values for the CfM-based switching case is better than the exact-switching case. Instead, monitoring the $CfM^{R_m}$ information carefully may be the appropriate trigger for the pilot to assume control.

### Scenario 2: Mild Anomaly

The averaged $e_{rms}$ (scaled by $10^4$) and $CfM^{R_m}$ values for this case are given in Table 6. The averaged bumpless transfer metric $\rho$, its standard error $\sigma_M(\rho)$, the average reaction times $t_{RT}$, and the average total reaction times $t_{TRT}$ are presented in Table 7. The statistical variations of these metrics over the three subjects are depicted in Table 8.

### Observations

Similar to the harsh anomaly case, pure autopilot control results in the largest tracking error in the case of a mild anomaly. Also similar to the harsh anomaly case, pilot engagement based on $CfM^{R_m}$ information produces the smallest tracking error (although the difference between the exact switching and CfM-based switching is not significant). However, in terms of preserving $CfM^{R_m}$, Table 8 demonstrates that no significant differences between different switching times can be detected due to the wide spread (high standard error) of the results. The same conclusion can be drawn for the bumpless transfer metric. One reason for this could be the low sample size (three) in this experiment. One conclusion that can be drawn from these results is that although CfM-based switching shows smaller tracking errors (compared to the alternatives), the advantage of the proposed SCA in the presence of mild anomaly is not as prominent as in the case of harsh anomaly.

## *Validation of Shared Control Architecture 2*

Unlike SCA1 (where the pilot took over control from the autopilot when an anomaly occurred), in SCA2, the pilot plays more of an advisory role, directing the autopilot that remains operational throughout. Specifically, the pilot provides appropriate values $\mu$ and, sometimes, $\hat{\Lambda}_{f_p}$.

### Experimental Setup

Contrary to the SCA1 case, the subjects use the joystick only to enter the $\mu$ and $\hat{\Lambda}_{f_p}$ values using the joystick lever. The

**TABLE 7** The averaged bumpless transfer metric $\rho$; standard error $\sigma_M(\rho) = \sigma(\rho)/\sqrt{n}$, where $\sigma$ is standard deviation of $\rho$ and $n$ is the subject size; averaged reaction time $t_{RT}$; and total reaction time $t_{TRT}$ for the experiment with a mild anomaly $S_{mild}$. The nature of the anomaly has a considerable effect on the bumpless transfer metric, that is, the difference between the exact and capacity for maneuver (CfM)-based alert timings is not as readily noticeable as the one of the harsh anomaly.

| Switch | $\rho$ | $\rho_M(\rho)$ | $t_{RT}$(s) | $t_{TRT}$(s) |
|---|---|---|---|---|
| Late | 196 | 3.23 | 1.06 | 11.06 |
| Exact | 209 | 8.88 | 1.02 | 1.02 |
| CfM based | 203 | 2.87 | 0.95 | 7.20 |

**TABLE 8** The mean $\mu$, and standard error $\sigma_M = \sigma/\sqrt{n}$, where $\sigma$ is the standard deviation and $n$ is the subject size of $e_{rms}$ (on the left) and capacity for maneuver ($CfM^{R_M}$) (on the left) for scenario 2 with a mild anomaly.

| Experiment | $\mu$ | $\sigma_M$ | Experiment | $\mu$ | $\sigma_M$ |
|---|---|---|---|---|---|
| $Exp_{late}$ | 259 | 19 | $Exp_{late}$ | 7.75 | 0.15 |
| $Exp_{exact}$ | 236 | 21 | $Exp_{exact}$ | 7.55 | 0.17 |
| $Exp_{CfM\text{-}based}$ | 218 | 14 | $Exp_{CfM\text{-}based}$ | 7.61 | 0.19 |

flight screen that the subjects see is illustrated in Figure 6. There are three subplots, which are normalized $c_i(t)$ (a), reference command ($r_0$) tracking (b), and the evolution of GCD (c). The horizontal black line in the normalized $c_i(t)$ subplot corresponds to the upper bound of the virtual buffer $[0, \delta] = [0, 0.25]$. The small rectangle in the upper right shows the amount of the $\mu$ input entered via the joystick lever. In this rectangle, the title "Antilocking" is used to emphasize the purpose of the $\mu$ input, which is preventing the saturation/locking of the actuators. There is also another region in this rectangle called the "Range," which shows the limits of this input. The snapshot of the pilot screen in Figure 6 displays a scenario with two anomalies introduced at $t_{a_1} = 32$ s and $t_{a_2} = 68$ s. The instants of anomaly occurrences are marked with vertical lines, the colors of which indicate the severity of the anomaly. The subjects are trained to understand and respond to the severity and the effect of the anomalies by monitoring the colors, CfM information, and tracking performance.

The details of subject training are provided in the "Experimental Procedure" section.

