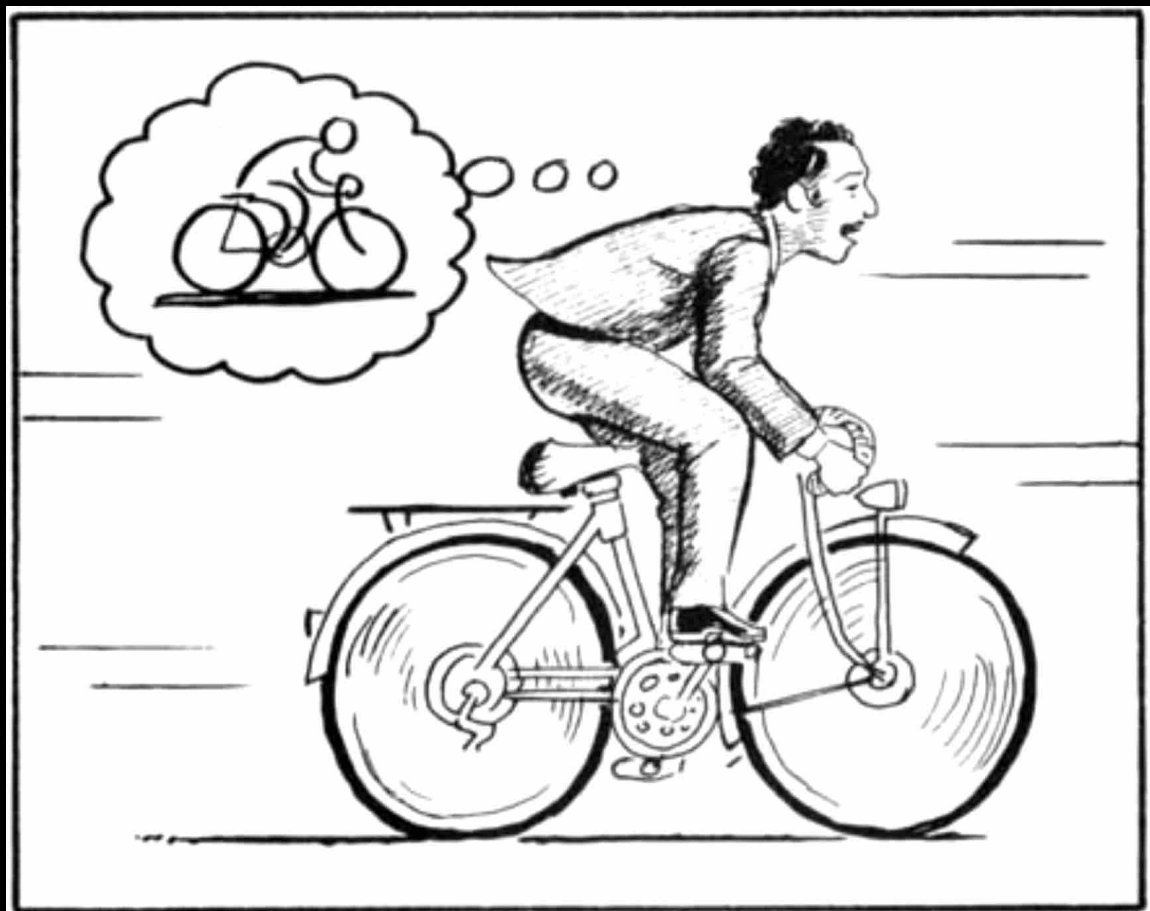


Adaptive Motor Control

A Predictive Coding Approach

Lorenzo Terenzi

MSc Thesis



Adaptive Motor Control

A Predictive Coding Approach

MSc THESIS

Lorenzo Terenzi

to obtain the degree of Master of Science
Faculty of Aerospace Engineering - Delft University of Technology

Supervised by:
Peter M. T. Zaal
D. M. Pool
M. Mulder

March 22, 2021

Faculty of Aerospace Engineering
Delft University of Technology

Preface

This report contains both a scientific article and a literature study on the topic of adaptive human motor control. This has been the most difficult project I conducted so far and I hope that these findings will be useful for future researchers on the topic.

I would like to thank my supervisors Daan, Max and Peter for the quality of supervision and their availability at all times. I also want to apologize to Peter for having woken him up early way too many times. Without them, this work would have not been possible. I would also like to thank my friends here in Delft, in the US and in Italy. They helped me stay positive even in a difficult time. A special thanks to Alejandro, with whom I always talk about the mysterious and amazing aspects of our universe and in particular about the topic of intelligence. I would like to further thank my family for their full support during these five years that allowed me to get this degree. Finally, I would like to thank Laura, now a de-facto member of my family, for the everyday support and love that she always shows for me.

Table of Contents

Preface	i
I Scientific article	1
II Literature Study	31
1 Introduction	32
2 Literature Study: Sensorimotor Control	33
2.1 Neocortex structure	33
2.2 Predictive coding	34
2.2.1 Computational Framework	35
2.2.2 Microcircuit	37
2.2.3 Evidence	37
2.3 Adaptive Sensorimotor Control	38
2.4 Internal Models	39
2.4.1 Young's manual control experiment	39
2.4.2 Cerebellum's motor models	40
2.4.3 State Estimation and Planning	40
2.5 Optimal Feedback Control	42
2.6 Hierarchical control	43
3 Literature Study: Models	45
3.1 Steady State Models	45
3.2 Models based on system identification's methods	46
3.2.1 Zaal model	46
3.2.2 Model based on Kalman Filter	46
3.2.3 Autoregressive models	47
3.3 Rule based models	47
3.3.1 Supervisory Phatak model	48
3.3.2 Fuzzy controllers	49
3.3.3 Hess Model	50
3.3.4 Young and Stark model	50
3.4 Adaptive control models	51
3.4.1 Incremental Non-Linear Dynamics Inversion (INDI)	51
3.4.2 Model Reference Adaptive Control (MRAC)	53
3.5 Machine learning methods	54
3.5.1 End to end Learning	54
3.5.2 Neural ODEs	54
3.5.3 Weight Agnostic Neural Networks (WANNs)	56
4 Research Objectives	58

5	Model Reference Adaptive Control	60
5.1	MRAC main ideas	60
5.1.1	MRAC for second-order SISO systems	61
5.1.2	Remarks	63
5.1.3	Comparison with classical pursuit architectures	64
5.1.4	Delay in MRAC	65
5.2	Selection Internal Model	66
5.3	Control under changing dynamics	67
5.3.1	Time-varying systems	68
5.3.2	Initial conditions and sensitivity to learning rate	68
5.3.3	Double to single integrator dynamics	69
5.3.4	Single to double integrator dynamics	69
6	Future Directions and Conclusion	72
6.1	Method's developments and simulations	72
6.2	Preliminary experimental plan	72
6.3	Conclusions	74
A	Time Traces of System Outputs	80
B	Time Traces of Participants and MRAC Control Outputs	82
C	Time Traces of Estimated MRAC Gains	84
D	Human Research Ethics Committee Checklist	85
E	Experiment Briefing	90
F	Experiment Consent Form	93
G	COVID19 Protocol	96
H	Runnable Example	103

List of Figures

2.1	Overview of the cortical layers, visualized using three different stain techniques. The Nissi stain mostly highlights the central body, the Weigert stain the synapses and the Golgi both.	34
2.2	Canonical microcircuit and established information flow in the cortex [1]	35
2.3	Hierarchical model that generates expectations from concepts [1]	36
2.4	Hierarchical model for predictive coding [6]	37
2.5	Detection times of the change in dynamics and their standard deviation for the three different groups [16]	40
2.6	Possible control scheme where the cerebellum encodes the forward dynamics (FM) of a system. The input arrives and together with the current information from the sensory system (SS) and prediction of the cerebellum, are sent to the motor cortex (MT) which elaborate the signals to be sent to the peripheral motor system (CO). The FM is trained by minimizing the difference the output of the model itself (b) with the feedback information (a) gathered by the sensory system (SS). This comparison is thought to happen in the inferior olive (IO), a specific component of the cerebellum [17].	41
2.7	On the left column you can observed the participants' measured limb position (A), velocity (B) and lateral deviation from the target (C) while on the are shown the optimal computed trajectories. Subfigure D, F, G show the same information but for policy generated by the optimal feedback system. Subfigure E shows how the different computed gains (shown in Equation 2.4) of the system change [28] . . .	43
3.1	Control diagram for a typical reference tracking task	45
3.2	Scheme of the Dual Extended Kalman Filter [34]	47
3.3	Example of convex partitioning of the state space into decision regions [38] . . .	48
3.4	Fuzzy system architecture [37]	49
3.5	Hess's model of human adaptive behavior [40]	50
3.6	Block diagram representation of the Young and Stark model for a compensatory display	50
3.7	Block diagram representation of an INDI controller	52
3.8	MRAC controller scheme[43] for a pursuit configuration	53
3.9	End-to-End learning architecture [45]	55
3.10	Reconstruction and extrapolation task comparison between a RNN and Neural ODE [46]	55
3.11	WANN architecture for the cart-pendulum task[46]	56
5.1	MRAC scheme for a pursuit display [43]	61
5.2	Diagram of the MRAC controller. The gains of the controller are K_x and k_r , the controlled plant is H_c , the closed loop reference model H_{mCL} and the prediction error e_p . Signals and diagrams related to predictions using the reference models are in blue while the ones directly used to adapt the gains are in red.	63

5.3	three channel model for modelling human behavior in a tracking task (objective is to follow signal f_t) on a pursuit display. The transfer function H_{o_t} represent the feedforward component of the controller, H_{o_x} the state feedback component while H_{o_e} is the fed the tracking error.	64
5.4	Two channel model for modelling human behavior in a tracking task (objective is to follow signal f_t) on a pursuit display. The transfer function H_{o_t} represent the feedforward component of the controller, H_{o_x} the state feedback component. . . .	64
5.5	MSE of reference model fitted over the averaged output of a participant of Zaal's reference tracking experiment [31].	67
5.6	Bode plots showing magnitude and phase of two of the controlled systems. . . .	68
5.7	Time series of the output (left column) and gains (right column) of the MRAC model while controlling system H_{21} . y_{ref} , y and y_m stand for the reference signal, the actual output and the reference model output.	70
5.8	Time series of the output of the reference model, y_m , operator (MRAC) model, y , the reference signal, y_{ref} (top row) and the gains (bottom row) of the MRAC model for controlling system H_{21} (left column) and H_{12} (right column)	71
6.1	Example of a display that could be used for the experiment	73
6.2	Sketch of the experimental procedure controlling system H_{r21} . The system switches between the H_1 and H_2 , indicated in the sketch respectively with the numbers 1 and 2	73
A.1	Time traces of the output of the system controlled by the participants. Each line represent the system's output for a single participants. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.	80
A.2	Time traces of the output of the system controlled by the MRAC controllers fitted on the participants data. Each line represent the system's output for a single MRAC controller. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.	81
B.1	Time traces of the control output of the participants. Each line represent the system's output for a single participants. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.	82
B.2	Time traces of the output of the MRAC controllers fitted on the participants data. Each line represent the system's output for a single MRAC controller. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.	83
C.1	Time series of the MRAC gains found for each participant.	84

List of Tables

2.1	Time, in seconds, for the detection in the change of dynamics of the controlled element. The controlled elements switches from first integrator to double integrator and vice-versa	40
-----	---	----

Part I

Scientific article

Adaptive Manual Control: a Predictive Coding Approach

Lorenzo Terenzi
Delft University of Technology
Delft, the Netherlands

Improved understanding of human adaptation can be used to design better autonomous systems and control systems that can support the human controller when the dynamics of the system that is being controlled suddenly change. This paper evaluates the effectiveness of a model-based adaptive control technique, Model-Based Reference Control (MRAC), for predicting the adaptive control policy shown by human operators while controlling a time-varying system in a pursuit-tracking task. Ten participants took part in an experiment, where they were asked to control a time-varying system whose dynamics changed twice and approximated a single and double integrator dynamics. A MRAC controller is composed of a feedforward and a feedback controller and an internal model that is used to drive the adaptive control policy. The active gains, the internal model parameters and the learning rates, have been estimated via a non-linear optimization aimed at maximizing quality of fitting of the participants' control output. The participant's control behavior rapidly changed when the dynamics of the controlled system changed, in particular when going from controlling a first to second order system. The MRAC model was able to accurately capture the transient dynamics exhibited by the participants when the system changed approximately from a first to a double integrator while it failed to do so when the system changed from double to first integrator. In the latter case the MRAC gains changed too slowly. Therefore MRAC can be used to approximate human adaptations in pursuit tracking tasks when a change in the dynamics of the controlled system requires an increase in the rate feedback controller to ensure accurate tracking of the reference signal.

Nomenclature

A	= State space matrix, –
A_m	= Reference model state space matrix, –
A_n	= Amplitude of the i^{th} sine of the disturbance signal, rad
B	= Control input matrix, –
B_m	= Reference model control input matrix, –
DYN	= Controlled element dynamics label
e	= Roll attitude error, rad
H_c	= Controlled dynamics transfer function
H_p	= Human operator transfer function
H_p	= Human operator transfer function
H_{mCL}	= Internal model closed loop transfer function
H_{oCL}	= Internal model open loop transfer function
$k_c(t)$	= Time varying gain of the controlled element, –
k_r	= MRAC feedforward gain, –
K_x	= Vector of state gains, –
k_{x1}	= MRAC state gain, –
k_{x2}	= MRAC state derivative gain, s
lr	= Learning Rate, –
r	= Forcing function, rad s^{-1}
$RMSE$	= Root mean square of the roll error, deg
$RMSU$	= Root mean square of the control input, deg
P	= Positive definite matrix, –
t	= Time, s

x_m	=	Reference model state
u_h	=	Human operator control output, rad
u_m	=	MRAC model control output, rad
VAF	=	Variance accounted for, %
ζ_{nm}	=	Human operator neuromuscular damping, –
τ	=	Human operator time delay, s
θ	=	Vector of the MRAC parameters, –
ϕ_n	=	Phase of the i^{th} sine of the disturbance signal, rad
γ_r	=	Learning rate of the feedforward gain, –
Γ_x	=	Vector of state gain's learning rate, –
γ_{x1}	=	Learning rate of the state gain, –
γ_{x2}	=	Learning rate of the state derivative gain, –
ω	=	Frequency, rad s ⁻¹
$\omega_b(t)$	=	Time-varying pole for the controlled element, –
ω_c	=	Crossover frequency of the internal model, rad s ⁻¹
ω_n	=	Frequency of the i^{th} sine of the disturbance signal, rad s ⁻¹

I. Introduction

Automation in the last decades has taken an increasingly important role in systems with humans in the loop, such as aircraft or road vehicles. Tasks that are structured in a relatively controlled environment are easiest to automate. In aerospace, for example, most parts of an actual flight have been automated because it is possible to rely on precise positioning systems, sensor readings and lack of obstacles and preestablished navigation plans. The roles of pilots have become more of a supervisor of automated sub-systems.

Current artificial systems can also achieve a human level of performance in perception, or better, in many tasks. For example, superhuman performance in image classification on the ImageNet dataset was achieved five years ago [1]. These systems can also make use of sensory data not available for humans, such as Lidar.

Nonetheless humans are still driving billions of cars and piloting aircraft, thanks of their ability to quickly adapt to a changing environment [2], carry out tasks that are complex and difficult to automate. A good understanding of how humans adapt their control policy could serve as inspiration for machine learning researchers and automation engineers to develop more advanced controllers, capable of exhibiting more adaptable behavior. Furthermore, it could help design better training programs for human controllers and more suitable control systems to increase the safety of the systems they operate [3].

This paper addresses the problem of understanding human motor adaptation for simple manual control tasks and it proposes a novel approach inspired by a neuroscience framework called "Predictive Coding" [4]. At the core of this approach lies an adaptive control technique called model reference adaptive control (MRAC), which defines a controller that imitates another controller taken as a reference. This work was inspired by the fact that McRuer found that the open-loop dynamics of a system controlled manually, with a compensatory display, has the structure of a first-order stable system with delay, independent of the dynamics of the controlled system.

The paper first presents the MRAC controller, the internal reference model chosen and the parameter estimation procedure. The model output stability and sensitivity to its parameters is assessed. To further assess the ability of MRAC to approximate time-varying dynamics of human controllers, an experiment was organized. The collected data are useful for facilitate future studies on time-varying models and to understand how MRAC responds to different transitions in the dynamics of the controlled system.

In Sec. II it is introduced and reviewed part of the work done on three fundamental areas, upon which this paper is based on: internal models, predictive coding and time-varying/adaptive human operators models. In Sec. III a basic introduction and analysis of the MRAC controller is presented together with a description of the control task performed during the experiment. In Sec. IV the experimental conditions, dependent measures, procedures and setup of the experiment are explained. In Sec. V the experimental results are analysed and presented and their discussion follows in Sec. VI. Finally, the conclusions of the paper are drawn in Sec. VII.

II. Background

A. Role of internal models

In this section, an overview of internal models is presented, and we try to justify their use in the development of adaptive and optimal control systems.

Internal models have been recognized to be an essential component for motor control in biological systems [5]. Animals have internal models of their body and world dynamics that can be used for motor control and learning [6] [7].

It is commonly accepted that the cerebellum learns and encodes models of the world [8]. These models likely encode the dynamics and properties of the human motor system and of objects that humans interact with. Patients with cerebellum damage experience very poor motor control and jittering in their movement and have problems learning new skills [9, 8]. The imprecise motor control in these patients suggests that their brains cannot make use of internal models of the motor system to find optimal motor commands and compute the set of probable future states. In this case, the brain will have to rely only on sensory feedback, which is severely delayed (such as visual feedback) and does not allow for precise and fast motor control [10].

Anatomically, the cerebellum contains the vast majority of the neurons present in the central nervous system, with estimates ranging from 70% to 80%. The cerebellum contains around 69 billion neurons, and an equivalent number of glial cells [11]. The human brain has only 19% of the neurons on the cerebral cortex, similar to the composition found in other mammals. However, while the cellular composition is equivalent, the size of human networks is substantially larger than those of other primates. The work done at OpenAI provides evidence for the hypothesis that what is commonly understood as intelligence increases with the network size. A set of scaling laws for language models were found and they highlighted that the model size is the principal parameter that affects performance: larger models are more sample-efficient and expressive [12].

The exact nature of the models used for sensorimotor control is not currently known, but is hypothesized they could be inverse and/or forward dynamical models [13]. Inverse models output the control action needed to achieve the desired output, while forward models predict the output given a control action. On the other hand, given a probability distribution of the current states and a control input, forward internal models can generate a distribution over future states. There is substantial anatomical and behavioral evidence that points towards their presence in the cerebellum [14, 15].

Internal forward models have two main uses. The first one is to plan complex actions without acting them out in the real world, by simulating the results of a control policy through the forward model. This way of controlling is compatible with Optimal Feedback Control (OFC), a control method that makes use of the dynamics of a system to find the set of optimal actions to define a feedback controller [16, 17, 18]. The OFC framework has shown predictive strength in many experiments testing sensorimotor control and in particular hand-eye coordination [5]. Currently, it is the dominant framework to model sensorimotor control [19, 20]. The second main use for internal forward models is state estimation: the distribution over the current states (given the previous state and a control action) is compared with sensory measurement to achieve the best estimate of the current state of the system [5]. A diagram that illustrates how internal models can be used for state estimation and planning is shown in Fig. 1. The state estimator makes use of the sensory measurements, z , and the predicted state x_p by the internal model to obtain a more reliable estimation. The expected state estimate, x_e can be fed back to the internal model. The internal model can be used to generate rollout (sequence of predicted states x_p) given of a policy (set of control inputs u) as part of an optimal control scheme.

The last piece of evidence we present here in favour of internal models is related to a manual control experiment with a compensatory display task performed by Young [2]. Participants were asked to control and detect the change of dynamics in a time-varying system. The participants of the experiment were divided into three groups: active controllers, passive controllers and observers. The active controllers actively controlled the system. The passive controllers were led to believe that they controlled the system, but the system was controlled by a different controller. The observers just observed the screen. All participants had to press a button when they recognized the change in dynamics. It was observed that the detection time across all conditions for the observers was significantly higher (mean of 3.74 s) than the other two classes of participants, who on the other hand showed similar times (mean of 1.30 s for active controllers and of 1.50 s for passive ones). The most probable cause for this delay for the observer is the absence of an internal model for the dynamics of the controlled system. The participants that only observed engage in no motor actions and therefore had a harder time estimating how the system would respond to input. This highlights the role that internal models have in motor control tasks.

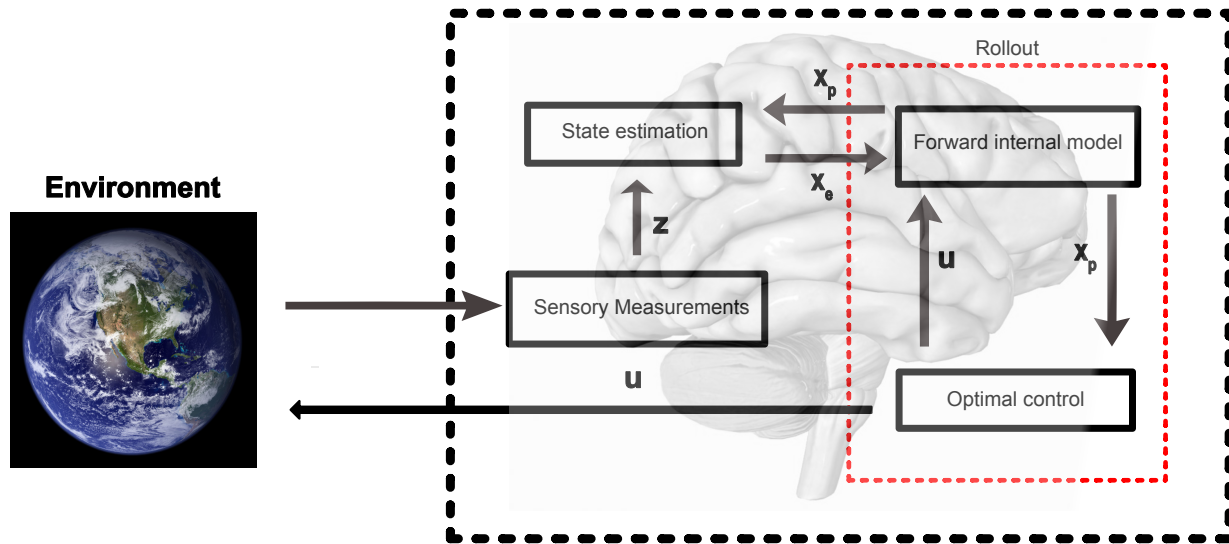


Fig. 1 Diagram illustrates how sensory data and the internal model can be used for state estimation, control and planning.

B. Predictive Coding

Predictive coding (PC) is a general framework to explain how the brain processes information. PC states that the human brain continuously generates predictions about the stream of information that is coming from the world, and its objective is to minimize the prediction error [18], and thus surprise. The minimization of the prediction error can be achieved by either changing the internal model that lead to the predictions (learning) or the actions in the world of the agent. The theory, very formal in nature, is further explained by Friston in a series of papers [21, 22].

It is postulated that the brain has an internal generative model of the world. This is a model that can generate a rich and hierarchical representation of the expected/predicted sensory information from a latent, non-observable variable. [22]. For example, from a single "idea" (the latent variable) of an object, it is possible to generate richer representations in the lower hierarchical cortical area, like its shape, sound, expected dynamic properties. The process that generates the sensory predictions is "top-down", from the higher level of the cortical areas to the lower ones, while the sensory information is processed "bottom-up" as illustrated in Fig. 2.

The sensory predictions at the very bottom of the abstraction hierarchy are compared with the actual sensory data. The unexplained sensory inputs, i.e., the "errors" are propagated up the hierarchy for further processing. These prediction errors are minimized either by acting on the world or by changing the internal model over time [24]. An essential part of the PC theory is *precision weighting*, i.e., the estimation of the error's reliability, which also depends on the level of signal noise [4].

This theory of neural computation is supported by anatomical evidence and there are canonical microcircuit models of PC [25]. As PC requires, there are different cells that are responsible for feedforward and feedback information [25, 4], and these cells show interactions at different frequencies [26]. Computational models of the visual cortex explain many of the unsolved visual phenomena like end-stopping and non-classical surround effects [27]. Similar results are reported for the auditory cortex [28, 29]. Nonetheless, while it is recognized that the ability to predict is a key element of the human type of intelligence, there is still not a consensus in the scientific community about PC: criticism highlights that the theory is quite imprecise, difficult to test, and therefore susceptible to *ad-hoc* changes [30].

C. Time varying models

There have been successful attempts at identifying time-varying policies that human controllers exhibit in response to a change in controlled system dynamics. A very flexible approach is based on state-estimation via Kalman Filters, to recursively estimate parameters of human operator models [31, 32, 33]. The main drawback of this method is the speed of convergence of the filter: in the event of a sudden change in the operator, the estimated parameters could be unreliable during the transient phase. Another promising line of research uses ARX-based identification techniques [34].

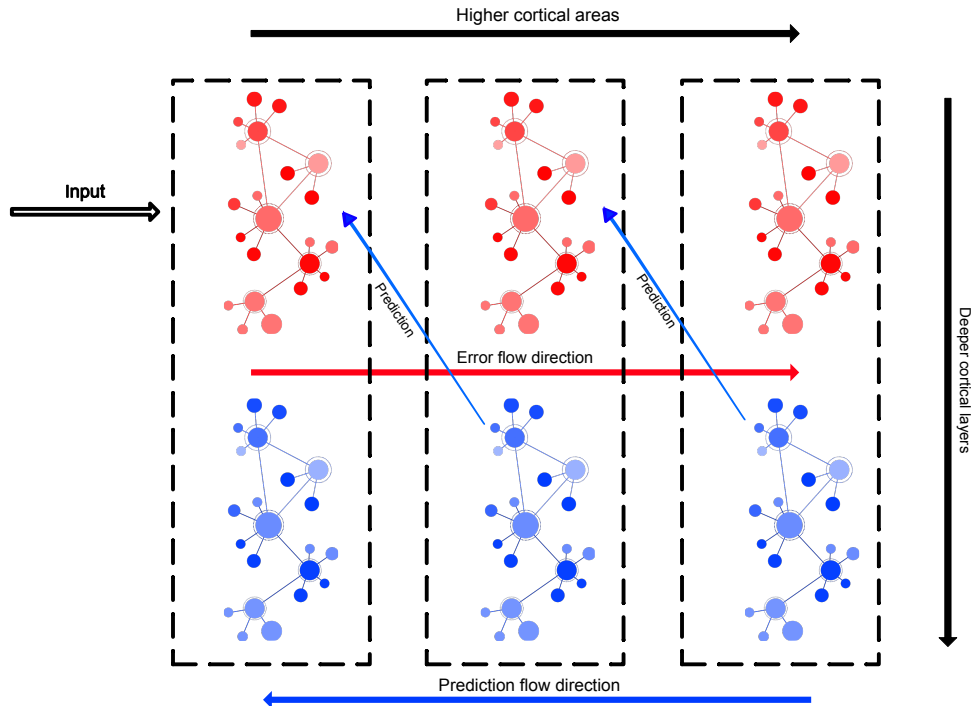


Fig. 2 Very simplified scheme of a predictive coding framework. The input enter a network region into the superficial cortical layers (in red). There the input is processed together with the prediction, coming from a deep cortical layer (blue) next in the hierarchy. The resulting error is further propagated up the hierarchy and the process repeats. Figure adapted from [23].

The ARX models rely on minimal assumptions about the human operator behaviour and can be extended for multimodal operators identification, such as control with the presence of motion feedback [35].

To identify human operators controlling a time-varying system, Zaal proposed a parametric model, the parameters of which are all a time-varying function of sigmoidal shape centred at the time of the transition of the dynamics form [36]. The model parameters were found by maximum likelihood estimation (including the transition time of the dynamics) augmented with a genetic algorithm [37]. The main drawback of the model is that the transient behavior is prescribed once, and thus cannot handle more than one transition in dynamics.

The Hess model is also a popular rule-based system that relies on a two-loop description of the human operator. The model relies on a hand-crafted set of gains and learning rules to ensure a representative adaptive behavior [38].

III. Methodology

This study uses Model Reference Adaptive Control (MRAC), an adaptive control technique, to predict a time-varying human operator adaptations in tracking tasks in a pursuit display. MRAC is first introduced, then the selection of the internal model, the architecture of the controller, and the considered parameter estimation method are presented. A description of the tracking task follows, which consisted on controlling a time-varying system. The participants' objective was to minimize the error between the indicated attitude and the current attitude, which could be inferred w.r.t. the artificial horizon line. This task was chosen because the MRAC model uses the reference signal as an input and it is simple enough to test the fundamental validity of the approach.

A. Model Reference Adaptive Control

MRAC was further researched to model time-varying human-operator policies for the following reasons:

- 1) Humans are adaptable: MRAC is an adaptive control algorithm for which the controlled system to track the trajectory of a reference model. The gains of the controller are time-varying.

- 2) Humans have internal models: while controlling the system, and receiving mostly visual feedback from the screen, humans can learn the dynamics of the controlled element. This new model is used for control. Similarly, MRAC uses its internal model for control by being the main drive behind the adaptation.
- 3) Humans in the loop show consistent open-loop dynamics when controlling a wide range of dynamical systems on a compensatory display: McRuer et al. found that the open-loop dynamics in the crossover region resemble those of a single integrator with delay [39]. In an analogous way MRAC uses an internal reference model to define the ideal control policy independently of the controlled dynamics.
- 4) Human adapt in the presence of mismatched predictions: PC states that the errors, i.e., the difference between the predicted and observed states, are propagated up the neuronal hierarchy to drive actions and/or change the encoded world models in the neural network. MRAC works similarly, since the adaptation is driven by the difference between the predicted output of the internal model and the observed one.

The mathematical details of MRAC and its formulation are explained here. More examples and further derivations of MRAC controllers can be found in Nguyen's work [40].

We first derive an expression for the dynamics of the error between the output of the reference model and the output of the controlled system, see Fig. 3. Let's assume that the reference model can be written in a state space form as:

$$\dot{x}_m = A_m x_m + B_m r \quad (1)$$

where x_m is the state of the reference model and r is the reference signal. The controlled element can also be expressed in state space form as:

$$\dot{x} = Ax + Bu \quad (2)$$

where u is the time-varying control input to be found. It is assumed a trivial observation equation such that the output $y = x$ and that the state space matrix A and the control input matrix B are unknown, which reflects the uncertainty about the dynamics of the system. The aim of the control action is to minimize the error $e_p = x_m - x$. The feedback control signal is assumed to have the following form:

$$u = K_x(t)x + k_r(t)r \quad (3)$$

where $K_x(t) = [k_{x1}, k_{x2}]$ and $k_r(t)$ are time-varying gains. MRAC assumes the existence of ideal gains, K_x^* and k_r^* , that can be found if Eq. (3) is substituted into Eq. (2). These ideal gains must satisfy:

$$\begin{aligned} A + BK_x^* &= A_m \\ Bk_r^* &= B_m \end{aligned} \quad (4)$$

To simplify the derivation, two new quantities \tilde{K}_x and \tilde{k}_r can be defined, which are the gain deviations from their optimal values:

$$\begin{aligned} \tilde{K}_x &= K_x(t) - K_x^* \\ \tilde{k}_r &= k_r(t) - k_r^* \end{aligned} \quad (5)$$

By combining Eq. (2), Eq. (5) and Eq. (3) the following expression for the error can be found:

$$\dot{e}_p = \dot{x}_m - \dot{x} = A_m e_p - B\tilde{K}_x x - B\tilde{k}_r r \quad (6)$$

MRAC relies on the Lyapunov stability theory to prove the stability of the system. To ensure tracking of the reference model and therefore that $\lim_{t \rightarrow \infty} e_p(t) = 0$ the following conditions need to be satisfied:

- 1) Find a Lyapunov function $V(t, e_p, \tilde{K}_x, \tilde{k}_r) > 0, \forall t > 0$
- 2) $\frac{dV(t, e_p, \tilde{K}_x, \tilde{k}_r)}{dt} < 0$ for all $t > 0$
- 3) $\frac{dV(t, e_p, \tilde{K}_x, \tilde{k}_r)}{dt} \in L_\infty$ norm, i.e., $\frac{d^2 V(t, e_p, \tilde{K}_x, \tilde{k}_r)}{dt^2}$ must be bounded

There are no real guidelines to choose the Lyapunov function, but usually it is a quadratic function with respect to the variables of interest [40]. It can also be interpreted as an energy function. In this case the following function was chosen

$$V(e_p, \tilde{K}_x, \tilde{k}_r) = e_p^T P e_p + |b|(\tilde{K}_x \Gamma_x^{-1} \tilde{K}_x^T + \frac{\tilde{k}_r^2}{\gamma_r}) > 0 \quad (7)$$

where b is the only entry of the control effectiveness matrix B , with P , Γ_x^{-1} , $\gamma_r > 0$, the function is also bigger than zero at all times. Notice that this derivation is valid for a second order system as we assumed the control matrix is of the form $B = [0; b]$. By taking the time derivative of the Lyapunov function and using Eq. (6), the following is obtained

$$\begin{aligned} \dot{V}(e_p, \tilde{K}_x, \tilde{k}_r) = & -e^T (PA_m + A_m^T P)e_p + 2|b|\tilde{K}_x(-xe_p^T \bar{P} \text{sgn}(b) + \Gamma_x^{-1} \dot{\tilde{K}}_x^T) \\ & + 2|b|\tilde{k}_r(-re_p^T \bar{P} \text{sgn}(b) + \frac{\dot{\tilde{k}}_r}{\gamma_r}) \end{aligned} \quad (8)$$

and by selecting P to satisfy the Lyapunov equation

$$PA_m + A_m^T P = -Q \quad (9)$$

it can be found that

$$\begin{aligned} \dot{V}(e_p, \tilde{K}_x, \tilde{k}_r) = & -e_p^T Q e_p + 2|b|\tilde{K}_x(-xe_p^T \bar{P} \text{sign}(b) + \Gamma_x^{-1} \dot{\tilde{K}}_x^T) \\ & + 2|b|\tilde{k}_r(-re_p^T \bar{P} \text{sign}(b) + \frac{\dot{\tilde{k}}_r}{\gamma_r}) \end{aligned} \quad (10)$$

where Q is a negative definite matrix and \bar{P} the second column of the matrix P (since it is assumed that matrix B has only one entry equal b). For the Lyapunov function derivative to be negative at all time the following conditions need to be imposed

$$\begin{aligned} -xe_p^T \bar{P} \text{sign}(b) + \Gamma_x^{-1} \dot{\tilde{K}}_x^T &= 0 \\ -re_p^T \bar{P} \text{sign}(b) + \frac{\dot{\tilde{k}}_r}{\gamma_r} &= 0 \end{aligned} \quad (11)$$

which implies

$$\begin{aligned} \dot{\tilde{K}}_x &= \Gamma_x x e_p^T \bar{P} \text{sgn}(b) \\ \dot{\tilde{k}}_r &= \gamma_r r e_p^T \bar{P} \text{sgn}(b) \end{aligned} \quad (12)$$

In this way an expression was derived for the rate of change of the feedback gain that would ensure the tracking of the reference model. Finally, it is necessary to show that the derivative of the Lyapunov function is bounded. The derivative of the Lyapunov function satisfies the following inequality

$$\dot{V}(e_p, \tilde{K}_x, \tilde{k}_r) = -e_p^T Q e_p \leq -\lambda_{\min} \|e_p\|_2^2 \quad (13)$$

where λ_{\min} is the smallest eigenvalue of the matrix Q . Since $\|e_p\|_2^2 \in L_\infty$ it can be shown that $\dot{V}(e_p, \tilde{K}_x, \tilde{k}_r)$ is bounded. By using Barbalat's lemma [40] it can be concluded that $\lim_{t \rightarrow \infty} e_p(t) = 0$.

Fig. 3 shows the diagram representing the used implementation of the MRAC controller, where the adaptive laws are given in Eq. (12). The gains of the controller are $K_x = [k_{x1}, k_{x2}]$ and k_r , the controlled plant is H_c , the closed-loop reference model H_{mCL} and the prediction error e_p . For readers more accustomed to the parameters used by El et al. [41], the MRAC coefficients are related to them in the following way:

$$\begin{aligned} O(k_e) &= O(k_{x1}) \\ O(K_f) &= O\left(\frac{k_r}{k_{x1}}\right) \\ O(T_L) &= O\left(\frac{k_{x2}}{k_{x1}}\right) \end{aligned} \quad (14)$$

where K_f is the *target weighting gain*, T_L the *lead time gain* and k_e the *proportional error gain* [41]. In the previous derivation, it was assumed that there was no delay in the control input to derive a simple adaptive control law. Humans have quite considerable delays in perception and actuation delays that should be accounted for. The

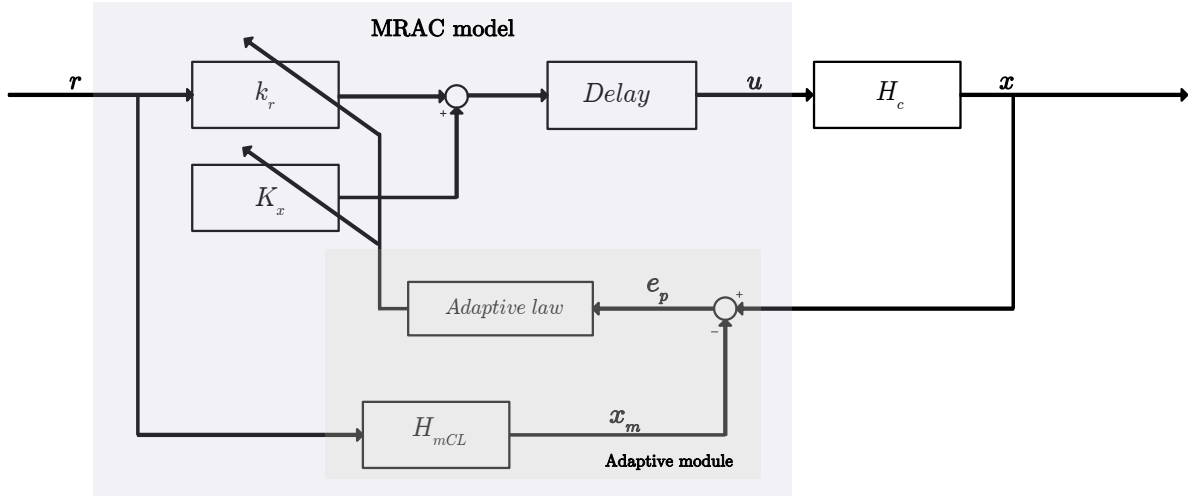


Fig. 3 Diagram of the MRAC controller.

presence of delay can severely affect the performance and stability of a linear state feedback controller. An adaptive state feedback controller, such as MRAC, is also affected by such delay.

The system controller by the human operator can be written more formally as

$$\dot{x}(t) = A(t)x(t) + B(t)u(t - \tau) \quad (15)$$

$$u(t) = K_x x(t) + k_r r(t) \quad (16)$$

where τ stands for the input delay. The reference system, which is also delayed in the input, is given by

$$\dot{x}_m(t) = A_m x_m(t) + B_m r(t - \tau_m) \quad (17)$$

This paper does not provide any formal guarantees on the stability of the considered MRAC controller, but there are some techniques that establish bounded (not asymptotic) stability, such as BLAS (Bounded Linear Stability Analysis) [42] and other techniques for time delay margin estimation [43]. More generally, the design of stable controllers and their analysis in presence of delays and uncertainties has been reported with Lyapunov-Krasovskii functional techniques [44]. A comprehensive tutorial and review on Lyapunov-methods for time-delayed system was written by Fridman [45].

