

# **Fuzzy Modelling of Human Psycho-Physiological State and Fuzzy Adaptive Control of Automation in Human-Machine Interface**

A thesis submitted to The University of Sheffield for the Degree of Doctor of  
Philosophy

**Changjiang He**

Department of Automatic Control and Systems Engineering

January 2021



*To my family  
for their unconditional love and support.*



## **Acknowledgements**

First of all, I would like to express my sincere gratitude to my supervisor, Prof. Mahdi Mahfouf for his continuous guidance and support throughout my PhD process. I am always grateful for his encouragement and patience that motivated me to achieve higher.

Many thanks to Dr Luis A. Torres-Salomao for his continuously help on all my technical questions throughout my study. Furthermore, I would like to thank my colleagues, Dr. Wafa' H. AlAlaweena and Dr. Olusayo Obajemu, for all their generous support in my academic endeavours.

Special thanks to all my friends that kindly participated in the experiments and enabled this study to be possible. Finally, I would also like to thank all the ACSE staff for their support during these four years.

## **Abstract**

This research aims at proposing a new modelling and control framework that monitors the human operators' psychophysiological state in the human-machine interface to prevent performance breakdown. This research started with the exploration of new psychophysiological state assessment approaches to the adaptive modelling and control method for predicting human task performance and balancing the engagement of the human operator and the automatic system. The results of this research may also be further applied in developing advanced control mechanisms, investigating the origins of human compromised performance and identifying or even remedying operators' breakdown in the early stages of operation, at least.

A summary of the current human psychophysiological studies, previous human-machine interface simulation and existing biomarkers for human psychophysiological state assessment was provided for simulation experiment design of this research. The use of newly developed facial temperature biomarkers for assessing the human psychophysiological state and the task performance was investigated. The research continued by exploring the uncertainty of the human-machine interface system through the use of the complex fuzzy logic based offline modelling approach. A new type-2 fuzzy-based modelling approach was then proposed to assess the human operators' psychophysiological states in the real-time human-machine interface. This new modelling technique integrated state tracking and type-2 fuzzy sets for updating the rule base with a Bayesian process. Finally, this research included a new type-2 fuzzy logic-based control algorithm for balancing the human-machine interface systems via adjusting the engagement of the human operators according to their psychophysiological state and task performance. This innovative control approach combined the state estimation of the human operator with the type-2 fuzzy sets to maintain the balance between the task requirements (i.e. difficulty level) and the human operator feasible effort (i.e. psychophysiological states). In addition, the research revealed the impacts of multi-tasking and general fatigue on human operator's performance.

# Contents

<b>Nomenclature</b>	<b>xv</b>
<b>Acronyms</b>	<b>xvii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Human-Machine Interface . . . . .	1
1.1.1 Human-Machine Interface . . . . .	1
1.1.2 Psychophysiological State . . . . .	2
1.1.3 Psychophysiological State Measurement . . . . .	3
1.1.4 Human-Machine Interface Simulation . . . . .	3
1.2 Biomarkers for Psychophysiological State . . . . .	4
1.2.1 Importance of New Biomakers . . . . .	4
1.2.2 Facial Temperature Biomarkers . . . . .	4
1.3 Modelling and Control of Human-Machine Interface . . . . .	5
1.3.1 Adaptive Online Modelling . . . . .	5
1.3.2 Balancing Control . . . . .	5
1.4 Thesis Structure . . . . .	6
<b>2 Introduction to Human-Machine Interface</b>	<b>8</b>
2.1 Human-Machine Interface . . . . .	8
2.2 Psychophysiological State . . . . .	11
2.2.1 Core Affective . . . . .	12
2.2.2 Core Affective Assessment . . . . .	12
2.3 Psychophysiological State Measurement . . . . .	15