### Anomaly

The anomaly considered in this experiment is loss of actuator effectiveness, indicated by $\Lambda_f$ in (1). $\Lambda_f$ is a $(2 \times 2)$ diagonal matrix with equal entries that are between zero and one, where zero corresponds to complete actuator failure and one corresponds to no failure. In the experiments, two anomalies are introduced at times $t = t_{a_1}$ and $t = t_{a_2}$. Consequently, the diagonal entries of $\Lambda_f$ vary as

$$\lambda_{fi} = \begin{cases} 1, & t < t_{a_1} \\ 0 < \lambda_{f_1} < 1, & t_{a_1} \le t < t_{a_2}, \\ 0 < \lambda_{f_2} < 1, & t_{a_2} \le t \end{cases} \tag{29}$$

where $\lambda_{fi}$ refers to the diagonal entries for the $i$th anomaly introduction. The anomaly injections are communicated to the subjects with colored vertical lines appearing on the pilot screen (as illustrated in Figure 6), together with sound alerts. Three anomalies with different severities are used during the experiments. The anomaly severity quantities are $\Lambda_f = 0.3$ (low), 0.2 (medium), and 0.15 (high), which are communicated through green-, purple-, and maroon-colored vertical lines, respectively (see Figure 6, where a sample experiment is shown with medium- and high-severity anomalies).

### Experimental Procedure

As in SCA1, the experimental procedure consists of three parts: *Pilot Briefing*, *Input Training*, and *Performance Test* (see Figure 9 for a schematic). The first part of the procedure is the *Pilot Briefing*, in which the subjects are required to read a pilot briefing to acquire knowledge about the experiment. The briefing consists of four main sections: *Overall Purpose*, *Flight Scenario*, *Flight Screen Interface*, and *Instructions*. In these sections, the main concepts and the experimental setup are covered by the experiment designer. The second part is the *Input Training*, in which the subjects are first introduced to the joystick lever, which they are required to use to enter the $\mu$ input [see Figure 4(b)]. By properly moving the lever, an integer value of $\mu$ (ranging from one to 20) can be given to the controller. Following this, a demonstration test is conducted by the experiment designer. In this test, a sample scenario with two anomalies is run, where the input $\mu$ is fixed to its nominal value of $\mu = 1$ throughout the flight. It is explicitly shown that 1) the actuators reach their saturation limit (and, thus, the CfM becomes zero) many times during the flight, which jeopardizes the aircraft stability; 2) it accepts the altitude $h$ a long time to recover to follow the reference command; and 3) a certain GCD occurs to relax the performance goals.

Following the demonstration test, another sample scenario is run by the designer, in which suitable $\mu$ values are provided upon occurrence of the anomalies. Unlike the previous demonstration, it is noted that 1) with a proper $\mu$, CfM can be kept away from zero, and 2) GCD is kept minimal. By this demonstration, the subjects are expected to appreciate that suitable $\mu$ inputs help the autopilot recover from severe anomalies in an efficient and swift manner.

Finally, a $\mu$-input training is performed on the subjects. Six scenarios are introduced to the participants in an

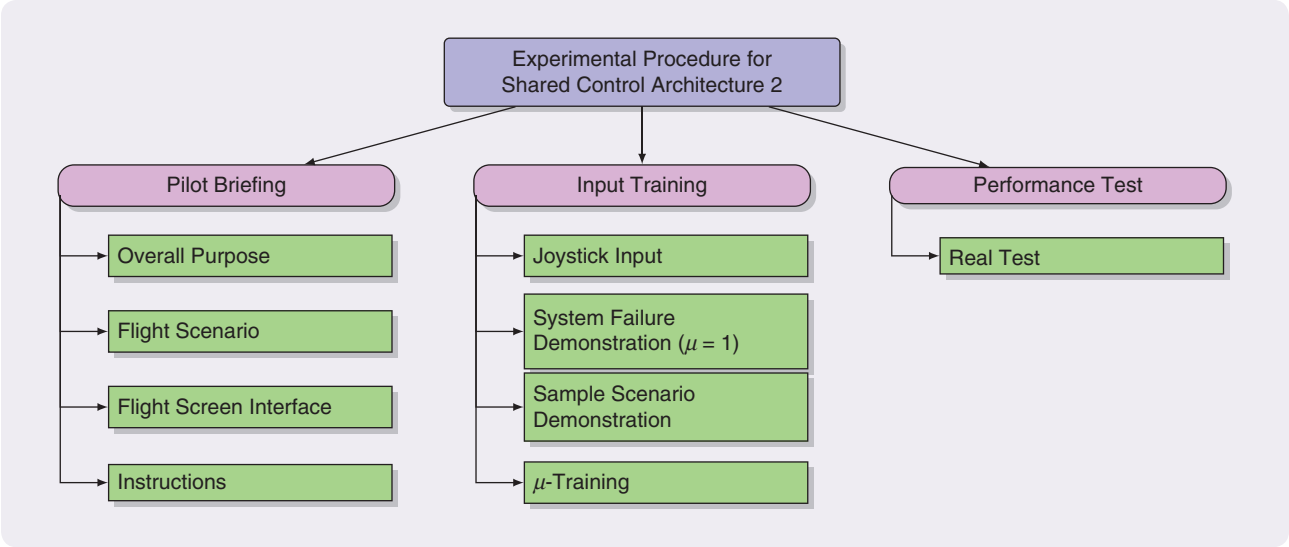



**FIGURE 9** The experimental procedure breakdown for shared control architecture 2. Three main tasks constituting the procedure are observed. In *Pilot Briefing*, the subjects read the pilot briefing, review it with the experiment designer, and have a question and answer session. In *Input Training*, the subjects learn how to provide the auxiliary inputs to the autopilot using the joystick. They also undergo $\mu$-input training, the details of which are covered in the "Experimental Procedure" section. Finally, in *Performance Test*, a real test with two successive anomalies $\Lambda_{fi}$ is conducted.

interactive manner, by which they learn how to be involved in the overall control architecture.