B. Selection of internal model

A key choice for every MRAC controller is the function of its internal model. The selected reference model proposed in this paper is the open-loop model proposed by McRuer [46], which describes the open loop behavior of a skilled human operator for a compensatory task. While the current model works with pursuit display (where the reference signal is available), a single integrator open loop model still is able to approximate the behavior of a trained controller. It is assumed that this model is also valid to describe the quasi-linear behavior of human operators for pursuit tasks. Further information on how pursuit tracking differs from compensatory tracking can be found in the review by Mulder et al. [47]. The crossover model in the Laplace domain has two free parameters: the effective time delay, τ , and the crossover frequency, ω_c :

$$H_{mOL}(s) = \frac{\omega_c}{s} e^{-\tau s} \quad (18)$$

MRAC makes use of its closed loop version:

$$H_{mCL}(s) = \frac{H_{mOL}(s)}{1 + H_{mOL}(s)} \quad (19)$$

The time delay and the crossover frequency of the model are estimated together with the other parameters necessary in an optimization routine.

C. Parameter estimation

The value of the following parameters is needed in order to specify the MRAC controller:

$$\theta = [k_{x1}, k_{x2}, k_r, \omega_c, \tau, \gamma_{x1}, \gamma_{x2}, \gamma_r] \quad (20)$$

k_{x1} is the state gain, k_{x2} is the state derivative gain, k_r is the feedforward gain, ω_c and τ are the parameter of the internal reference model, and γ_{x1} , γ_{x2} , and γ_r are the learning rate of k_{x1} , k_{x2} and k_r , in this order. These parameters can be found solving the following non-linear optimization problem:

$$\min_{\theta} J(\theta) = \frac{1}{N} \left(\sum_{i=0}^{N-1} (u_h - u_m)^2 \right) \quad (21)$$

where u_m stands for the MRAC controller output and u_h the measured output of the human controller. This procedure has been applied in the past to estimate the gains of human operators [36]. Since this is a non-linear optimization scheme care must be taken in initializing the parameters to minimize to chance of finding a local minimum. For each case, the optimization was run fifteen times, each time with a different initial condition. The set of parameters that resulted in the minimum value of the objective function were selected. Additionally since the parameters have different scales, they were normalized to improve the conditioning of the optimization problem. Finally, an upper bounds on the learning rate (in this case the value of 50 was used for all learning rates) improves the chances of the optimization converging. The optimization could otherwise diverge if very high learning rates are selected by the optimizer.

D. Control Task

This paper focuses on how the control policy of humans adapts in a single-axis pitch control task, done with a pursuit display. Participants performed a pitch tracking control task. Fig. 4 shows the block diagram of the task. Participants minimized the error e between the target attitude and the current attitude through a pursuit display, while controlling the pitch dynamics of the system $H_c(t)$. The target attitude was generated using the forcing function $r(t)$, which was a sum of ten sinusoids to make the task challenging and the signal unpredictable. The task was performed on the HMI (Human Machine Interaction) lab at the TU Delft faculty of Aerospace Engineering [48].

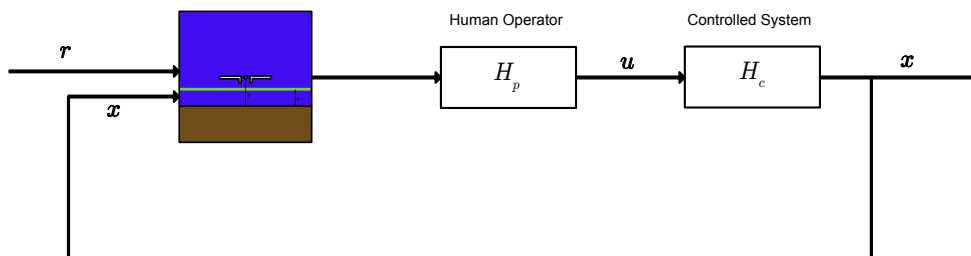


Fig. 4 Control task block diagram.

The controlled system dynamics $H_c(t)$ can change through time and had the following structure

$$H_c(t) = \frac{k_c(t)}{s(s + \omega_b(t))} \quad (22)$$

where the parameters $k_c(t)$ and $\omega_b(t)$ were time-varying. In this task, participants controlled the system $H_c(t)$ across multiple and significant changes in the system dynamics. By modifying the parameter ω_c the system response changed either towards an single-integrator response (for $\omega_c \gg 1$ rad/sec) or a double-integrator response (for $\omega_c \ll 1$ rad/sec). The k_c was also varied over time to keep the level of control activity approximately constant. The mathematical representation of ω_c and k_c was the following

$$\omega_b(t) = \begin{cases} \omega_{b1} + \frac{\omega_{b2} - \omega_{b1}}{1 + e^{-G(t-M_1)}}, & \text{for } t \leq t_0 + \frac{T}{2} \\ \omega_{b2} + \frac{\omega_{b1} - \omega_{b2}}{1 + e^{-G(t-M_2)}}, & \text{for } t \geq t_0 + \frac{T}{2} \end{cases} \quad (23)$$

$$k_c(t) = \begin{cases} k_{c1} + \frac{k_{c2}-k_{c1}}{1+e^{-G(t-M_1)}}, & \text{for } t \leq t_0 + \frac{T}{2} \\ k_{c2} + \frac{k_{c1}-k_{c2}}{1+e^{-G(t-M_2)}}, & \text{for } t \geq t_0 + \frac{T}{2} \end{cases} \quad (24)$$

where $M_1 = \frac{T}{3}$, $M_2 = \frac{2T}{3}$, and the measurement time was $T = 90$ s. The value of G which controlled the transition speed, is kept constant at a value of 100/sec to simulate a "step-like" change in the controlled system's parameters. The reference attitude signal was defined by

$$r(t) = \sum_{n=1}^{10} A_n \sin(\omega_n t + \phi_n) \quad (25)$$

where A_n is the amplitude of the n^{th} sine wave, ϕ_n the phase, and ω_n the frequency. While it was not necessary to have a sinusoidal tracking signal, it was important to have an unpredictable signal to prevent humans from learning its structure. The signal was also chosen to be periodic, with a period equal to $P = \frac{T}{3}$, one-third of the measurement time. The periodicity preserved the task difficulty, as participants were exposed to the same signal when transitioning from one dynamics to the other and vice-versa. The parameters of the sinusoidal waves that make up the signal $r(t)$ are shown in Table 1.

Table 1 Parameters of the sinusoidal function used to construct the reference signal.

n	A_n [rad]	ω_n [rad/s]	ϕ_n [rad] Testing	ϕ_n [rad] Validation
1	$2.905 \cdot 10^{-2}$	0.419	2.841	3.006
2	$1.916 \cdot 10^{-2}$	1.047	3.319	6.037
3	$1.020 \cdot 10^{-2}$	1.885	0.718	4.544
4	$6.032 \cdot 10^{-3}$	2.722	0.768	2.811
5	$3.356 \cdot 10^{-3}$	3.979	2.925	5.917
6	$1.983 \cdot 10^{-3}$	5.655	5.145	1.842
7	$1.230 \cdot 10^{-3}$	8.188	2.085	3.401
8	$9.331 \cdot 10^{-4}$	10.681	0.383	2.998
9	$7.541 \cdot 10^{-4}$	14.032	0.763	4.614
10	$6.674 \cdot 10^{-4}$	17.383	3.247	2.888

E. Stability of the estimated coefficients

This paragraph is used to explain how it was verified that the parameters estimated by MRAC do not vary excessively in time, in particular when it is trying to represent a steady state controller.

To test the reliability of the obtained coefficients we first solved the optimization problem using data generated from controllers with known, arbitrarily chosen and fixed values of the coefficients. The steady controllers controlled systems with steady dynamics and a time varying system that alternates between two dynamics labelled $DYN = 1$ and $DYN = 2$. It is important to test if the estimated coefficients will not vary much in time even in the case where a dynamical transition in the controlled system happens.

Fig. 5 shows the a time-series plot of the estimated MRAC gains for controlled systems with steady and time-varying dynamics. The controller chosen to control a time invariant system has as gains $k_{sx} = 0.15$, $k_{sx2} = 0.09$, $k_{sr} = 0.15$, while the values of the found MRAC coefficients are $k_{x1} = 0.150$, $k_{x2} = 0.089$, $k_r = 0.151$. The VAF u of the model is 0.997. The VAF between two vectors, x and y , is defined as:

$$VAF(x, y) = 1 - \frac{\text{var}(y - x)}{\text{var}(y)} \quad (26)$$

and it's a measures of the quality of fit similar to the R^2 measure. For controlling the time-varying system a controller was chosen with gains $k_{x1} = 0.16$, $k_{x2} = 0.04$, $k_r = 0.16$, while the values of the found initial MRAC active gains are $k_x = 0.167$, $k_{x2} = 0.039$, $k_r = 0.179$. The VAF of the model is 0.987.

Therefore when the controlled system is time-invariant we expect the estimates of the MRAC coefficient to be reliable and that any change in their value reflects a real change in the underlying controller. When the controlled system is time-varying, the coefficient that adapts the most is k_r , which at the end of the run has increased to a value of 0.16, an increase of approximately 5% over the true parameter. This change though barely affects the VAF u of the MRAC model.

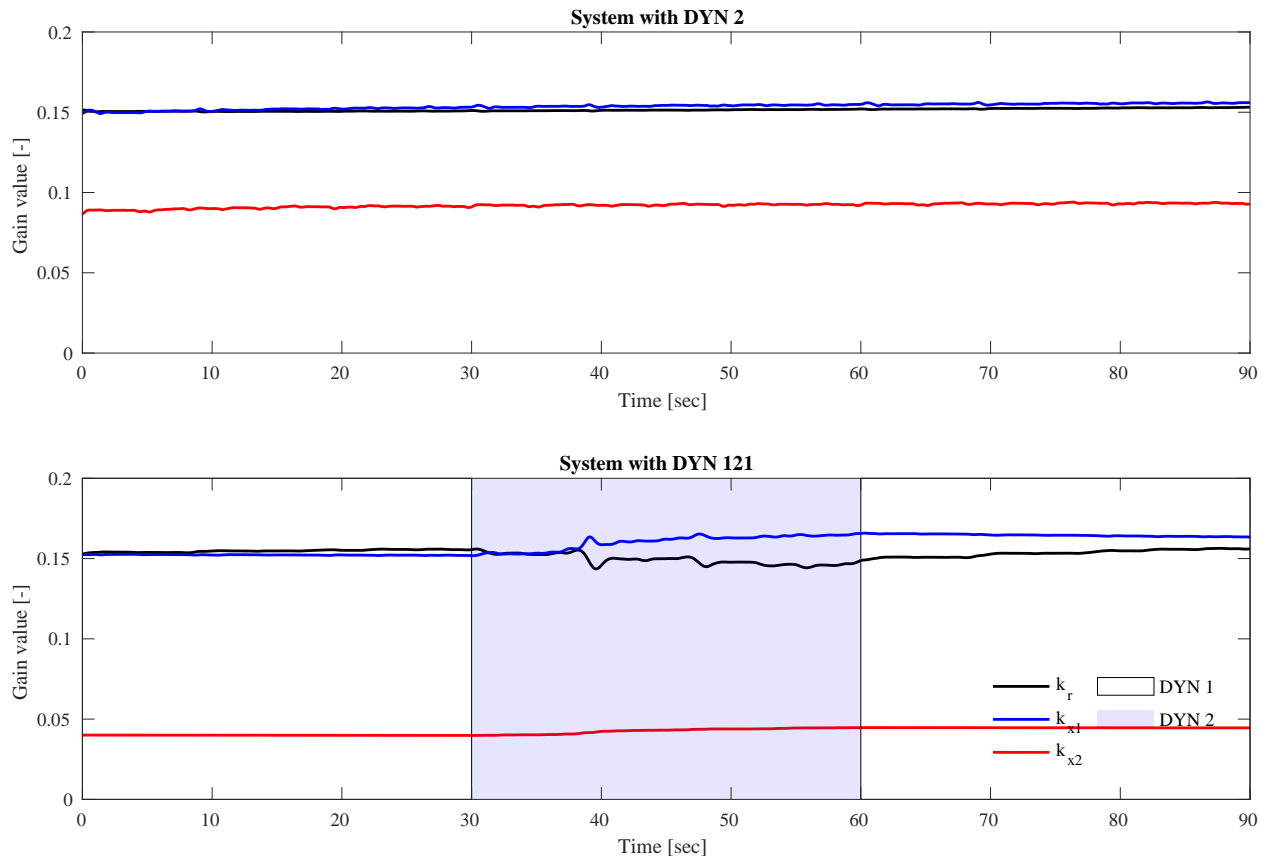


Fig. 5 Time series of the estimated MRAC gains using data from two steady state controllers.

F. Model Sensitivity

This paragraph is focused on the sensitivity of the MRAC control gains to the selected learning rates for their adaptation. The learning rates control how quickly the model adapts to changes in the controlled dynamics, which are detected through large errors between the expected and the observed output of the plant. In the current MRAC formulation there are three learning rates γ_{x1} , γ_{x2} and γ_r , i.e. one for each control gain, that can freely vary with respect to each other. The results shown here are purely illustrative and should serve the reader in gaining an intuition about the effect of the learning rates. Their effect strongly depends on the initial parameters of the controller, the internal model, the present delay and the controlled system. A controlled system is considered whose dynamics change between the ones of approximately a first order system ($DYN = 1$) and a second order system ($DYN = 2$) at frequencies close to the human crossover. The label $DYN = 121$ indicates the order of the dynamics' transition: first, second and first order. Using the learning rates $lr = [0.5, 1, 5, 10, 30]$, we checked how the output of the controlled system, with $DYN = 121$, and the gains of the MRAC controller varied through time.

Fig. 6 shows the output of the model for different learning rates. The lower the learning rate, the sharper the oscillation observed after the change in dynamics at 30 seconds. In a similar way, when the dynamics of the plant are reverted back, at 60 seconds, controllers with higher learning rates undershoot the reference signal. Fig. 7 depicts the time traces for k_r , k_{x1} and k_{x2} . The most pronounced change is observed in the k_{x2} gain, which sharply increases after the dynamics are changed from 1 to 2. The decline of the gain after the dynamics are reverted back to 1 is less pronounced, and for small learning rates almost null. The gains change faster when the dynamics change from 1 to 2, since the controller initially is not able to stabilize the system and large tracking error occur. On the other hand when switching the dynamics from 2 to 1, the closed loop system is already stable and the error between the internal model output and the system output remains smaller, which leads to slower rate of change in the gains. In addition, the parameters' adaptation is also controlled by the value of the state and its derivative (the higher their value the faster the higher the signal for adaptation) which are higher in absolute value in the transition from DYN 1 to 2. A relatively large increase of the k_r gain can also be observed at 60 seconds for larger values of learning rate. The increase in value is likely caused

by the still elevated value of the k_{x2} gain which decreases the control output of the controller.

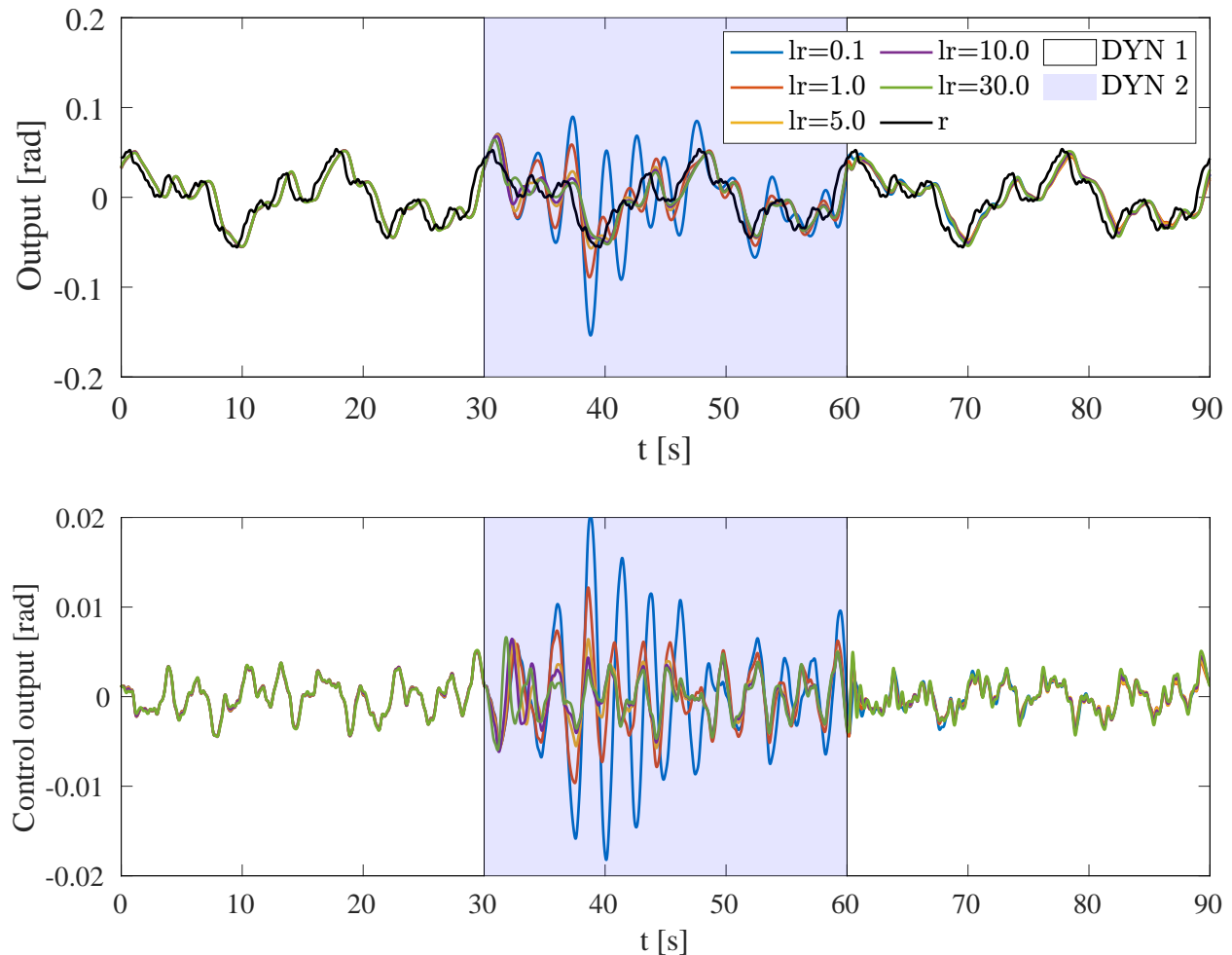


Fig. 6 Output and control output sensitivity to the learning rate (lr).

IV. Experiment Setup

A. Independent Variables

The experiment has one independent variable: controlled dynamics (DYN). Additionally, the data are separated into two datasets, one used to train the MRAC model and one used to validate it. Since the estimated parameters, whose uncertainty is also unknown, are found by solving a non-convex optimization problem it is necessary to assess the real performance of MRAC on a separated dataset as over-fitting is of real concern.

B. Controlled Dynamics

This study focused on testing whether the proposed model-based control algorithm, MRAC, is able to represent how humans adapt to changes in the dynamics of the system they are controlling. Therefore most of the runs had two transitions in the controlled dynamics. To make the runs more unpredictable, some of them kept the controlled dynamics constant. A controlled system is studied whose dynamics change between approximately a first order system ($DYN = 1$) and a second order system ($DYN = 2$) at frequencies close to the human crossover frequency range. The label $DYN = 121$ indicated the order of the dynamics' transition: first, second and first order.

The controlled dynamics (DYN) levels were: 121, 212, 1, 2. The dynamics labelled as $DYN = 212$ are approximately

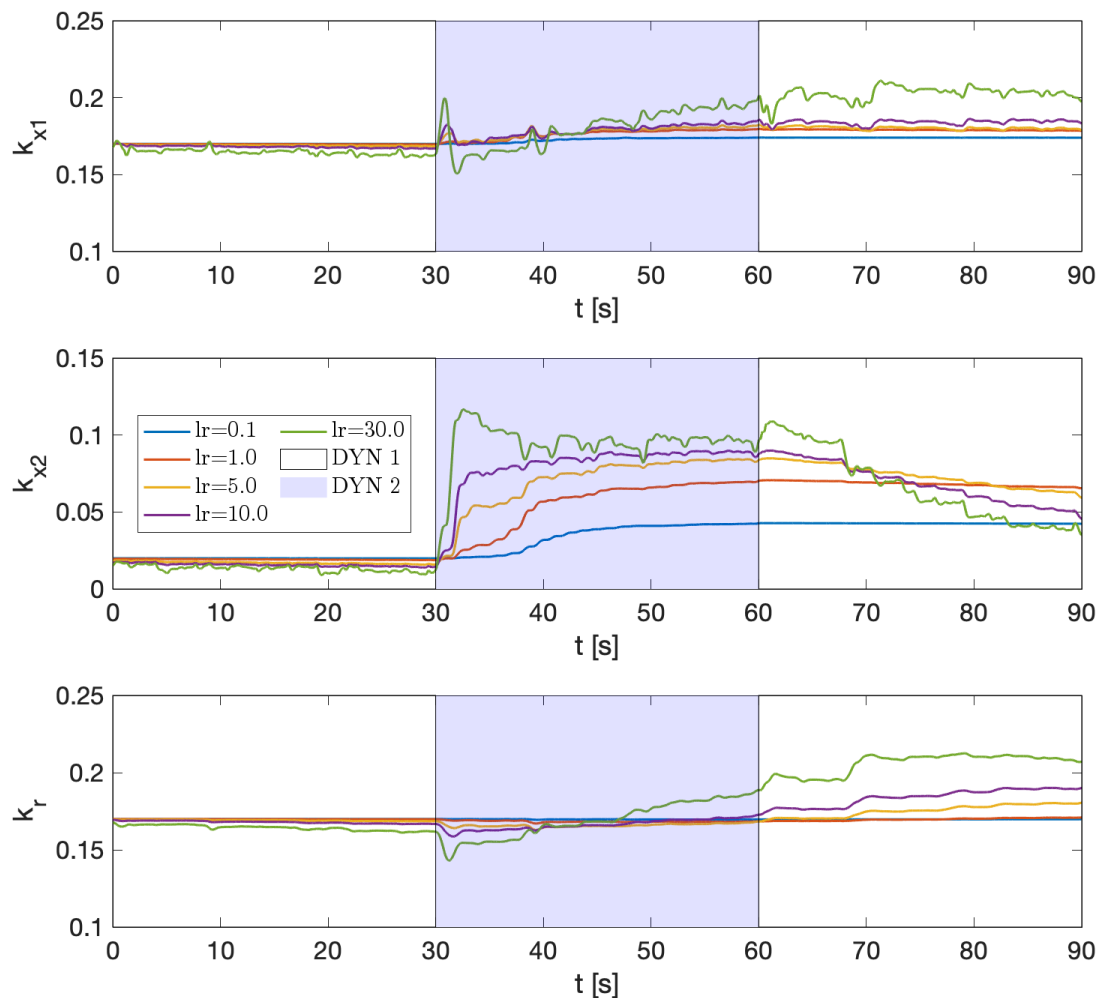


Fig. 7 Sensitivity of MRAC gains to the learning rate (lr).

equal to the ones of a first order system, for frequencies close to the crossover frequency. On the other hand the dynamics labelled as $DYN = 2$ are approximately equal to the ones shown by a second order system, for frequencies close to the crossover frequency. Sequences of these two basic labels, such as $DYN = 121$ and $DYN = 212$, represent the order in which the time-varying system dynamics vary over time.

In the condition $DYN = 121$ the participant controlled a time-varying system that for the first 30 seconds behaved approximately like a first order integrator, in the next 30 approximately as a double integrator and in the final 30 seconds it returned to the initial dynamics. Condition $DYN = 212$, was essentially the same but with the order of the dynamics swapped. In conditions $DYN = 212$ and $DYN = 2$ the participants controlled a steady state system for the whole duration of the run.

C. Dataset type

The objective was to gather two datasets: a testing dataset and a validation dataset. The testing dataset is used to estimate the parameters of the MRAC model, for it to approximate as well as possible the measured the human operator response. The validation dataset is used to verify that the MRAC model is not overfitting and to gain an understanding of the overall quality of the model predictions.

For convenience, the dataset type was included as an independent variable in the experiment. The validation dataset tracking signal used the same frequencies and amplitudes of the one in the training dataset but different phases. In Table 1 the amplitudes, frequencies and phases of the sinusoidal signal tracking signal for the testing and validation conditions are listed.

Table 2 Experimental settings across the main two independent variables for the testing dataset

DYN	ω_{b1} [rad/s]	ω_{b2} [rad/s]	k_{c1} [-]	k_{c2} [-]	G [1/s]	r_t
121	6	0.2	90	30	100	$r(t)$
212	0.2	6	30	90	100	$-r(t)$
1	6	6	90	90	100	$r(t)$
2	0.2	0.2	30	30	100	$r(t)$

The variables that controlled the time-varying behavior of the controlled system across the different conditions are shown in Table 2 for the testing dataset. The tracking signals were transformed by mirroring across the time and variable axis across conditions. This was done to make the signal as unpredictable as possible while keeping it as close as possible to the original in terms of tracking difficulty.

D. Apparatus

The experiment was performed in the Human-machine Interaction Laboratory (HMILab), in the faculty of Aerospace of the Delft University of Technology. The facility uses the software DUSIME for real-time simulations [49]. The participants sit behind a simple artificial horizon display as shown in Fig. 9. The display was updated at a frequency of 60 Hz and had a lag of approximately 20-25 milliseconds [48]. The control inputs were given through an hydraulically actuated side-stick, that could rotate only around the pitch axis, located at the right hand side of the participant. The stick torsional stiffness is 25 N m rad^{-1} , the damping coefficient is $0.22 \text{ N s rad}^{-1}$ and its inertia is 0.01 kg m^2 . The setup of the experiment is shown in Fig. 9.

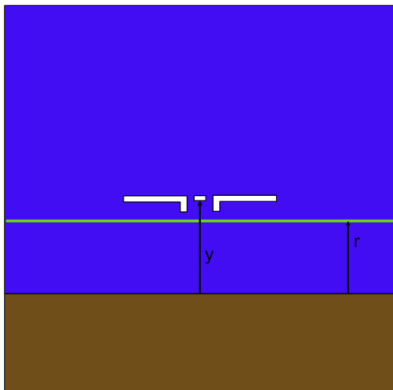


Fig. 8 Pursuit display used for the experiment.

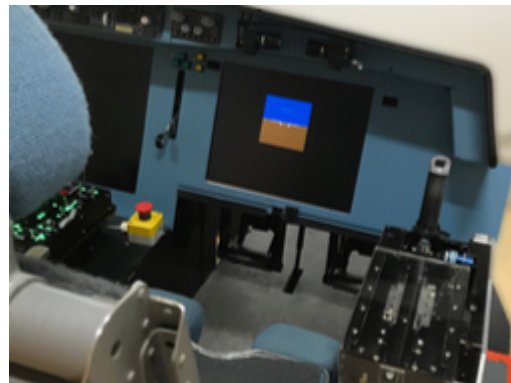


Fig. 9 Lab setup.

E. Participants

Ten participants performed the experiment. Since this study was focused on understanding the general adaptability of human controllers, participants were not required to be professional pilots. Participants were seven males and three females. Only two participants had experience with tracking tasks, prior to their participation in the experiment.

F. Procedures

The participants were given a briefing about the experiment and the manual control task, prior to performing the experiment. They were told that the goal of the experiment was to understand how humans adapt their control policy when controlling a time-varying experiment. They were made aware of the applicability of the study for producing safer

automation and inspire machine learning researchers to create always more adaptable artificial agents. The participants were encouraged to provide continuous control inputs at all time to follow the always moving target attitude indicator. After the briefing the demographics data were collected.

The participants started with a familiarization phase, which was short and ended when the participants were able to stabilize the system across the different conditions. It was followed by a training phase, with the objective to bring participants to asymptotic performance in tracking the signal. During the training phase the participants were exposed first to conditions 1 and 2 and then to the conditions with varying dynamics. The training phase ended when the participant could achieve consistent performance for three consecutive runs in the time-varying conditions. The measurement phase consisted of 22 runs. In the testing dataset five runs each were collected for the conditions with $DYN = 121$ and $DYN = 212$ and three each for the validation dataset. Furthermore three additional runs for both conditions $DYN = 1$ and $DYN = 2$ were added to make the experiment less predictable, i.e., to not have always expose participants to time-varying control systems. The order of runs was determined using an incomplete Latin square: considering each of the 22 runs as a separate entities, 22 different experimental sequences were obtained. Of these 10, randomly chosen were used in this experiment.

G. Dependent Measures

The following dependent measures were collected or estimated:

- The control outputs of the participants,
- The output of the system,
- The root mean square of the tracking error and the control input: RMSE and RMSU.
- The MRAC controller parameters are estimated using the control input,
- A time-windowed Variance Accounted For (VAF) measure of the MRAC model for the participant control output data is computed.

The MRAC controller parameters are also estimated on a subset of the measured data, that included a single transition in the dynamics of the controlled element. The subset of the data was obtained by excluding the last thirty seconds of each run.

H. Hypotheses

The following hypotheses were formulated:

- 1) MRAC is expected to fit equally well the steady state conditions with $DYN = 1$ and $DYN = 2$. So far parametric time-varying models of human control behavior fit better participants controlling systems that behave like a second order systems at frequencies close to crossover [36].
- 2) MRAC is expected to fit equally the controlled system dynamics's transition from $DYN = 1 \rightarrow 2$ and $2 \rightarrow 1$.
- 3) Since both conditions $DYN = 121$ and $DYN = 212$ have present both types of dynamics' transitions, MRAC is expected to have non significantly different values of learning rates across the two conditions.

Hypotheses 1 is used to test the baseline behavior of the MRAC controller and possibly expose limitation in using a time-invariant reference model, whose open-loop dynamics resemble the ones of a single integrator. Furthermore if *both* hypotheses 2 and 3 are accepted, then there would be evidence in favor of MRAC being a good model for human control policy adaptation

Most of the work in the paper though does not aim at testing a specific hypothesis but rather at investigating in depth the abilities and limitation of the MRAC controller.

V. Results

Firstly, the tracking performance and behavior of the participants is described, then it follows a statistical analysis on the data and a description of the estimated MRAC controllers. A linear-mixed effect model, using the subject ID as random factor, is fit to all the continuous variables averaged time-series. The orthogonal contrasts for the models are:

- "Steady vs Time-Varying": compares the steady state conditions, $DYN = 1$ and $DYN = 2$, against the time-varying ones, $DYN = 121$ and $DYN = 212$.
- $DYN = 1$ vs $DYN = 2$.
- $DYN = 121$ vs $DYN = 212$.

A. Main Results

Table 3 shows the number of datasets used for each of the DYN conditions and the tracking performance in terms of the Root Mean Square Error and the control output measured with Root Mean Square of the output signal (*RMSU*). Each measurement corresponds to the average of 5 runs for the conditions $DYN = 121$ and $DYN = 212$ and of 2 runs for the conditions $DYN = 1$ and $DYN = 2$. The number of measurements, present in the final dataset, is not equal to the number of participants for the conditions $DYN = 2$ and $DYN = 212$, since three and two measurements, respectively, have been excluded because those were considered outliers. An explanation for the detection of these outliers is presented in subsection V.B.

The *RMSE* is not significantly different between steady state conditions and time-varying ones, the same holds for conditions with $DYN = 121$ ($M = 1.008$) and $DYN = 212$ ($M = 1.029$). Instead, the *RMSE* is significantly lower between the condition $DYN = 1$ ($M = 0.981$) and $DYN = 212$ ($M = 1.075$) is significantly lower, $b = -0.077$ ($SE = 0.0189$), $t(22) = -4.115$, $p = 0.0005$. This has also been found repeatedly in the literature [41, 36].

The *RMSU* is not significantly different between steady state conditions and time-varying ones. It was significantly lower for condition $DYN = 1$ ($M = 0.1$ deg) than for condition $DYN = 2$ ($M = 0.14$), $b = -0.016$ ($SE = 0.004$), $t(22) = -4.04$, $p = 0.005$ and suggestively lower for condition $DYN = 121$ ($M = 0.117$) compared to condition $DYN = 212$ ($M = 0.137$), $b = -0.007$ ($SE = 0.004$), $t(22) = -1.84$, $p = 0.0778$. This weak trend is probably a consequence of the higher amount of time spent controlling $DYN = 2$ for condition $DYN = 212$ compared to condition $DYN = 121$.

Table 3 Summary of tracking performance and control output levels across the different conditions in the testing dataset

Dynamics	Num. Meas.	Mean <i>RMSE</i> [deg]	Std <i>RMSE</i> [deg]	Mean <i>RMSU</i> [deg]	Std <i>RMSU</i> [deg]
1	10	0.980	0.175	0.108	0.027
2	7	1.075	0.208	0.148	0.049
121	10	1.028	0.191	0.117	0.029
212	8	1.009	0.183	0.137	0.034

The *RMSE* is not significantly different between steady state conditions and time-varying ones, the same holds for conditions with $DYN = 121$ ($M = 1.008$) and $DYN = 212$ ($M = 1.029$). Instead the *RMSE* is significantly lower between the condition $DYN = 1$ ($M = 0.981$) and $DYN = 212$ ($M = 1.075$), $b = -0.077$ ($SE = 0.0189$), $t(22) = -4.115$, $p = 0.0005$. This was expected as second order dynamics are more difficult to control than first order ones.

Fig. 10 shows a box plot of the Variance Accounted For (VAF) of the aggregated control output for the found MRAC controllers. There are not significant changes in the VAF across steady and time-varying conditions. The VAF is significantly lower for the condition $DYN = 1$ ($M = 0.601$) compared to condition $DYN = 2$ ($M = 0.722$), $b = -0.053$ ($SE = 0.013$), $t(22) = -3.901$, $p < 0.001$. It is also significantly lower for $DYN = 121$ ($M = 0.644$) compared to $DYN = 212$ ($M = 0.71$), $b = -0.027$, $t(22) = -2.105$, $p = 0.047$.

Therefore overfitting of the model can be identified, as the VAF in the validation dataset for both time-varying DYN conditions is lower compared to their counterparts in the testing dataset.

Because of the nature of the adaptive controller, it is more informative to look at the time series data to analyse the ability of MRAC to capture how humans adapt. Fig. 11a and Fig. 11a show the time-series data of the VAF u of the MRAC controller computed over a moving time window of ten seconds, respectively, for the conditions with $DYN = 121$ and $DYN = 212$. The time windows of 10 seconds is able to show the ability of MRAC to capture the transient dynamics without having too much noise in the signal. It's possible to notice two things. First, the model fits the data better during the periods when $DYN = 2$ (with a mean of approximately 0.75) compared to the periods when $DYN = 1$ (with a mean of approx 0.6). This phenomenon is consistent with the VAFs shown in Fig. 10 where an equivalent difference is observed between $DYN = 1$ and $DYN = 2$. Second, while the VAF tends to raise when DYN changes from 1 to 2, it drops when DYN changes from 2 to 1. This indicates that the MRAC controller can capture how the participants adapt their control strategy when the order of the dynamics of the controlled element changes from first to second order, while the opposite is not true. After the DYN changes from 2 to 1 the VAF tends to increase to the baseline levels, established in the time-invariant condition with $DYN = 1$.

To understand why this phenomenon happens it is useful to compare the control output and the related system output for

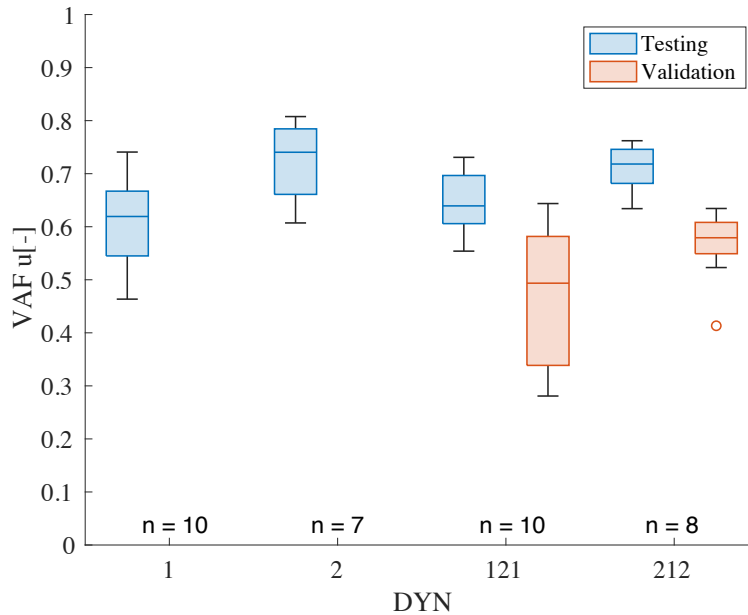
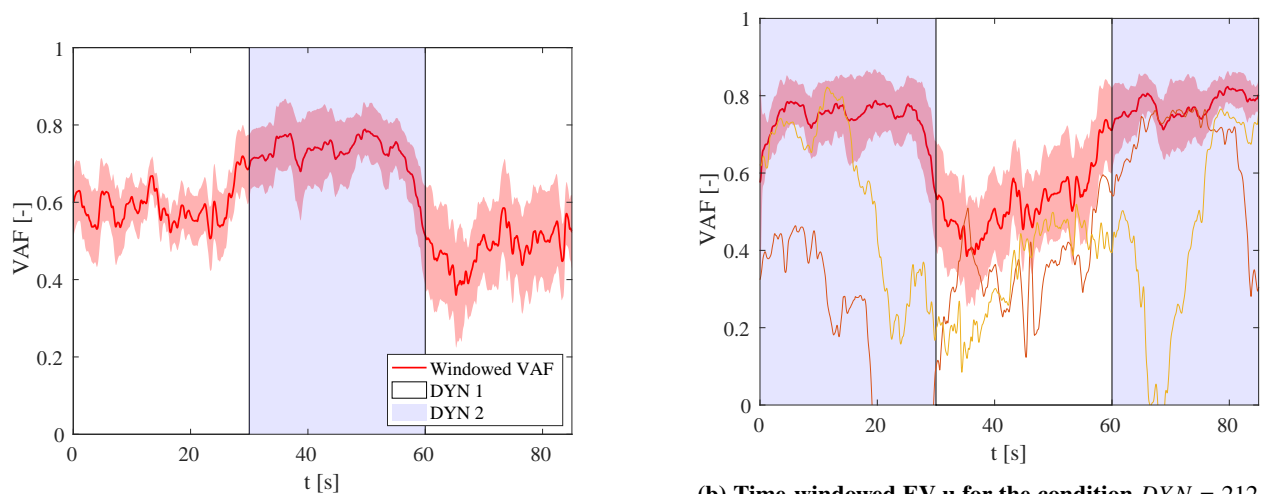


Fig. 10 VAF u values for the MRAC controller's output for the different conditions in the controlled dynamics and dataset types.



(a) Time-windowed VAF u for the condition $DYN = 121$

(b) Time-windowed EV u for the condition $DYN = 212$. In the figure, the VAFs of the two outlier MRAC controllers are shown with solid lines.

Fig. 11 Time-windowed VAF u of the MRAC controller estimated using a window size of 10 seconds. The solid line represents the mean value, while the shaded areas are one standard deviation away from the mean.

the MRAC controllers and the participants.

Since MRAC is a deterministic controller, the average control output data of each participant for each different dynamics' conditions were used to estimate the parameters of the controller. In this way, four different controllers were fit per participant. The validation dataset was used to benchmark and assess the generalization ability of MRAC, especially in conditions with $DYN = 121$ and $DYN = 212$.