2.3.1	Peripheral Physiology . . . . .	17
2.3.2	Electrocardiogram . . . . .	18
2.3.3	Heart Rate Variability . . . . .	20
2.3.4	Central Physiology . . . . .	24
2.3.5	Electroencephalogram . . . . .	25
2.3.6	Task Load Index . . . . .	28
2.3.7	Affect Modulated Startle . . . . .	29
2.3.8	Pupil Size . . . . .	31
2.3.9	Pupil Diameter Marker . . . . .	32
2.3.10	Behaviour . . . . .	32
2.3.11	Facial Temperature . . . . .	34
2.4	Human-Machine Interface Simulation . . . . .	35
2.4.1	Stroop Colour-Word Interference . . . . .	37
2.4.2	Virtual Vehicle Operation . . . . .	38
2.4.3	Management System Simulation . . . . .	39
2.4.4	Mental Arithmetic Experiment . . . . .	41
2.4.5	Experimental Configuration . . . . .	42
2.5	Control of the Human-Machine Interface . . . . .	49
2.6	Summary . . . . .	51
<b>3</b>	<b>Facial Thermal Imaging for Psychophysiological State Detection</b>	<b>52</b>
3.1	Background . . . . .	53
3.1.1	Psychophysiological State Detection . . . . .	53
3.1.2	Facial Temperature . . . . .	54
3.2	Experimental Setup . . . . .	55
3.2.1	Participants . . . . .	55
3.2.2	Simulation Experiments . . . . .	56
3.2.3	Temperature Data Acquisition . . . . .	57
3.3	Developed Facial Temperature Biomarkers . . . . .	57
3.4	Evaluation of Facial Temperature Biomarkers . . . . .	59
3.4.1	Maximum Facial Temperature . . . . .	59

3.4.2	Mean Nasal Temperature . . . . .	66
3.4.3	Mean Forehead Temperature . . . . .	68
3.4.4	Differential Energy between Philtrum and Forehead . . . . .	73
3.4.5	Comparison of Facial Temperature Biomarkers with Conventional Biomarkers . . . . .	75
3.5	Discussion . . . . .	78
3.6	Summary . . . . .	80
<b>4</b>	<b>Adaptive General Type-2 Fuzzy Modelling for Psychophysiological State Prediction</b>	<b>81</b>
4.1	Background . . . . .	82
4.1.1	Fuzzy Logic Development . . . . .	82
4.1.2	Fuzzy Logic . . . . .	82
4.1.3	Fuzzy System . . . . .	83
4.1.4	Type-2 Fuzzy Logic Modelling . . . . .	86
4.2	Experimental Setup . . . . .	87
4.2.1	Data Acquisition . . . . .	87
4.2.2	Modelling Experiment Configuration . . . . .	89
4.3	Offline Modelling . . . . .	89
4.3.1	Complex Fuzzy Sets and Logic . . . . .	91
4.3.2	Single Partition Complex Fuzzy Inference Modelling . . . . .	91
4.3.3	Evaluation of the Single Partition Complex Fuzzy Inference Model . . . . .	94
4.4	Adaptive General Type-2 Fuzzy Modelling . . . . .	97
4.5	Evaluation of Adaptive General Type-2 Fuzzy Modelling . . . . .	104
4.5.1	Model Configuration . . . . .	105
4.5.2	Modelling Results Comparison . . . . .	106
4.5.3	Adaptive Learning of General Type-2 Fuzzy Model . . . . .	111
4.5.4	Computation of General Type-2 Fuzzy Model . . . . .	116
4.6	Summary . . . . .	119

<b>5 Adaptive General Type-2 Fuzzy Controller for Balancing Human-machine Interface</b>	<b>120</b>
5.1 Background . . . . .	121
5.1.1 Multitasking . . . . .	121
5.1.2 Fatigue . . . . .	123
5.2 Experimental Setup . . . . .	124
5.2.1 Human-Machine Interface Simulation . . . . .	124
5.2.2 Controlling Experiment Configuration . . . . .	125
5.3 Human-Machine Interface Balance Control . . . . .	126
5.4 Evaluation of Impaired Task Performance and Controller . . . . .	128
5.4.1 Controller Configuration . . . . .	128
5.4.2 Controller Evaluation . . . . .	130
5.4.3 Controller Comparison . . . . .	132
5.4.4 Multitasking . . . . .	134
5.4.5 Fatigue . . . . .	139
5.4.6 Fatigue Management . . . . .	142
5.5 Summary . . . . .	146
<b>6 Conclusion and Future Work</b>	<b>148</b>
6.1 Conclusion . . . . .	148
6.2 Future Recommendations . . . . .	150
<b>A Summary of Two Sample T Test for Facial Temperature Experiment</b>	<b>152</b>
<b>B Summary of Means and Standard Deviations of the Psychophysiological Biomarkers for Multitasking</b>	<b>156</b>
<b>C Summary of Means and Standard Deviations of the Psychophysiological Biomarkers for Fatigue</b>	<b>158</b>
<b>References</b>	<b>160</b>