As a first step, three scenarios with single anomalies are considered. In each severity of the anomaly, an optimal $\mu$ (which trades CfM for GCD) is conveyed to the subjects. By doing this, attention is drawn to the fact that each anomaly causes a certain sharp drop in the CfM variation. In other words, more drastic drops occur in CfM with the increase in the severity of anomaly. Following this, three other scenarios with two successive anomalies are studied. It is noted by the combination of different anomalies that nonlinear effects are present in the flight simulation; that is, the $\mu$ values corresponding to single anomalies are no longer effective when the anomalies are combined. In both the training and the performance tests, the anomaly severity estimation $\hat{\Lambda}_{f_p}$ input is automatically fed to the controller with an error of 0.2 in absolute value. Then, to compare the effect of this estimation input, two pilot types are compared (where the estimation is not provided in one case and provided in the other).

The third part is the *Performance Test*, in which the subjects are expected to handle a combination of highly severe anomalies ($\Lambda_{f_1} = 0.20$ and $\Lambda_{f_2} = 0.15$) in a flight simulation. They are expected to give suitable $\mu$ values on the basis of the training they obtain. In both the training and the performance tests, the introduction of anomaly times are chosen to be completely random to prevent the subjects from making a guess as to whether or not the anomaly is about to happen. The set of tasks to be tackled by the subjects is summarized as a flowchart in Figure 10. These tasks are explained interactively during the sample scenario demonstration (see Figure 9), where the subjects have the opportunity to practice the steps to control the flight simulation.

### Details of the Human Subjects

The experiment was performed by 10 subjects, all of whom attended the previous experiment. This was done deliberately since the subjects of the previous experiment were acquainted with the autopilot and shared control concepts. For this reason, the pilot briefing regarding this experiment was written in a tone in which the subjects were already familiar with these basic notions. The average age of the participants was 24.2, with a standard deviation of 1.8.

### Results and Observations

The autopilots consist of an adaptive controller as in [13] (which has no apparent interface with the autopilot), a $\mu$-mod adaptive controller with a fixed value of $\mu = 50$, and an optimal controller. The numerical results averaged over 10 subjects are presented in Table 9. The indexes in curly brackets, $\{h, v\}$, denote the altitude and velocity, respectively. SAP and SUP refer to the "situation-aware pilot" and the "situation-unaware pilot," respectively. SUP provides only a $\mu$ input, and SAP provides both $\mu$ and $\hat{\Lambda}_{f_p}$ to the autopilot. The results for the SUP are obtained by simulating the performance tests using only the $\mu$ inputs provided by the subjects, without their severity estimation inputs $\hat{\Lambda}_{f_p}$. Therefore, (16) is used to obtain the SUP results instead of (17) while still incorporating the same $\mu$ values that were entered by the participants. Furthermore, the tracking performance $\gamma_i$ is calculated as

$$\gamma_i = \mathrm{RMSE}_i^+ - \mathrm{RMSE}_i^-, \tag{30}$$

where

$$\mathrm{RMSE}_i^- = \mathrm{rms}(e_i)\Big|_0^{t_{a1}}, \quad \mathrm{RMSE}_i^+ = \mathrm{rms}(e_i)\Big|_{t_{a1}}^{T}. \tag{31}$$

In Figure 11, a comprehensive comparison of SCA2 with autopilot-only cases is given as a matrix of 4 × 4 plots. Each column presents the results of a specific controller, whereas each row shows the comparison of these controllers based on altitude tracking $h$, velocity tracking $V$, CfM, and elevator control input $\delta_{\mathrm{el}}$, respectively. The horizontal dashed red line in the third row shows the buffer limit. The horizontal dashed red and green lines in the fourth row show the limitations posed by $u_{\max}$ and

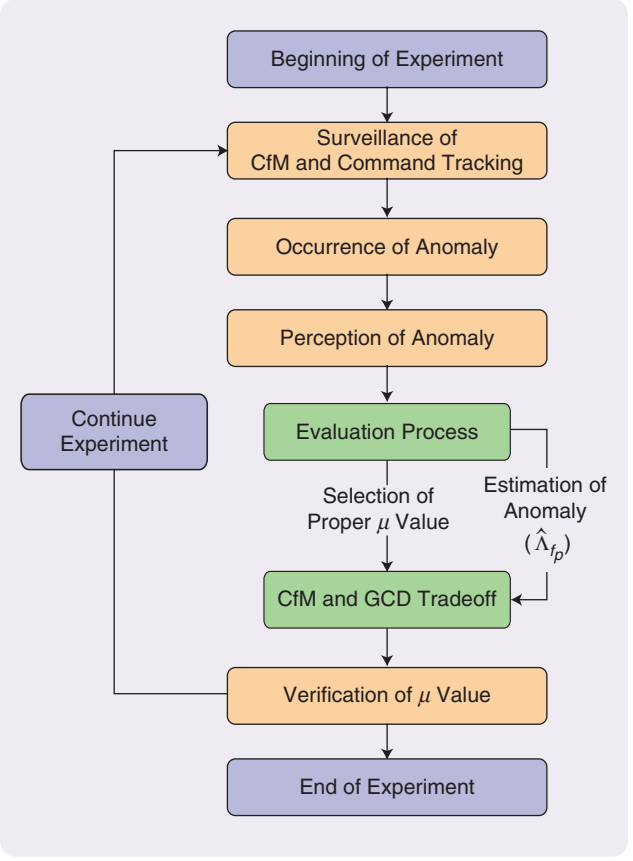



**FIGURE 10** An algorithm of pilot tasks. This presents the step-by-step procedure that should be followed by the subject. GCD: graceful command degradation; CfM: capacity for maneuver.