Fig. 12 depicts the reference signal and the system's output aggregated across the participants for each of the different conditions in the testing dataset. The aggregated output shows the mean response in a solid line and the standard deviation at all times with a shaded area. The output of the participants is shown in blue while the output of the MRAC models in red. The participants controlled the steady-state system with $DYN = 1$ and $DYN = 2$ consistently. On the

other hand, the tracking ability of the participants decreases just after a change in dynamics in the controlled system. In particular, they tend to overshoot the target, with a signal in average 35% than the reference signal, when the dynamics change from 1 to 2 and undershoot it when the opposite happens, the output signal is 30% lower than the reference signal. For the system with $DYN = 121$, after the 30-second mark in Fig. 12 (c), this phenomenon can be observed: the next two stationary points in the reference signal are both significantly overshoot. The same can be observed for $DYN = 212$, in Fig. 12 (d) after the dynamics transition at 60 seconds. The standard deviation of the output also increases after a dynamical transition. The tracking performance gradually improves after the transition, as participants adapt to the new dynamics.

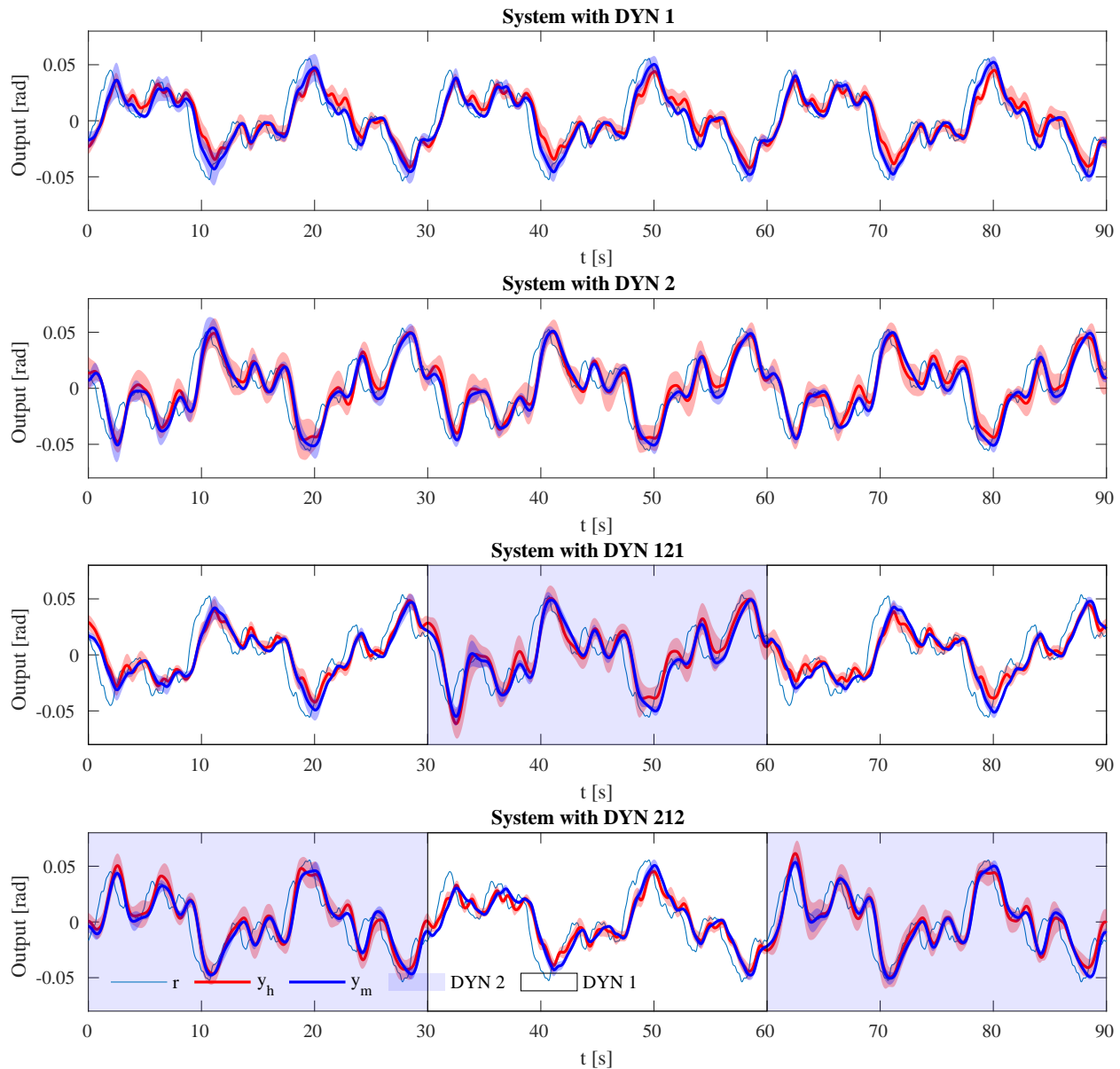


Fig. 12 Standard deviation shaded plot of the aggregated output of the participants, y_h and of the MRAC controllers, y_m , for the four DYN conditions

Fig. 13 compares the aggregated control output of the participants and the estimated MRAC controller. For condition $DYN = 121$, between $t = 30$ seconds and $t = 40$ seconds, MRAC captures the transient control behavior shown by participants, who have a higher control activity and tend to overshoot the target reference signal as shown in Fig. 12. On the other hand, MRAC is not equally able to capture the $DYN = 2$ to $DYN = 1$ transition: as it can be observed in

Fig. 13 (c) and (d) the control output is lower and it has higher frequency component compared to the signal generated by the participants. To better understand why the MRAC controller behaves this it is necessary to have a look at its gains and at the prediction error.

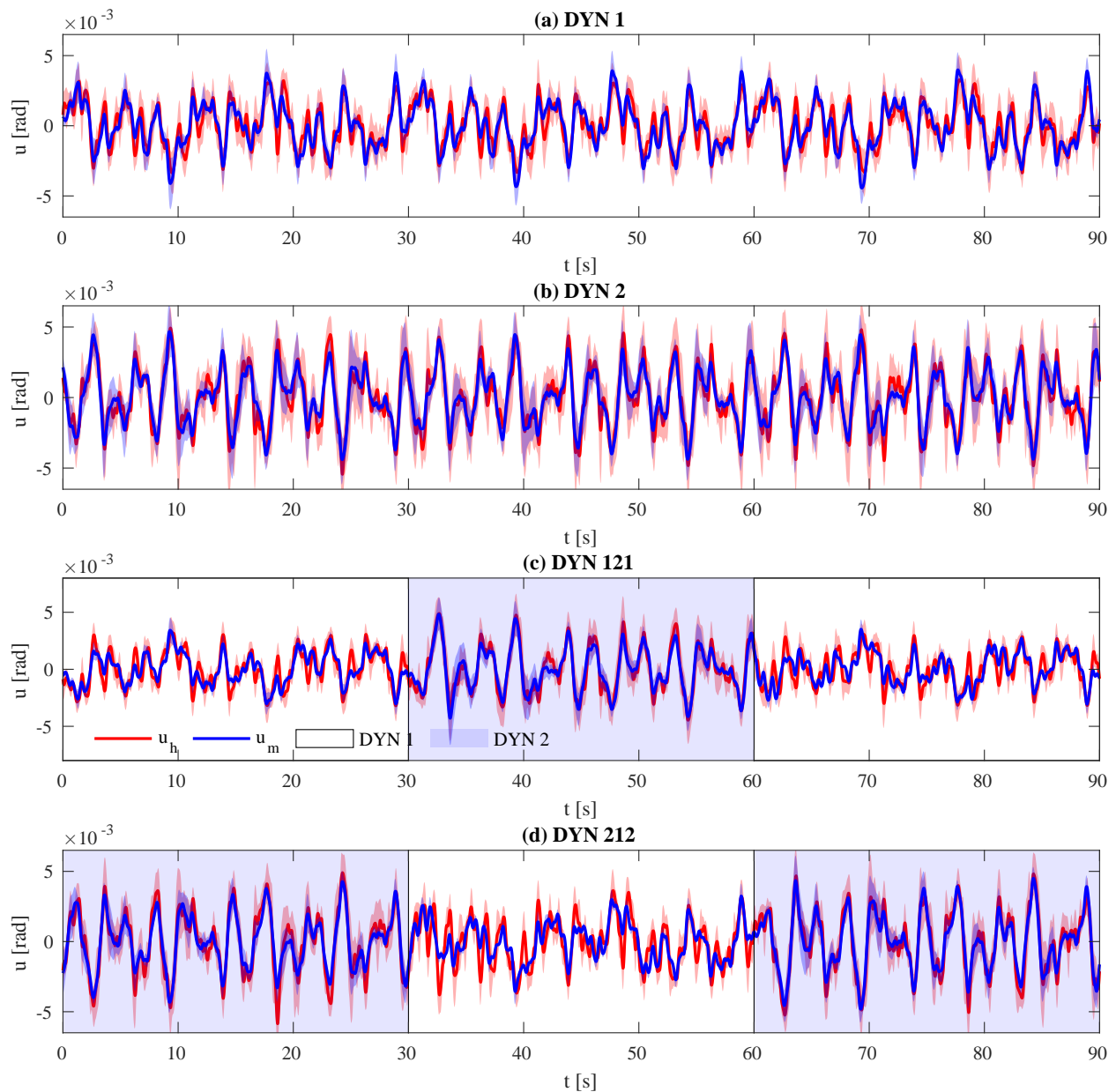


Fig. 13 Standard deviation shaded plot for the aggregated control output of the participants u_h and the MRAC models u_m

Fig. 14 shows the aggregated prediction error, e_p , for the estimated controllers. The prediction error is the difference between the observed output the one generated by the internal model. This parameter, together with the state x , r and learning rates drive the adaptation of the active MRAC gains. The prediction error is highest after the two dynamical transitions at $t = 30$ seconds and $t = 60$ seconds. As expected the error never converges to zero, due to the presence of the delay in the controller. Even though the magnitude of the prediction error is comparable for the transitions $DYN 1 \rightarrow 2$ and $2 \rightarrow 1$ it is expected that the gains change more for the transition $DYN 1 \rightarrow 2$ since the values of the state x and its derivative are higher.

The gains related to the steady state conditions are analyzed first. The gain k_{x1} is significantly higher in condition

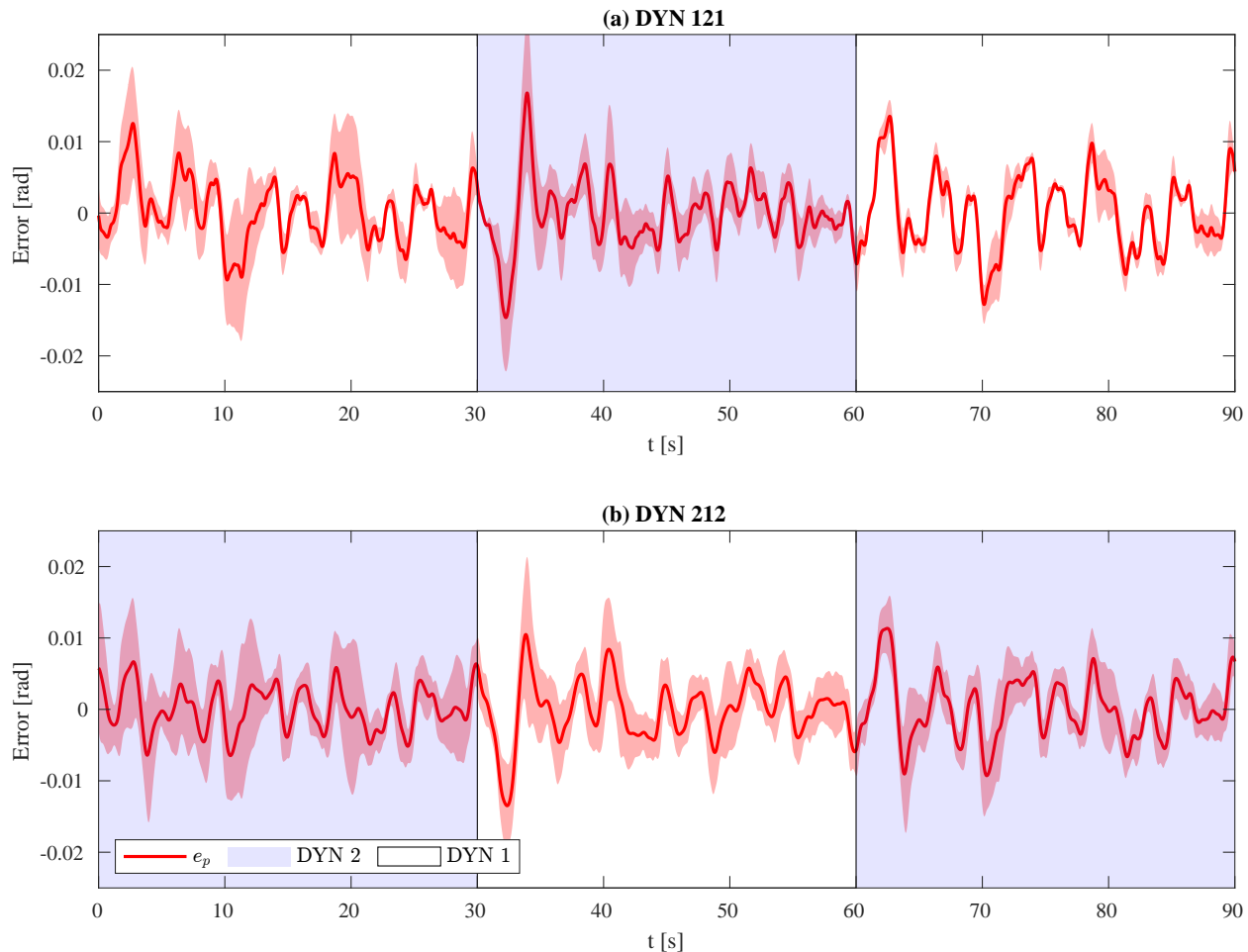


Fig. 14 Time series of the prediction error for $DYN = 121$ and $DYN = 212$. The prediction error is the difference between the system output and the expected output of the internal model

$DYN = 1$ ($M = 0.179$) then in condition $DYN = 2$ ($M = 0.145$), $b = 0.037$ ($SE = 0.01$), $t(6) = 3.696$, $p = 0.01$. The gain k_r is suggestively higher in condition $DYN = 1$ ($M = 0.163$) then in condition $DYN = 2$ ($M = 0.153$), $b = 0.033$ ($SE = 0.015$), $t(6) = 2.24$, $p = 0.06$. The gain k_{x2} is significantly lower in condition $DYN = 1$ ($M = 0.045$) then in condition $DYN = 2$ ($M = 0.070$), $b = 0.024$ ($SE = 0.002$), $t(6) = 11.81$, $p < 0.001$. These results, i.e., higher state derivative gains and lower state feedback and feedforward gains for condition $DYN = 2$ compared to condition $DYN = 1$, are in line with the findings of El et al., who modelled human control behavior in pursuit tasks [41].

In Fig. 15 it is shown a time-series of the aggregated MRAC gains: in solid lines are shown the mean and the colored area encloses values that less than one standard deviation away from the mean. Fig. 16 also shows the time-series of the MRAC gains for each participant. As it can be observed most of the variance in the gain values is across different participants. While the feedforward gain k_r (in black) and the state gain k_{x1} (in blue) do not show significant changes in their mean values, the gain related to the state derivative k_{x2} (in red) changes significantly with after a change in the controlled dynamics. For the conditions $DYN = 121$ after the 30 second mark and for the conditions $DYN = 212$ after the 60 seconds mark, k_{x2} increases from a value of 0.05 to 0.085. This means that the contribution of "lead" increases as expected. On the other hand when the dynamics of the system change from second to first order the gain k_{x2} gradually decreases from a the value of 0.085 to 0.06. As we noted before in the sensitivity analysis, see subsection III.F, the rate of change of the gain is not symmetric to the order of the transition in dynamics: a change from $DYN = 1$ to $DYN = 2$ leads to a much faster adaptation in the response compared to a change from $DYN = 2$ to $DYN = 1$. The other two gains change mostly across participants, depending on the control activity and tracking ability.

To have a complete picture of the how the gains adapt through time it is useful to analyse the found learning rates for the MRAC controllers.

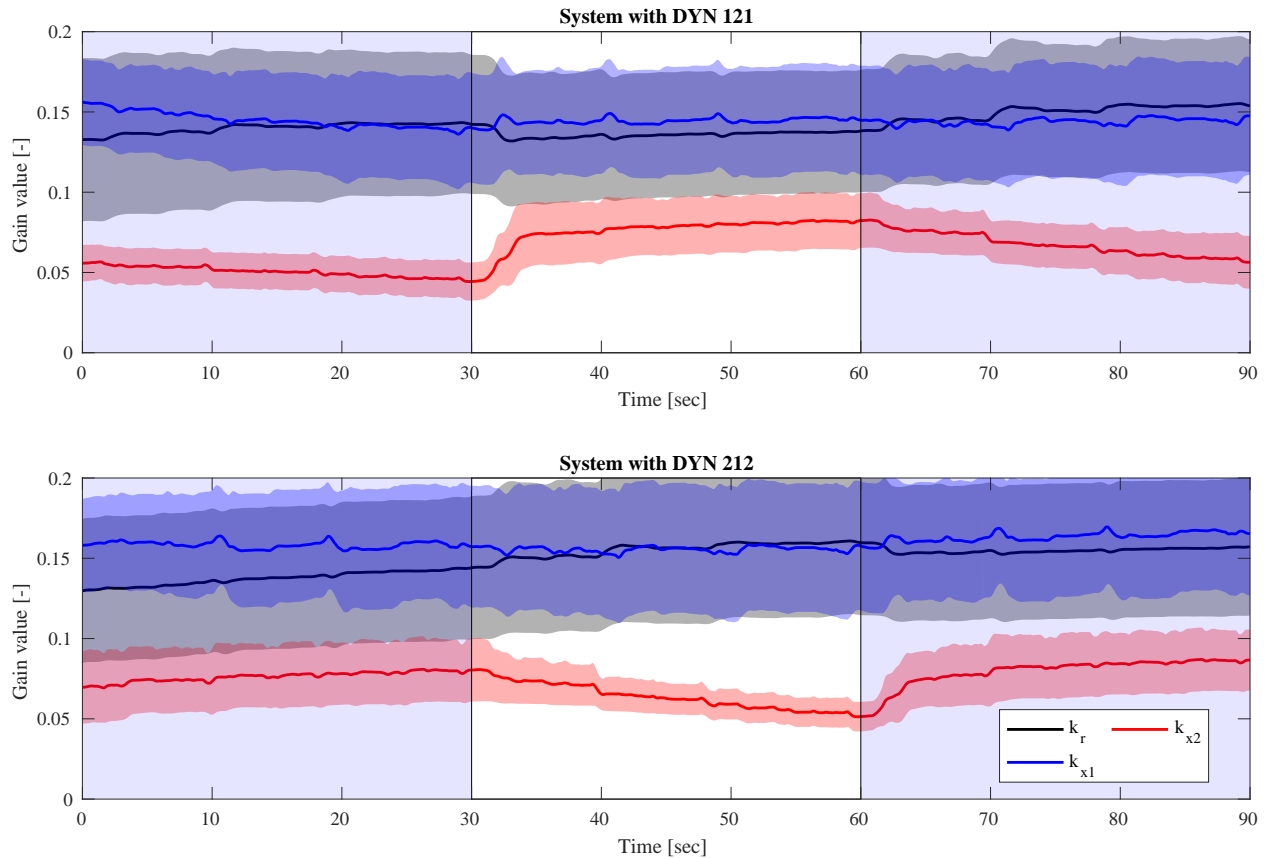


Fig. 15 Time-series of the aggregated MRAC gains.

Fig. 19 depicts the learning rates associated with the different gains. As expected, the learning rates are not significantly different across the two conditions $DYN = 121$ and $DYN = 212$, since both types of transitions are present in the both the time-varying conditions.

Fig. 17a and Fig. 17b show the parameters ω_c (crossover frequency) and τ (delay) of the internal model and of the MRAC controller, respectively.

No significant changes across conditions have been found for the internal model crossover frequency parameter. The value delay is only significantly different between $DYN = 1$ ($M = 0.228$ s) and $DYN = 2$ ($M = 0.207$ s), $b = 0.009$ ($SE = 0.003$), $t(22) = 2.64$, $p = 0.014$. Notice though that the effect size is very small. No other significant changes across conditions have been found. Overall these results in line with the findings of [36].

B. Outlier analysis

The following data have been excluded from the aggregated analysis of the MRAC gains: data related to $DYN = 2$ and $DYN = 4$ for "subject02", $DYN = 2$ for "subject05" and $DYN = 2$ and $DYN = 4$ for "subject07". These sets of runs were excluded because the windowed VAF, resulting from the fitted MRAC controller, dropped to 0 for certain periods. Therefore the estimated parameters are not indicative of the controller behavior and are therefore excluded from further analysis. The VAF corresponding to these set of runs is still shown in Fig. 11b. The inability of MRAC to model the data is caused by the poor and inconsistent tracking performance that the participants exhibited. At the same time we can deduce that MRAC is relatively brittle when fitting it to participants that are not expert controllers.

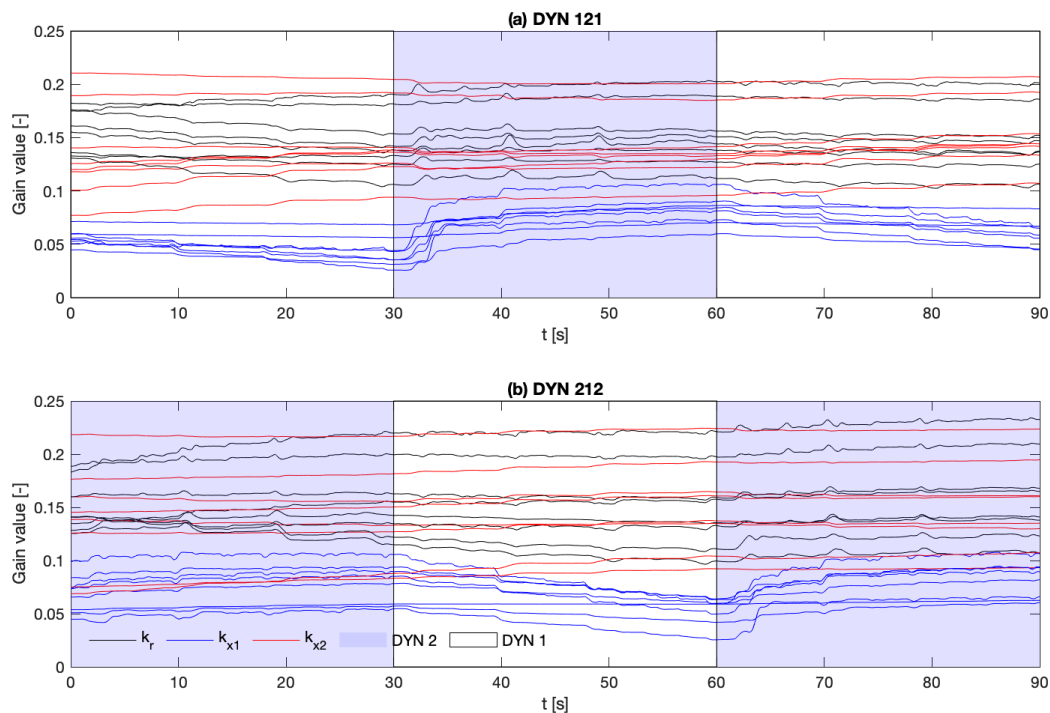
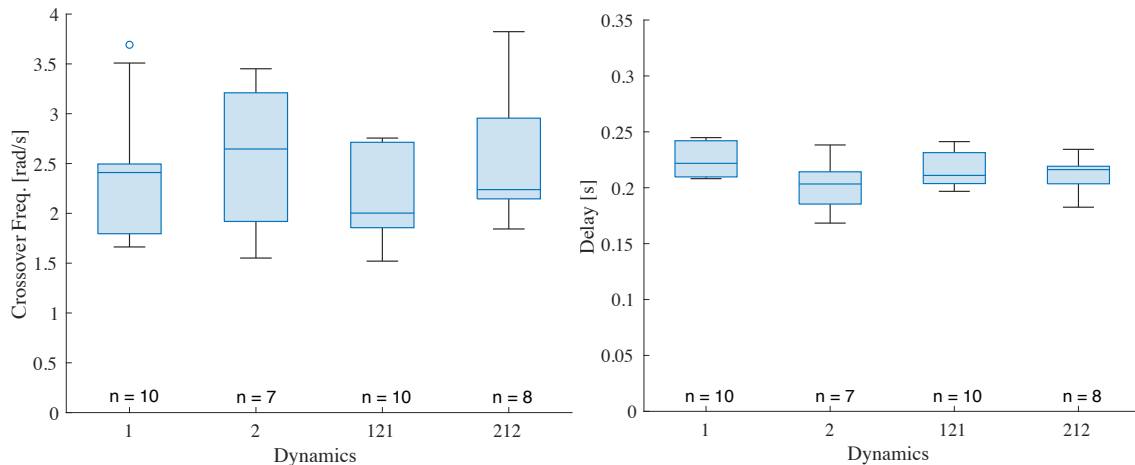


Fig. 16 Time series of the MRAC gains found for each participant.



(a) Box plot of the crossover frequency parameter ω_c across the different dynamics conditions **(b)** Box plot of the delay parameter τ across the different dynamics conditions

Fig. 17 Box plots of the internal model parameters.

C. MRAC controller fitted to one dynamical transition

In the previous section, we observed that MRAC is unable with the same learning rate to fit the human data equally well the time-varying runs. This section investigates whether fitting MRAC to each transition separately can increase the quality of fit. This is achieved by fitting the MRAC controller using only the first 70 seconds of every run for conditions with $DYN = 121$ and $DYN = 212$. Since there is only one transition of the controlled dynamics we relabel the datasets to $DYN = 12$ and $DYN = 21$, respectively.

The MRAC gains found in this way are shown in Fig. 18, the aggregated mean is shown with a solid line while the shaded areas include values closer than one standard deviation from the mean. It is useful to compare these time-series with the ones shown in Fig. 15. Comparing condition $DYN = 12$ with $DYN = 121$, there are no significant differences in the mean of the gains, nor in the rate of change of gain k_{x2} , which transition from a value of 0.05 to 0.09 in approximately 4 seconds for both conditions. The main difference resides in the standard deviation of the gains which is reduced from 0.04 to 0.02 from the gain k_{x1} .

If we compare conditions with $DYN = 212$ and $DYN = 21$ we can instead observe a difference in the time evolution of the mean of the parameter k_{x2} . In both cases the value of $k_{x2} \approx 0.08$ while the controlled system has $DYN = 2$, but after the transition to $DYN = 1$ the mean of the gain drops after 30 seconds to a value of approximately 0.03, in condition $DYN = 21$ while in the condition $DYN = 121$ the mean has a final value value of 0.06.

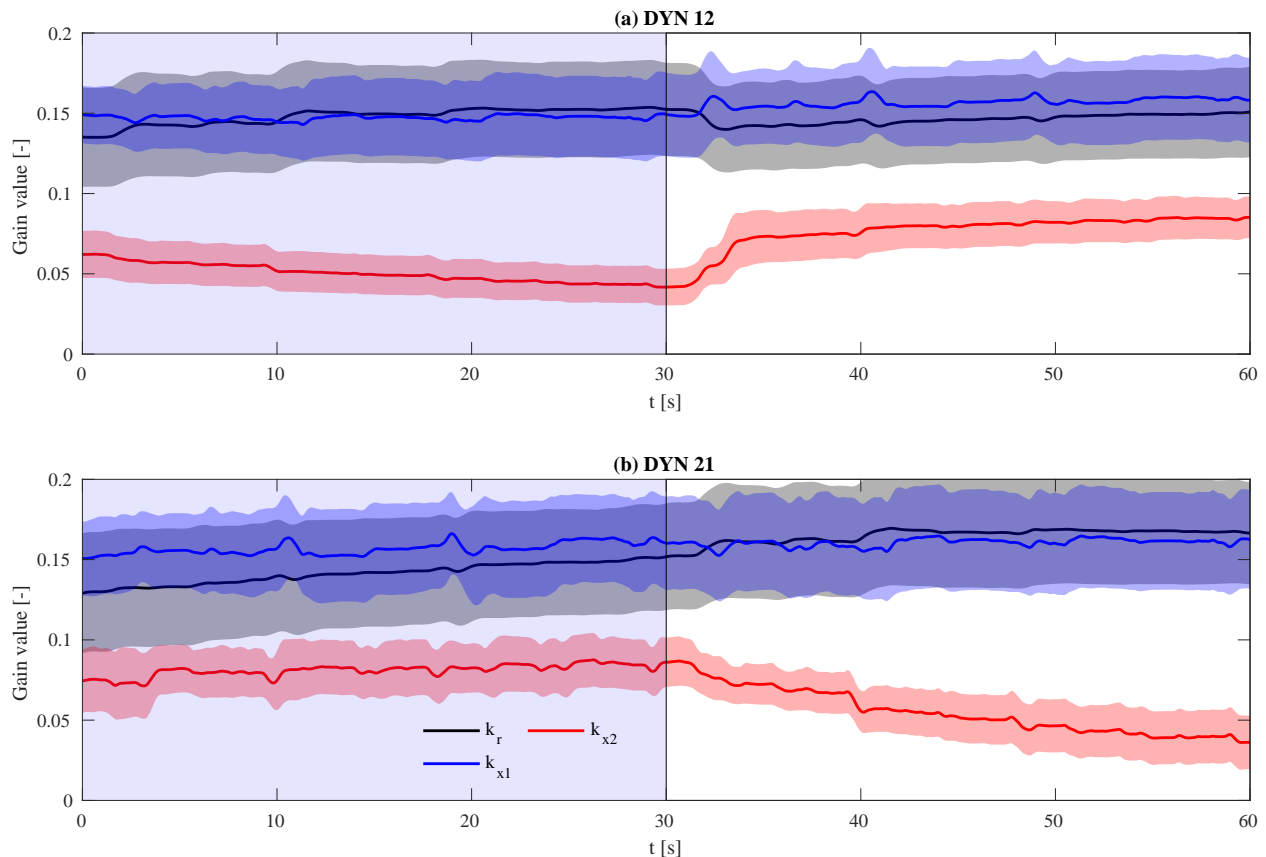


Fig. 18 Standard deviation shaded plot of the aggregated MRAC gains, estimated using only the first sixty seconds of the each run

It is therefore useful to analyse the learning rates of the MRAC controller for conditions $DYN = 12$ and $DYN = 21$. In Fig. 20 is shown a box plot of the MRAC learning rates for the two conditions. All the median learning rates for condition $DYN = 12$ are lower than ones for $DYN = 21$. This is expected, as the transition $DYN = 1 \rightarrow 2$ induces stronger adaptation compared to $DYN = 2 \rightarrow 1$ while keeping the learning rates constant. In particular for $DYN = 12$ the median $\gamma_{x1} = 12 \text{ rad}^{-1}$ and $\gamma_{x2} = 5 \text{ s}^{-1} \text{ rad}^{-1}$ while for $DYN = 21$ the median $\gamma_{x1} = 22 \text{ rad}^{-1}$ and $\gamma_{x2} = 18 \text{ rad}^{-1} \text{ s}^{-1}$. These values are also higher compared to ones found for $DYN = 121$ and $DYN = 212$ for which the median $\gamma_{x1} = 13.8 \text{ rad}^{-1}$ and $\gamma_{x2} = 6 \text{ rad}^{-1} \text{ s}^{-1}$.

Fig. 21a and Fig. 21b show the windowed VAF of the obtained MRAC controllers on the reduced datasets. Notwithstanding an increase in the learning rate, the mean windowed VAF at the transition for condition $DYN = 12$ is not significantly different from its value for condition $DYN = 121$. The same holds for conditions $DYN = 212$ and $DYN = 21$.

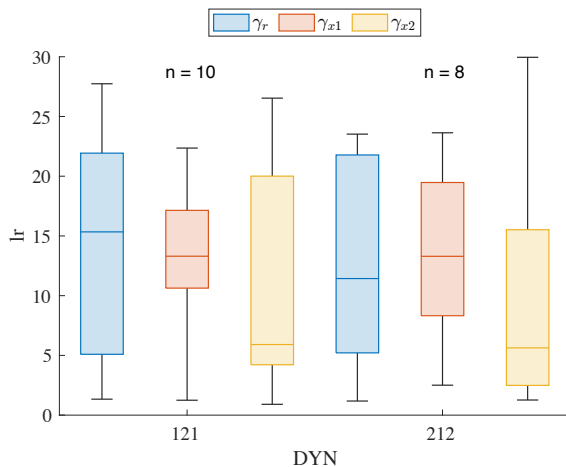


Fig. 19 Learning rates found using the entire run

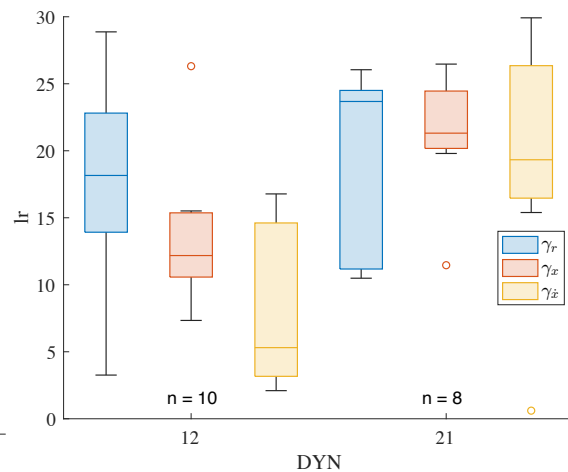
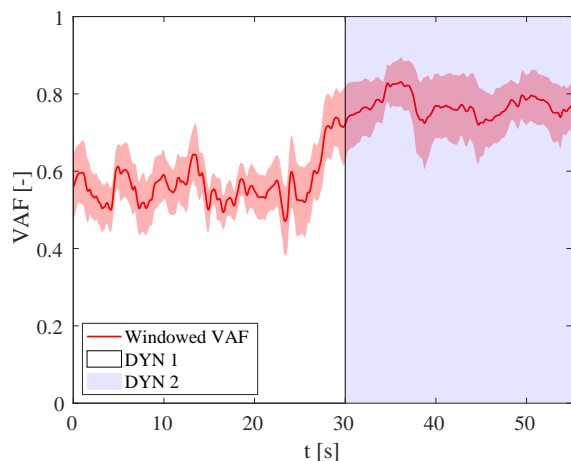
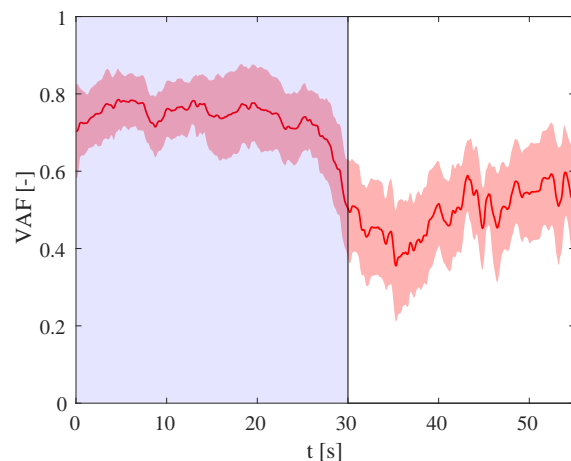


Fig. 20 Learning rates parameters found using only the first sixty seconds of the measured run



(a) Time-windowed VAF u for the condition $DYN = 12$



(b) Time-windowed VAF u for the condition $DYN = 21$

Fig. 21 Time windowed VAF u of the MRAC controller estimated using a window size of 10 seconds and a reduced dataset.

VI. Discussion

The purpose of this paper was to evaluate whether a model-based adaptive controller, such as MRAC, would be a viable candidate to model and predict human adaptive behavior under changes of the controlled dynamics. In particular, two dynamical transitions were chosen to assess the validity of the adaptive law. In subsection VI.A it is discussed the ability of MRAC of modelling human control behavior and adaptation. In subsection VI.B are presented the main findings and implications of the comparison of the MRAC coefficients found on the truncated dataset (only one transition) with respect to the ones found using the full dataset, whose runs have two show two transitions in the controlled dynamics. Finally, in subsection VI.C we summarize the limitations of MRAC and propose future research directions.

A. MRAC ability to model human adaptation

We first assess the MRAC ability to capture the steady-state control behavior of the participants. MRAC unfortunately is not very effective in capturing steady-state behavior. It has median VAF of 0.64 and 0.75 for conditions with $DYN = 1$ and $DYN = 2$ which are lower compared to the state of the art values of 0.77 and 0.85 reported by a time-varying parametric model [36]. It also implies that MRAC is inherently less able to approximate human control behavior for the first-order system that we used in the experiment. This finding is compatible with the literature [36]. Therefore

hypotheses 1 is rejected.

This limitation stems probably from the limited amount of possible controllers that can be effectively represented by having a first-order integrator as an open-loop internal model. Furthermore, the internal model is time-invariant and assumes the same asymptotic control behavior for both first and second-order systems. Second-order systems are arguably harder to control and especially not expert controllers make often mistakes and have a lower control performance.

MRAC ability to approximate human time-varying behavior strongly depends on the type of transition that the controlled system goes through. The transition from first to second order dynamical systems are well approximated. On the contrary, there is a sharp drop in the aggregated VAF of the control output of the participants when the controlled system transition from second to first order. This could be caused by a transient behavior that cannot be approximated by MRAC. On the other hand it also possible, that while ideally this transient behavior could be captured by the set of gains k_r, k_{x1}, k_{x2} , they don't change fast enough. This implies that hypothesis 2 must be rejected. Overall the experimental results provide evidence that the current adaptive law is not a good approximation of the program that humans operators use to adapt their control strategy.

B. MRAC fit to partial data

The MRAC learning rates across conditions $DYN = 121$ and $DYN = 212$ do not significantly vary which implies that hypotheses 3 is accepted. Nonetheless, the found learning rates are not appropriate to model both types of transition in the controlled dynamics. To investigate this issue further, the MRAC controllers were fit only using the data of the first 70 seconds of conditions $DYN = 121$ and $DYN = 212$. This was done to test if the transient control behavior, shown while the controlled system transitions from first to second-order dynamics, could be better approximated. Unfortunately, the results suggest that while the derivative gain changed twice as fast in condition $DYN = 21$ compared to condition $DYN = 212$, the increased learning rate was still not sufficient to improve the VAF of the control output immediately after the transition. Higher learning rates are, unfortunately, not possible since the MRAC controller can easily become unstable. A nominal MRAC controller has theoretical convergence guarantees but for the one considered here, these guarantees do not hold because of the presence of a delay. The presence of the delay makes the optimization itself harder to perform. Without proper bounds on the parameter search, an optimization algorithm can easily pick values of delay or learning rate that make the optimization diverge.

C. Limitations and future research

The use of a MRAC controller to model human time-varying behavior, as investigated in this paper, is limited by the adaptive control law, by the time-invariant internal reference model and by the fact that it requires a reference signal. It is unable to generate complex plans, possibly subject to constraints, that humans can carry out such as picking up objects, parking a vehicle or driving avoiding obstacles.

Future experiments should include conditions where the human controller is asked to perform control task with a large space of possible solutions, where a complex plan, that unrolls over many seconds, is a necessary condition to solve it. In the aerospace field landing without instruments, performing an obstacle avoidance maneuver, reaching a hovering target under disturbances with an helicopter are all tasks that fit those criteria. Any model that tries that has the potential to model real-world human behavior must be able to deal with a high dimensional input (such as images) and to generate a long sequence of planned actions.

Going back to MRAC, to improve the current formulation it could be worthwhile to couple it with an estimator of the currently controlled system dynamics that could drive the adaptation of the internal model. Furthermore a different adaptation law could be investigated, potentially causing a stronger adaptation even in absence of large errors but it is strongly discouraged tinkering with adaptation laws that are not backed by a strong mathematical justification. There are adaptive laws for the main MRAC gains that make the adaptation more stable, such as the σ or the MIT rules [40], they could help achieving higher learning rates while maintaining stability.

A promising line of research focuses on controllers based on model-based optimal control. These controllers have been already used in capturing complex dynamical behavior shown by animals [17]. There is substantial evidence that animals use internal dynamical models to move and plant their actions, furthermore, they can do open-loop planning [5]. In particular, if they aim is to identify the behavior of human controllers, inverse reinforcement learning [50] and probabilistic inference are promising frameworks to work with. Inverse reinforcement learning aims at estimating the rewards function (or intention) of the observed control behavior and from that derive an explicit control policy for the controller. Probabilistic inference relaxes the optimality condition and allows to estimate suboptimal controllers [51].

Neural networks could be also used as function approximations to learn non-linear systems and for the policy itself to allow for a wide range of complex behaviors.

VII. Conclusion

This paper evaluated the use of a model-based adaptive control algorithm, Model-Based Reference Control (MRAC), to model the adaptive control policy shown by a human operator while controlling a time-varying system in a pursuit tracking task. An experiment was conducted where participants ($n = 10$) performed a pursuit tracking task while controlling a system that changed its dynamics twice. The dynamics are approximately the ones shown by either first order and second order system at frequencies close to crossover.

The choice of this model was motivated by the fact that biological systems make use of internal models of the worlds for sensorimotor control and by "Predictive Codings", a theory in neuroscience that states that brains constantly try to predict incoming sensory input and mostly process prediction errors that drives the change in behavior shown by the organism. The reference model of the MRAC controller was designed by closing the loop on single integrator with delay, aimed at approximating the open-loop dynamics of a proficient controller. The parameters of the model were estimated with a non-linear optimization techniques aimed at minimizing the error between the MRAC output and the response of the participant averaged across multiple runs.

MRAC has shown the ability to approximate the transient control policy of the participants while the controlled dynamics transitioned from approximately first to second order, an adaptation that requires an increase in rate feedback control. On the other hand, it cannot effectively, with the same learning rates, approximate the change in the control policy of the participants for the opposite transition from second to first-order dynamics. It is hypothesized that the gains responsible for lead changes too slowly compared to the speed of human adaption. Furthermore, MRAC is not able to achieve state of the art results in approximating the steady steady-state dynamics of the human controllers, especially while they are controlling first-order systems.

While MRAC was only partially successful at predicting human control behavior under different changes of the controlled system, its highlighted limitations, innovations in machine learning (aimed at encoding always richer internal models), in the neuroscience of predictive coding and in optimal control can be useful in creating model-based controllers with human-like abilities to adapt their control policies.

References

- [1] K. He et al. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification". In: *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015 IEEE International Conference on Computer Vision (ICCV). Dec. 2015, pp. 1026–1034. doi: 10.1109/ICCV.2015.123.
- [2] L. R. Young. "On Adaptive Manual Control". In: *IEEE Transactions on Man-Machine Systems* 10.4 (Dec. 1969), pp. 292–331. ISSN: 2168-2860. DOI: 10.1109/TMMS.1969.299931.
- [3] M. Mulder et al. "Manual Control Cybernetics: State-of-the-Art and Current Trends". In: *IEEE Transactions on Human-Machine Systems* 48.5 (Oct. 2018), pp. 468–485. ISSN: 2168-2305. DOI: 10.1109/THMS.2017.2761342.
- [4] K. Friston. "Does Predictive Coding Have a Future?" In: *Nature Neuroscience* 21.8 (Aug. 1, 2018), pp. 1019–1021. ISSN: 1546-1726. DOI: 10.1038/s41593-018-0200-7. URL: <https://doi.org/10.1038/s41593-018-0200-7>.
- [5] D. McNamee and D. Wolpert. "Internal Models in Biological Control". In: (Oct. 3, 2018). ISSN: 2573-5144. DOI: 10.17863/CAM.30459. URL: <https://www.repository.cam.ac.uk/handle/1810/283097>.
- [6] R. Miall and D. Wolpert. "Forward Models for Physiological Motor Control". In: *Four Major Hypotheses in Neuroscience* 9.8 (Nov. 1, 1996), pp. 1265–1279. ISSN: 0893-6080. DOI: 10.1016/S0893-6080(96)00035-4. URL: <http://www.sciencedirect.com/science/article/pii/S0893608096000354>.
- [7] M. Kawato. "Internal Models for Motor Control and Trajectory Planning". In: *Current Opinion in Neurobiology* 9.6 (Dec. 1, 1999), pp. 718–727. ISSN: 0959-4388. DOI: 10.1016/S0959-4388(99)00028-8. URL: <http://www.sciencedirect.com/science/article/pii/S0959438899000288>.
- [8] A. J. Bastian. "Moving, Sensing and Learning with Cerebellar Damage". In: *Current opinion in neurobiology* 21.4 (Aug. 2011), pp. 596–601. ISSN: 0959-4388. DOI: 10.1016/j.conb.2011.06.007. pmid: 21733673. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3177958/> (visited on 01/14/2021).

- [9] C. J. Stoodley. “The Cerebellum and Neurodevelopmental Disorders”. In: *Cerebellum (London, England)* 15.1 (Feb. 2016), pp. 34–37. ISSN: 1473-4222. DOI: 10.1007/s12311-015-0715-3. pmid: 26298473. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4811332/> (visited on 01/13/2021).
- [10] D. Wolpert, Z. Ghahramani, and M. Jordan. “An Internal Model for Sensorimotor Integration”. In: *Science* 269.5232 (Sept. 29, 1995), p. 1880. DOI: 10.1126/science.7569931. URL: <http://science.sciencemag.org/content/269/5232/1880.abstract>.
- [11] S. Herculano-Houzel. “The Human Brain in Numbers: A Linearly Scaled-up Primate Brain”. In: *Frontiers in Human Neuroscience* 3 (2009). ISSN: 1662-5161. DOI: 10.3389/neuro.09.031.2009. URL: <https://www.frontiersin.org/articles/10.3389/neuro.09.031.2009/full> (visited on 01/13/2021).
- [12] J. Kaplan et al. *Scaling Laws for Neural Language Models*. Jan. 22, 2020. arXiv: 2001.08361 [cs, stat]. URL: <http://arxiv.org/abs/2001.08361> (visited on 01/14/2021).
- [13] D. M. Wolpert, R. C. Miall, and M. Kawato. “Internal Models in the Cerebellum”. In: *Trends in Cognitive Sciences* 2.9 (Sept. 1, 1998), pp. 338–347. ISSN: 1364-6613. DOI: 10.1016/S1364-6613(98)01221-2. URL: <http://www.sciencedirect.com/science/article/pii/S1364661398012212> (visited on 01/14/2021).
- [14] H. Tanaka et al. “The Cerebro-Cerebellum as a Locus of Forward Model: A Review”. In: *Frontiers in Systems Neuroscience* 14 (2020). ISSN: 1662-5137. DOI: 10.3389/fnsys.2020.00019. URL: <https://www.frontiersin.org/articles/10.3389/fnsys.2020.00019/full> (visited on 01/14/2021).
- [15] S. Shipp. “Neural Elements for Predictive Coding”. In: *Frontiers in Psychology* 7 (2016). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2016.01792. URL: <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01792/full>.
- [16] E. Todorov and M. I. Jordan. “Optimal Feedback Control as a Theory of Motor Coordination”. In: *Nature Neuroscience* 5.11 (Nov. 1, 2002), pp. 1226–1235. ISSN: 1546-1726. DOI: 10.1038/nn963. URL: <https://doi.org/10.1038/nn963>.
- [17] E. Todorov. “Optimality Principles in Sensorimotor Control”. In: *Nature Neuroscience* 7.9 (Sept. 1, 2004), pp. 907–915. ISSN: 1546-1726. DOI: 10.1038/nn1309. URL: <https://doi.org/10.1038/nn1309>.
- [18] K. Friston. “What Is Optimal about Motor Control?” In: *Neuron* 72.3 (Nov. 3, 2011), pp. 488–498. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2011.10.018. URL: <http://www.sciencedirect.com/science/article/pii/S0896627311009305>.
- [19] J. Diedrichsen, R. Shadmehr, and R. B. Ivry. “The Coordination of Movement: Optimal Feedback Control and Beyond”. In: *Trends in Cognitive Sciences* 14.1 (Jan. 1, 2010), pp. 31–39. ISSN: 1364-6613. DOI: 10.1016/j.tics.2009.11.004. URL: <http://www.sciencedirect.com/science/article/pii/S1364661309002587>.
- [20] S. H. Scott. “The Computational and Neural Basis of Voluntary Motor Control and Planning”. In: *Trends in Cognitive Sciences* 16.11 (Nov. 1, 2012), pp. 541–549. ISSN: 1364-6613. DOI: 10.1016/j.tics.2012.09.008. URL: <https://doi.org/10.1016/j.tics.2012.09.008> (visited on 09/02/2020).
- [21] K. J. Friston et al. “Action and Behavior: A Free-Energy Formulation”. In: *Biological Cybernetics* 102.3 (Mar. 2010), pp. 227–260. ISSN: 1432-0770. DOI: 10.1007/s00422-010-0364-z. pmid: 20148260.
- [22] K. Friston. “The Free-Energy Principle: A Unified Brain Theory?” In: *Nature Reviews Neuroscience* 11.2 (Feb. 1, 2010), pp. 127–138. ISSN: 1471-0048. DOI: 10.1038/nrn2787. URL: <https://doi.org/10.1038/nrn2787>.
- [23] W. Jiahui. *Python Interactive Network Visualization Using NetworkX, Plotly and Dash*. URL: <https://towardsdatascience.com/python-interactive-network-visualization-using-networkx-plotly-and-dash-e44749161ed7>.
- [24] M. Heilbron and M. Chait. “Great Expectations: Is There Evidence for Predictive Coding in Auditory Cortex?” In: *Sensory Sequence Processing in the Brain* 389 (Oct. 1, 2018), pp. 54–73. ISSN: 0306-4522. DOI: 10.1016/j.neuroscience.2017.07.061. URL: <http://www.sciencedirect.com/science/article/pii/S030645221730547X>.
- [25] A. Bastos et al. “Canonical Microcircuits for Predictive Coding”. In: *Neuron* 76 (Nov. 21, 2012), pp. 695–711. DOI: 10.1016/j.neuron.2012.10.038.

- [26] J. F. Mejias et al. “Feedforward and Feedback Frequency-Dependent Interactions in a Large-Scale Laminar Network of the Primate Cortex”. In: *Science Advances* 2.11 (Nov. 1, 2016), e1601335. doi: 10.1126/sciadv.1601335. URL: <http://advances.sciencemag.org/content/2/11/e1601335.abstract>.
- [27] R. P. N. Rao and D. H. Ballard. “Predictive Coding in the Visual Cortex: A Functional Interpretation of Some Extra-Classical Receptive-Field Effects”. In: *Nature Neuroscience* 2.1 (1 Jan. 1999), pp. 79–87. ISSN: 1546-1726. doi: 10.1038/4580. URL: https://www.nature.com/articles/nn0199_79.
- [28] M. I. Garrido et al. “Dynamic Causal Modeling of the Response to Frequency Deviants”. In: *Journal of Neurophysiology* 101.5 (May 1, 2009), pp. 2620–2631. ISSN: 0022-3077. doi: 10.1152/jn.90291.2008. URL: <https://doi.org/10.1152/jn.90291.2008>.
- [29] A. Todorovic and F. P. de Lange. “Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields”. In: *The Journal of Neuroscience* 32.39 (Sept. 26, 2012), p. 13389. doi: 10.1523/JNEUROSCI.2227-12.2012. URL: <http://www.jneurosci.org/content/32/39/13389.abstract>.
- [30] N. Kogo and C. Trengove. “Is Predictive Coding Theory Articulated Enough to Be Testable?” In: *Frontiers in computational neuroscience* 9 (Sept. 8, 2015), pp. 111–111. ISSN: 1662-5188. doi: 10.3389/fncom.2015.00111. PMID: 26441621. URL: <https://pubmed.ncbi.nlm.nih.gov/26441621>.
- [31] A. Popovici, P. Zaal, and D. M. Pool. “Dual Extended Kalman Filter for the Identification of Time-Varying Human Manual Control Behavior”. In: *AIAA Modeling and Simulation Technologies Conference*. American Institute of Aeronautics and Astronautics, June 2, 2017. doi: 10.2514/6.2017-3666. URL: <https://arc.aiaa.org/doi/10.2514/6.2017-3666>.
- [32] J. Rojer et al. “UKF-Based Identification of Time-Varying Manual Control Behaviour”. In: *IFAC-PapersOnLine*. 14th IFAC Symposium on Analysis, Design, and Evaluation of Human Machine Systems HMS 2019 52.19 (Jan. 1, 2019), pp. 109–114. ISSN: 2405-8963. doi: 10.1016/j.ifacol.2019.12.120. URL: <http://www.sciencedirect.com/science/article/pii/S240589631931955X> (visited on 01/18/2021).
- [33] E. R. Boer and R. V. Kenyon. “Estimation of Time-Varying Delay Time in Nonstationary Linear Systems: An Approach to Monitor Human Operator Adaptation in Manual Tracking Tasks”. In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 28.1 (Jan. 1998), pp. 89–99. ISSN: 1558-2426. doi: 10.1109/3468.650325.
- [34] A. van Grootheest et al. “Identification of Time-Varying Manual Control Adaptations with Recursive ARX Models”. In: *2018 AIAA Modeling and Simulation Technologies Conference*. AIAA SciTech Forum. American Institute of Aeronautics and Astronautics, Jan. 7, 2018. doi: 10.2514/6.2018-0118. URL: <https://arc.aiaa.org/doi/10.2514/6.2018-0118>.
- [35] M. Linssen. “Identifying Time-Varying Multimodal Manual Control Using Recursive ARX Model Techniques”. In: (2020). URL: <https://repository.tudelft.nl/islandora/object/uuid%3A442f4308-0ea2-41a5-b38c-ee6b1a289f78> (visited on 01/18/2021).
- [36] P. M. Zaal. “Manual Control Adaptation to Changing Vehicle Dynamics in Roll–Pitch Control Tasks”. In: *Journal of Guidance, Control, and Dynamics* 39.5 (2016), pp. 1046–1058. doi: 10.2514/1.G001592. URL: <https://doi.org/10.2514/1.G001592>.
- [37] P. M. T. Zaal et al. “Modeling Human Multimodal Perception and Control Using Genetic Maximum Likelihood Estimation”. In: *Journal of Guidance, Control, and Dynamics* 32.4 (July 1, 2009), pp. 1089–1099. doi: 10.2514/1.42843. URL: <https://arc.aiaa.org/doi/10.2514/1.42843> (visited on 09/22/2020).
- [38] R. A. Hess. “Modeling Human Pilot Adaptation to Flight Control Anomalies and Changing Task Demands”. In: *Journal of Guidance, Control, and Dynamics* 39.3 (2016), pp. 655–666. doi: 10.2514/1.G001303. URL: <https://doi.org/10.2514/1.G001303>.
- [39] D. T. McRuer, R. E. Magdaleno, and G. P. Moore. “A Neuromuscular Actuation System Model”. In: *IEEE Transactions on Man-Machine Systems* 9.3 (Sept. 1968), pp. 61–71. ISSN: 2168-2860. doi: 10.1109/TMMS.1968.300039.
- [40] N. T. Nguyen. *Model-Reference Adaptive Control: A Primer*. Advanced Textbooks in Control and Signal Processing. Springer International Publishing, 2018. ISBN: 978-3-319-56392-3. doi: 10.1007/978-3-319-56393-0. URL: <https://www.springer.com/gp/book/9783319563923>.

- [41] K. van der El et al. “Effects of Preview on Human Control Behavior in Tracking Tasks With Various Controlled Elements”. In: *IEEE Transactions on Cybernetics* 48.4 (Apr. 2018), pp. 1242–1252. ISSN: 2168-2275. DOI: 10.1109/TCYB.2017.2686335.
- [42] N. Nguyen et al. “Bounded Linear Stability Analysis - A Time Delay Margin Estimation Approach for Adaptive Control”. In: Aug. 1, 2009. DOI: 10.2514/6.2009-5968.
- [43] N. Nguyen and E. Summers. “On Time Delay Margin Estimation for Adaptive Control and Robust Modification Adaptive Laws”. In: Aug. 1, 2011. DOI: 10.2514/6.2011-6438.
- [44] E. Fridman and S.-I. Niculescu. “On Complete Lyapunov–Krasovskii Functional Techniques for Uncertain Systems with Fast-Varying Delays”. In: *International Journal of Robust and Nonlinear Control* 18.3 (Feb. 2008), pp. 364–374. ISSN: 10498923, 10991239. DOI: 10.1002/rnc.1230. URL: <http://doi.wiley.com/10.1002/rnc.1230> (visited on 01/12/2021).
- [45] E. Fridman. “Tutorial on Lyapunov-Based Methods for Time-Delay Systems”. In: *European Journal of Control* 20.6 (Nov. 1, 2014), pp. 271–283. ISSN: 0947-3580. DOI: 10.1016/j.ejcon.2014.10.001. URL: <http://www.sciencedirect.com/science/article/pii/S0947358014000764>.
- [46] D. McRuer and H. Jex. “A Review of Quasi-Linear Pilot Models”. In: *IEEE Transactions on Human Factors in Electronics* HFE-8.3 (Sept. 1967), pp. 231–249. ISSN: 2168-2852. DOI: 10.1109/THFE.1967.234304.
- [47] M. Mulder et al. “Manual Control with Pursuit Displays: New Insights, New Models, New Issues”. In: *IFAC-PapersOnLine*. 14th IFAC Symposium on Analysis, Design, and Evaluation of Human Machine Systems HMS 2019 52.19 (Jan. 1, 2019), pp. 139–144. ISSN: 2405-8963. DOI: 10.1016/j.ifacol.2019.12.125. URL: <http://www.sciencedirect.com/science/article/pii/S2405896319319603>.
- [48] O. Stroosma et al. “Measuring Time Delays in Simulator Displays”. In: AIAA Modeling and Simulation Technologies Conference. Aug. 20, 2007. ISBN: 978-1-62410-160-1. DOI: 10.2514/6.2007-6562.
- [49] M. M. Van Paassen, O. Stroosma, and J. Delatour. “DUECA - Data-Driven Activation in Distributed Real-Time Computation”. In: AIAA Modeling and Simulation Technologies Conference. Aug. 14, 2000. DOI: 10.2514/6.2000-4503.
- [50] A. Y. Ng and S. Russell. “Algorithms for Inverse Reinforcement Learning”. In: *In Proc. 17th International Conf. on Machine Learning*. Morgan Kaufmann, 2000, pp. 663–670.
- [51] S. Levine. *Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review*. May 20, 2018. arXiv: 1805.00909 [cs, stat]. URL: <http://arxiv.org/abs/1805.00909> (visited on 03/03/2021).

Part II

Literature Study

Introduction

Understanding and modeling the human motor system has a large variety of applications ranging from manual control of aerospace vehicles, to driving, prosthetic engineering and robotic tele-operations among many.

One framework to model the human motor system uses classical control theory to generate representative controllers. This approach had success in modeling humans that were performing restricted and simple tasks that were easy to formalize. Examples of these tasks are following a reference signal on a display with the help of a joystick. This approach though presents currently some limitations: the neuromuscular system description and the understanding of learning and adaptation is quite limited. The neuromuscular system is a redundant system of actuators and joints that is very difficult to accurately model and current descriptions (in a low dimensional state space) are not accurate enough to capture many phenomena. Learning is also hard to study as experiments require a high number of participants that have to be measured at large intervals of time. Finally adaptation is also an open problem and arguably is the most important feature of human's controllers. The ability to adapt and apply generalized learned policies of actions to new situations is a key feature of intelligence. This feature has kept human in the loop either as supervisor, as is it the case for most commercial aerospace applications, or directly controlling the vehicle as it is the case for cars.

In this thesis we address the problem of understanding human motor adaptation for simple manual control tasks and we propose a novel approach inspired by a neuroscience framework called "Predictive Coding". At the core of this approach lies an adaptive control technique called model reference adaptive control (MRAC), which defines a controller that imitates another controller taken as reference. The key insight that made this work possible is that McRuer found that the open loop dynamics of a system controlled manually has the structure of a first order stable system with delay.

This report is structured as follows. In chapter 2 we present the literature study done on the field of motor control in biological systems. In chapter 3 are presented a series of models that, either have been used in the past to model the control strategy of humans in very simple tasks, or that have certain features that are promising to model adaptive behavior. Follow in chapter 4 an overview is presented of the research objective and related questions that we hope are going to be answered. The theoretical foundation and simulation results of the selected model, "Model Reference Adaptive Control", are presented in chapter 5. Finally the report concludes in chapter 6 describing future developments and the experimental plan.

Literature Study: Sensorimotor Control

The literature study focuses on sensorimotor control, i.e how control is achieved by biological systems by means of perception and action. We often abbreviate the term sensorimotor control for motor control, taking it for granted that biological systems gather information from the world, through their sensors, to plan and act. In this chapter, we give an overview of two fundamental and interrelated aspects: internal models and predictive coding. We present evidence that indicates that humans use an internal model of the dynamics of the system to general actions. Furthermore, we briefly summarize the most important aspect of predictive coding. Predictive coding is a paradigm in neuroscience that recently received more support and that tries to explain many aspects of human intelligence.

2.1 Neocortex structure

We layout here a basic overview of the structure of the cortex. In the end, every theory of intelligence must be mapped and explained causally in terms of its basic computational operations that are carried out by the canonical microcircuits in the brain.

The neocortex is a convoluted structure on the surface of the brain that it is suspected to be the key component for higher-order brain functions in human beings. The neocortex, which only recently evolved, is present only on mammals brains and contains the vast majority of the neurons in the brain. The cortex structurally is a "quasi" 2D surface, or sheet with a relatively small thickness of 2-4 mm. It is efficiently organized in 3D space, inside the skull, with a highly convoluted structure, rich of ridges and sulci.

The cortex has an underlying basic topological structure that is repeated throughout it: the micro-column. It can be seen a fundamental computational unit constituted by 100-110 neurons, which share similar inputs and outputs and strictly interconnected. For example in the cortical area V1, the primary visual cortex, different micro-columns are activated by edges with different orientations. The fundamental unit of computation for a micro-column is a neuron which can have both excitatory and inhibitory synapses.

The cortex is globally divided into 4 lobes: temporal, parietal, visual and frontal. In humans have been distinguished hundreds of different areas. These subdivisions are based on both the function of that specific area of the cortex and their connectivity. The cortex depth-wise is subdivided into six different layers, each layer has different cell types and connectivity patterns. In particular from an anatomical perspective:

- Layer 1: It is a thin layer with a lower than average density of neurons
- Layer 2-3: Thick layers rich in deep pyramidal cells
- Layer 4: It is the layer rich in different kinds of neuronal cells that receive cortical input from the thalamus

- Layer 5: Layer where the largest pyramidal neural cells are found. Parts of cells are connected to layer 1 and 2-3 while others, mostly in the motor cortex, project directly in the motor area in the basal ganglia
- Layer 6: This layer sends feedback signals to the Thalamus

In Figure 2.1 you can observe a schematic representation of the layers in the cortex.

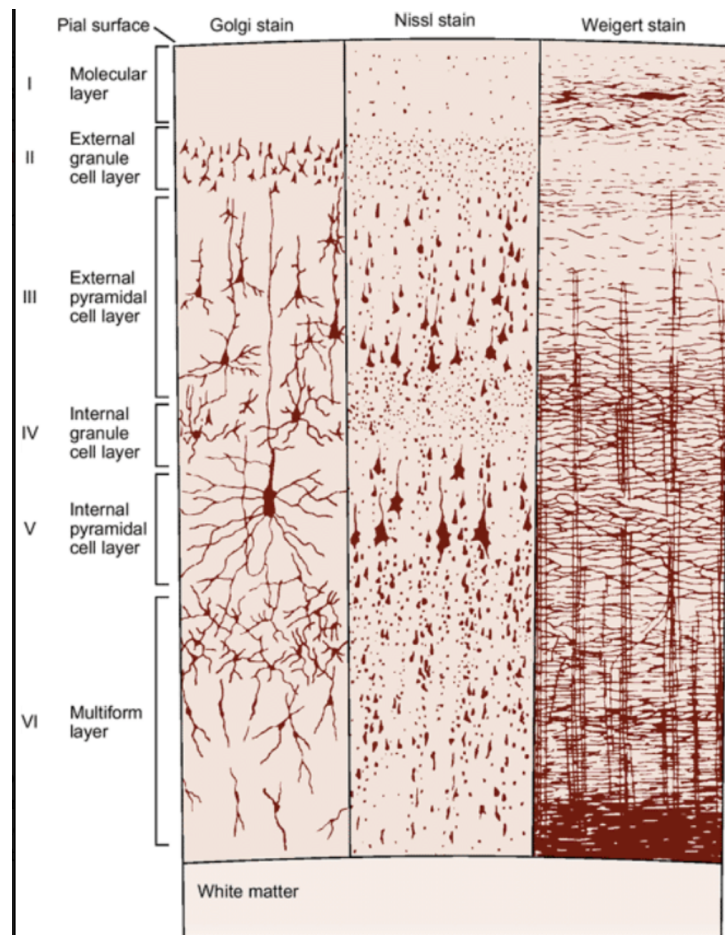


Figure 2.1: Overview of the cortical layers, visualized using three different stain techniques. The Nissl stain mostly highlights the central body, the Weigert stain the synapses and the Golgi both.

In Figure 2.2 is shown the flow of information through the cortical layers. Information flows from the Thalamus to mostly layer 4. Layer 4 sends feed-forward information in terms of both an excitatory and inhibitory stimulus to the higher layers. Layer 2/3 is in contact with the dendrites of the pyramidal cells in layer 5/6 to which the information continues to flow in a forward fashion. Feed-forward connections seem to be driving, stimulating spiking responses to the efferent neurons. On the other hand feedback responses, i.e. the flow of information from regions of higher hierarchy to lower ones, seem to be mostly modulatory and inhibitory [1].

2.2 Predictive coding

During the last decades, our understating of the brain has significantly improved and changed. In the past we were mostly focused on understating how the brain extracted information from the sensory information and in building a functional map of the different areas of the organ. Today there is more attention and research dedicated to which general learning programs and principles are employed/embodyed in our brains [2]. This shift was made possible by a series

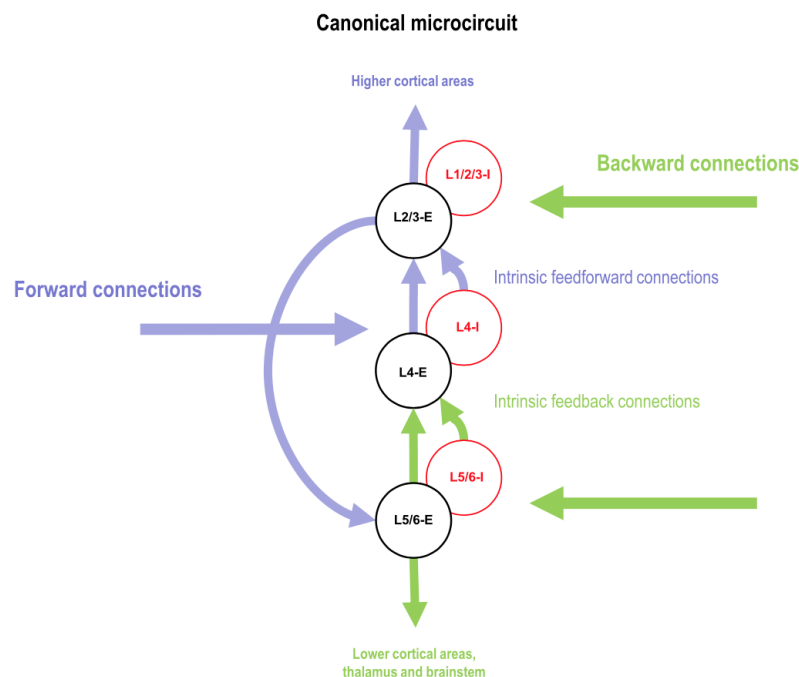


Figure 2.2: Canonical microcircuit and established information flow in the cortex [1]

of new empirical observations, thanks also to advances in neuroimaging, that propelled the theoretical development of "predictive coding". We lay first the basic principle of predictive coding, we then proceed to lay out how the canonical microcircuits can produce the type of computation required for predictive coding and finally we present more evidence in its favor.

2.2.1 Computational Framework

How do we process information coming from the world? Understanding the computations and programs that define intelligent systems has been the dream of both neuroscientists and AI researchers for decades.

Predictive coding (PC) lays a framework to explain how the brain processes information and how we make sense of the world.

PC states that our brains continuously generate predictions about the stream of information that is coming from the world. The sensory information is used to update the prediction system and continuously "tune it" to reality.

In particular, the brain has an internal generative model of the world. By generative model we mean a model that can generate a richer and hierarchical representation of the expected sensory information [3]. The process that generates the sensory predictions is "top-down" while the sensory information is processed "bottom-up". Given an abstract and compressed representation of the world and a generative model (implemented in the neural network), a prediction about the expected information can be generated continuously at all the levels of the hierarchy: starting from the very abstract representation at the top until the lower levels where the brain predicts the exact raw sensory information.

In Figure 2.3 you can observe how a concept, i.e. a compressed representation, can generate hierarchical expectations in the cortex down to the raw sensory data. From the concept of the bird, the brain can generate expectations over its shape or the series of sounds that it makes.

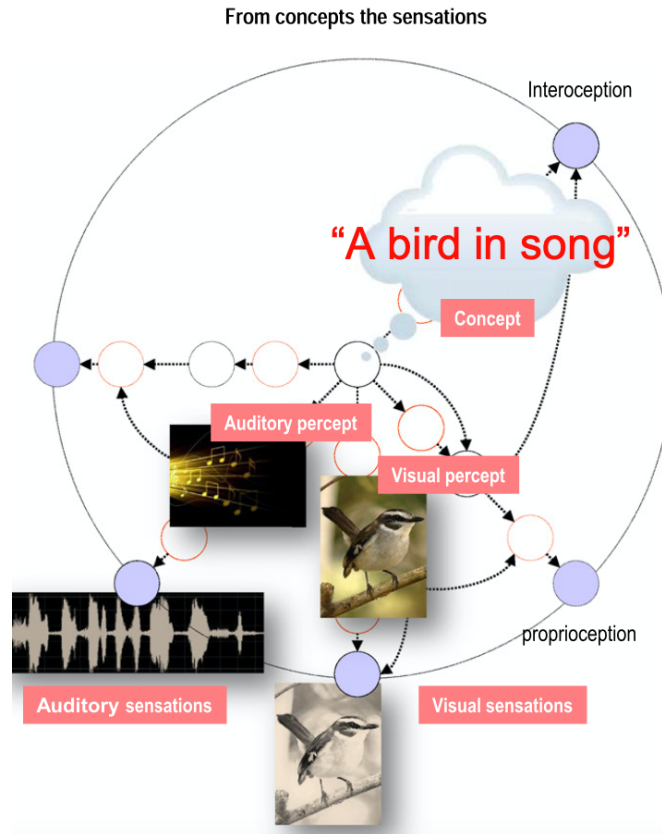


Figure 2.3: Hierarchical model that generates expectations from concepts [1]

The sensory predictions, at the very bottom of the abstraction hierarchy, are compared with the actual sensory data. The unexplained sensory inputs, i.e. the "errors", are propagated up the hierarchy for further processing. These predictions errors are minimized either by acting on the world or overtime changing the internal model[4]. An essential part of the PC theory is precision weighting, i.e. the estimation of the error's reliability, which also depends on the level of signal noise [2].

Friston developed also a formal theory of PC based on Bayesian inference. Each biological, self-organizing, adaptive system tends to resist disorder and maintain itself in a state of homeostasis. The set of internal and sensory states is rather limited and the probability that an organism is in one of these states is high. Now let's recall the notion of entropy, as the logarithmic measure of the number of states that have a probability p_i of being occupied, in particular $S = k_b \sum_{i=0}^N p_i \log(p_i)$, where k_b is the Boltzmann constant. Friston states that the entropy of functioning biological systems is low and that agents must minimize the average long term entropy, or surprise, to continue keeping the entropy low, defined from the amount of information present in the sensory data. In a Bayesian framework, entropy or surprise is equivalent to the evidence variational lower bound (ELBO) on the model evidence. ELBO is a technique to make tractable the computation of posterior high dimensional distributions of models over the evidence. By minimizing ELBO the neural systems would be maximizing probability for the current generative internal models given the current evidence gathered from the world. We refer to Friston [3] for more details.

This theory of neural computation breaks away with the tradition viewed that saw perception mainly as feed-forward computational phenomena [2]. Certainly, there are concerns that part of the theory is based on untested assumptions [5], or that the PC is not the all grad theory of perceptions and inference that Friston proclaims [2]. Nonetheless, there is a great amount of evidence, both physiological and in term of neural population dynamics, that support the view

that the brain generates an internal model of the world.

2.2.2 Microcircuit

With this section, we try to explain the abstract concepts presented in the previous section by presenting a plausible biological neural implementation of predictive coding. The sensory cortex is a layered structure of neurons that is hierarchically organized. For example, the visual cortex region V1 receives feed-forward connections from the lateral geniculate nucleus (LGN), a part of the thalamus that receives retinal input. The information is then passed through feed-forward connections to higher cortical areas such as V2, V3, V4, V5 and V6. Nonetheless, there are strong feedback connections from higher to lower cortical areas and between the layers of the cortex. A schematic representation is shown in Figure 2.2. Forward connections arrive at layer 4, information is then passed the superficial cells in layer 2/3 and then to the deep-pyramidal cells in layer 5-6, which send feedback to lower cortical areas and to layer 4. Feedback connections appear to be mostly modulating and inhibitory [1].

There is not a definite answer yet on how the predictive coding computation maps on onto the canonical microcircuit. Here we explore one way in which the brain does the necessary computations but there are other possibilities [4] [1] [5]. The scheme is shown in Figure 2.4 show where the computations required by PC are carried out. The error is computed in layers 2/3 by comparing the received signals with the predicted one. The predictions happen in the deep cortical layers 5/6 and are sent to cortical areas of lower hierarchy. The errors are propagated up the hierarchy and are used to update the conditional expectations.

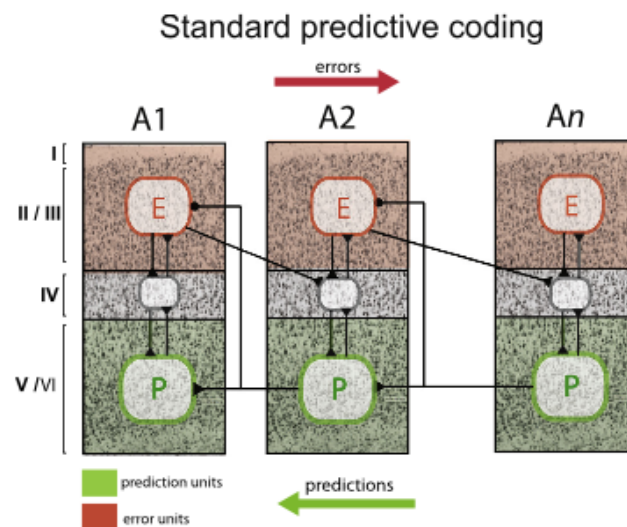


Figure 2.4: Hierarchical model for predictive coding [6]

2.2.3 Evidence

In this section, we present the experimental evidence that supports PC. Nonetheless, there is still substantial controversy in the field, with many pointing out parts of the theory that are not fully testable or that lack evidence [5]. A deeper understanding of the visual cortex made predictive the predicting coding framework relatively popular. It was discovered that, in the visual cortex, the feed-forward connections across the cortical area transmit residual errors, the difference between the predicted activity and the actual one at lower cortical areas. The feedback connection, in particular, helped to modulate the activity of neurons lower in the hierarchy. A neural model of this kind showed phenomena like end-stopping and non-classical surround effects [7].

As for the auditory cortex, the evidence for predictive coding is not as strong. Several studies presented evidence that a bottom-up model of neuronal dynamics cannot account for phenomena like "Mismatch Negativity" (MMN) [4]. MMN is a deviant response that appears in the event-related potential to an odd stimulus in a series of repetitive patterns. Predictive coding models of auditory perception could explain the neural suppression after a repeated stimulus and activation for an odd stimulus [8] [9]. Such phenomena appear also in case of omission of an expected stimulus, i.e silence [10]. In the field of natural language processing, there is also evidence that predicts incoming stimuli and that letters are more easily identified when embedded in a word [4]. On a physiological level, as predicted by PC, different cells are responsible for feedforward and feedback information [1] [2]. Furthermore these neuronal populations show interactions at different frequencies [11].

2.3 Adaptive Sensorimotor Control

In this section, we give an overview of the current understanding of adaptive motor control. We start by listing what are the current problems that a framework for adaptive motor control should address. We proceed in explaining what is the role of internal models in motor control. We then describe what is the current view on the internal model and of what is the state of the art framework "optimal feedback control" that best describes and predicts the adaptive motor control of biological systems.

Sensory-motor control is a very difficult control task to solve, in particular, the brain faces many challenges: uncertainty, redundancy, noise, delays and time-varying dynamics [12].

Uncertainty There is always uncertainty about the state and future states of external objects and of parts of the motor-system itself, like position and velocity of limbs, etc ... We may ask what is a natural framework to reason, plan and act given a set of random variables that come from a range of statistical distributions?

Redundancy Biological systems are incredibly redundant in the set of actuators that can be used. There are hundreds of muscular fibres that can be controlled individually in addition to numerous joints and ligaments. How does the brain solve this control allocation problem? The brain *learns* to solve these problems, as infants start to walk only around at the age of 12 months. Furthermore, there is another problem to tackle, out of the possible infinite trajectories to solve a task why do we pick a specific one? The answer to both these questions probably lies in optimization, since it's a way to deal with these higher dimensional problems.

Noise Noise is present in all the feedback information that we receive from the world. It could be external noise, present in the objects we observe or interact with, or internal. Internal noise is chemical noise, electrical noise that is inherent in the way biological systems propagate information inside their bodies. Given that certain observation has higher reliability than others, how does the brain *weight* different information based on their precision? How is information from multiple senses properly integrated?

Delays Biological systems live in the past. Visual and sensory information has inherent delays due to the upper bound on the transmission speed of information across nerves. Furthermore, extra computation time is needed to process such information. How can we then produce smooth and fast movements without much effort? Part of the answer to this problem is having internal models.

Time-varying dynamics Our sensory-motor system is highly time-varying, the impedance of our muscles, change very rapidly depending on the task at hand. Furthermore, we continue to update our control policies as growing and our muscular-skeletal system change or becomes affected by ageing.

2.4 Internal Models

In this section, we present an overview on internal models. Internal models have been recognized to be an essential component for motor control in biological systems [13]. Animals have internal dynamics models that can be used to control the body and act in the environment [14] [15]. Internal models can be used in many ways: to predict future actions, to do state estimation and filtering by comparing the predicted state with the stream of data generated by biological sensors, to try-out in simulation the possible outcomes of an action.

In this section, we shortly introduce internal models and present an old one of the first experiments that indicated the role that internal models play in motor control. We proceed then to give a more updated overview of the current view on internal models.