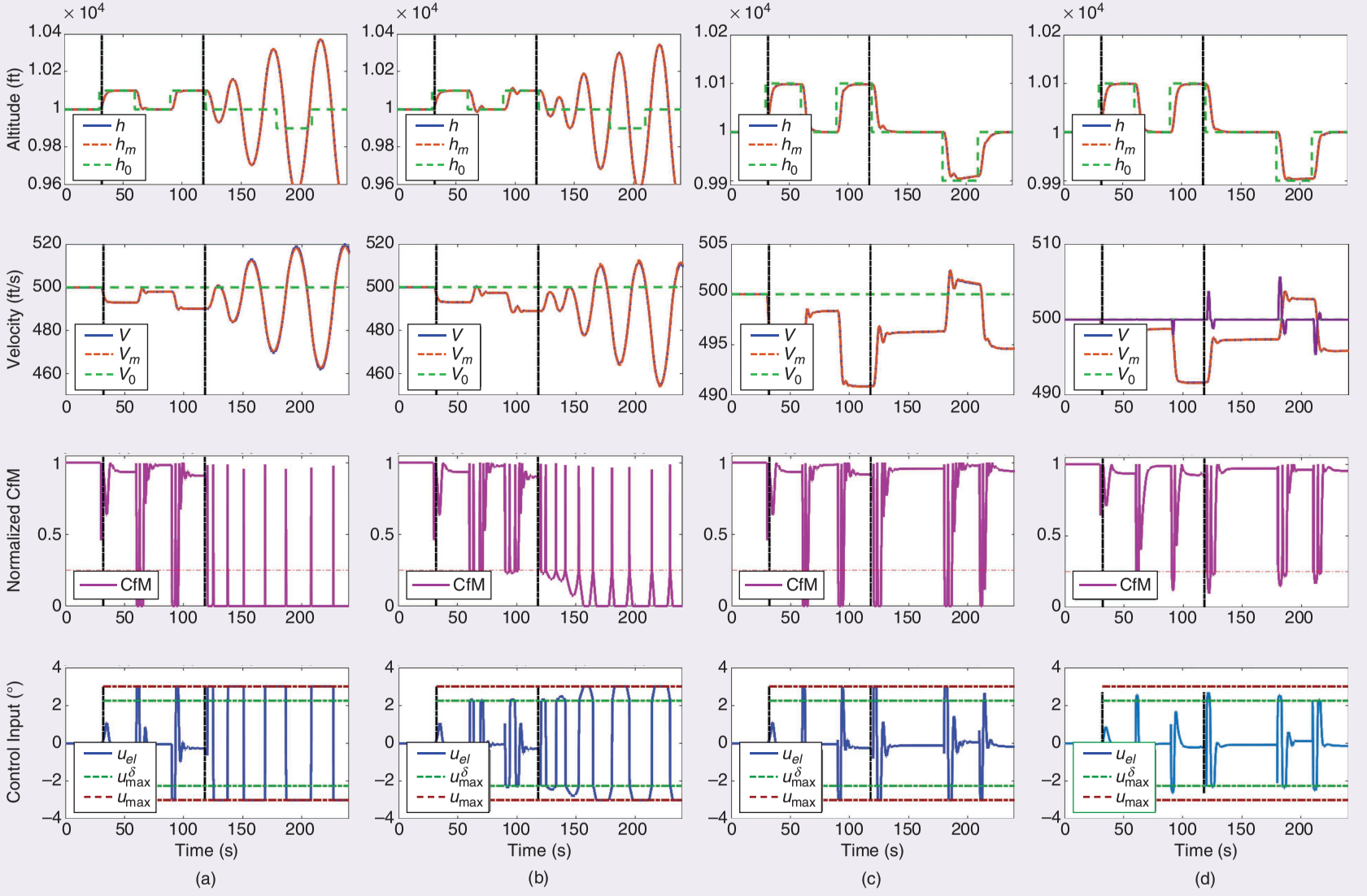



**FIGURE 11** The shared control architecture 2 versus autopilot-only cases. (a) Optimal autopilot. (b) μ-mod autopilot. (c) Adaptive autopilot. (d) Situation-aware pilot (SAP). Each column presents the results of a specific controller, whereas each row shows the comparison of these controllers based on altitude tracking $h$, velocity tracking $V$, capacity for maneuver, and elevator control input $\delta_{el}$, respectively. The horizontal dashed red line in the third row shows the buffer limit ($\delta = 0.25$). The horizontal dashed red and green lines in the fourth row show the limitations posed by $u_{max_{el}} = 3$ and $u^{\delta}_{max_{el}} = 2.25$. CfM: capacity for maneuver.

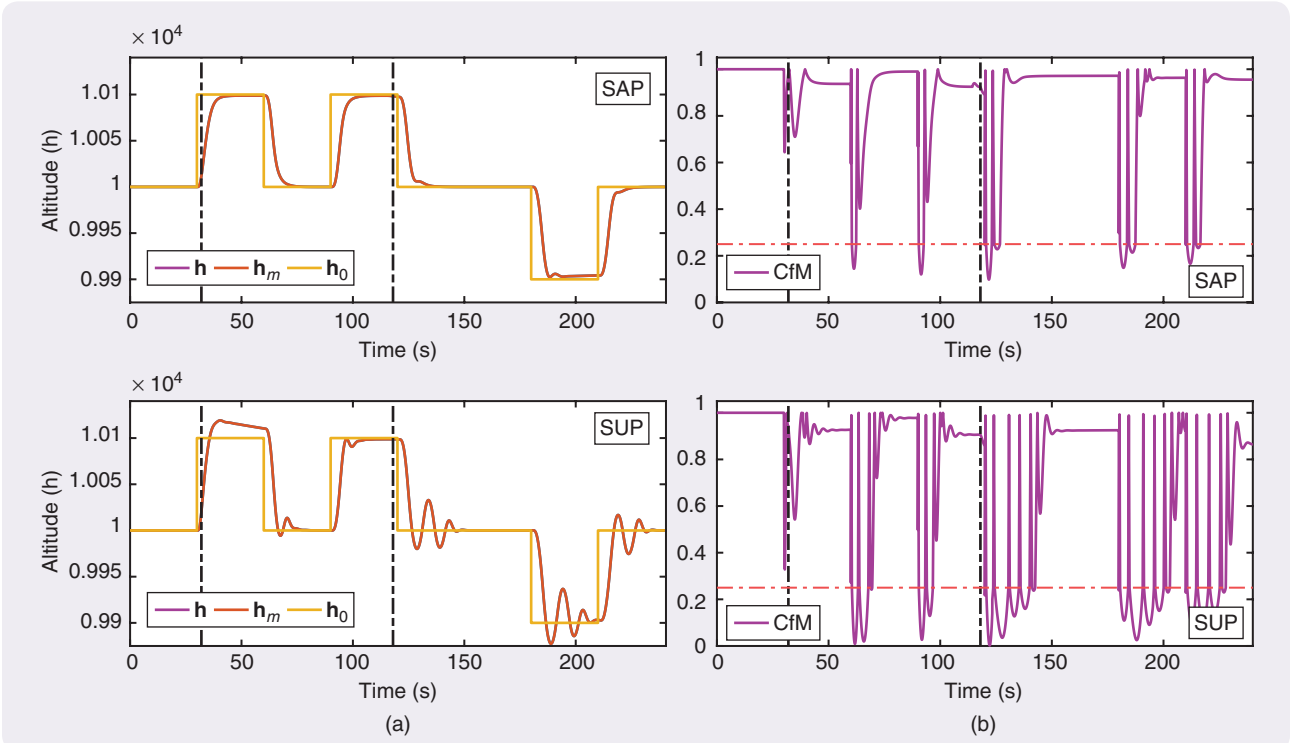



**FIGURE 12** A comparison of situation-aware pilot (SAP) and situation-unaware pilot (SUP) for the same scenario. The pure difference between these two pilots is based on the additional anomaly estimation input $\hat{\Lambda}_{f_p}$. It greatly attenuates the oscillatory tracking behavior around the reference model and contributes to a fast recovery from the buffer region. (a) Tracking performance for pilots. (b) Capacity for maneuver (CfM) graphs for pilots.

$u_{\max}^{\delta}$ introduced in (7). The results given in Table 9 and Figure 11 can be summarized as follows. First, it is observed that SCA2 (whether with SAP or SUP) outperforms all of the other autopilot-only cases. SCA2 provides a higher tracking performance $\gamma$, a higher CfM, and a lower GCD. Second, SAP shows a better performance than other controllers (including SUP) throughout the simulation run by not only showing a higher tracking performance but also preventing CfM from reaching zero (saturation point). Third, a quick inspection of the first row of Figure 11 shows that both optimal and $\mu$-mod autopilots fail to respond to the second anomaly. This can be explained by the fact that CfM reaches zero (saturation point) many times, especially in the case of the optimal controller. Fourth, the adaptive autopilot shows an acceptable performance up to the second anomaly, but it demonstrates degraded behavior with repeated saturation. It is noted that the elevator in the case of adaptive autopilot also hits the saturation point many times and shows a more oscillatory response than SAP.

To emphasize the effect of anomaly estimation input $\hat{\Lambda}_{f_p}$, the performances of SCA2 with SAP and SUP are presented separately in Figure 12. It is seen that SAP performs better than SUP in terms of both tracking and CfM metrics. The selection of a suitable $\mu$ and the introduction of anomaly estimation (even with an error of 0.2) contribute dramatically to resilient performance, which can also be numerically verified by performance metrics given in Table 9.

**TABLE 9** **The shared control architecture 2 (SCA2) versus autopilot-only cases. SCA, both with a situation-aware pilot (SAP) and a situation-unaware pilot (SUP), results in higher tracking performance $\gamma$, higher capacity for maneuver (CfM) values, and smaller graceful command degradation (GCD), which is based on altitude $h$. Alternately, SAP performs better than SUP in all metrics. NA: not applicable.**

| Method | $\text{RMSE}_{\{h,v\}}$ | $\gamma_{\{h,v\}}$ | CfM | GCD |
|---|---|---|---|---|
| SAP | $\{60, 23\}\times 10^{-4}$ | $\{36, 135\}\times 10^{-4}$ | 1.21 | $24.8\times 10^{-4}$ |
| SUP | $\{60, 23\}\times 10^{-4}$ | $\{51, 152\}\times 10^{-4}$ | 1.16 | $24.8\times 10^{-4}$ |
| Adaptive | $\{24.1, 0.4\}$ | $\{0.51, 1.06\}$ | 0.92 | NA |
| $\mu$-mod | $\{60, 23\}\times 10^{-4}$ | $\{0.10, 0.27\}$ | 0.81 | $342\times 10^{-4}$ |
| Optimal | $\{24.1, 0.4\}$ | $\{160.8, 13.8\}$ | 0.84 | NA |