We can distinguish several types of internal models [13]:

- Prior models: encode the priors over observations $p(z)$ and states of the world $p(x)$, where z represent a measurement and x the state of a system.
- Precision models: encode the reliability of a particular source of information
- Forward dynamics: estimates the future state of a system $p(x_{k+1}|x_k, u)$ given the current state, x_k and a control action, u .
- Hierarchical latent models: these can be graphical models that encode hierarchical relationships between different states. These models can be used to direct planning across large time-scales.

2.4.1 Young's manual control experiment

Young proposed several models that would account for human adaptation. In subsection 3.3.4 we explore the qualitative model that he proposed to model a human controller. The scheme assumes that humans develop internal models of the dynamics of the system that they control. This assumption was tested in a later experiment where participants, using a compensatory display, were asked to control and detect the change of dynamics in a time-varying system. The participants of the experiment were divided into three groups: active controller, passive controllers and observers. The active controllers actively controlled the systems. The passive controllers were led to believe that they controlled the system but the system was controlled by an active controller. A scheme of the experiment is shown in Figure 2.5. The observers on the contrary just observed the screen. All participants had to press a button when they recognized the change in dynamics. The measured time for detection is shown in Table 2.1.

As we can observe the time for the observers is significantly higher than the other two classes of participants, who on the other hand show similar times. The most probable cause for this delay on the observation is the lack of an internal model for the dynamics of the controlled elements. The participants that only observe, engage in no motor actions and therefore have a harder time in estimating how the system responds to input. We claim this is evidence for the role that internal models have in motor control tasks. Furthermore since both active and passive controllers have a similar detection time we speculate that the open loop dynamics might be a good model for the internal model of the human controllers.

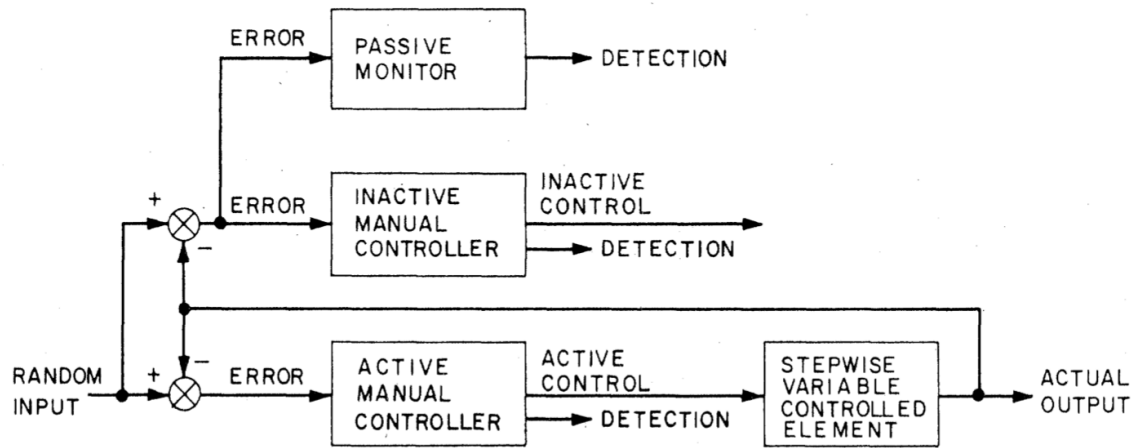


Figure 2.5: Detection times of the change in dynamics and their standard deviation for the three different groups [16]

Subject	2nd to 1st	1st to 2nd
Active Controller	1.23	1.36
Passive Controller	1.45	1.56
Passive Observer	2.48	5

Table 2.1: Time, in seconds, for the detection in the change of dynamics of the controlled element. The controlled elements switches from first integrator to double integrator and vice-versa. The participants could only observe the error between the target state and their state.

2.4.2 Cerebellum's motor models

The cerebellum seems to create internal models of the world that can be used to more effectively operate in it. In particular, we can produce models of the dynamics of body parts or systems that we interact with. Internal models present several advantages: reduced sampling from our sensors and lower reliance from sensory feedback signals. The cerebellum could produce either an inverse (feed-forward) model or forward model, that encodes the dynamics of the system in consideration. The feedback signals, produced by the visual cortex and sensory-motor system, play still an important role in error-correcting. A possible schema, in which the cerebellum produces a forward-model of the dynamics is shown in Figure 2.6.

The ability to create internal models is also necessary to create efficient learning *adaptive* programs. This is what happens to a human when they learn to drive, ride a bicycle or other motor tasks. The creation of internal models affects also the way we process information. The voluntary movements commands are generated in the premotor cortex or anterior cingulate cortex, then the information flows to the motor cortex and the cerebellum. The output of the cerebellum, in the possible presence of feedback, can further be processed by the cortex and then sent to the motor system. The exact information flow has not yet been established [17]. The computation that happens in the cerebellum is not consciously accessible. In fact, in general, while learning a new motor skill or the cognitive ability the amount of conscious computation will decrease with time.

2.4.3 State Estimation and Planning

State estimation is essential to accurately estimate the current state and integrate all the sensory information to predict future state of the motor system. Internal forward models of the dynamics

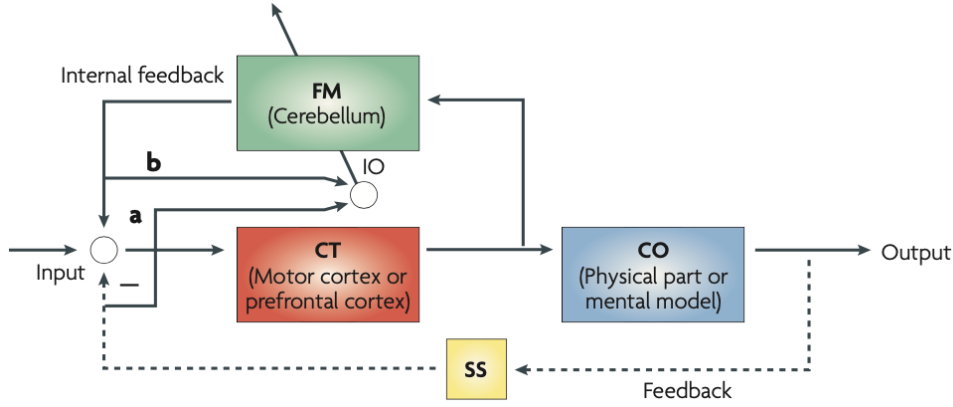


Figure 2.6: Possible control scheme where the cerebellum encodes the forward dynamics (FM) of a system. The input arrives and together with the current information from the sensory system (SS) and prediction of the cerebellum, are sent to the motor cortex (MT) which elaborate the signals to be sent to the peripheral motor system (CO). The FM is trained by minimizing the difference the output of the model itself (b) with the feedback information (a) gathered by the sensory system (SS). This comparison is thought to happen in the inferior olive (IO), a specific component of the cerebellum [17].

of a system make state estimation and filtering possible through the comparison of the predicted state with the actual information coming from the sensory system. The way state estimation could be implemented could be related to Bayesian inference [13].

Bayesian inference is a framework that could explain how the neural system makes decisions and plans in an uncertain world and the presence of noise [3]. Bayesian inference relies on the application of Bayes' rules to find the most probable posterior state of the world, $p(x|z)$ given observations with likelihood $p(z|x)$ and a set of priors $p(x)$:

$$p(x|z) = \frac{p(z|x)p(x)}{p(z)} \quad (2.1)$$

where $p(z) = \sum_y p(z|x)p(x)$ is the prior probability of an observation. This framework can also be used to infer the current state of a system given a series of observations and priors about the precision of the measurement. The Kalman filter is a specific implementation of the general class of Bayesian filters. There is evidence that humans behave like optimal bayesian estimators in visual tasks [18] [19] and with other sensory cues [12].

Bayesian inference can be used also to select the best model that explains the observed data. This is particularly relevant when the dynamics of a system change and we have to infer the change of model. In the case of multiple models M_i and observations x the posterior distribution of the new state given the previous state is the weighted average by the priors over the possible models

$$p(x_k|x_{k-1}) = \sum_{i=0}^N p(x_k|x_{k-1}, M_i)p(M_i|x_{k-1}) \quad (2.2)$$

where

$$p(M_i|x_{k-1}) = \frac{p(x_{k-1}|M_i)p(M_i)}{p(x)} \quad (2.3)$$

in this case the prior probability $p(x)$ is not particularly important and can be neglected as it is the same for all models.

2.5 Optimal Feedback Control

Optimal feedback control (OFC) is currently the framework able to better predict and explain motor behavior [20] [13] [21] [22]. OFC is the dominant framework to model voluntary control [23][24]. Optimal feedback control solves the redundancy in the number of actuators and combines optimal control with sensory feedback. An optimal control policy computes the optimal solutions to complete a certain task and continuously corrects for disturbances. One of key difference from normal feedback control lies in the way disturbances are corrected. Disturbances are only corrected if actually needed to optimize the given task. An example is given below: consider the task of reaching a target point with a limb, the target point could be either stationary or perturbed after 100, 200 or 300 ms from the start of the task. The controller has the following form:

$$u(t) = k_p(t)(p^* - p(t)) - k_v(t)v(t) - k_a(t)a(t) \quad (2.4)$$

where p^* is the target position and $p(t), v(t), a(t)$ are the position, velocity and acceleration of the limb at time t . The gains has been found by solving the optimization problem minimizing the final position error and control effort:

$$J \approx \|p^* - p(t_f)\|^2 + k \int_{t_0}^{t_f} \|u(t)\|^2 dt \quad (2.5)$$

The results of the experiment are further described in Figure 2.7. An OFC model matches extremely well the findings of the experiment, in particular, you can notice that if the perturbation is done early enough the followed trajectory does not immediately change. If the perturbation is not yet relevant (i.e. final objective can be reached at not extra cost) there no change in the chosen trajectory. This phenomenon cannot be explained with a simple feedback system. OFC has strong experimental evidence: it is successful in predicting eye movements [25] and limb trajectories in several other tasks [26] [12].

The optimal control problem is formulated in terms of a cost function that penalized the endpoint variance and the energy used [21]. The assumption that the control policy minimizes energy consumption can be justified from an evolutionary perspective. In this review, we will not go into the details of the implementation of optimal feedback control, which can be found in [20]. Briefly, the trajectory is computed, with an infinite time-horizon, using the classical linear quadratic Gaussian framework (LQG). The main drawback of a basic LQG optimal feedback methods is that the control time horizon is infinite (or until the end of the reference signal) and cannot anticipate changes. Also, the open-loop solution is computed once and feedback is introduced by optimizing a different objective function [20].

Finally the internal models are essential for OFC since they provide a model of dynamics that can be used to predict the future (with MPC for example) and they provide a way to reduce the effect of sensory delay and to perform state estimation.

Summarizing, the key components of OFC are three [27]:

1. Control actions are generated by optimizing for a function
2. Perturbations from the planned trajectory are corrected in an optimal fashion prioritizing what is important for the task
3. Internal models are used for the optimization routine, to compensate for sensory delays and to perform state estimation

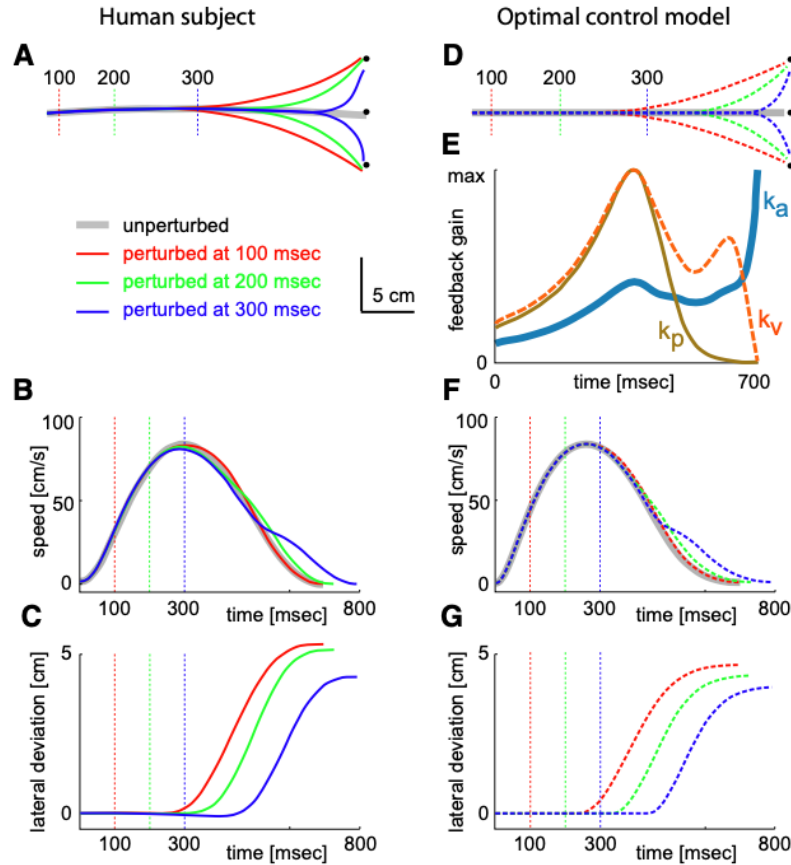


Figure 2.7: On the left column you can observe the participants' measured limb position (A), velocity (B) and lateral deviation from the target (C) while on the right are shown the optimal computed trajectories. Subfigure D, F, G show the same information but for policy generated by the optimal feedback system. Subfigure E shows how the different computed gains (shown in Equation 2.4) of the system change [28]

2.6 Hierarchical control

The motor control in biological systems is hierarchical. The cortex, hypothalamus, basal ganglia, cerebellum and spine are all involved in generating movements. In particular the cortex is a high level controller responsible for a broad and adaptive repertoire of behaviors. The basal ganglia and hypothalamus are thought to be involved in the selection of the appropriate motor program while the cerebellum and spine effectively send low level commands to actuate the muscle fibers [27]. The advantages of having a hierarchical motor control system are numerous [27]:

- **Information factorization:** only relevant information is sent to the specific subsystems. For example, the low level motor control system is not fed high level objectives, therefore it's able to generalize better since once it's trained it can be used for a large variety of objectives.
- **Partial autonomy:** lower level systems are somewhat autonomous and can operate robustly. Many studies have been performed with decorticated cats showing that they preserve most of the locomotion patterns.
- **Temporal abstraction:** allows different systems to plan and act at different time-scales. The cortex might plan to make dinner, while the a lower level system like the basal ganglia

might take care of actually walking to the kitchen. Finally the cerebellum, hippocampus and spinal circuits are responsible to activate the specific muscle fibers that are necessary to perform a step.

We believe that a hierarchical motor control structure is key to produce ever more complex movements across large time scales. In the context of a pursuit or compensatory manual control task the agent is told to either follow a signal or perform disturbance rejection. Since the agent cannot look into the future there is no opportunity to do medium-long term planning and since the task is already tightly specified there is also not much room to plan complex operations. Not tightly constrained experiments, that require planning across time scales of approximately 10 seconds are more appropriate to study the hierarchical and possibly optimal (in some sense) control structure of the human motor system. An example of such experiment could be picking up objects and repositioning them, obstacle avoidance while driving or other locomotor tasks. For our task what is most relevant is creation of internal models, encoded in the cerebellum and the activity of the cortex that is utilizing new incoming data to either act, probably directing the basal ganglia and spinal circuit to continue their execution program, or to modify the internal model.

Literature Study: Models

This chapter presents a series of models aimed at capturing the human behavior in motor tasks. We aimed at finding a model that is adaptive, learns from predictions errors, has an internal model itself, is related to Optimal Feedback Control and is hierarchical. These are some of the key features of intelligent biological systems.

3.1 Steady State Models

This is a brief overview of steady-state models usually considered for the human operators. These models describe the human steady-state behavior in only relatively restrictive control tasks and controlled plant dynamics. The tasks are usually carried out on preview, pursuit and compensatory displays. One of the first and most used models of human control behavior is McRuer's crossover model [29]. Given the plant transfer function H_c and the human controller transfer function H_p , as shown in Figure 3.1, McRuer model states that the *open-loop* response of the system with the human in the loop can be approximated by

$$H_{OL} = H_p H_c = \frac{\omega_c}{i\omega} e^{-i\omega\tau}, \omega \approx \omega_c \quad (3.1)$$

where ω_c and τ correspond to the crossover frequency and delay of the human operator model. The model is valid for frequencies close to the crossover frequency.

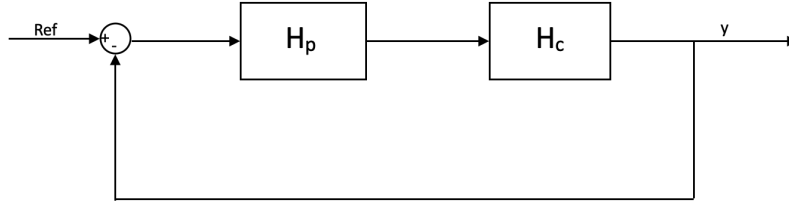


Figure 3.1: Control diagram for a typical reference tracking task

This model presents few interesting features. First of all the model can be used for different types of dynamical systems as the model prescribes the open loop behavior response. It also implies that humans form an internal model of the plant dynamics and cancel the poles of the plant via a partial dynamic inversion. McRuer also modeled explicitly the human operator, including in the model the dynamics of the neuro-muscular actuators [30]. The updated model (not in his notation) can be written as

$$H_p = k_p \frac{1 + T_{lead}(i\omega)}{1 + T_{lag}(i\omega)} \frac{\omega_{nm}^2}{(i\omega)^2 + 2\zeta\omega_{nm}(i\omega) + \omega_{nm}^2} e^{-i\omega\tau} \quad (3.2)$$

This model contains the dynamics of a classical phase compensator and of an harmonic oscillator, that represents the neuro-muscular actuators dynamics.

The stochastic and non-linear behavior of human controllers in simple tracking tasks cannot be captured by a steady-state linear model. The part that cannot be directly modeled in this way

is referred to as the remnant. To simplify the problem of modeling human adaptation we will ignore the contribution of the remnant.

3.2 Models based on system identification's methods

In this section, we listed and briefly described time-varying identification approaches.

3.2.1 Zaal model

The underlying idea behind this method is to create a time-varying parametric model of the human operator and then find the parameters that lead to the best fit of the observed control behavior. P. Zaal proposed a parametric model with similar form to the one in Equation 3.2, but where each parameter is a time-varying function of the form [31]

$$P(t) = P_0 + \frac{P_1 - P_0}{1 + e^{G(t-M)}} \quad (3.3)$$

where G and M represent the speed of the parameter change and the time at which adaptation happens. The model parameters were found by maximum likely-hood estimation augmented with a genetic algorithm [32].

Pros The advantages of these methods are the relative stability of the algorithm and the "sigmoid-like" parameters that guarantee the convergence to appropriate steady-state models. The sigmoid shape of the parameters allows for both a fast (step-like) and a more slow adaptation. The parameters can be tuned specifically for each controller.

Cons The behavior during the adaptation is prescribed and it is the same for all parameters. Therefore rather than estimating the transient behavior a sigmoid structure is imposed on it. It is hard to validate this assumption since other verification methods, like Kalman Filters, do not guarantee the estimated parameters are optimal. The model can only be used off-line since the optimizer needs the whole time-trace of the experiment.

3.2.2 Model based on Kalman Filter

Another attempt to identify the parameters of a human operator model is through state estimation via a Kalman filter [33]. Popovici used a Dual Extended Kalman Filter to estimate separately the compensator gains k_p and k_v and the neuromuscular parameters and delay ξ , ω_{nm} and τ of the following model [34]

$$H_p = (k_p + k_v(i\omega)) \frac{\omega_{nm}^2}{(i\omega)^2 + 2\xi\omega_{nm}(i\omega) + \omega_{nm}^2} e^{-i\omega\tau} \quad (3.4)$$

The parameters to be estimated are augmented into the state space corresponding to Equation 3.4, and their dynamics are driven by white noise. A representation of the two filters is shown in Figure 3.2.

Pros This model is extremely flexible and does not prescribe any adaptation dynamics. The estimated parameters are also specific to an individual human operator.

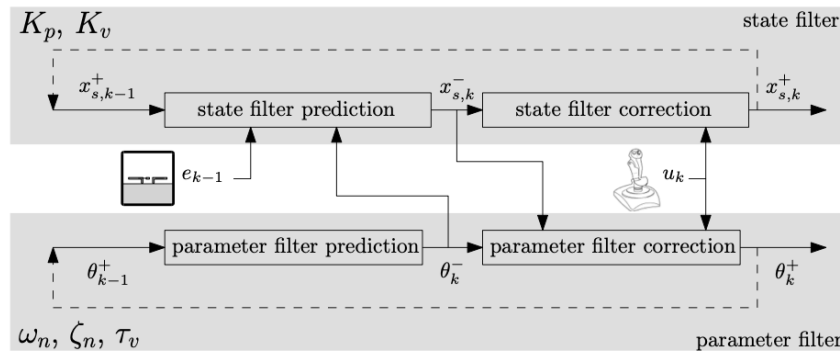


Figure 3.2: Scheme of the Dual Extended Kalman Filter [34]

Cons The major drawback of this model is the lack of convergence guarantees on the estimated parameters. During adaptation, the model parameters can vary relatively fast and it's important to have a responsive yet stable estimation of those.

3.2.3 Autoregressive models

The ARX (autoregressive) model is a general linear model. To describe the model we need to introduce the discrete time shift operator q , defined as $q^{-1}u(t_k) = u(t_{k-1})$. The ARX model can be written in the following form

$$u(t_k) = \frac{B(q)}{A(q)}e(t_{k-n}) + \frac{1}{A(q)}\epsilon(t_k) \quad (3.5)$$

where $B(q)$, $A(q)$ are polynomials of the linear operator q , e represent an error signal and $\epsilon(t_k)$, white noise. This model has already been successfully applied to identify human time-varying behavior in simple control tasks [35]. The model can be fitted with a variety of techniques including ordinary least-squares and regressive least square, for online estimation.

Pro This method depends less on the initial conditions to a stable convergence of the estimated parameters. The method can also be used online using a regressive least square as a fitting method. The ARX also is very flexible in its model structure since the order of the polynomial can be freely chosen.

Cons ARX estimated parameters have some unknown amount of bias. In the simulation, it has been shown that bias can be minimized by an appropriate choice of remnant filter for the human operator noise [36].

3.3 Rule based models

In the '60s great effort has been placed in creating qualitative and quantitative human adaptation models that are rule-based. By rule-based, we mean that the conditions for the start and type of adaptation are set by an expert in the forms of explicit rules. Most information about these models and experimental observation about human adaptation under change of dynamical systems are compiled in a review article by L.R. Young [16] and more recently by Shutting et al [37].

We start the section with older proposed models and we conclude it with an overview of fuzzy systems and a recent rule-based system proposed by Hess.

3.3.1 Supervisory Phatak model

Phatak proposed a model of human adaptation for tasks on a compensatory display that is a supervisory adaptive switching controller [38]. He identified three main discrete phases of adaptation:

- change of dynamics detection
- identification of the dynamics
- explicit adaptation

Phatak organized an experiment to verify his model. In Phatak's experiment participants had to perform a tracking task with one dynamical system (out of four available) that suddenly (or slowly) changed dynamics during the run. The design of the supervisory algorithms directly reflects the experiment design: the supervisory system chooses between four types of steady-state human operator models according to the error trajectory in the phase plane of the closed-loop system. In Figure 3.3 is shown the state space used to detect the change in dynamics of the system and switch controller. On one axis is reported the error rate and on another the error. The space \mathcal{R}^2 is subdivided into convex sub-spaces, called decision regions. The decision regions are predetermined and set to multiples of the maximum error rate, that the participants achieve when controlling a dynamical system without change in dynamics. When the dynamical system passes from a lower to a higher decision region (as numbered in Figure 3.3) it triggers a switch in the used human operator's steady-state model used.

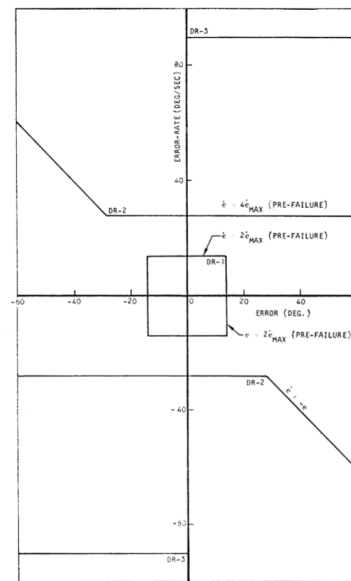


Figure 3.3: Example of convex partitioning of the state space into decision regions [38]

To avoid constant switching, which is a common problem in switching systems [SwitchingSystemsControla], the supervisory system makes a binary choice between two possible controllers at the time. The supervisory algorithm can be summarized as follows:

- check current decision region based on the error and error rate
- if the state trajectory crossed the decision boundary, switch the current controller with the next one in the queue
- remove the previous controller from the queue
- repeat

Pros The advantage of this model is that it can be customized and made relatively reliable for a specific task with some engineering effort. Furthermore, the models at which the algorithm converges are guaranteed to be realistic of human behavior.

Cons The system is not guaranteed to be stable. The partitions in decision regions depend both on the order, type of dynamical system and number of possible human controller models. It becomes quickly unfeasible to handle many different conditions. The supervisory algorithm does sequential binary choices, therefore if the "right" controller is at the end of the queue, the system might become unstable. There is no guarantee that the transient behavior is representative at all of what humans do. If more than one controller can make the system stable the supervisory algorithm will choose the one that is in front of the queue, independently of the actual human behavior.

3.3.2 Fuzzy controllers

Fuzzy controllers were introduced to better capture nuances and a non-precise assessment made by humans. Let's assume for example that humans use a function of the error rate and error to determine if the system has changed. Putting simple thresholds for the values of the considered variable is probably too simplistic, and one could define decision regions as in subsection 3.3.1 as fuzzy sets.

A simple schematics of the model is shown in Figure 3.4. The inputs are first fuzzified to establish the degree their membership for each fuzzy set. Then if-then rules, designed by engineers, are applied to the fuzzy sets, the results are then combined and defuzzified. The defuzzification is an interpolation procedure. The output now is a scalar number and it represents the output of the controller.

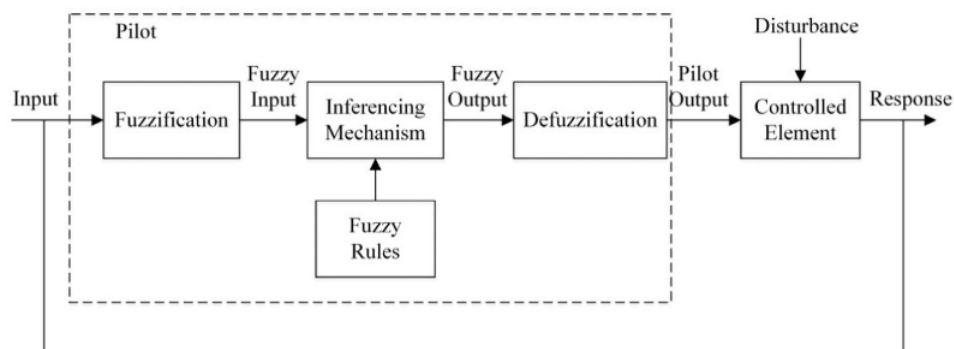


Figure 3.4: Fuzzy system architecture [37]

An example of a fuzzy system applied to model a human controller can be found in [39], few more example in the review article [37]. The fuzzy controller can also be combined with neural networks to generate fuzzy rules, which are interpretable to humans.

Pros They can capture the uncertainty present in humans in categorizing objects or signals.

Cons They are usually very laborious for engineers to design. Great care must be taken in defining the fuzzy sets, then the many fuzzy rules and the defuzzification scheme. Part of these problems could be alleviated by using neuro-fuzzy systems.

3.3.3 Hess Model

The Hess model is a popular rule-based system that relies on a two-loop description of the human operator. Details of the inner and outer loop are shown in Figure 3.5. While the outer loop is a simple proportional controller, the inner loop works with the derivative of the state of the system.

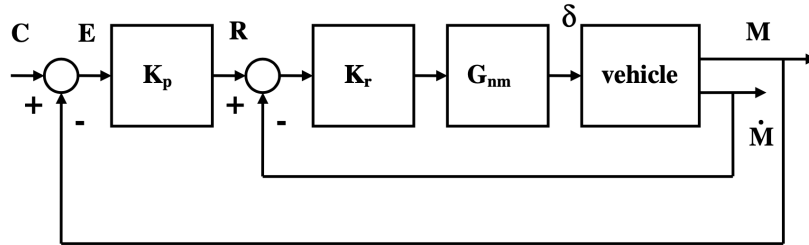


Figure 3.5: Hess's model of human adaptive behavior [40]

Hess hand-crafted a set of gains and learning rule to ensure a representative adaptive behavior [40]. The learning rule is active only after the change in the dynamics of the controlled element.

Pros The model is relatively simple, intuitive and can be used and adapted to several cases.

Cons One of the disadvantages of the model is the lack of convergence guarantees. The final values of the gains of the system both depend on the controlled dynamics and the input. Depending on the input the final controller can end up in a configuration far from representative of a human controller. Furthermore, the rules are found ad-hoc and lack theoretical justification. Finally, this model requires knowledge about the time at which the dynamics of the controlled plant change.

3.3.4 Young and Stark model

Young and Stark proposed a qualitative model for the human controller adaptation [41]. Their model relies on the mismatch between the observed error and error rate and the observed variables which we call δ_e . The expected error rate is generated by an internal model of the dynamics of the controlled elements. The signal δ_e is then fed to an adaptive control operator that changes both the internal model of the plant dynamics and the current human operator model. A control diagram of the model is shown in Figure 3.6.

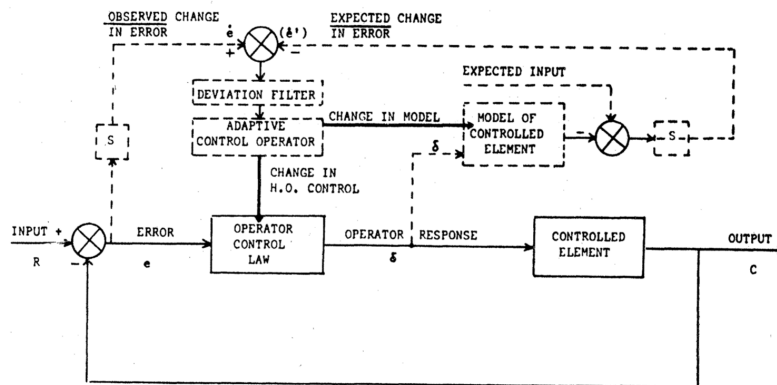


Figure 3.6: Block diagram representation of the Young and Stark model for a compensatory display

Pros This model has several desirable characteristics probably also present in biological systems: the adaptation is driven by the error between the expected and actual observed state of the controlled system, it has an internal model and it is hierarchical.

Cons This model is only qualitative and does not prescribe a way to implement the adaptive mechanism and what is the internal model that should be used.

3.4 Adaptive control models

In this section we describe two inherently adaptive algorithms that do not have explicit hand-crafted rules for adaptation.

3.4.1 Incremental Non-Linear Dynamics Inversion (INDI)

An INDI controller is an extension of a non linear dynamics inversion (NDI) controller, therefore we briefly introduced the latter in the following paragraph.

NDI The basic idea behind NDI is that if you have a model of the dynamics of your system, then you could invert to find a specific action to obtain for example a desired acceleration. This allows you to design an outer loop linear control system to control directly the desired acceleration and then map it to a control action through the inverted model.

For example consider the following system

$$\begin{cases} \dot{x} = f(x) + G\bar{u} \\ \bar{y} = h(x) \end{cases} \quad (3.6)$$

We would like to obtain an expression of the kind

$$\frac{d^N y}{dt^N} = v \quad (3.7)$$

because then the problem is reduced to the design of a *linear* control system with input v . A convenient formula for v is

$$v = \sum_{i=0}^N k_i \frac{d^i(y - y_{ref})}{dt^i} + \frac{d^N y_{ref}}{dt^N} \quad (3.8)$$

because substituting it in Equation 3.7 we can obtain an homogeneous ordinary differential equation of $e = y - y_{ref}$.

Now to obtain a mapping from v (virtual input) to u (actual input to the system) we want to differentiate the output y in Equation 3.6 until the control input u appears in it

$$\dot{\bar{y}} = \frac{\partial h}{\partial x} \dot{x} = \frac{\partial h}{\partial x} f(x) + \frac{\partial h}{\partial x} G\bar{u} := b(x) + a(x)\bar{u} \quad (3.9)$$

For this example we assume the order of the system is 1 which means we have only to differentiate the output once to express it as an explicit function of the control input u . In this way we find that the function that maps v unto u is

$$u = a(x)^{-1}(v - b(x)) \quad (3.10)$$

Therefore by designing a linear controller, w.r.t the system acceleration and with output v , as defined in Equation 3.8, we can then obtain the actual control input u using Equation 3.10.

INDI The INDI controller does not need an explicit model of the system $f(x)$ and relies only a very precise measurement of the state derivatives \dot{x}_m and the control effectiveness matrix G . A very important assumption is that the sampling rate of the sensors is high enough such that between two measurement the effect of the autonomous system dynamics can be neglected. We show here a simple derivation of the INDI controller.

If we consider a $\dot{x} = f(x, u)$ as our system to control and we Taylor expand around the current state x_0 and input u_0

$$\dot{x} = f(x, u) = f(x_0, u_0) + \frac{\partial f}{\partial x}_{x=x_0, u=u_0} (x - x_0) + \frac{\partial f}{\partial u}_{x=x_0, u=u_0} (u - u_0) + H.O.T \quad (3.11)$$

and we assume that the system dynamics related to u are more influential than the one related to x and that we can measure fast enough the state derivatives, we can simplify the equation further

$$\dot{x} = \dot{x}_{0m} + \frac{\partial f}{\partial u}_{x=x_0, u=u_0} (u - u_0) = \dot{x}_{0m} + G(x_0) \Delta u \quad (3.12)$$

As we did for in the previous chapter we can use a virtual input to linearize the system $\dot{x} = v$ and the actual control input change Δu can be obtained by

$$\Delta u = G(x_0)^{-1} (v - \dot{x}_{0m}) \quad (3.13)$$

A scheme of the INDI controller is shown in Figure 3.7. The virtual input can be designed using Equation 3.8.

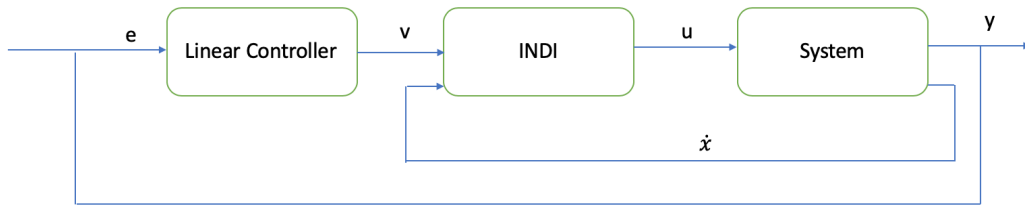


Figure 3.7: Block diagram representation of an INDI controller

Example We present here a simple example for a linear system (note that if the system is linear most of the benefits of NDI and INDI controller is lost), that should convince the reader that an INDI controller is definitively **not** right scheme to model human controllers.

Given the following system:

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ -k & -b \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \\ y = x \end{cases} \quad (3.14)$$

We could define the following outer loop linear controller

$$v = k_p(x - x_{ref}) + \frac{dx_{ref}}{dt} \quad (3.15)$$

Taking the pseudo inverse of the control action we can recover the actual input actual input

$$u = \begin{bmatrix} 0 & 1 \end{bmatrix} (v - \dot{x}_{0m}) \quad (3.16)$$

For second order system the vector \dot{x}_{0m} contains both the velocity and acceleration of the system. You can immediately spot two problems:

- the visual estimation of the acceleration of the state is very poor for humans
- humans are heavily delayed and INDI requires very fast sampling rates, even a small delay can make the system unstable.

Pros -

Cons While an INDI controller relies on a reliable and fast the measurement of its state derivatives. Given this constraint we can hope at best to model how humans control first order systems. Humans cannot easily perceive visually derivatives or order higher then one. Furthermore humans are also heavily delayed, which also violates the assumption of fast measurements for the INDI controller. There are more drawbacks from a theoretical prospective in using INDI to model human adaptive control behavior. All humans rely on a model of the dynamics of the vehicles they are controlling [42]. Furthermore, INDI adaptation is very limited and does not allow the system controlled to change its order or radically change its dynamics. The virtual input, which is pre-specified, would have to change when the order of the system changes. Given these observation we see not reason to further investigate INDI controllers as possible model of human manual control policy.

3.4.2 Model Reference Adaptive Control (MRAC)

In this subsection we briefly describe MRAC. This method was selected to be further explored and a whole chapter is dedicated to its in-depth description.

MRAC is an adaptive control scheme that requires minimal knowledge about the dynamics of the system. Only the sign of the input matrix must be known in advance. In its essence, the MRAC scheme creates a controller that converges to a specified reference controller, called the reference model.

McRuer observed that the structure of the open-loop dynamics of a system with a human in the loop is invariant to changes in the dynamics of the controlled plant. We use the dynamics specified in section 3.1 to define a reference model that should be followed independently of the controlled system.

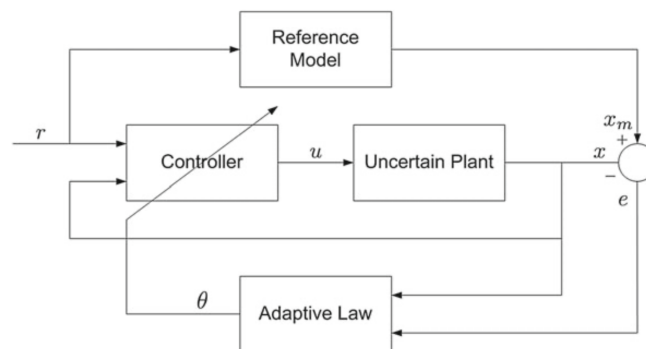


Figure 3.8: MRAC controller scheme[43] for a pursuit configuration

Pros The MRAC controller provides strong guarantees of convergence of the time-varying controller towards the reference model. It can handle a very wide range of dynamical systems of different orders. The MRAC controller can be extended to MIMO systems that could be used to describe the control policy of humans in more complex tasks.

Cons One of the disadvantages of MRAC is the fact that in its basic form it cannot capture changes in the crossover frequency or delay without knowing in advance when such changes should happen. Furthermore, the reference model and controlled plant should have the same structure and order to ensure asymptotic convergence.

3.5 Machine learning methods

In this section, we briefly describe a few ideas that have the potential to model human control behavior even in high dimensional input spaces, such as controlling a vehicle given a video as input.

3.5.1 End to end Learning

Neural Networks are extremely promising for learning to control time-varying dynamical systems with a high level of reliability. Tesla and Waymo are using mostly an end to end learning to create self-driving cars. This approach requires a lot of data that is provided by drivers and sensors on board of the car, about the correct trajectory and behavior. Driving a car is a highly non-linear control task and the car itself is a time-varying dynamical system. While the reliability of self-driving systems is not high enough for driving in all conditions and situations, these systems perform well in highway driving under nominal conditions. Highways driving is again a non-linear time varying-task that requires an adaptive control strategy. Given the relative simplicity of compensatory, pursuit or preview time-varying tasks, it is probable that Neural Network could learn extremely well the adaptive behavior of human operators. In Figure 3.9 is shown one possible architecture for end-to-end learning for a driving task. LSTMs (Long Short Term Memory) networks have been used for system identification of non-linear dynamical systems, a convex superposition of LSTMs seems particularly promising for online identification[44].