## SUMMARY AND FUTURE WORK

Although increased automation has made it easier to control aircrafts, ensuring a safe interaction between the pilots and the autopilots is still a challenging problem, especially in the

presence of severe anomalies. The domain of decision making is common to both human experts and feedback control systems. However, the process of detecting the anomaly and mitigating it, the speed of response, intrinsic latencies, and the overall decision-making procedure may vary between the pilot and the autopilot. Current approaches consist of autopilot solutions that disengage themselves in the face of an anomaly and the pilot assumes control. This may cause reengagement of the pilot at the worst possible time, which can result in undesired consequences and a bumpy transfer. This article proposes an SCA where the decision making of the human pilot is judiciously coordinated with that of an advanced autopilot and is highly attractive in flight control problems where severe anomalies are present. Two distinct architectures through which such coordination can take place are presented, which are validated through human-in-the-loop simulations. Though the type of coordination varies between these two architectures, both employ a common principle from cognitive engineering, namely, CfM. While these architectures were proposed in recent publications [11]–[13], a common framework to view them as well as their validation using human subject data are the main contributions of this article.

As automation increases in engineered systems, the creation of new cyberphysical and human systems is inevitable. There will be a variety of scenarios where humans and machines will need to interact and engage in combined decision making to lead to resilient autonomous behavior. These interactions will be complex and distinct and require new tools and methodologies. Deeper engagement with the social science community for better insight into human decision making and advanced modeling approaches is necessary. The results reported here should be viewed as the first of several steps in this research direction.

## ACKNOWLEDGMENTS

The first two authors would like to express their gratitude for the support of The Scientific and Technological Research Council of Turkey under grant 118E202. The third author would like to gratefully acknowledge the support of Boeing through the Boeing Strategic University Initiative.

## AUTHOR INFORMATION

***Emre Eraslan*** (emre.eraslan@bilkent.edu.tr) is a graduate student in the Department of Mechanical Engineering at Bilkent University, where he is a researcher in the Systems Laboratory. He received B.S. degrees in mechanical engineering and physics and a minor degree in German from the Middle East Technical University, Ankara, Turkey, in 2018. His research interests include adaptive control and human–machine interactions.

***Yildiray Yildiz*** is an assistant professor and director of the Systems Laboratory at Bilkent University, Ankara, Turkey. He received the B.S. degree (valedictorian) in mechanical engineering from the Middle East Technical University, Ankara, Turkey, in 2002; the M.S. degree in mechatronics engineering from Sabanci University, Istanbul, in 2004; and the Ph.D. degree in mechanical engineering with a mathematics minor from the Massachusetts Institute of Technology, Cambridge, in 2009. He held postdoctoral associate and associate scientist positions with NASA Ames Research Center, Mountain View, California, from 2009 to 2010 and from 2010 to 2014, respectively. During this time, he was employed by the University of California, Santa Cruz, through its University Affiliated Research Center. He is the recipient of the American Society of Mechanical Engineers Best Student Paper in Conference Award (2008); the NASA Group Achievement Award (2012) for outstanding technology development; the Turkish Science Academy's Young Scientist Award (2017); the Science Academy, Turkey's Young Scientist Award (2017); and Prof. Mustafa N. Parlar Foundation's Research Incentive Award (2018). He has been an associate editor for the IEEE Conference Editorial Board since 2015, *IEEE Control Systems Magazine* since 2016, and *European Journal of Control* since 2019. He has also been a member of the *International Journal of Control* Editorial Board since 2019. His research interests are the theory and applications of system dynamics and control. He is especially interested in human–machine interactions, adaptive control, reinforcement learning and game theory, and applications in the aerospace and automotive domains.

***Anuradha M. Annaswamy*** received the Ph.D. degree in electrical engineering from Yale University in 1985. She has been a member of the faculty at Yale, Boston University, and the Massachusetts Institute of Technology, where she is currently the director of the Active-Adaptive Control Laboratory and a senior research scientist in the Department of Mechanical Engineering. Her research interests pertain to adaptive control theory and applications to aerospace, automotive, and propulsion systems, cyberphysical systems science, and cyberphysical systems applications to smart grids, smart cities, and smart infrastructure. She is the author of 100 journal publications and numerous conference publications, coauthor of a graduate textbook on adaptive control (2004), and coeditor of several reports, including *Systems & Control for the Future of Humanity, Research Agenda: Current and Future Roles, Impact, and Grand Challenges*, *IEEE Vision for Smart Grid Control: 2030 and Beyond*, and *Impact of Control Technology*. She has received several awards, including the George Axelby and IEEE Control Systems Magazine Best Paper Awards from the IEEE Control Systems Society (CSS), the Presidential Young Investigator Award from the National Science Foundation, the Hans Fisher Senior Fellowship from the Institute for Advanced Study at the Technische University Munich, the Donald Groen Julius Prize from the Institute of Mechanical Engineers, a Distinguished Member Award, and a Distinguished Lecturer Award from IEEE CSS. She is a Fellow of IEEE and the International

Federation of Automatic Control. She is the president of the IEEE CSS. She is the deputy editor of *Annual Reviews in Control* (2016–present).

## REFERENCES


[1] R. A. Hess, "Modeling pilot control behavior with sudden changes in vehicle dynamics," *J. Aircr.*, vol. 46, no. 5, pp. 1584–1592, 2009. doi: 10.2514/1.41215.

[2] R. A. Hess, "A model for pilot control behavior in analyzing potential loss-of-control events," *Proc. Inst. Mech. Eng., G, J. Aerosp. Eng.*, vol. 228, no. 10, pp. 1845–1856, 2014. doi: 10.1177/0954410014531218.

[3] R. A. Hess, "Modeling human pilot adaptation to flight control anomalies and changing task demands," *J. Guid., Control, Dyn.*, vol. 39, no. 3, pp. 655–666, 2016. doi: 10.2514/1.G001303.

[4] D. D. Woods and E. Hollnagel, *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*. Boca Raton, FL: CRC Press, 2006.

[5] D. D. Woods and M. Branlat, "Basic patterns in how adaptive systems fail," in *Resilience Engineering in Practice*. C. P. Nemeth and E. Hollnagel, Eds. Farnham, U.K.: Ashgate., 2014, pp. 127–144.

[6] "In-flight icing encounter and loss of control, Simmons Airlines, d.b.a. American Eagle flight 4184 Avions De Transport Regional (ATR) model 72-212, N401am Roselawn, Indiana October 31, 1994, NTSB aircraft accident report," National Transportation Safety Board, Washington, D.C., NTSB/AAR-96/01, 1996.

[7] D. D. Woods, "The theory of graceful extensibility: Basic rules that govern adaptive systems," *Environ. Syst. Decis.*, vol. 38, no. 4, pp. 433–457, 2018. doi: 10.1007/s10669-018-9708-3.

[8] D. A. Abbink et al., "A topology of shared control systems: Finding common ground in diversity," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 5, pp. 509–525, 2018. doi: 10.1109/THMS.2018.2791570.

[9] T. B. Sheridan, *Monitoring Behavior and Supervisory Control*, vol. 1. New York: Springer-Verlag, 2013.

[10] M. Mulder, D. A. Abbink, and E. R. Boer, "Sharing control with haptics: Seamless driver support from manual to automatic control," *Hum. Factors*, vol. 54, no. 5, pp. 786–798, 2012. doi: 10.1177/0018720812443984.

[11] A. Farjadian, A. Annaswamy, and D. Woods, "Towards a resilient control architecture: A demonstration of Bumpless Re-engagement following an anomaly in flight control," in *Proc. Int. Symp. Sustainable Systems and Technologies*, vol. 4, 2016. [Online]. Available: http://aaclab.mit.edu/resources/Towards_A_Resilient_Control_Architecture.pdf

[12] B. T. Thomsen, A. M. Annaswamy, and E. Lavretsky, "Shared control between adaptive autopilots and human operators for anomaly mitigation," *IFAC-PapersOnLine*, vol. 51, no. 34, pp. 353–358, 2019. doi: 10.1016/j.ifacol.2019.01.018.

[13] A. B. Farjadian, B. Thomsen, A. M. Annaswamy, and D. D. Woods, "Resilient flight control: An architecture for human supervision of automation," *IEEE Trans. Control Syst. Technol.*, early access, 2020. doi: 10.1109/TCST.2019.2959542.

[14] J. Smisek, E. Sunil, M. M. van Paassen, D. A. Abbink, and M. Mulder, "Neuromuscular-system-based tuning of a haptic shared control interface for UAV teleoperation," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 4, pp. 449–461, 2016. doi: 10.1109/THMS.2016.2616280.

[15] D. W. van der Wiel, M. M. van Paassen, M. Mulder, M. Mulder, and D. A. Abbink, "Driver adaptation to driving speed and road width: Exploring parameters for designing adaptive haptic shared control," in *Proc. 2015 IEEE Int. Conf. Systems, Man, and Cybernetics*, pp. 3060–3065. doi: 10.1109/SMC.2015.532.

[16] J. Smisek, W. Mugge, J. B. Smeets, M. Van Paassen, and A. Schiele, "Adapting haptic guidance authority based on user grip," in *Proc. 2014 IEEE Int. Conf. Systems, Man, and Cybernetics (SMC)*, pp. 1516–1521. doi: 10.1109/SMC.2014.6974131.

[17] E. J. Rossetter and J. C. Gerdes, "Lyapunov based performance guarantees for the potential field lane-keeping assistance system," *Trans. ASME, J. Dyn. Syst. Meas. Control*, vol. 128, no. 3, pp. 510–522, 2006. doi: 10.1115/1.2192835.

[18] K. Ackerman et al., "Flight envelope information-augmented display for enhanced pilot situational awareness," in *Proc. AIAA Infotech Aerospace*, 2015, p. 1112. doi: 10.2514/6.2015-1112.

[19] N. Tekles, F. Holzapfel, E. Xargay, R. Choe, N. Hovakimyan, and I. M. Gregory, "Flight envelope protection for NASA's transport class model," in *Proc. AIAA Guidance, Navigation, and Control Conf.*, 2014, p. 0269. doi: 10.2514/6.2014-0269.

[20] J. Chongvisal, N. Tekles, E. Xargay, D. Talleur, A. Kirlik, and N. Hovakimyan, "Loss-of-control prediction and prevention for NASA's transport class model," in *Proc. AIAA Guidance, Navigation, and Control Conf.*, 2014, p. 0784. doi: 10.2514/6.2014-0784.