Pros With enough training data, the Neural Network model should be able to learn both the steady-state and transient behavior of human operators. If the aim is solely to produce a good model of human adaptation we suggest Neural Networks as the best information processing system for the task.

Cons The main drawback of this method is the interpretability of the results. A large number of parameters in neural networks makes it almost impossible for humans to understand or find a non-trivial explanation for the network behavior.

3.5.2 Neural ODEs

A promising new approach to solve ordinary differential equations (ODEs) is called Neural ODEs. The main idea behind this method is that residual network and recurrent neural networks hidden layers follow an update rule can be seen as a step in an Euler integration scheme [46]. The hidden is updated as follows

$$h_{t+1} = h_t + f(h_t, \theta_t) \quad (3.17)$$

where h_t stands for the hidden layer at time step t and θ_t are the parameters of the network. The objective with Neural ODEs is not to parametrize and learn the response or output of a system but rather to learn it's dynamics. In practice, this means to parametrize with a neural network the differential equation that describes the system behavior. The main advantages of this method are the lower memory costs, lower amount of parameters and the fact that the model

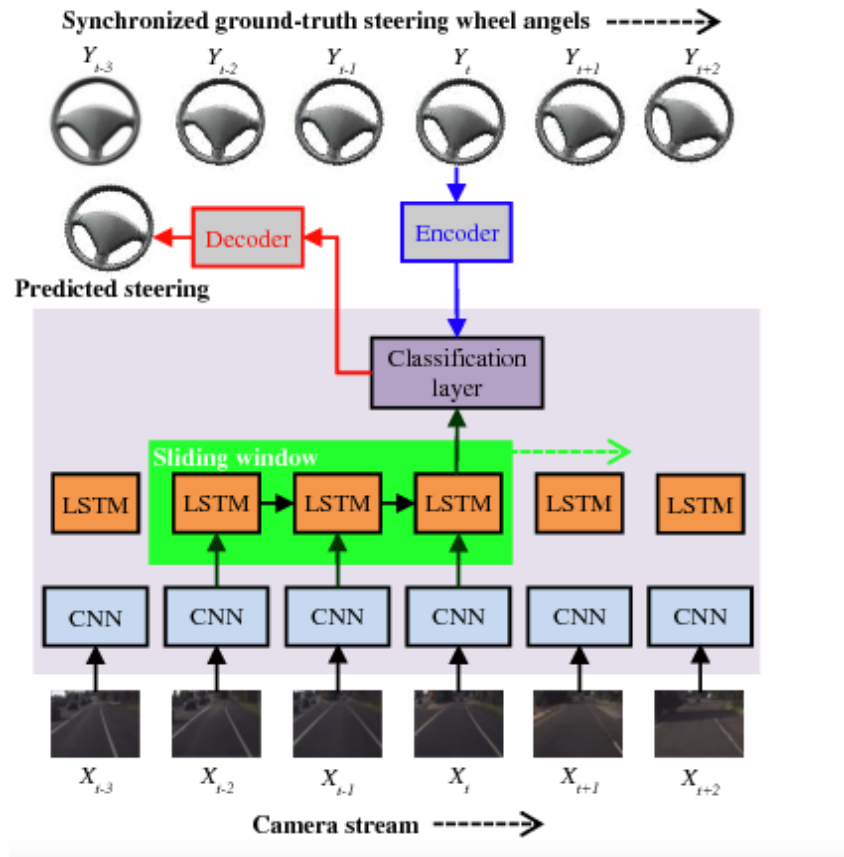


Figure 3.9: End-to-End learning architecture [45]

is a continuous-time model. In Figure 3.10 is shown the reconstruction of the spiral trajectory of a system done by a Recurrent Neural Network (RNN) and a Neural ODE.

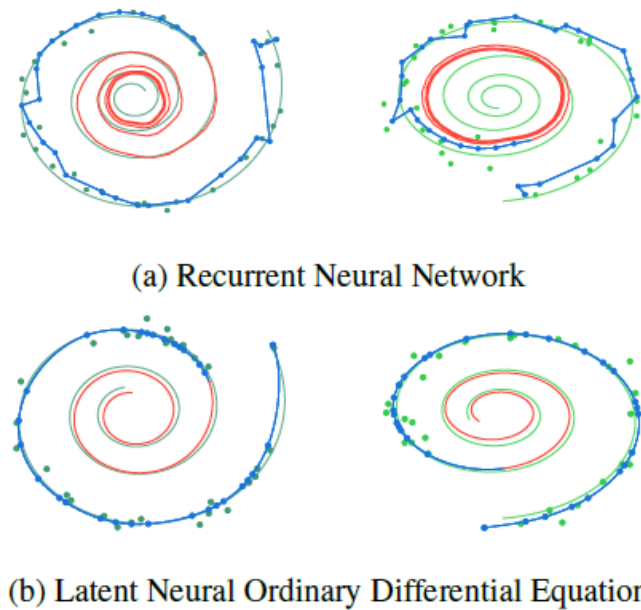


Figure 3.10: Reconstruction and extrapolation task comparison between a RNN and Neural ODE [46]

Pros It is a very natural way to solve ODEs using neural networks and it has been shown to work effectively for fluid problems [46].

Cons One of the main drawbacks of this method is that, as of now, Neural ODEs can model only autonomous systems, which are not driven by control inputs.

3.5.3 Weight Agnostic Neural Networks (WANNs)

The main idea behind this architecture is that humans have learning algorithms directly encoded in the topology of their neural networks. The topology of our brain has been optimized and transformed by evolution throughout hundreds of millions of years and throughout of billions of sentient living information processing beings like ourselves [47]. Humans learn extremely fast even from a few examples while current neural networks are much less *efficient*. Ducks after just hatched can swim and snakes and lizards show innate behavior in avoiding predators [48]. These are just of the few examples of behavior that is caused by the topology of the brains of these systems. Therefore instead of just focusing on training the weights of the network, we should investigate what neural architectures can improve learning itself. The invention of the Convolutional Neural Network is one of the examples where the topology highly facilitates the processing of information coming from images. WANNs are evolved by running an evolutionary network topology optimization algorithm called NEAT [49], that gradually improves the network topology based on how it performs on a task. The weights of the network are all kept equal to 1 so that all the observed behavior is a consequence of the network's topology. The authors of the original paper managed to create simple networks able to solve cart-inverted pendulum problem and to control a bipedal robot in simulation [49]. In Figure 3.11 is shown the topology of the best network for the cart-inverted pendulum control problem.

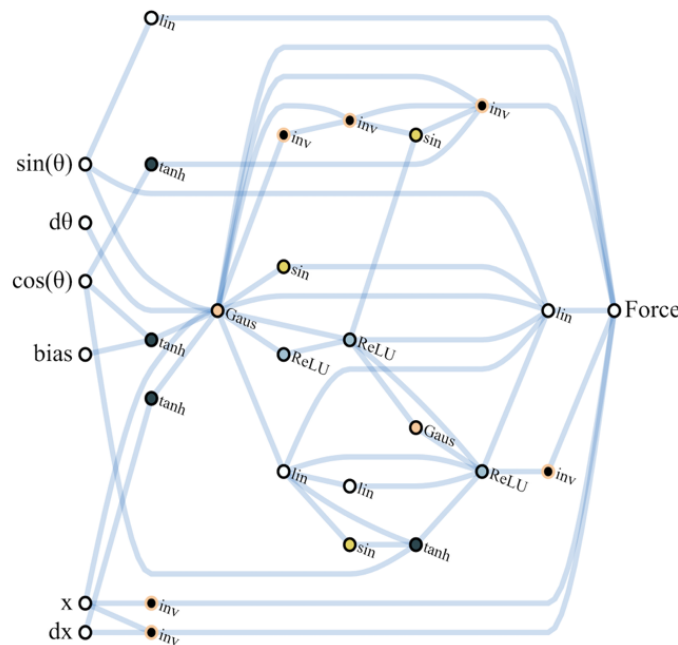


Figure 3.11: WANN architecture for the cart-pendulum task[46]

This model could be useful for the current research because it can evolve adaptive control strategies with the functional building blocks and sensor data available to humans. The fitness of the model could be related to both minimizing the tracking error and following, for example, McRuer model. The sensor data available to human can be restricted, for a compensatory display to the error and its derivative.

Pros This model matches what we currently know about the brain of most animals: the topology of the brain already encodes learning algorithms to solve problems. It is an evolutionary model that given the right task and building blocks could perform well in a large variety of control problems.

Cons It might be very difficult to evolve such a system for complex tasks. Furthermore the system that evolves to solve a specific task, effectively lives in another universe and not necessarily will show the characteristic of a biological controller such as humans. Loosely speaking, we evolved to solve many different tasks therefore the objective or fitness function that has been optimized to produce our motor system is very different from the one set up to solve a simple control task.

Research Objectives

The research project aims at identifying human in the loop adaptive behavior for changes in dynamics of the controlled vehicle. It started initially by investigating if a model-free control technique called INDI could be used to develop an adaptive controller that would mimic the way humans adapt. Unfortunately the way the INDI works prevents it from adapting to strong changes in the dynamics of the controlled system. Furthermore INDI adapts using all the variables present on the state space of a system, this means that potentially it could use derivative of higher degree than two, which cannot be perceived by humans.

By investigating several other adaptive control techniques we found that Model Reference Adaptive Control (MRAC), an adaptive control technique, could effectively be used to model human adaptation. MRAC can be heuristically justified since it implements some of the important principles, like predictive coding, that neuroscientists discovered are at the base of human cognition. Nonetheless MRAC distances itself from current state-of-the-art models for the field of human sensorimotor control. The state of the art is currently a flavor of optimal feedback control, which is supported by a large body of evidence. OFC does not require the choice of internal model, like in MRAC, and can do some basic motor and planning tasks. It's not trivial though to adapt optimal control to changes in the dynamics of the controlled system. Another line of research that looks very promising is the implementation in state space of "Active Inference" (a form of predictive coding), presented in section 2.2 and supported by Friston [50].

Therefore the research question of the project is:

Do humans use internal models and are prediction errors key elements for sensory-motor adaptation under change of dynamics of the controlled elements?

To answer this research question we decided to model the human control policy with an adaptive control scheme called MRAC (Model Reference Adaptive Control). This choice is justified in the following chapter.

The overall research questions can be broken in further theoretical and experimental sub-questions.

Theoretical sub-questions:

- What is the most appropriate model reference?
- How can you approximate non-linear dynamics in the model reference?
- The existence of optimal feedback parameters is guaranteed only if the model reference has the same structure as the controlled element, is there still convergence when this assumption is relaxed?
- What is the influence of the learning rate on the convergence of the model?
- Do the estimated gains drift in absence of a persistent excitation of the system?

-
- How should the system identification be performed? What parameters should be optimized to match the observed data?
 - Can a time-varying reference model be used in MRAC?
 - What is the relation between MRAC and predictive coding?
 - What datasets can be used to validate the method before performing an experiment?

Experimental questions:

- Do humans use prediction error to drive the adaptation?
- Does the rate of change of changing controlled dynamics affect how fast human controllers adapt?
- Is the speed of human adaptation symmetric with respect to switches in the controlled plant between two dynamical systems?

Experimental design questions:

- What kind of experiment will allow to determine if MRAC is indeed a good model of human motor adaptation?
- What kind of experiment will disprove the MRAC hypothesis?
- What are the dynamics that the operator should control?
- What experiment will allow to determine if MRAC also models correctly the transient dynamics exhibited by humans during adaptation?

The rest of this report will address most of the theoretical sub-questions and formulate a potential plan for the experiment to be conducted.

Model Reference Adaptive Control

This chapter is dedicated to an in-depth explanation of MRAC, why it was selected to pursue the research objective, its theoretical justification and its ties to predictive coding.

5.1 MRAC main ideas

We start by explaining the main ideas behind MRAC, how predictive coding and the presence of internal models provide a justification for it. We will explain how it possibly relates to the accepted view that human motor control can be best described and explain in light of optimal feedback control.

Humans are highly adaptable MRAC is an adaptive control algorithm that allows the controlled element to track the trajectory of a reference model. The direct MRAC modifies the controller gains online and can adapt very rapidly to significant changes in the measured dynamics. MRAC can show asymptotic convergence for the parameters and bounded convergence in the presence of delays/non-minimum phase behavior.

Humans have internal models We believe that MRAC can represent how humans adapt their motor behavior because humans also have created internal models of the expected dynamics. While controlling the object, and receiving mostly visual feedback from the screen, humans can learn the dynamics of the controlled element. This new model is used together with the already present internal neuromuscular model. Let's consider the example of steering a bike: the cortex can make use of an internal model of a bike to plan it's next steps to follow the road. It can test whether under the current policy (center of gravity shift and handle rotation) the can be achieved and can simulate the next expected state of the bike. The gravity shift and handle rotations are also transformed into lower level commands thanks to internal models of our neuromuscular system. The process of predictions and state estimation can continue down the hierarchy until arriving at the level of regulating muscle fibers activations. This open loop predictive model is then used by the human create an expectation about the future states of the system

Open loop dynamics McRuer measured that the structure of open loop dynamics of humans controllers in tracking or compensatory tasks tend to be approximately the same independently of the controlled element close to the crossover region. As we reviewed in section 3.1, humans' open loop dynamics resemble a first order system or "integrator-like" dynamics.

$$H_{OL} = \frac{\omega_c}{i\omega} e^{-i\omega\tau}, \omega \approx \omega_c \quad (5.1)$$

The fact that the open-loop dynamics always resemble single integrator dynamics, makes the problem particularly suited for MRAC whose reference model (closed-loop) is invariant with respect to the controlled dynamics. In this formulation the crossover frequency and delay are kept fixed. This possibly limits the ability of the model to predict phenomena like crossover regression [29].

Adaptation is triggered by mismatched expectations Predictive coding states that our brain creates an internal model of the world and continually predicts the future states across all the neural hierarchies. The errors, the difference between the predicted and observed states, are propagated up the hierarchy to drive actions and/or change the encoded world models in the neural network. The aim is to reduce surprise and minimize errors. MRAC works in a similar way: the adaptation is driven by the difference between the predicted behavior of the internal model and the observed one. The adaptation continues until the observed behavior matches the expected models, by minimizing a Lyapunov function, that can be interpreted as measuring the level of surprise experienced.

Figure 5.1 shows the block diagram of the MRAC controller. You can observe that the error between the expected output and the measured one is feedback to change the controller.

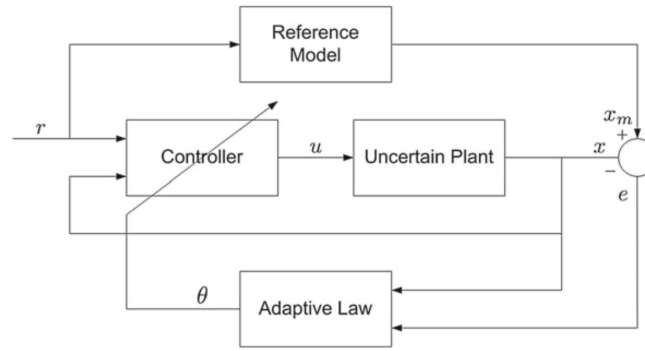


Figure 5.1: MRAC scheme for a pursuit display [43]

5.1.1 MRAC for second-order SISO systems

In this subsection, we go into the mathematical details of MRAC and derive its formulation for second-order dynamical systems. These are the models of the kind of systems that are usually controlled by humans. Let's assume that the reference model can be written in state space form as

$$\dot{x}_m = A_m x_m + B_m r \quad (5.2)$$

where x_m is the state of the reference model and r is the reference signal. The controlled element can also be expressed in state space form as

$$\dot{x} = Ax + Bu(t) \quad (5.3)$$

where $u(t)$ is the time-varying control input to be found. We assume that the state space matrix A and the control input matrix B are unknown, which reflect the uncertainty about the dynamics of the system. The aim of the control action is to minimize the error $e(t) = x_m - x$. We assume the form the feedback control signal to have the following form

$$u = K_x(t)x + k_r(t)r \quad (5.4)$$

where $K_x(t)$ and $k_r(t)$ are time-varying gains. If we substitute Equation 5.4 into Equation 5.3 we find that the ideal gains K_x^* and k_r^* to obtain perfect tracking of the reference signal are

$$\begin{aligned} A + BK_x^* &= A_m \\ Bk_r^* &= B_m \end{aligned} \quad (5.5)$$

This set of equations is solvable only if A and A_m share same structure. To simplify the derivation we define two new quantities \tilde{K}_x and \tilde{k}_r , which are the gain deviations from their optimal values

$$\begin{aligned} \tilde{K}_x &= K_x(t) - K_x^* \\ \tilde{k}_r &= k_r(t) - k_r^* \end{aligned} \quad (5.6)$$

By expressing the feedback law in Equation 5.4 in terms of \tilde{K}_x and \tilde{k}_r and substituting it into Equation 5.3 we obtain

$$\dot{x} = (A + BK_x^* + B\tilde{K}_x)x + (Bk_r^* + B\tilde{k}_r)r \quad (5.7)$$

Now we can find an expression for the error (which we want minimize) by making use of Equation 5.7 and Equation 5.5

$$\dot{e} = \dot{x}_m - \dot{x} = A_m e - B\tilde{K}_x x - B\tilde{k}_r r \quad (5.8)$$

MRAC relies on the Lyapunov stability theory to prove the stability of the system. To ensure tracking of the reference model and therefore that $\lim_{t \rightarrow \infty} e(x) = 0$ we need to satisfy the following conditions:

1. Find a Lyapunov function $V(t, e, \tilde{K}_x, \tilde{k}_r) > 0 \forall t > 0$
2. $\frac{dV(t, e, \tilde{K}_x, \tilde{k}_r)}{dt} < 0$ for all $t > 0$
3. $\frac{dV(t, e, \tilde{K}_x, \tilde{k}_r)}{dt} \in L_\infty$ norm, i.e $\frac{d^2 V(t, e, \tilde{K}_x, \tilde{k}_r)}{dt^2}$ must be bounded

There are no real guidelines to choose the Lyapunov function, but usually it is a quadratic function with respect to the variables of interest. It can also be interpreted as an energy function. In this case we can choose the following function [43]

$$V(e, \tilde{K}_x, \tilde{k}_r) = e^T P e + |b|(\tilde{K}_x \Gamma_x^{-1} \tilde{K}_x^T + \frac{\tilde{k}_r^2}{\gamma_r}) > 0 \quad (5.9)$$

where b is the only entry of the control effectiveness matrix B , with $P, \Gamma_x^{-1}, \gamma_r > 0$, the function is also bigger than zero at all time. Now by taking the time derivative of the Lyapunov function we obtain

$$\dot{V}(e, \tilde{K}_x, \tilde{k}_r) = \dot{e}^T P e + e^T P \dot{e} + |b| \left(2\tilde{K}_x \Gamma_x^{-1} \dot{\tilde{K}}_x + 2\frac{\dot{\tilde{k}}_r \tilde{k}_r}{\gamma_r} \right) \quad (5.10)$$

By substituting for the expression for the error we obtain

$$\begin{aligned} \dot{V}(e, \tilde{K}_x, \tilde{k}_r) = & -e^T (PA_m + A_m P)e + 2|b|\tilde{K}_x(-xe^T \bar{P} \text{sgn}(b) + \Gamma_x^{-1} \dot{\tilde{K}}_x^T) \\ & + 2|b|\tilde{k}_r(-re^T \bar{P} \text{sgn}(b) + \frac{\dot{\tilde{k}}_r}{\gamma_r}) \end{aligned} \quad (5.11)$$

and by selecting P to satisfy the Lyapunov equation

$$PA_m + A_m^T P = -Q \quad (5.12)$$

we obtain

$$\begin{aligned} \dot{V}(e, \tilde{K}_x, \tilde{k}_r) = & -e^T Q e + 2|b|\tilde{K}_x(-xe^T \bar{P} \text{sgn}(b) + \Gamma_x^{-1} \dot{\tilde{K}}_x^T) \\ & + 2|b|\tilde{k}_r(-re^T \bar{P} \text{sgn}(b) + \frac{\dot{\tilde{k}}_r}{\gamma_r}) \end{aligned} \quad (5.13)$$

where Q is negative definite matrix. In order for the Lyapunov function derivative to be negative at all time we have to impose the following conditions

$$\begin{aligned} -xe^T \bar{P} \text{sgn}(b) + \Gamma_x^{-1} \dot{\tilde{K}}_x^T &= 0 \\ -re^T \bar{P} \text{sgn}(b) + \frac{\dot{\tilde{k}}_r}{\gamma_r} &= 0 \end{aligned} \quad (5.14)$$

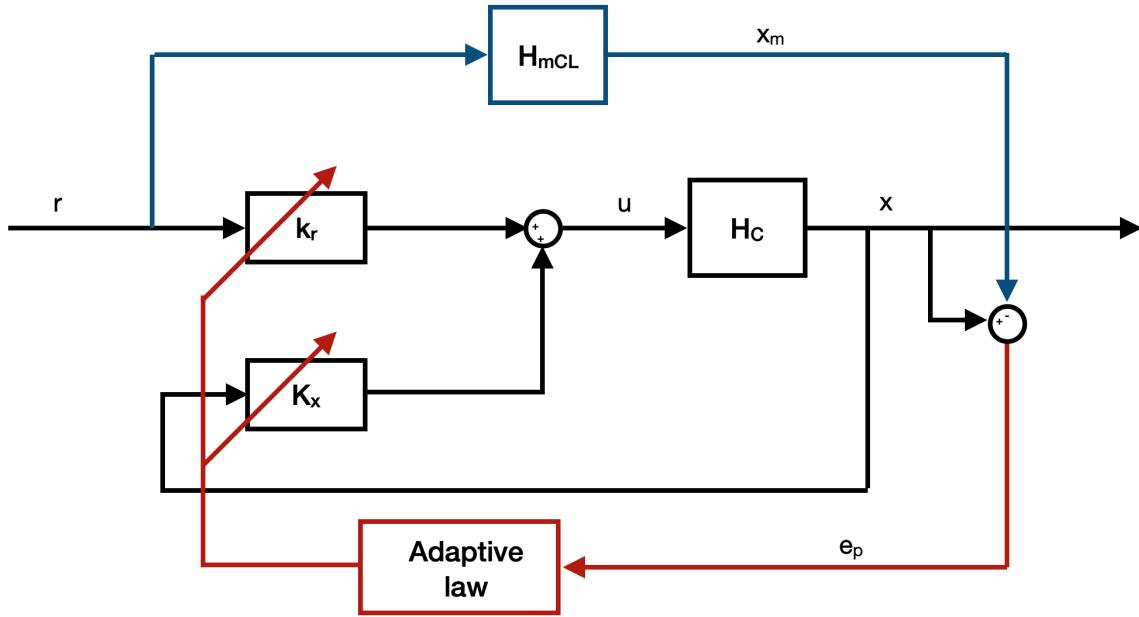


Figure 5.2: Diagram of the MRAC controller. The gains of the controller are K_x and k_r , the controlled plant is H_c , the closed loop reference model H_{mCL} and the prediction error e_p . Signals and diagrams related to predictions using the reference models are in blue while the ones directly used to adapt the gains are in red.

which implies

$$\begin{aligned}\dot{\tilde{K}}_x &= \Gamma_x x e^T \bar{P} \operatorname{sgn}(b) \\ \dot{\tilde{k}}_r &= \gamma_r r e^T \bar{P} \operatorname{sgn}(b)\end{aligned}\tag{5.15}$$

In this way we found an expression for the rate of change of the feedback gain to ensure the tracking of the reference model. Finally we just have to show that the derivative of the Lyapunov function is bounded. The derivative of the Lyapunov function satisfies the following inequality

$$\dot{V}(e, \tilde{K}_x, \tilde{k}_r) = -e^T Q e \leq -\lambda_{\min} \|e\|_2^2\tag{5.16}$$

where λ_{\min} is the smallest eigenvalue of the matrix Q . Since $\|e\|_2^2 \in L_\infty$ we can conclude that $\dot{V}(e, \tilde{K}_x, \tilde{k}_r)$ is bounded. By using Barbalat's lemma [43] we conclude that $\lim_{t \rightarrow \infty} e(x) = 0$.

In Figure 5.2 you can see the a diagram representing the used implementation of the MRAC controller.

5.1.2 Remarks

The model error $e_p(t) = x_m(t) - x(t)$ can be interpreted as difference between the expected model error and the actual error that is observed, if we add and subtract the reference signal r to the model error:

$$e(t) = x_m(t) - x(t) = r - x(t) - (r - x_m(t)) = e_{act} - e_{exp}\tag{5.17}$$

this quantity was used by L.R. Young to qualitatively design an adaptive controller [16]. The Lyapunov function can also be interpreted as a measure of surprise since it's an always positive quantity that measures the deviation from the expected response. As in predictive coding neural system is adjusting the gains of the feedback control to minimize ($\dot{V} < 0$) surprise [50].

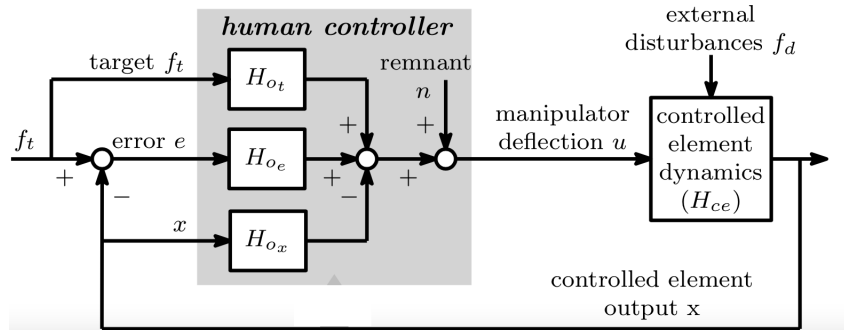


Figure 5.3: three channel model for modelling human behavior in a tracking task (objective is to follow signal f_t on a pursuit display). The transfer function H_{ot} represent the feedforward component of the controller, H_{ox} the state feedback component while H_{oe} is the fed the tracking error.

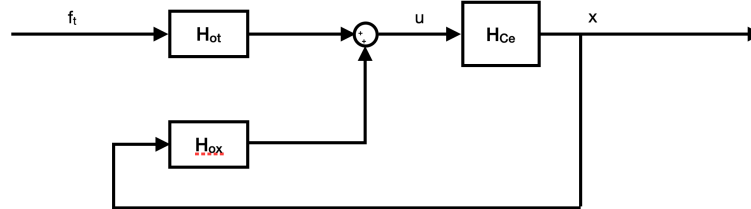


Figure 5.4: Two channel model for modelling human behavior in a tracking task (objective is to follow signal f_t on a pursuit display). The transfer function H_{ot} represent the feedforward component of the controller, H_{ox} the state feedback component.

5.1.3 Comparison with classical pursuit architectures

Models of human operators for manual control tracking task using a pursuit display have been already developed. These models though are often linear and time-invariant, characteristic that strongly limit their usability. Mulder et al. give an overview of current challenges and state of the art for manual pursuit tracking [51]. A general controller that makes use of three available features is a three channel model as shown in Figure 5.3. The three channel model can be further simplified since the the tracking error, the state and the signal are linearly related. A simplified version is a two channel model with only the feedforward and state feedback components as shown in Figure 5.4.

Interpretation of MRAC coefficients At every instant in time the MRAC controller is equivalent to a two channel model, with a state feedback and a state feedforward component. For the moment we ignore the time delay but we will analyse it late. The state feedback transfer function is related ot the MRAC gain $K_x = [k_{x1}, k_{x2}]$ by

$$H_{ox} = k_{x1} + k_{x2}s \quad (5.18)$$

while the feedforward transfer function corresponds to the MRAC feedforward gain

$$H_{ot} = k_r \quad (5.19)$$

This correspondence is exact if the delay is explicitly introduced in the MRAC controller. The interpretation of the gains in therefore straightforward: higher k_r and k_{x1} imply high level of proportional control and high value of k_{x2} imply that model is using lead (or derivative control).

Otherwise if the delay is not explicit, the MRAC controller can still approximate the delay by changing the feedback gains and feedforward gains in a non trivial way. For example if the reference model is chosen to be a simple integrator with delay, we can use the Padé approximation of a first order to create a second order system, that approximate the desired behavior and can be followed by the MRAC controller

$$H_{ref} = \frac{\omega_c}{s} \exp(-\tau s) \approx \frac{\omega_c(2 - \tau s)}{s(2 + \tau s)} \quad (5.20)$$

5.1.4 Delay in MRAC

In the previous derivation we assumed no delay in the control input to derive a simple adaptive control law. Unfortunately humans have delays in perception and actuation delays that should be accounted for. The presence of delay can severely affect the performance and stability of a linear state feedback controller. An adaptive state feedback controller, such as MRAC, is also affected by such delay. In this section we briefly reformulate the problem including the input delay and direct the reader to possible improvements on a the model considered in this thesis.

More formally we can write the system controller by the human operator as

$$\dot{x}(t) = A(t)x(t) + B(t)u(t - \tau) \quad (5.21)$$

$$u(t) = K_x x(t) + k_r r(t) \quad (5.22)$$

where τ stands for the input delay. The reference system, which is also delayed in the input, is given by

$$\dot{x}_m(t) = A_m x_m(t) + B_m r(t - \tau) \quad (5.23)$$

More generally the design of stable controllers and their analysis in presence of delays and uncertainties can be done with Lyapunov-Krasovskii functional techniques [52]. A comprehensive tutorial and review on Liapunov-methods for time-delayed system was written by Fridman [53]. For example consider a system with a constant input time delay τ

$$\dot{x}(t) = Ax(t) + A_1 x(t - \tau) \quad (5.24)$$

where $A_1 = BK$. We can write a functional of the form

$$V(x_t) = x^T(t)Px(t) + \int_{t-\tau}^t x^T(s)Qx(s)ds \quad (5.25)$$

with P and Q positive semidefinite, that satisfies the condition $V(x) > 0 \forall x \text{ and } \forall t > 0$. By differentiating the Lyapunov function and applying the stability conditions we can derive the following two conditions, which guarantee stability [53]

- A and $A + A_1$ are Hurwitz
- $A^{-1}A_1$ is Schur matrix, i.e. all eigenvalues are inside the unit circle.

Therefore for each choice of feedback matrix K we can check with the above conditions if the system is stable.

Specifically for evaluating the stability and robustness of MRAC controllers, Nguyen proposes three different methods that make use of [54]:

- Padé approximation for delay
- Lyapunov-Krasovskii functionals with sum of squares optimization

- Matrix measure method

Having a more robust MRAC controller or more information on its stability is highly valuable when doing system identification. In particular, if an optimization routine is selected to find parameters such as delay, gains and the reference model then having a robust controller can prevent the optimization steps from diverging.

5.2 Selection Internal Model

We first examined the open-loop model proposed by McRuer. The model in the Laplace domain has two free parameters: the time delay, τ , and the crossover frequency, ω_c . These two parameters can be found by a general non-linear optimization procedure or using the McRuer rules [29].

To check the validity of the model, we fitted it on data provided by Zaal, that was collected in a previous experiment [31]. In that experiment, participants were asked to perform a disturbance rejection task, controlling the first order, a second-order and other 6 time-varying dynamical systems.

In Figure 5.5a are shown the reference signal in yellow r , the time-averaged response of the human operator in blue x and the predicted response by the model x_m in red for the control task of a first-order system. The same can be observed in Figure 5.5b where this time the system was of second order. The parameters of the model have been optimized to minimize the mean square error with the actual pilot response. As you can see, while McRuer original model fits relatively well the human operator's response for the control of a first-order system, it fails to capture many of the oscillations present in the control of a second-order system. This is also reflected in the much higher MSE of the model for the case in which the controlled dynamics were of second order.

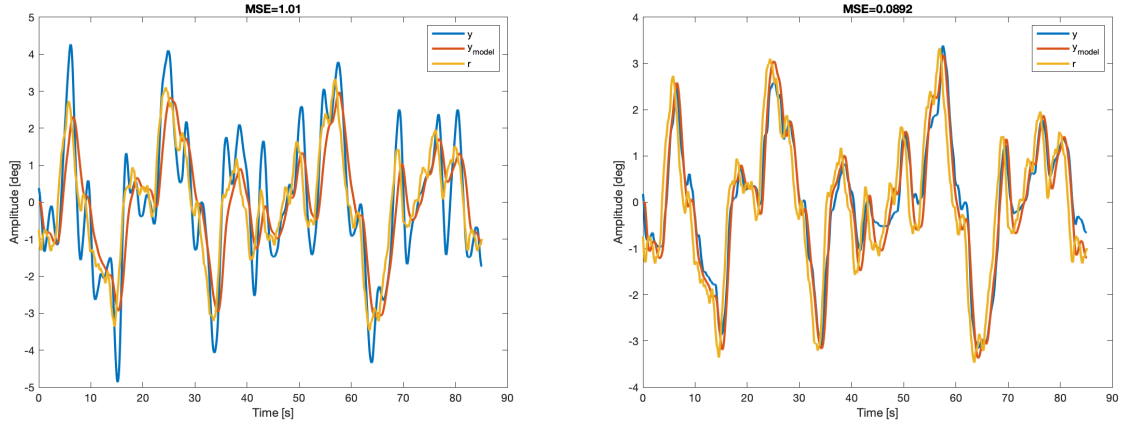
To solve this problem we propose to introduce "harmonic-oscillator" like dynamics that should allow the model to better capture oscillations. From a physical point of view, these new dynamics approximate the behavior of the arm of the controller. The proposed model is

$$H_{OL} = H_p H_c = \frac{\omega_c}{i\omega} \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2} e^{-i\omega\tau} \quad (5.26)$$

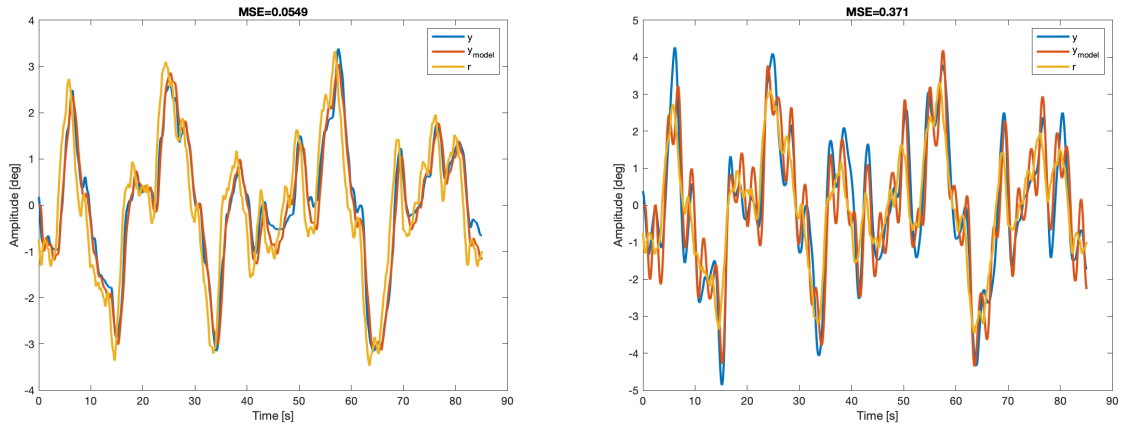
where ω_n and ξ are the natural frequency and damping of the arm dynamics. These coefficients can either be found with an optimization technique or set to $\omega_n \approx 8$ and $\xi \approx 0.7$ (critically damped), as Hess suggests [40].

The results for the new model can be observed in Figure 5.5c and Figure 5.5d. You can observe that the overall fitting improved for both conditions. In particular, the MSE is significantly lower for fitting the control behavior for a second-order system. The higher level of oscillation present in the tracking task for a second-order system is probably due to the higher bandwidth required to control the system. The neuromuscular system could be therefore excited close to its resonance frequency usually between 5/10 rad/s.

Therefore we consider the new proposed model as an alternative to the original McRuer model for the reference controller in the MRAC algorithm. To be able to actually use it, we need to transform it into a state-space representation. This is accomplished by using a Pade approximation for the time delay.



(a) McRuer's original model with a first order system controlled. (b) McRuer's original model with second order system controlled.



(c) New reference model with a first order system controlled. (d) New reference model with a second order system controlled.

Figure 5.5: MSE of reference model fitted over the averaged output of a participant of Zaal's reference tracking experiment [31].

5.3 Control under changing dynamics

In this section, we examine how the MRAC algorithms behave, in a pursuit task, with drastic changes in the dynamics of the controlled plant. In the following section, we consider sharp changes in the dynamics all happening at time $t = 25$ seconds. We tested how the algorithm responded for two classes of dynamical systems: first and second-order systems. For we used the transfer function in Equation 5.27, Equation 5.28 as representative respectively of second, first order systems. Figure 5.6 shows the bode plots for system H_1 and H_2 .

$$H_2(s) = \frac{1}{s(s + 0.1)} \quad (5.27)$$

$$H_1(s) = \frac{1}{(s + 10)(s + 0.1)} \quad (5.28)$$

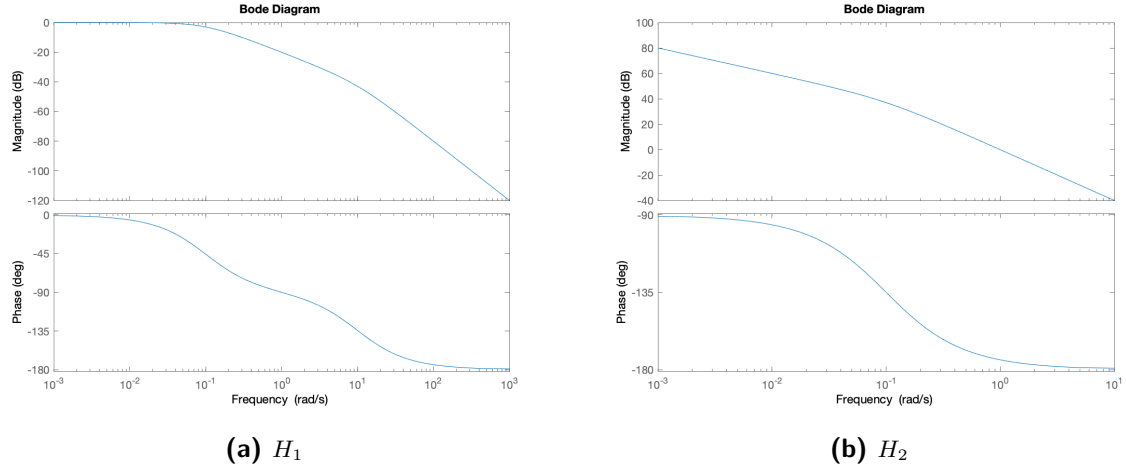


Figure 5.6: Bode plots showing magnitude and phase of two of the controlled systems.

5.3.1 Time-varying systems

To assess the adaptive capacities of MRAC we select the following two time-varying systems H_{21} , H_{12} , to control

$$H_{21}(t) = \begin{cases} H_1, & \text{for } t \leq 25 \\ H_2, & \text{for } t > 25 \end{cases}$$

$$H_{12}(t) = \begin{cases} H_2, & \text{for } t \leq 25 \\ H_1, & \text{for } t > 25 \end{cases}$$

All cases are interesting and we expect a different transient response when going from controlling a double to a single integrator w.r.t single to a double integrator. In the latter case, the control policy of the human operator adapted to control a single integrator will lead to an unstable system when the dynamics are changed.

The open loop reference model used to assess the transitions is

$$H_{OL} = \frac{3}{i\omega} \frac{8^2}{s^2 + 20.78 + 8^2} e^{-i0.3\omega} \quad (5.29)$$

The values of the crossover frequency and delay are chosen equal to 3 rad/s and 0.3 s. These values were chosen because they are in the range of values usually observed in these type of experiment.

5.3.2 Initial conditions and sensitivity to learning rate

To perform the simulations we had to select initial values for the feedback gains k_r and K_x and for the learning rates.

Given the task of controlling the system H_{ij} , the initial value of the gains k_{r0} and K_{x0} are equal to the steady state gains k_r and K_x that MRAC converges to while controlling system H_i . In this way, the initial gains do not change much before the dynamics of the controlled elements are modified.

The value of the learning rate is also important. A higher learning rate can have several effects: it could the convergence towards the optimal gains faster but at the same time, it could make

the controller less robust towards noise and change of dynamics. A learning rate that is too high effectively prevents the system from learning and converging to the reference model dynamics.

We examined the effect of the learning on the control of H_{21} , which involves a rapid change of gains to prevent system output to diverge. In Figure 5.7a, Figure 5.7c, Figure 5.7e are shown the outputs of the MRAC system for an increasing value of the learning rate. A higher value leads to faster convergence and more moderate oscillations during the transient control regime, just after the change of dynamics. In Figure 5.7b, Figure 5.7d, Figure 5.7f you can see the corresponding gains associated to the previous conditions. It is remarkable how fast the gains can change and how small differences in the gains can already lead to very different transient behaviors as shown in Figure 5.7c and Figure 5.7e.

5.3.3 Double to single integrator dynamics

In this subsection we consider the task of controlling the time-varying system H_{12} . The learning rates for the gains are set to $\gamma_{k_r} = 0.2$ and $\gamma_{k_x} = [0.05; 0.1]$. The initial values of the gains were $k_x = [-15, 5]$, $k_r = 15$. In Figure 5.8a you can see the time series of the output of the reference model, of the MRAC operator model and of the reference signal. You can see that at time $t = 25$ seconds.

In Figure 5.8a you can see that just after $t = 25$ seconds the MRAC controller undershoots the reference signal and deviates from the reference model. The system adapts overtime: in Figure 5.8c you can see the change in sign of the gain k_{x1} over a span of 10 seconds approximately. With time the system learns to follow better the signal by increasing the gains k_r and k_{x1} .

5.3.4 Single to double integrator dynamics

In this subsection we consider the task of controlling the time-varying system H_{21} . The learning rates for the gains are set to $\gamma_{k_r} = 0.2$ and $\gamma_{k_x} = [0.05; 0.1]$. The initial values of the gains were $k_x = [-15, 5]$, $k_r = 15$. In Figure 5.8a you can see the time series of the output of the reference model, of the MRAC operator model and of the reference signal. You can see that at time $t = 25$ seconds.

In Figure 5.8a you can see that just after $t = 25$ seconds the MRAC controller overshoots the reference signal and the closed-loop system is unstable. Nonetheless, the system quickly adapts: in Figure 5.8c you can see the change in sign of the gain k_{x1} . With time the system learns to follow better the signal by increasing slightly the gains k_r and k_{x1} .

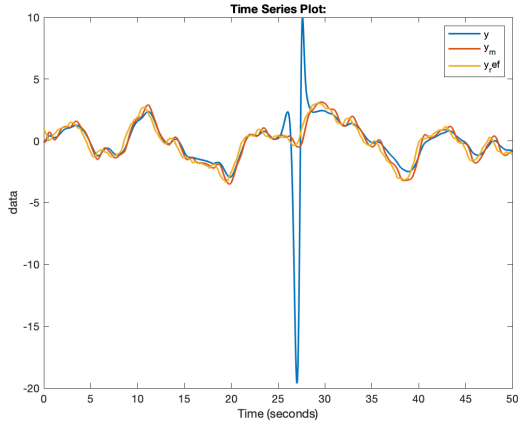
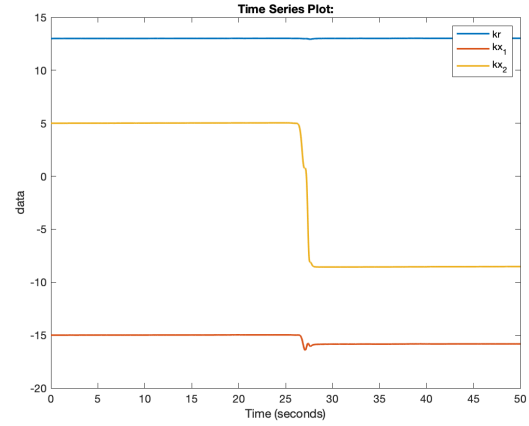
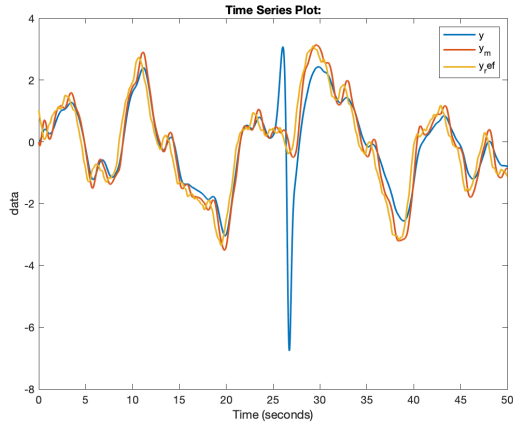
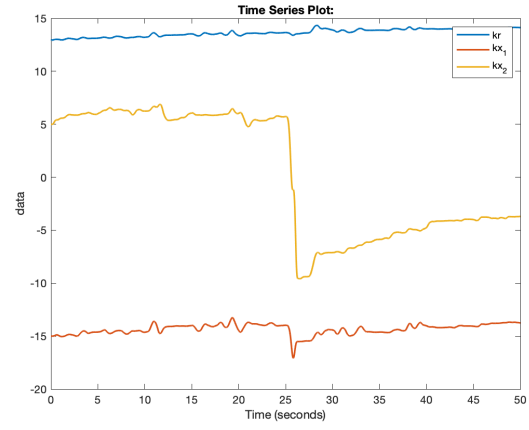
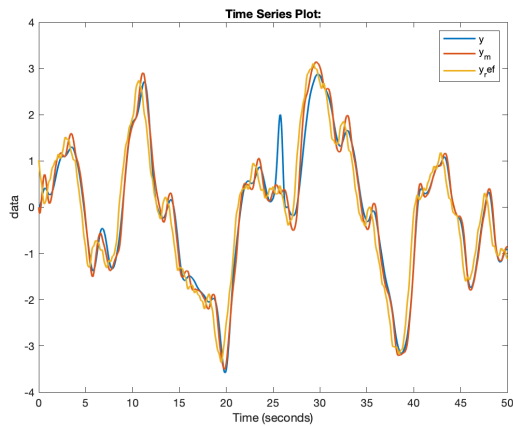
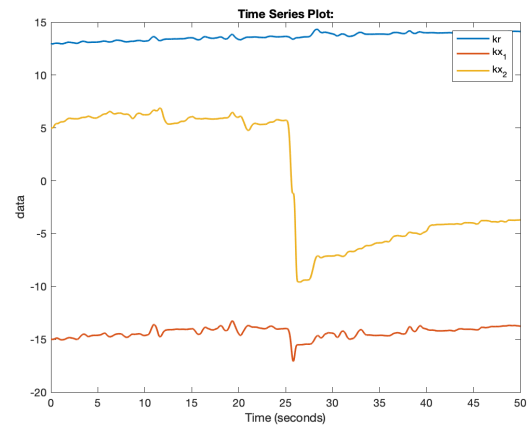
(a) $\gamma_r = 0.001$ and $\gamma_x = [0.0005, 0.0005]$ (b) $\gamma_r = 0.001$ and $\gamma_x = [0.0005, 0.0005]$ (c) $\gamma_r = 0.011$ and $\gamma_x = [0.0055, 0.0055]$ (d) $\gamma_r = 0.011$ and $\gamma_x = [0.0055, 0.0055]$ (e) $\gamma_r = 0.1$ and $\gamma_x = [0.05, 0.05]$ (f) $\gamma_r = 0.1$ and $\gamma_x = [0.05, 0.05]$

Figure 5.7: Time series of the output (left column) and gains (right column) of the MRAC model while controlling system H_{21} . y_{ref} , y and y_m stand for the reference signal, the actual output and the reference model output.

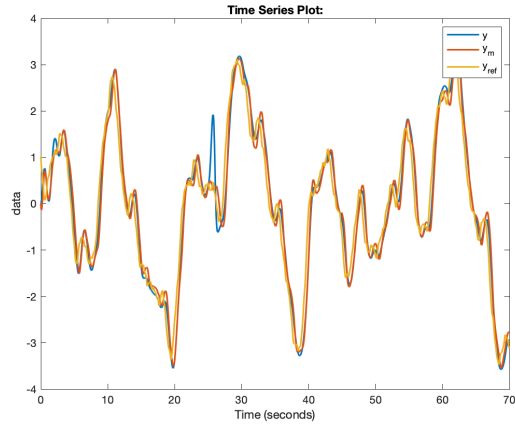
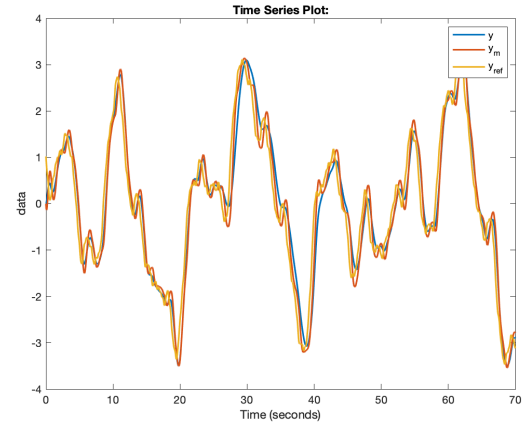
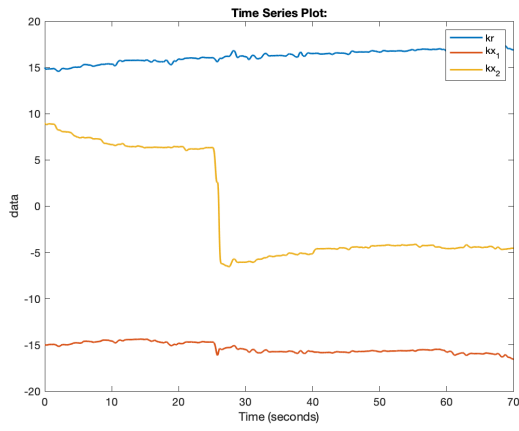
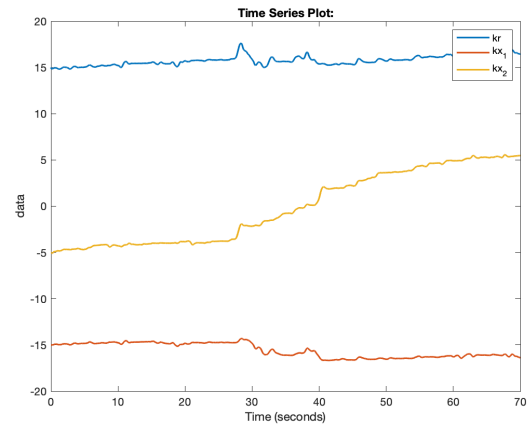
(a) Outputs for H_{21} (b) Outputs for H_{12} (c) Gains for H_{21} (d) Gains for H_{12}

Figure 5.8: Time series of the output of the reference model, y_m , operator (MRAC) model, y , the reference signal, y_{ref} (top row) and the gains (bottom row) of the MRAC model for controlling system H_{21} (left column) and H_{12} (right column)

Future Directions and Conclusion

In this chapter, we discuss the next steps in the research project.

6.1 Method's developments and simulations

There are some limitations to the current methods that could potentially be addressed. To obtain asymptotic convergence towards the reference model the matrices A_m and A must be of the same order and structure. A higher-order for the matrix A_m is desirable to increase the order of the padé approximation as well as to include the effect of the neuromuscular system. There is a formulation of MRAC that is robust and guarantees bounded convergence of the adaptive parameters even in case of a mismatch between the structure of A_m and A . It would be also interesting to include the optimal control modification of MRAC [43] to be consistent with the discovered optimal nature of motor control [55].

We plan to verify the current method using the data collected by Zaal in a tracking experiment with a plant with time-varying dynamics [31]. In particular, we plan to verify the quality of fitting of the data and find the optimal value of the learning rate of the MRAC controller and the delay and crossover frequency of the reference model.

6.2 Preliminary experimental plan

With the experimental plan, we plan to test whether MRAC is a good model of the human motor control system. The experiment design is simple to reduce the number of participants for the experiment.

Participants will be asked to perform a roll compensatory tracking task using a display, similar to Figure 6.1, where the error, between the current state and the reference signal, is visualized. The current state can be modified with the help of a joystick. The task of the participant is to minimize the difference between the reference signal and the current state. The reference signal is be a sum of sinusoidal functions at several frequencies.

The task will consist in controlling the system H_{r21} , a time-varying transfer function that alternates between H_2 and H_1 . A single run will last 100 seconds, the participant will have to control system H_1 for the first 40 seconds, then system H_2 until the 70 seconds mark and finally system H_1 again until the end of the run. The first 10 seconds of the run will be discarded.

In total there are two conditions. The first condition corresponds to an abrupt change in the controlled dynamics in "step-like" fashion. In the second condition the change will happen more gradually during an interval of 10 seconds. We chose these two conditions to test weather the current adaptive scheme is able to model both situations.

The measured run will be ten per condition. Five runs will be used to train the model and the other five to validate it. The five validation runs will be randomly distributed across the 10

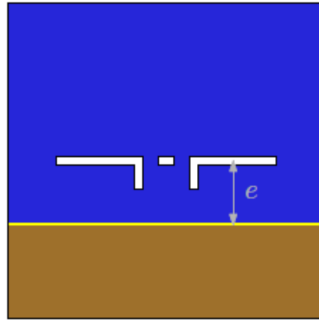


Figure 6.1: Example of a display that could be used for the experiment

runs. The validation and training set will have two different forcing functions. The sinusoidal forcing functions will have the same set of frequencies and amplitudes but different phases. The forcing function across different runs of the same set will be the same. This condition is necessary to obtain an average policy by averaging across multiple runs the participant response. The average or expected policy will then either be used to either train or validate the model.

Sufficient training will be given to the participants until they reach asymptotic performance. For each training run a forcing function with different phases will be used to ensure that the participants don't memorize the moment in time at which the system changes.

A sketch of the experimental procedure for the case H_{r21} is shown in Figure 6.2.

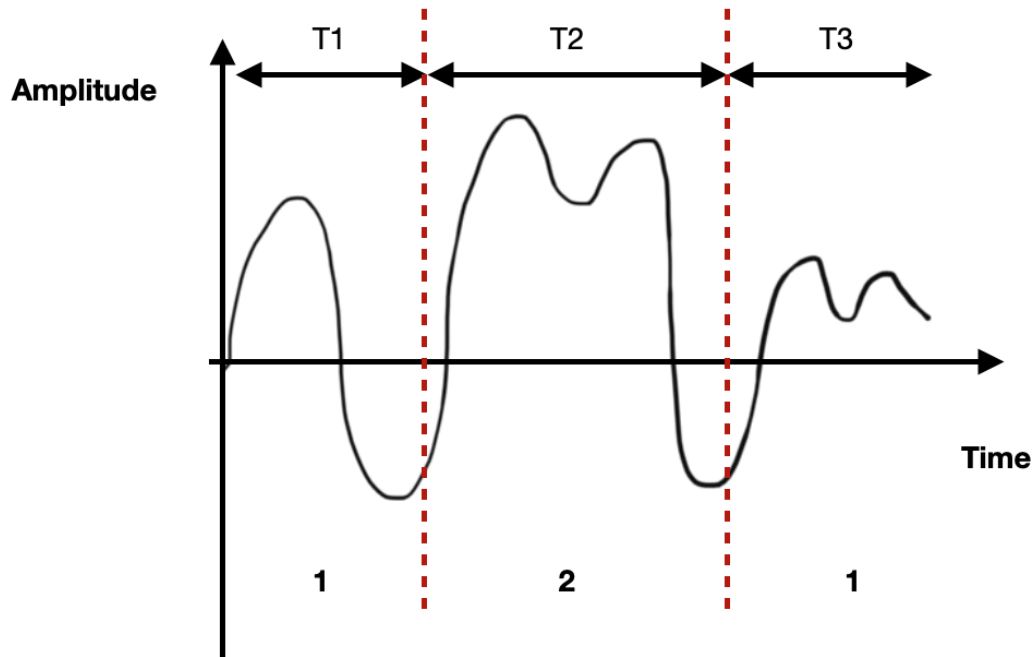


Figure 6.2: Sketch of the experimental procedure controlling system H_{r21} . The system switches between the H_1 and H_2 , indicated in the sketch respectively with the numbers 1 and 2

We would like to test whether MRAC with a constant learning rate can be a good model of human adaptive motor behavior going from controlling H_1 to H_2 and vice-versa.

6.3 Conclusions

In this report, we addressed the problem of human adaptation under changes in the controlled dynamics in the context of manual control with a joystick.

With the literature review, we surveyed the state of the art human motor control and predictive coding, which is a well-supported theory of how the brain process information about the world. Currently, the mainstream opinion about human motor control is that human behave like optimal controllers to complete a task. The optimal control framework solves the problem of redundancy, noise, uncertainty, delays and time-varying dynamics that are properties of the human sensorimotor system. The optimal control framework has substantial experimental evidence and it's motivated by evolutionary ideas. Furthermore, it has been found that humans optimally perform motor tasks making use of an internal model of the motor system. The internal model allows the brain to predict the future, find the optimal policy to accomplish the task and mostly rely on feedback for disturbance rejection rather than planning. There is evidence that the cerebellum's network encodes the integral model of the dynamics of the human body.

The predictive coding framework is instead of a more general framework about intelligence. It broadly states that intelligent systems have internal generative models about the world and make a prediction about its future states. These predictions are then compared to the actual observed sensory input and the resulting errors are minimized. The errors, the mismatch between the expected state of the world and reality, are further propagated up the hierarchical neural network for further processing. They can either drive actions or change the internal model so that the errors are minimized.

Inspired by predictive coding we selected the adaptive control technique "Model Reference Adaptive Control" (MRAC) to model human adaptation, as it leverages McRuer's insight that the open-loop dynamics of a system controlled manually has the structure of a first-order stable system with delay. In its essence, the MRAC scheme creates a controller that converges to a specified reference controller, called the reference model. The reference model used is the McRuer model for open-loop dynamics. The reference model corresponds to the internal model that human development while learning to control a new dynamical system.

To validate the MRAC framework we plan first to test it in simulation using data collected by P. Zaal [31]. We aim at establishing the best learning rate, delay and crossover frequency, which are the free parameter of the MRAC controller, by using an optimization procedure. After the validation in the simulation, we plan to perform a the reference tracking experiment where participants are asked to control a time-varying system. We hope in this way to establish if MRAC can predict the transient behavior of the human operators.

Bibliography

- [1] A. Bastos et al. “Canonical Microcircuits for Predictive Coding”. In: *Neuron* 76 (Nov. 21, 2012), pp. 695–711. DOI: [10.1016/j.neuron.2012.10.038](https://doi.org/10.1016/j.neuron.2012.10.038).
- [2] K. Friston. “Does Predictive Coding Have a Future?” In: *Nature Neuroscience* 21.8 (Aug. 1, 2018), pp. 1019–1021. ISSN: 1546-1726. DOI: [10.1038/s41593-018-0200-7](https://doi.org/10.1038/s41593-018-0200-7). URL: <https://doi.org/10.1038/s41593-018-0200-7>.
- [3] K. Friston. “The Free-Energy Principle: A Unified Brain Theory?” In: *Nature Reviews Neuroscience* 11.2 (Feb. 1, 2010), pp. 127–138. ISSN: 1471-0048. DOI: [10.1038/nrn2787](https://doi.org/10.1038/nrn2787). URL: <https://doi.org/10.1038/nrn2787>.
- [4] M. Heilbron and M. Chait. “Great Expectations: Is There Evidence for Predictive Coding in Auditory Cortex?” In: *Sensory Sequence Processing in the Brain* 389 (Oct. 1, 2018), pp. 54–73. ISSN: 0306-4522. DOI: [10.1016/j.neuroscience.2017.07.061](https://doi.org/10.1016/j.neuroscience.2017.07.061). URL: <http://www.sciencedirect.com/science/article/pii/S030645221730547X>.
- [5] N. Kogo and C. Trengove. “Is Predictive Coding Theory Articulated Enough to Be Testable?” In: *Frontiers in computational neuroscience* 9 (Sept. 8, 2015), pp. 111–111. ISSN: 1662-5188. DOI: [10.3389/fncom.2015.00111](https://doi.org/10.3389/fncom.2015.00111). PMID: [26441621](https://pubmed.ncbi.nlm.nih.gov/26441621). URL: <https://pubmed.ncbi.nlm.nih.gov/26441621>.
- [6] S. Shipp. “Neural Elements for Predictive Coding”. In: *Frontiers in Psychology* 7 (2016). ISSN: 1664-1078. DOI: [10.3389/fpsyg.2016.01792](https://doi.org/10.3389/fpsyg.2016.01792). URL: <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01792/full>.
- [7] R. P. N. Rao and D. H. Ballard. “Predictive Coding in the Visual Cortex: A Functional Interpretation of Some Extra-Classical Receptive-Field Effects”. In: *Nature Neuroscience* 2.1 (1 Jan. 1999), pp. 79–87. ISSN: 1546-1726. DOI: [10.1038/4580](https://doi.org/10.1038/4580). URL: https://www.nature.com/articles/nn0199_79.
- [8] M. I. Garrido et al. “Dynamic Causal Modeling of the Response to Frequency Deviants”. In: *Journal of Neurophysiology* 101.5 (May 1, 2009), pp. 2620–2631. ISSN: 0022-3077. DOI: [10.1152/jn.90291.2008](https://doi.org/10.1152/jn.90291.2008). URL: <https://doi.org/10.1152/jn.90291.2008>.
- [9] A. Todorovic and F. P. de Lange. “Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields”. In: *The Journal of Neuroscience* 32.39 (Sept. 26, 2012), p. 13389. DOI: [10.1523/JNEUROSCI.2227-12.2012](https://doi.org/10.1523/JNEUROSCI.2227-12.2012). URL: <http://www.jneurosci.org/content/32/39/13389.abstract>.
- [10] S. Chennu et al. “Silent Expectations: Dynamic Causal Modeling of Cortical Prediction and Attention to Sounds That Weren’t”. In: *The Journal of Neuroscience* 36.32 (Aug. 10, 2016), p. 8305. DOI: [10.1523/JNEUROSCI.1125-16.2016](https://doi.org/10.1523/JNEUROSCI.1125-16.2016). URL: <http://www.jneurosci.org/content/36/32/8305.abstract>.
- [11] J. F. Mejias et al. “Feedforward and Feedback Frequency-Dependent Interactions in a Large-Scale Laminar Network of the Primate Cortex”. In: *Science Advances* 2.11 (Nov. 1, 2016), e1601335. DOI: [10.1126/sciadv.1601335](https://doi.org/10.1126/sciadv.1601335). URL: <http://advances.sciencemag.org/content/2/11/e1601335.abstract>.
- [12] D. W. Franklin and D. M. Wolpert. “Computational Mechanisms of Sensorimotor Control”. In: *Neuron* 72.3 (Nov. 3, 2011), pp. 425–442. ISSN: 0896-6273. DOI: [10.1016/j.neuron.2011.10.006](https://doi.org/10.1016/j.neuron.2011.10.006). URL: <http://www.sciencedirect.com/science/article/pii/S0896627311008919>.

- [13] D. McNamee and D. Wolpert. “Internal Models in Biological Control”. In: (Oct. 3, 2018). ISSN: 2573-5144. DOI: [10.17863/CAM.30459](https://doi.org/10.17863/CAM.30459). URL: <https://www.repository.cam.ac.uk/handle/1810/283097>.
- [14] R. Miall and D. Wolpert. “Forward Models for Physiological Motor Control”. In: *Four Major Hypotheses in Neuroscience* 9.8 (Nov. 1, 1996), pp. 1265–1279. ISSN: 0893-6080. DOI: [10.1016/S0893-6080\(96\)00035-4](https://doi.org/10.1016/S0893-6080(96)00035-4). URL: <http://www.sciencedirect.com/science/article/pii/S0893608096000354>.
- [15] M. Kawato. “Internal Models for Motor Control and Trajectory Planning”. In: *Current Opinion in Neurobiology* 9.6 (Dec. 1, 1999), pp. 718–727. ISSN: 0959-4388. DOI: [10.1016/S0959-4388\(99\)00028-8](https://doi.org/10.1016/S0959-4388(99)00028-8). URL: <http://www.sciencedirect.com/science/article/pii/S0959438899000288>.
- [16] L. R. Young. “On Adaptive Manual Control”. In: *IEEE Transactions on Man-Machine Systems* 10.4 (Dec. 1969), pp. 292–331. ISSN: 2168-2860. DOI: [10.1109/TMMS.1969.299931](https://doi.org/10.1109/TMMS.1969.299931).
- [17] M. Ito. “Control of Mental Activities by Internal Models in the Cerebellum”. In: *Nature Reviews. Neuroscience* 9.4 (Apr. 2008), pp. 304–313. ISSN: 1471-0048. DOI: [10.1038/nrn2332](https://doi.org/10.1038/nrn2332). pmid: [18319727](https://pubmed.ncbi.nlm.nih.gov/18319727/).
- [18] M. O. Ernst and M. S. Banks. “Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion”. In: *Nature* 415.6870 (Jan. 1, 2002), pp. 429–433. ISSN: 1476-4687. DOI: [10.1038/415429a](https://doi.org/10.1038/415429a). URL: <https://doi.org/10.1038/415429a>.
- [19] K. P. Körding and D. M. Wolpert. “Bayesian Integration in Sensorimotor Learning”. In: *Nature* 427.6971 (6971 Jan. 2004), pp. 244–247. ISSN: 1476-4687. DOI: [10.1038/nature02169](https://doi.org/10.1038/nature02169). URL: <https://www.nature.com/articles/nature02169> (visited on 09/22/2020).
- [20] E. Todorov. “Optimality Principles in Sensorimotor Control”. In: *Nature Neuroscience* 7.9 (Sept. 1, 2004), pp. 907–915. ISSN: 1546-1726. DOI: [10.1038/nn1309](https://doi.org/10.1038/nn1309). URL: <https://doi.org/10.1038/nn1309>.
- [21] E. Todorov and M. I. Jordan. “Optimal Feedback Control as a Theory of Motor Coordination”. In: *Nature Neuroscience* 5.11 (Nov. 1, 2002), pp. 1226–1235. ISSN: 1546-1726. DOI: [10.1038/nn963](https://doi.org/10.1038/nn963). URL: <https://doi.org/10.1038/nn963>.
- [22] K. Friston. “What Is Optimal about Motor Control?” In: *Neuron* 72.3 (Nov. 3, 2011), pp. 488–498. ISSN: 0896-6273. DOI: [10.1016/j.neuron.2011.10.018](https://doi.org/10.1016/j.neuron.2011.10.018). URL: <http://www.sciencedirect.com/science/article/pii/S0896627311009305>.
- [23] J. Diedrichsen, R. Shadmehr, and R. B. Ivry. “The Coordination of Movement: Optimal Feedback Control and Beyond”. In: *Trends in Cognitive Sciences* 14.1 (Jan. 1, 2010), pp. 31–39. ISSN: 1364-6613. DOI: [10.1016/j.tics.2009.11.004](https://doi.org/10.1016/j.tics.2009.11.004). URL: <http://www.sciencedirect.com/science/article/pii/S1364661309002587>.
- [24] S. H. Scott. “The Computational and Neural Basis of Voluntary Motor Control and Planning”. In: *Trends in Cognitive Sciences* 16.11 (Nov. 1, 2012), pp. 541–549. ISSN: 1364-6613. DOI: [10.1016/j.tics.2012.09.008](https://doi.org/10.1016/j.tics.2012.09.008). URL: <https://doi.org/10.1016/j.tics.2012.09.008> (visited on 09/02/2020).
- [25] C. M. Harris and D. M. Wolpert. “The Main Sequence of Saccades Optimizes Speed-Accuracy Trade-Off”. In: *Biological Cybernetics* 95.1 (July 1, 2006), pp. 21–29. ISSN: 1432-0770. DOI: [10.1007/s00422-006-0064-x](https://doi.org/10.1007/s00422-006-0064-x). URL: <https://doi.org/10.1007/s00422-006-0064-x>.
- [26] D. A. Braun et al. “Learning Optimal Adaptation Strategies in Unpredictable Motor Tasks”. In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29.20 (May 20, 2009), pp. 6472–6478. ISSN: 1529-2401. DOI: [10.1523/JNEUROSCI.3075-08.2009](https://doi.org/10.1523/JNEUROSCI.3075-08.2009). PMID: [19458218](https://pubmed.ncbi.nlm.nih.gov/19458218/). URL: <https://pubmed.ncbi.nlm.nih.gov/19458218>.

- [27] J. Merel, M. Botvinick, and G. Wayne. “Hierarchical Motor Control in Mammals and Machines”. In: *Nature Communications* 10.1 (Dec. 2019), p. 5489. ISSN: 2041-1723. DOI: [10.1038/s41467-019-13239-6](https://doi.org/10.1038/s41467-019-13239-6). URL: <http://www.nature.com/articles/s41467-019-13239-6> (visited on 09/02/2020).
- [28] D. Liu and E. Todorov. “Evidence for the Flexible Sensorimotor Strategies Predicted by Optimal Feedback Control”. In: *The Journal of Neuroscience* 27.35 (Aug. 29, 2007), p. 9354. DOI: [10.1523/JNEUROSCI.1110-06.2007](https://doi.org/10.1523/JNEUROSCI.1110-06.2007). URL: <http://www.jneurosci.org/content/27/35/9354.abstract>.
- [29] D. McRuer and H. Jex. “A Review of Quasi-Linear Pilot Models”. In: *IEEE Transactions on Human Factors in Electronics* HFE-8.3 (Sept. 1967), pp. 231–249. ISSN: 2168-2852. DOI: [10.1109/THFE.1967.234304](https://doi.org/10.1109/THFE.1967.234304).
- [30] D. T. McRuer, R. E. Magdaleno, and G. P. Moore. “A Neuromuscular Actuation System Model”. In: *IEEE Transactions on Man-Machine Systems* 9.3 (Sept. 1968), pp. 61–71. ISSN: 2168-2860. DOI: [10.1109/TMMS.1968.300039](https://doi.org/10.1109/TMMS.1968.300039).
- [31] P. M. Zaal. “Manual Control Adaptation to Changing Vehicle Dynamics in Roll–Pitch Control Tasks”. In: *Journal of Guidance, Control, and Dynamics* 39.5 (2016), pp. 1046–1058. DOI: [10.2514/1.G001592](https://doi.org/10.2514/1.G001592). URL: <https://doi.org/10.2514/1.G001592>.
- [32] P. M. T. Zaal et al. “Modeling Human Multimodal Perception and Control Using Genetic Maximum Likelihood Estimation”. In: *Journal of Guidance, Control, and Dynamics* 32.4 (July 1, 2009), pp. 1089–1099. DOI: [10.2514/1.42843](https://doi.org/10.2514/1.42843). URL: <https://arc.aiaa.org/doi/10.2514/1.42843> (visited on 09/22/2020).
- [33] E. Boer and R. Kenyon. “Estimation of Time-Varying Delay Time in Nonstationary Linear Systems: An Approach to Monitor Human Operator Adaptation in Manual Tracking Tasks”. In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 28.1 (Jan. 1998), pp. 89–99. ISSN: 1558-2426. DOI: [10.1109/3468.650325](https://doi.org/10.1109/3468.650325).
- [34] A. Popovici, P. Zaal, and D. M. Pool. “Dual Extended Kalman Filter for the Identification of Time-Varying Human Manual Control Behavior”. In: *AIAA Modeling and Simulation Technologies Conference*. American Institute of Aeronautics and Astronautics, June 2, 2017. DOI: [10.2514/6.2017-3666](https://doi.org/10.2514/6.2017-3666). URL: <https://arc.aiaa.org/doi/10.2514/6.2017-3666>.
- [35] N. Roggenkämper et al. “Objective ARX Model Order Selection for Multi-Channel Human Operator Identification”. In: *AIAA Modeling and Simulation Technologies Conference*. American Institute of Aeronautics and Astronautics. DOI: [10.2514/6.2016-4299](https://doi.org/10.2514/6.2016-4299). URL: <https://arc.aiaa.org/doi/abs/10.2514/6.2016-4299> (visited on 09/22/2020).
- [36] A. van Grootheest et al. “Identification of Time-Varying Manual Control Adaptations with Recursive ARX Models”. In: *2018 AIAA Modeling and Simulation Technologies Conference*. AIAA SciTech Forum. American Institute of Aeronautics and Astronautics, Jan. 7, 2018. DOI: [10.2514/6.2018-0118](https://doi.org/10.2514/6.2018-0118). URL: <https://arc.aiaa.org/doi/10.2514/6.2018-0118>.
- [37] S. Xu et al. “Review of Control Models for Human Pilot Behavior”. In: *Annual Reviews in Control* 44 (Jan. 1, 2017), pp. 274–291. ISSN: 1367-5788. DOI: [10.1016/j.arcontrol.2017.09.009](https://doi.org/10.1016/j.arcontrol.2017.09.009). URL: <http://www.sciencedirect.com/science/article/pii/S136757881730024X>.
- [38] A. V. Phatak and G. A. Bekey. “Model of the Adaptive Behavior of the Human Operator in Response to a Sudden Change in the Control Situation”. In: *IEEE Transactions on Man-Machine Systems* 10.3 (Sept. 1969), pp. 72–80. ISSN: 2168-2860. DOI: [10.1109/TMMS.1969.299886](https://doi.org/10.1109/TMMS.1969.299886).

- [39] K. Zaychik, F. Cardullo, and G. George. “A Conspectus on Operator Modeling: Past, Present and Future”. In: *AIAA Modeling and Simulation Technologies Conference and Exhibit*. Guidance, Navigation, and Control and Co-Located Conferences. American Institute of Aeronautics and Astronautics, Aug. 21, 2006. DOI: [10.2514/6.2006-6625](https://doi.org/10.2514/6.2006-6625). URL: <https://arc.aiaa.org/doi/10.2514/6.2006-6625>.
- [40] R. A. Hess. “Modeling Human Pilot Adaptation to Flight Control Anomalies and Changing Task Demands”. In: *Journal of Guidance, Control, and Dynamics* 39.3 (2016), pp. 655–666. DOI: [10.2514/1.G001303](https://doi.org/10.2514/1.G001303). URL: <https://doi.org/10.2514/1.G001303>.
- [41] L. Stark and L. R. Young. *Biological Control System - a Critical Review and Evaluation, Developments in Manual Control*. In collab. with NASA. Mar. 1, 1965. URL: http://archive.org/details/nasa_techdoc_19650009660.
- [42] C.-S. Poon and D. M. Merfeld. “Internal Models: The State of the Art”. In: *Journal of Neural Engineering* 2.3 (Aug. 2005). ISSN: 1741-2552. DOI: [10.1088/1741-2552/2/3/E01](https://doi.org/10.1088/1741-2552/2/3/E01). URL: <https://doi.org/10.1088/1741-2552/2/3/E01>.
- [43] N. T. Nguyen. *Model-Reference Adaptive Control: A Primer*. Advanced Textbooks in Control and Signal Processing. Springer International Publishing, 2018. ISBN: 978-3-319-56392-3. DOI: [10.1007/978-3-319-56392-3](https://www.springer.com/gp/book/9783319563923). URL: <https://www.springer.com/gp/book/9783319563923>.
- [44] Yu Wang. “A New Concept Using LSTM Neural Networks for Dynamic System Identification”. In: *2017 American Control Conference (ACC)*. May 2017, pp. 5324–5329. DOI: [10.23919/ACC.2017.7963782](https://doi.org/10.23919/ACC.2017.7963782).
- [45] H. M. Eraqi, M. N. Moustafa, and J. Honer. “End-to-End Deep Learning for Steering Autonomous Vehicles Considering Temporal Dependencies”. In: (Oct. 31, 2017). URL: https://openreview.net/forum?id=rys_hvLRW (visited on 09/22/2020).
- [46] R. T. Q. Chen et al. *Neural Ordinary Differential Equations*. Dec. 13, 2019. arXiv: [1806.07366](https://arxiv.org/abs/1806.07366) [cs, stat]. URL: <http://arxiv.org/abs/1806.07366> (visited on 09/22/2020).
- [47] A. M. Zador. “A Critique of Pure Learning and What Artificial Neural Networks Can Learn from Animal Brains”. In: *Nature Communications* 10.1 (1 Aug. 21, 2019), p. 3770. ISSN: 2041-1723. DOI: [10.1038/s41467-019-11786-6](https://doi.org/10.1038/s41467-019-11786-6). URL: <https://www.nature.com/articles/s41467-019-11786-6>.
- [48] J. Burger. “Antipredator Behaviour of Hatchling Snakes: Effects of Incubation Temperature and Simulated Predators”. In: *Animal Behaviour* 56.3 (Sept. 1, 1998), pp. 547–553. ISSN: 0003-3472. DOI: [10.1006/anbe.1998.0809](https://doi.org/10.1006/anbe.1998.0809). URL: <http://www.sciencedirect.com/science/article/pii/S0003347298908090>.
- [49] A. Gaier and D. Ha. *Weight Agnostic Neural Networks*. Sept. 5, 2019. arXiv: [1906.04358](https://arxiv.org/abs/1906.04358) [cs, stat]. URL: <http://arxiv.org/abs/1906.04358> (visited on 09/22/2020).
- [50] K. J. Friston et al. “Action and Behavior: A Free-Energy Formulation”. In: *Biological Cybernetics* 102.3 (Mar. 2010), pp. 227–260. ISSN: 1432-0770. DOI: [10.1007/s00422-010-0364-z](https://doi.org/10.1007/s00422-010-0364-z). PMID: [20148260](https://pubmed.ncbi.nlm.nih.gov/20148260/).
- [51] M. Mulder et al. “Manual Control with Pursuit Displays: New Insights, New Models, New Issues”. In: *IFAC-PapersOnLine*. 14th IFAC Symposium on Analysis, Design, and Evaluation of Human Machine Systems HMS 2019 52.19 (Jan. 1, 2019), pp. 139–144. ISSN: 2405-8963. DOI: [10.1016/j.ifacol.2019.12.125](https://doi.org/10.1016/j.ifacol.2019.12.125). URL: <http://www.sciencedirect.com/science/article/pii/S2405896319319603>.
- [52] E. Fridman and S.-I. Niculescu. “On Complete Lyapunov–Krasovskii Functional Techniques for Uncertain Systems with Fast-Varying Delays”. In: *International Journal of Robust and Nonlinear Control* 18.3 (Feb. 2008), pp. 364–374. ISSN: 10498923, 10991239. DOI: [10.1002/rnc.1230](https://doi.org/10.1002/rnc.1230). URL: <http://doi.wiley.com/10.1002/rnc.1230> (visited on 01/12/2021).