[21] D. B. Kaber, "Adaptive automation," in *The Oxford Handbook of Cognitive Engineering*, J. D. Lee and A. Kirlik, Eds. Oxford: Oxford Univ. Press, 2013, pp. 5–12.

[22] E. Lavretsky and K. A. Wise, *Robust and Adaptive Control*. London: Springer-Verlag, 2013.

[23] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*. Chelmsford, MA: Courier Corp., 2005.

[24] S. P. Karason and A. M. Annaswamy, "Adaptive control in the presence of input constraints," in *Proc. 1993 American Control Conf.*, pp. 1370–1374. doi: 10.23919/ACC.1993.4793095.

[25] E. Lavretsky and N. Hovakimyan, "Stable adaptation in the presence of input constraints," *Syst. Control Lett.*, vol. 56, nos. 11–12, pp. 722–729, 2007. doi: 10.1016/j.sysconle.2007.05.002.

[26] T.-G. Lee and U.-Y. Huh, "An error feedback model based adaptive controller for nonlinear systems," in *Proc. IEEE Int. Symp. Industrial Electronics (ISIE'97)*, 1997, pp. 1095–1100. doi: 10.1109/ISIE.1997.648892.

[27] T. E. Gibson, A. M. Annaswamy, and E. Lavretsky, "On adaptive control with closed-loop reference models: Transients, oscillations, and peaking," *IEEE Access*, vol. 1, pp. 703–717, Sept. 2013. doi: 10.1109/ACCESS.2013.2284005.

[28] A. E. Bryson, "Optimal control—1950 to 1985," *IEEE Control Syst. Mag.*, vol. 16, no. 3, pp. 26–33, 1996. doi: 10.1109/37.506395.

[29] D. T. McRuer and E. S. Krendel, "Mathematical models of human pilot behavior," Advisory Group for Aerospace Research and Development, Neuilly-sur-Seine, France, Tech. Rep. AGARDograph no. 188, 1974.

[30] M. Lone and A. Cooke, "Review of pilot models used in aircraft flight dynamics," *Aerosp. Sci. Technol.*, vol. 34, pp. 55–74, Apr. 2014. doi: 10.1016/j.ast.2014.02.003.

[31] S. Xu, W. Tan, A. V. Efremov, L. Sun, and X. Qu, "Review of control models for human pilot behavior," *Annu. Rev. Control*, vol. 44, pp. 274–291, Oct. 2017. doi: 10.1016/j.arcontrol.2017.09.009.

[32] D. T. McRuer and H. R. Jex, "A review of quasi-linear pilot models," *IEEE Trans. Human Factors Electron.*, vol. HFE-8, no. 3, pp. 231–249, 1967. doi: 10.1109/THFE.1967.234304.

[33] R. D. Wierenga, "An evaluation of a pilot model based on Kalman filtering and optimal control," *IEEE Trans. Man-Mach. Syst.*, vol. 10, no. 4, pp. 108–117, 1969. doi: 10.1109/TMMS.1969.299907.

[34] D. L. Kleinman, S. Baron, and W. Levison, "An optimal control model of human response Part I: Theory and validation," *Automatica*, vol. 6, no. 3, pp. 357–369, 1970. doi: 10.1016/0005-1098(70)90051-8.

[35] R. Hosman and H. Stassen, "Pilot's perception and control of aircraft motions," *IFAC Proc. Vol.*, vol. 31, no. 26, pp. 311–316, 1998. doi: 10.1016/S1474-6670(17)40111-X.

[36] R. A. Hess, "Unified theory for aircraft handling qualities and adverse aircraft-pilot coupling," *J. Guid., Control, Dyn.*, vol. 20, no. 6, pp. 1141–1148, 1997. doi: 10.2514/2.4169.

[37] D. D. Woods and N. B. Sarter, "Learning from automation surprises and going sour accidents," in *Cognitive Engineering in the Aviation Domain*, N. Sarter and R. Amalberti, Eds. Hillsdale, NJ: Erlbaum, 2000, pp. 327–353.

[38] A. B. Farjadian, A. M. Annaswamy, and D. Woods, "Bumpless reengagement using shared control between human pilot and adaptive autopilot," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 5343–5348, 2017. doi: 10.1016/j.ifacol.2017.08.925.

[39] S. Tohidi and Y. Yildiz, "Adaptive human pilot model for uncertain systems," in *Proc. 2019 18th European Control Conf. (ECC)*, pp. 2938–2943. doi: 10.23919/ECC.2019.8795847.

[40] B. L. Stevens, F. L. Lewis, and E. N. Johnson, *Aircraft Control and Simulation: Dynamics, Controls Design, and Autonomous Systems*. Hoboken, NJ: Wiley, 2015.

[41] L. T. Nguyen, M. E. Ogburn, W. P. Gilbert, K. S. Kibler, P. W. Brown, and P. L. Deal, "Simulator study of stall/post-stall characteristics of a fighter airplane with relaxed longitudinal static stability," Langley Research Center, NASA Tech. Paper 1538, Dec. 1979.

[42] R. A. Hess, "Simplified approach for modelling pilot pursuit control behaviour in multi-loop flight control tasks," *Proc. Inst. Mech. Eng., G, J. Aerosp. Eng.*, vol. 220, no. 2, pp. 85–102, 2006. doi: 10.1243/09544100JAERO33.