- [53] E. Fridman. “Tutorial on Lyapunov-Based Methods for Time-Delay Systems”. In: *European Journal of Control* 20.6 (Nov. 1, 2014), pp. 271–283. ISSN: 0947-3580. DOI: [10 . 1016 / j . ejcon . 2014 . 10 . 001](https://doi.org/10.1016/j.ejcon.2014.10.001). URL: <http://www.sciencedirect.com/science/article/pii/S0947358014000764>.
- [54] N. Nguyen and E. Summers. “On Time Delay Margin Estimation for Adaptive Control and Robust Modification Adaptive Laws”. In: *AIAA Guidance, Navigation, and Control Conference*. 0 vols. Guidance, Navigation, and Control and Co-Located Conferences. American Institute of Aeronautics and Astronautics, Aug. 8, 2011. DOI: [10.2514/6.2011-6438](https://doi.org/10.2514/6.2011-6438). URL: <https://arc.aiaa.org/doi/10.2514/6.2011-6438>.
- [55] D. M. Wolpert and M. Kawato. “Multiple Paired Forward and Inverse Models for Motor Control”. In: *Neural Networks: The Official Journal of the International Neural Network Society* 11.7-8 (Oct. 1998), pp. 1317–1329. ISSN: 0893-6080. DOI: [10.1016/s0893-6080\(98\) 00066-5](https://doi.org/10.1016/s0893-6080(98)00066-5). pmid: [12662752](https://pubmed.ncbi.nlm.nih.gov/12662752/).

Time Traces of System Outputs

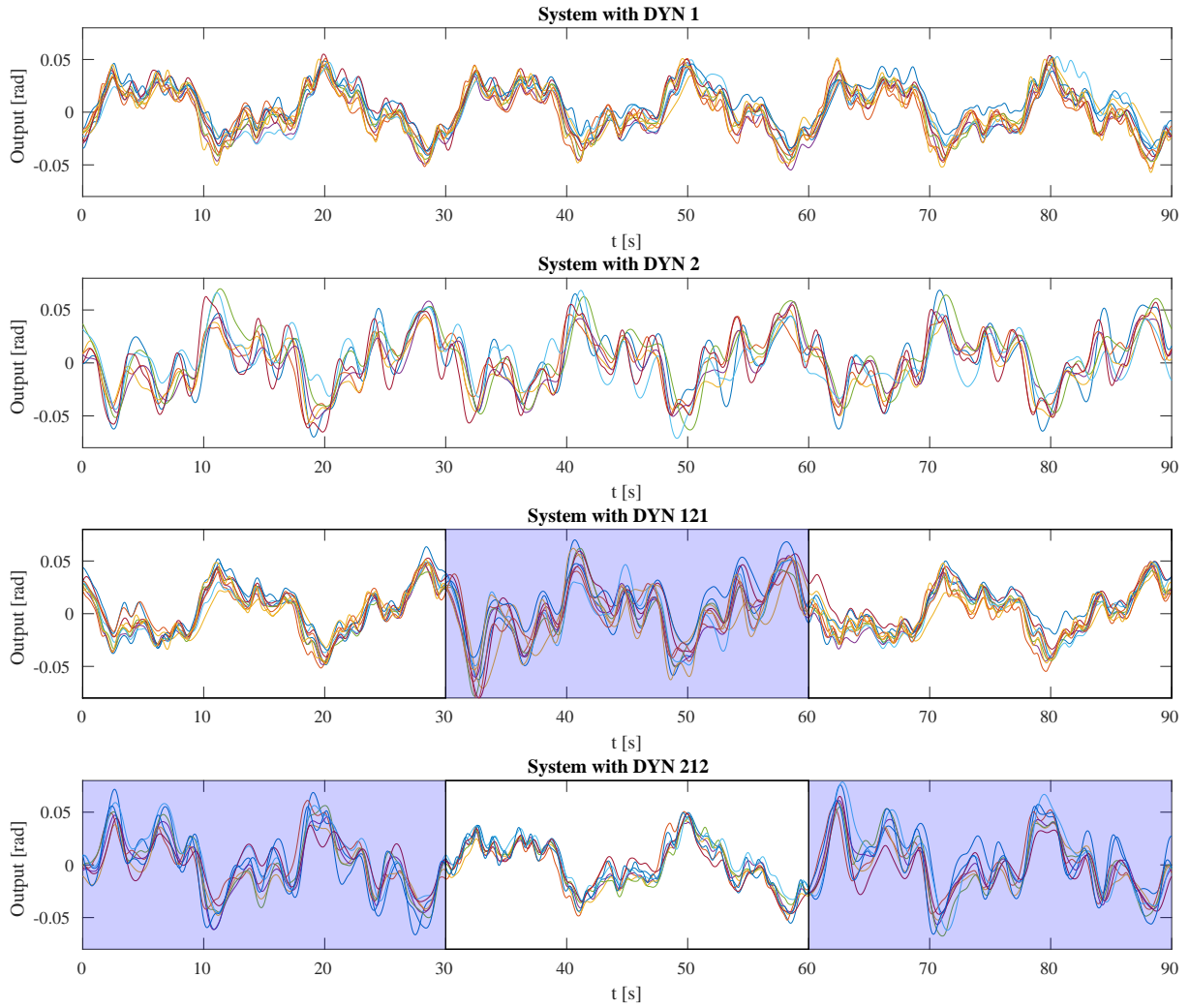


Figure A.1: Time traces of the output of the system controlled by the participants. Each line represent the system's output for a single participants. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.

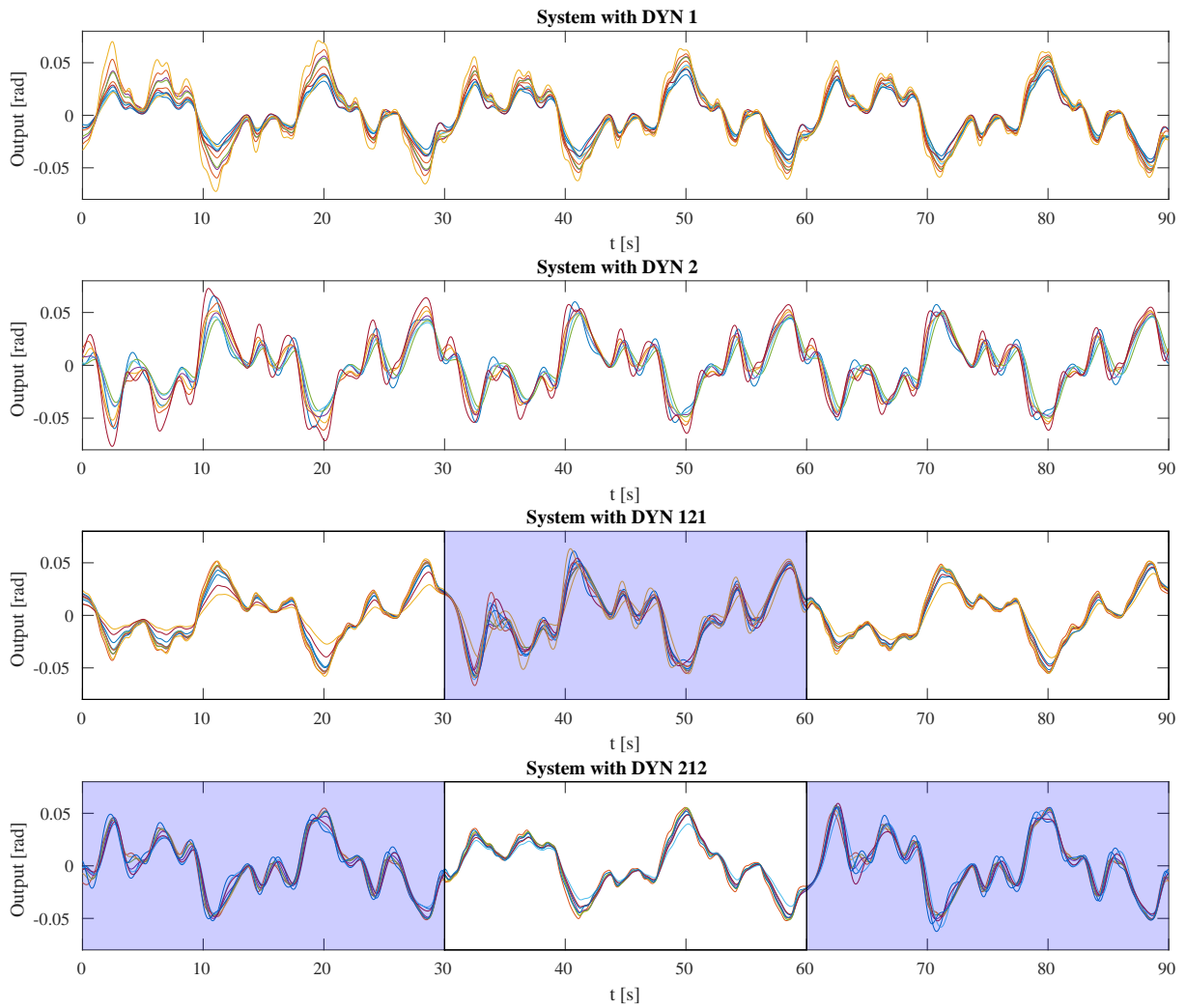


Figure A.2: Time traces of the output of the system controlled by the MRAC controllers fitted on the participants data. Each line represent the system's output for a single MRAC controller. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.

Time Traces of Participants and MRAC Control Outputs

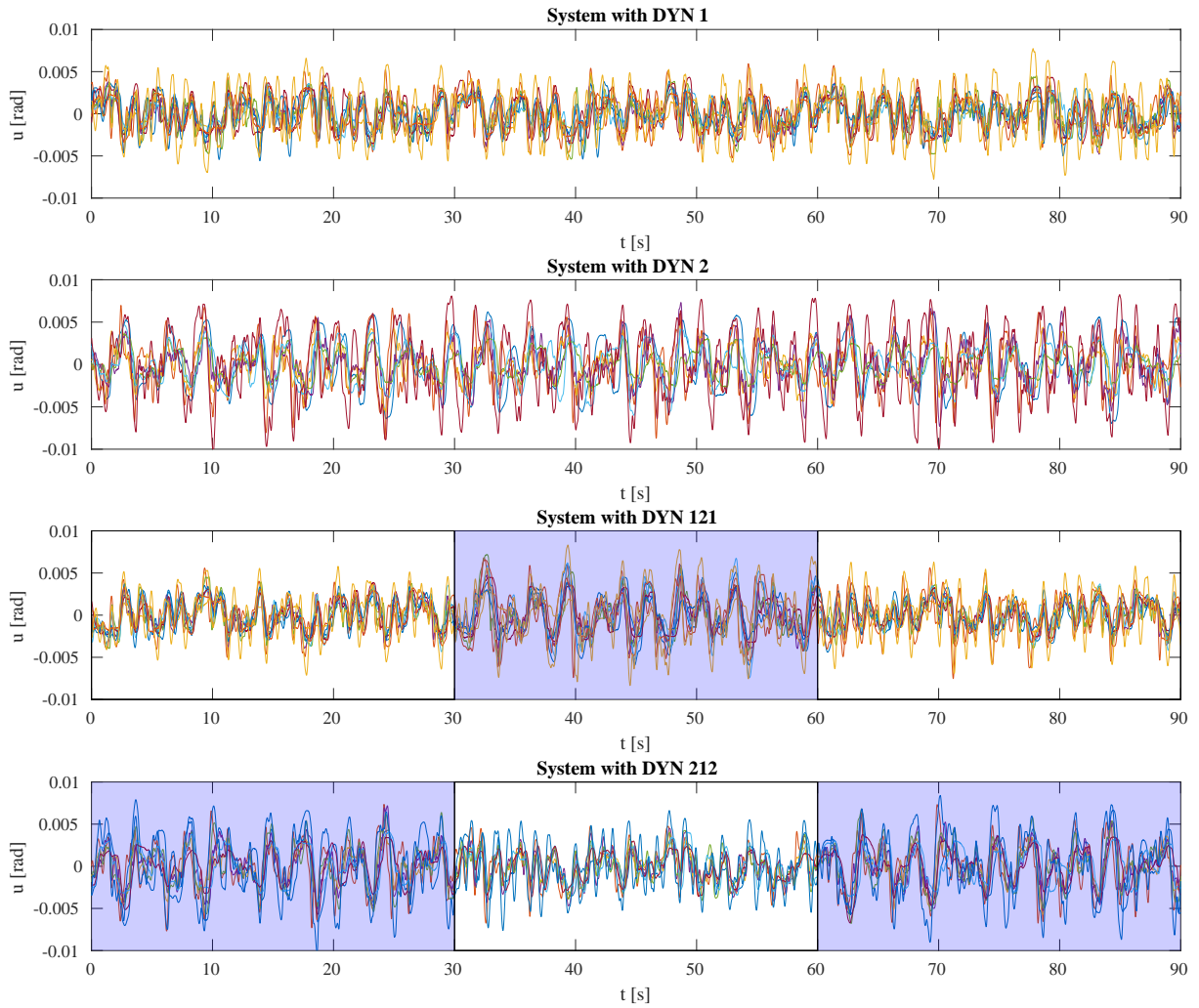


Figure B.1: Time traces of the control output of the participants. Each line represent the system's output for a single participants. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.

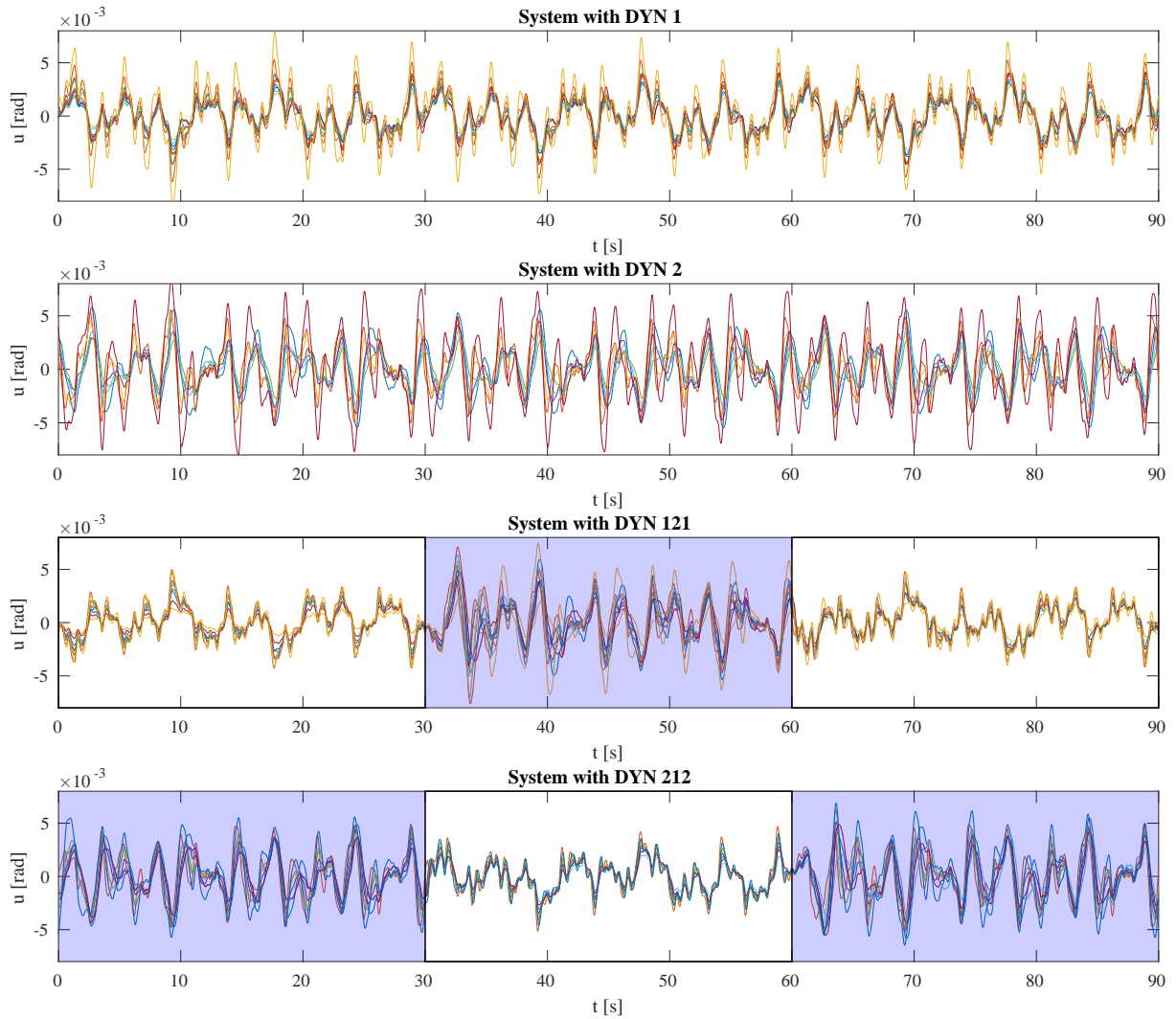


Figure B.2: Time traces of the output of the MRAC controllers fitted on the participants data. Each line represent the system's output for a single MRAC controller. For conditions $DYN = 212$ and $DYN = 121$, the controlled system in the areas colored in blue behaved approximately a double integrator while in the area left uncolored as a single integrator.

Time Traces of Estimated MRAC Gains

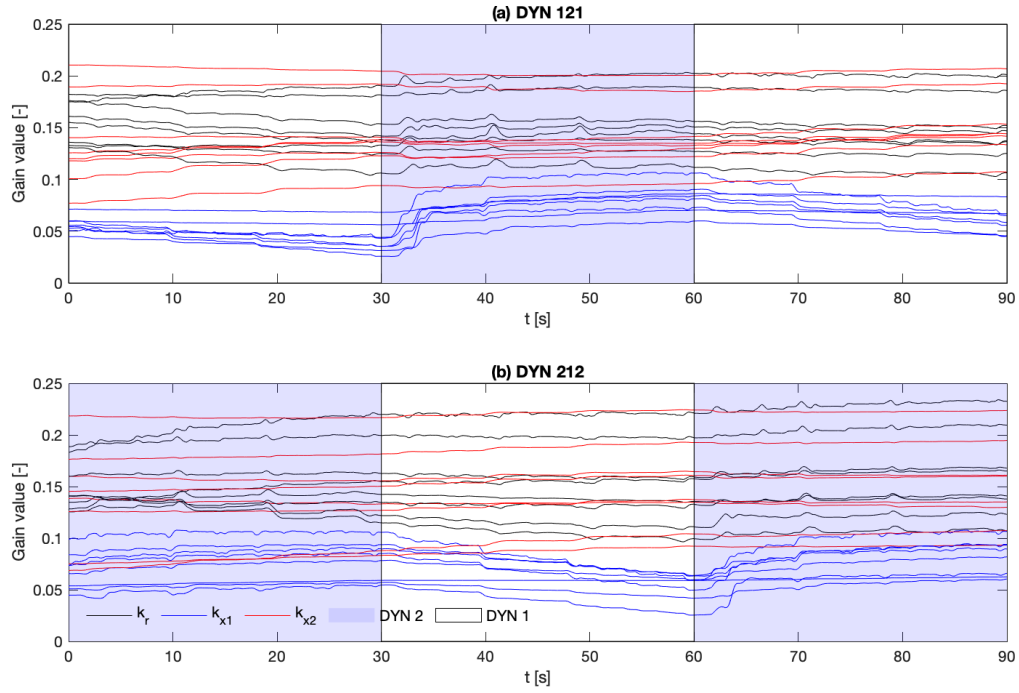


Figure C.1: Time series of the MRAC gains found for each participant.

Human Research Ethics Committee Checklist

Delft University of Technology
ETHICS REVIEW CHECKLIST FOR HUMAN RESEARCH
 (Version 18.06.2020)

This checklist should be completed for every research study that involves human participants and should be submitted before potential participants are approached to take part in your research study. This also applies for students doing their Master-thesis.

In this checklist we will ask for additional information if need be. Please attach this as an Annex to the application.

The data steward of your faculty can help you with any issues related to the protection of personal data. Please note that research related to medical questions/health may require special attention. See also the website of the [CCMO](#).

Please upload the documents (go to [this page](#) for instructions).

Thank you and please check our [website](#) for guidelines, forms, best practices, meeting dates of the HREC, etc.

I. Basic Data

Project title:	A predictive coding approach to adaptive manual control
Name(s) of researcher(s):	Lorenzo Terenzi
Research period (planning)	January 2021
E-mail contact person	lterenzi@student.tudelft.nl
Faculty/Dept.	Aerospace Engineering / C&S
Position researcher(s):	MSc Student
Name of supervisor (if applicable):	Dr. ir. D.M. (Daan) Pool
Role of supervisor (if applicable):	Assistant Professor

I.

II. A) Summary Research

This research investigates the abilities of human operators to adapt to changing controlled element dynamics. The research objective is to come up with a model that can predict the human adaptation process. Earlier research was used to find already existing models and extend them into a more general framework. Our model was firstly tested by performing computer simulations. The results of a simple tracking task experiment will be used to verify and validate the model and the computer simulation findings.

B) Risk assessment & risk management

Two potential risk factors are fully dealt with, as described below. Additionally, the COVID-19 pandemic requires a protocol which is described in the remainder of this section.

1. All stored data is anonymized and stored under only a numeric subject ID. In addition, only objective human control data is collected, no sensitive personal data is analysed.
1. Simulator sickness is highly unlikely during this experiment, due to the use of a simple abstract display that is located directly in front of the participant. However, any discomfort experienced by the participants may be reported, after which the experiment will be aborted. Participants will be briefed on this prior to the experiment (see attached Experiment Briefing).

COVID-19 protocol

Given the health risks related to the outbreak of COVID-19 ('coronavirus'), measures will be taken to reduce the risk of spreading this virus for experiment participants, the researcher and other people present at and around the facility. A detailed overview of these measures can be seen in the attached document "COVID-19 Protocols for Human Subject Experiments of the TU Delft Control & Simulation department". They are in line with the latest advice given by the Dutch health authorities (see <https://www.rijksoverheid.nl/coronavirus>). A combination of hygiene actions and social distancing measures is applied to reduce the risk of spreading the virus.

a. General measures

In the case that participants show symptoms of the virus, they will not be allowed to take part in the experiment. As part of the strict entrance policy at the aerospace faculty, all participants are obliged to swipe their campus card or inscribe in a dedicated form when entering the building. Therefore, an overview is available of every person entering the building. This helps contact tracing in the exceptional (and hopefully absent) case of a positive COVID-19 test. Furthermore both the researcher and the participant will wear a face mask while in the public spaces of the Faculty building.

b. Hygiene measures

All parts of the simulator that the experiment participants touches will be cleaned with Disinfectant before the experiment and in between any possible break. Moreover, both participant and researcher will wash or disinfect their hands before and after the experiment. Disinfectant will be made available.

a. Social distancing

In line with the advice by the Dutch health authorities, 1.5 meters distance will be kept between researcher and participant at all times. During the experiment, the participant will sit in the simulator, while the researcher is at his control desk. These are two different spaces, separated by a glass wall.

The following statements will be added to the informed consent form:

- I confirm that the researcher has provided me with detailed safety instructions to ensure my experiment session can be performed in line with current RIVM COVID-19 regulations at all times and that these instructions are fully clear to me.
- I understand that also for my travel to/from the experiment session I should adhere to all current RIVM COVID-19 regulations. I confirm that I have travelled to TU Delft's Faculty of Aerospace Engineering with either my own car, by bicycle, or on foot.

III. Checklist

Question	Yes	No
1. Does the study involve participants who are particularly vulnerable or unable to give informed consent? (e.g., children, people with learning difficulties, patients, people receiving counselling, people living in care or nursing homes, people recruited through self-help groups).		X
2. Are the participants, outside the context of the research, in a dependent or subordinate position to the investigator (such as own children or own students)?		X
3. Will it be necessary for participants to take part in the study without their knowledge and consent at the time? (e.g., covert observation of people in non-public places).		X
4. Will the study involve actively deceiving the participants? (For example, will participants be deliberately falsely informed, will information be withheld from them or will they be misled in such a way that they are likely to object or show unease when debriefed about the study).		X
5. Sensitive personal data <ul style="list-style-type: none"> Will the study involve discussion or collection of personal sensitive data (e.g., financial data, location data, data relating to children or other vulnerable groups)? Definitions of sensitive personal data, and special cases thereof are provided here. 		X
6. Will drugs, placebos, or other substances (e.g., drinks, foods, food or drink constituents, dietary supplements) be administered to the study participants?		X
7. Will blood or tissue samples be obtained from participants?		X
8. Is pain or more than mild discomfort likely to result from the study?		X
9. Does the study risk causing psychological stress or anxiety or other harm or negative consequences beyond that normally encountered by the participants in their life outside research?		X
10. Will financial inducement (other than reasonable expenses and compensation for time) be offered to participants?		X
Important: if you answered 'yes' to any of the questions mentioned above, please submit a full application to HREC (see: website for forms or examples).		
11. Will the experiment collect and store videos, pictures, or other identifiable data of human subjects?		X
12. Will the experiment involve the use of devices that are not 'CE' certified? Only, if 'yes': continue with the following questions:	X	

➤ Was the device built in-house?	X	
➤ Was it inspected by a safety expert at TU Delft? (Please provide device report, see: HREC website)	X	
➤ If it was not built in house and not CE-certified, was it inspected by some other, qualified authority in safety and approved? (Please provide records of the inspection).		
13. Has or will this research be submitted to a research ethics committee other than this one? (if so, please provide details and a copy of the approval or submission).		X

IV. Enclosures

Please, tick the checkboxes for submitted enclosures.

Required enclosures

X A data management plan reviewed by a data-steward.

Conditionally required enclosures

If you replied 'yes' to any of the questions 1 until 10:

- A full research application

If you replied 'yes' to questions 11:

- An Informed consent form

If you replied 'yes' to questions 12:

X A device report

If you replied 'yes' to questions 13:

- Submission details to the external HREC, and a copy of their approval if available.

Additional enclosures

X COVID-19 experiment protocols.

V. Signature(s)


Signature(s) of researcher(s)

Date: 23-10-2020



Signature (or upload consent by mail) research supervisor (if applicable)

Date: 20-12-2020

 Daan Pool
2020.12.20
21:15:06
+01'00'

Experiment Briefing

Experiment Briefing

ADAPTIVE MANUAL CONTROL

The experiment, conducted in the Human-Machine Interaction Laboratory (HMI-Lab), analyses human tracking behaviour. The experiment consists of a simple tracking task. This briefing will introduce you to the experiment and what is expected of you as a participant.

Experiment Goal

The goal of this experiment is to investigate human adaptive behaviour. A simple tracking task, containing variations in controlled element dynamics, will be performed to gather data on this adaptive behaviour.

Experiment Task

The task you will be carrying out is a tracking task with a pursuit display (i.e. you can only see the target signal and the output of the system you are controlling). It is your task to keep the error as low as possible by moving the side-stick on your right hand-side to left or right.

Experiment Procedures

Each tracking run lasts about 100 seconds. During each run the controlled element dynamics of the system might change. It is important that you continue to focus on keeping the error as low as possible by continuously controlling the system. The researcher will keep track of your performance and will announce when the experiment has been completed. You will start the experiment with a training phase, where you will be familiarised with the different scenarios and controlled element dynamics.

Short breaks can be taken between runs to alleviate any discomfort that might occur due to controlling the side-stick or after sitting in a fixed position for a prolonged period of time. A longer break will be taken after the first hour, where you will be taken out of the simulator for 5-10 minutes. The experiment will last approximately 2 hours.

For each driving trial, the subsequent procedure will be followed:

1. The researcher applies the settings for the next run.
2. The researcher checks whether the participant is ready to proceed and initiates the run after a countdown from 3 (3-2-1-go).
3. The participant performs the tracking task.

4. The participant will be notified of their performance in the run in terms of error score after the completed run.

COVID-19 protocol

Due to the ongoing COVID-19 ('coronavirus') pandemic, several measures are taken to reduce the risk of spreading it. First and foremost, researcher and participant will follow the guidelines as indicated on the Dutch government website¹ on the day of the experiment. Related to this experiment, the following four measures are taken:

- Both researcher and participants confirm they do not have symptoms related to COVID-19.
- 1.5 meter distance will be kept between researcher and participant at all times.
- All touched objects in the simulator will be disinfected by the researcher before and after the experiment.
- Before and after the experiment both researcher and participant will wash or disinfect their hands.

This experiment will be performed following the most recent "COVID-19 Protocols for Human Subject Experiments" of the Control and Simulation department.

Your rights

Participation in the experiment is voluntary. This means that you can terminate your cooperation at any time. By participating in the experiment you agree that the collected data may be published. Your data will remain confidential and anonymous, so only the experimenter can link the results to a particular participant. To make sure that you understand and comply with the conditions of the experiment, you will be asked to sign an informed consent form.

Contact information researcher:

Lorenzo Terenzi
lterenzi@student.tudelft.nl
+31 640660024

Contact information research supervisor

Dr. ir. Daan Pool
d.m.pool@tudelft.nl
+31 15 2789611

Thank you for participating!

¹ <https://www.rijksoverheid.nl/coronavirus>

Experiment Consent Form

Experiment Consent Form

Adaptive Manual Control Experiment

I hereby confirm, by ticking each box, that:

1. I volunteer to participate in the experiment conducted by the student researcher (**Lorenzo Terenzi**) under supervision of **Dr.ir. Daan Pool** from the Faculty of Aerospace Engineering of TU Delft. I understand that my participation in this experiment is voluntary and that I may withdraw and discontinue participation at any time, for any reason. ☐
2. I have read the experiment briefing and confirm that I understand the instructions and have had all remaining questions answered to my satisfaction. ☐
3. I understand that my participation involves performing a tracking task in a fixed-based simulator. ☐
4. I confirm that the researcher has provided me with detailed safety and operational instructions for the hardware (simulator setup, control-loaded stick, fire escape) used in the experiment. ☐
5. I understand that (though very unlikely) it is possible that I may develop some feelings of discomfort caused by focussing on the display. If this is the case, I will inform the experimenter. I also understand that the experiment may be discontinued for this reason. ☐
6. I understand that the researcher will not identify me by name in any reports or publications that will result from this experiment, and that my confidentiality as a participant in this study will remain secure. ☐
7. I give permission to publish the anonymised data gathered in this study in a Open Access repository ☐

My Signature

Date

My Printed Name

Signature of researcher

Contact information researcher:

Lorenzo Terenzi
lterenzi@student.tudelft.nl
+31 640660024

Contact information research supervisor

Dr. ir. Daan M. Pool
d.m.pool@tudelft.nl
+31 15 2789611

COVID19 Protocol

It's attached the part of the faculty COVID19 protocol that is relevant for this experiment.

Facility Report

COVID-19 Protocols for
Human Subject Experiments

by

Max Mulder,
Rene van Paassen,
Daan Pool,
Clark Borst,
Olaf Stroosma,
Olaf Grevenstuk,
Andries Muis,
Ferdinand Postema,
Harold Thung

1

General protocols

In this document we formalize “COVID-19” protocols for human subject experiments in three facilities available at TU Delft’s Faculty of Aerospace Engineering – the SIMONA Research Simulator, the Human-Machine Interaction Laboratory, and the Air Traffic Management Laboratory – to ensure experiments can still be performed safely. Our main goal is to ensure that the current RIVM regulations (at the time of writing of this document those dated October 20, 2020) as well as the advice from the TU Delft Human Research Ethics Committee will be followed *at all times*. This includes maintaining 1.5 meters distance, no shaking hands, ensuring proper ventilation, etc. Additionally, all experiments that would require close contact with participants (e.g., placing heart-rate sensors) are forbidden. The general protocols described in this chapter will be adhered to for human subject experiments in all three experiment facilities. **In addition, we intend to update this document, and our procedures, in line with any future changes of the RIVM regulations and our experiences running experiments under this protocol.**

1.1. Recruiting participants

Recruitment of participants for experiments shall be limited to TU Delft staff members, TU Delft students and/or external professionals (e.g., pilots) only. Upon recruitment we will explain all safety precautions taken before, during, and after the experiment (including travel requirements) to participants. In addition, for all experiments we will add two additional statements to our Informed Consent forms to also record participants’ awareness of, and agreement to meet, these COVID-19 safety requirements:

- I confirm that the researcher has provided me with detailed safety instructions to ensure my experiment session can be performed in line with current RIVM COVID-19 regulations *at all times* and that these instructions are fully clear to me.
- I understand that also for my travel to/from the experiment session I should *at all times* adhere to current RIVM COVID-19 regulations. I confirm that I have travelled to TU Delft’s Faculty of Aerospace Engineering with either my own car, by bicycle, or on foot.

Participants with symptoms cannot take part and should stay home. This will be checked and asked for during planning the experiment sessions, but also before entering the faculty (similar to rules for “contact-beroepen”). Facility-specific procedures are disclosed in the next chapters of this document.

Finally, we limit experiment participation to a maximum of two participants per facility per day, to avoid having large numbers of additional people in the Faculty building. If two participants perform an experiment in the same facility, the experimenter is responsible for planning in sufficient time between sessions to guarantee that:

1. the two participants will not be in the Faculty building at the same time nor be at the Faculty entrance at the same time
2. sufficient time is available between the two experiment sessions to ensure all surfaces can be properly cleaned and the room can be ventilated

1.2. Traveling

Due to the current RIVM guidelines, participants who will have to use the public transportation system for travelling to the faculty are excluded by default. Participants should only travel to their experiment session by either car, bicycle, or on foot. Traveling to and from the faculty shall be avoid the rush hours as much as possible.

1.3. Entering and leaving the faculty

Given that there are three experiment facilities (i.e., SIMONA, HMILab and ATMLab) that are at different locations in the faculty (at more than 15 meters apart!), a maximum of six experiment sessions can be performed per day. Scheduling the start and end time of an experiment will include at least a 15-minute interval between different experiments, such that only one participant will arrive, and enter the faculty, per time slot. This will avoid gatherings of researchers and participants at the entrance/exit of the faculty.

For participants to be able to enter the faculty for your experiment, it is required to register them beforehand. This has to happen in two steps:

1. **Registration of the experiment at C&S.** To ensure we can coordinate who is experimenting when across the different simulators, please send an email to Daan Pool (d.m.pool@tudelft.nl). This email should include:
 - Name of researcher (typically the student)
 - Name of the responsible experiment supervisor (**an employee who is present at the Faculty on the day(s) of your experiment sessions**)
 - Simulator facility you will be using
 - The dates and times at which you will test participants
2. **Registering participants at the Servicepoint (main entrance).** **NOTE: This step should only be done after getting approval from the experiment from C&S (point 1 above)!** Participants can only be allowed to enter the faculty if they have been registered prior to their arrival at the Servicepoint. For this the researcher (i.e., the person who will run the experiment) has to send the following information to the Servicepoint (Servicepunt-LR@tudelft.nl), **with a "cc" to Joyce ten Berge (K.J.tenBerge@tudelft.nl):**
 - Name of researcher (typically the student) and contact phone number (cell phone)
 - Name of the responsible experiment supervisor (an employee)
 - A list of participant names and the dates and times at which they will be at the faculty

If your list is accepted, the Servicepoint will then receive your participants at the main entrance, sign them in on the faculty attendance list, and call the researcher. Registering new participants with the Servicepoint can be done until the morning of the day of the experiment. **Please make sure that you instruct your participants to contact you themselves upon their arrival at the Faculty, so that you pick them up as quickly as possible.**

After entering the faculty building, the researcher will pick up participants at the main entrance and lead them (while keeping 1.5 meters distance) to the experiment facility, following all guiding indications and signage for moving through the buildings safely. **Under the current RIVM regulations, that means that both the researcher and the participant will wear a face mask while in the public spaces of the Faculty building.** So, overall we will follow exactly the same procedures as used for graduating MSc. students entering the building. All experiment sessions will be scheduled during the regular opening hours of the faculty during normal week days (thus, no sessions in the evening and/or during the weekend).

1.4. Before, during and after an experiment

Before an experiment, the researcher and participant will both wash their hands according to the RIVM guidelines.

During the experiment, the distance norm of 1.5 meters will be respected at all times, and direct physical contact with the participant will be avoided (e.g., no shaking of hands). Sharing objects, such as keyboards, mouse devices, coffee cups, pens and papers will be avoided. When a paper briefing is required, a new set of

papers will be printed for each individual participant.

After each break during an experiment, the researcher and participant will both wash their hands before resuming the experiment session. Additionally, the researcher will sanitise the hardware equipment (using disinfecting spray and dispensable paper towels) that will be in direct contact with the participant.

After each experiment, the researcher will guide the participant towards the building exit. Afterwards, the researcher will again sanitise the hardware equipment. Other facility-specific procedures are defined in the chapters that follow.

1.5. Evacuation

In case of an emergency situation (e.g., fire), any social distancing requirements are temporarily suspended in favour of a safe and expeditious evacuation.

1.6. Experiment supervisors

The experiment supervisors will always be TU Delft staff members. In cases where the experimenter is an MSc student, the experiment supervisor has to be “on call” in the Faculty building as the responsible supervisor.

3

Human-Machine Interaction Laboratory

The Human-Machine Interaction Laboratory (HMI Lab), see Fig. 3.1, is located in low-rise part of the Aerospace Engineering building in room LB037. There are two connected rooms, see Fig. 3.2: a control room for the experimenter and a room with the real experiment setup, consisting of car driving and aircraft control stations. Access to the experiment room is only possible through the control room.



Figure 3.1: The Human-Machine Interaction Laboratory (HMI Lab).

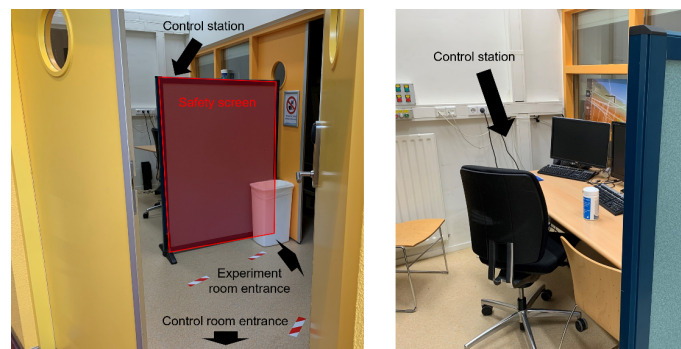


Figure 3.2: COVID-19 access measures for the HMI Lab.

3.1. Recruiting participants

The HMI Lab recruiting protocol is the same as the general recruiting protocol.

3.2. Traveling

The HMI Lab traveling protocol is the same as the general traveling protocol.

3.3. Entering and leaving the faculty and facility

The HMI Lab entry protocol is the same as the general entry protocol. At the laboratory, the experimenter and participant will avoid being both in the same room, see Fig. 3.2. This will be achieved by having the participant enter first, and continue directly to the experiment room. The experimenter enters second, keeping 1.5 m distance at all times. Exiting the HMI Lab happens in the reverse order.

3.4. Briefing

The briefing and debriefing of the participant will be performed at standing tables outside the HMI Lab, see Fig. 3.3, while maintaining sufficient distance between experimenter and participant (e.g., separate tables for experimenter and participant).



Figure 3.3: Briefing space for HMI Lab experiments.

3.5. Experiment

The rooms are too small to easily maintain the required 1.5 m distance, therefore the participant will use the experiment room and only use the control room to enter or leave the experiment room, the experimenter only uses the control room, and vacates that room when the participant wants to leave or enter the experiment room.

All surfaces and objects the participant and experimenter handle during the experiment, shall be disinfected between participants, after each break and after the experiment. This includes:

- Control inceptors used by the experiment (side stick, steering wheel, throttle handle)
- Other interfaces used during the experiment (MCP, CDU)
- Simulator surfaces (emergency buttons, door handles, chair and armrest)
- Control room devices and surfaces (mouse, keyboard, door handles)
- Tables used for experiment briefing and debriefing.

3.6. Evacuation

The evacuation protocol is the same as the general evacuation protocol.

Runtable Example

In appendix is shown an example of a run table used for one of the participants. The left most column is used to fill the run number, the middle column states the experimental condition to be tested and the third columns leaves space for possible comments on the runs.

schedule_5

[illegible]