# List of Figures

2.1	Combination of Dimensional and Discrete Affective State (data based on [11, 41, 49, 68]) . . . . .	14
2.2	The Relationship between Heart Rate and Autonomic Nervous System [13] . . . . .	19
2.3	Example of ECG Signal [21] . . . . .	21
2.4	ECG Electrodes [1] . . . . .	24
2.5	EEG Electrodes [1] . . . . .	30
2.6	Model of aCAMS [80, 81] . . . . .	40
2.7	Image of ActiView Software by Biosemi [1] . . . . .	43
2.8	EOG Electrodes [1] . . . . .	44
2.9	Gazepoint™ GP3 Eye Tracker Software . . . . .	45
2.10	Thermal Camera Software . . . . .	46
2.11	Mental Arithmetic Test GUI . . . . .	47
3.1	Difficulty levels for facial temperature experiment session 1 (progressive increase) and session 2 (random) . . . . .	56
3.2	Regions of interest: periorbital, nasal and forehead . . . . .	57
3.3	Maximum Facial Temperature . . . . .	64
3.4	Mean Nasal Temperature . . . . .	69
3.5	Mean Forehead Temperature . . . . .	72
3.6	Differential Energy between Philtrum and Forehead . . . . .	76
4.1	Diagram of type-1 fuzzy sets and type-2 fuzzy sets . . . . .	83
4.2	Example of a simple fuzzy logic system . . . . .	84

4.3	A Picture of a participant during the HMI experiment . . . . .	88
4.4	Difficulty level for prediction experiment session 1 (progressive increase) and session 2 (random) . . . . .	90
4.5	The structure of single partition complex fuzzy inference model . . . . .	93
4.6	Accuracy performance from the participant 08 and from the single partition complex fuzzy inference model prediction in training session . . . . .	96
4.7	Accuracy performance from the participant 08 and from the single partition complex fuzzy inference model prediction in testing session . . . . .	97
4.8	Example rule of GT2FM for the first state . . . . .	98
4.9	Diagram of the GT2FM for the HMI simulation experiment (controller and other sections in - - will be the subject of the next chapter, chapter 5) . . . . .	105
4.10	Membership for the performance function $f_A(\bar{T}_n, \bar{T}_f)$ . . . . .	107
4.11	Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM, $\bar{T}_n$ , $\bar{T}_f$ , $T_{maxf}$ ) for the participant 08 in session 2 . . . . .	110
4.12	Accuracy performance from the participant 08 and from the adaptive general type-2 fuzzy model prediction in session 2 . . . . .	111
4.13	Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM, $\bar{T}_n$ , $\bar{T}_f$ , $T_{maxf}$ ) for the participant 08 in session 1 . . . . .	113
4.14	Accuracy performance from the participant 08 and from the adaptive general type-2 fuzzy model prediction in session 1 . . . . .	114
4.15	Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM, $\bar{T}_n$ , $\bar{T}_f$ , $T_{maxf}$ ) for the participant 03 in session 1 . . . . .	115
4.16	Accuracy performance from the participant 03 and from the adaptive general type-2 fuzzy model prediction in session 1 . . . . .	116
4.17	Example of uncertainty convergence with the learning algorithm . . . . .	118
5.1	Diagram of the controller for the HMI simulation experiment (GT2FM and other sections in - - is the subject of the previous chapter, chapter 4) . . . . .	129

---

5.2	Real-time experiment results for the participant 08 with and without the HMIBC . . . . .	132
5.3	Real-time experiment results for the HMIBC and the E-MBC . . . . .	133
5.4	Real-time experiment results for the participant 10 in Single-tasking and Multitasking . . . . .	135
5.5	Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PMD, $\bar{T}_n$ , $\bar{T}_f$ , $T_{maxf}$ ) for the participant 02 in multitasking session . . . . .	138
5.6	Accuracy performance and HMIBC control action from the participant 02 in multitasking session . . . . .	139
5.7	Real-time experiment results for the participant 03 . . . . .	140
5.8	Accuracy performance and HMIBC control action from the participant 06 in fatigue session (shade green area represented self-report fatigue) . . . . .	142
5.9	Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM, $\bar{T}_n$ , $\bar{T}_f$ , $T_{maxf}$ ) for the participant 06 in fatigue session (shade green area represented self-report fatigue) . . . . .	143
5.10	Real-time HMIBC experiment results for the participant 01 without vs. with the fatigue management . . . . .	144

# List of Tables

2.1	Mental Arithmetic Difficulty Level . . . . .	48
3.1	Sample Sizes of Different Phase within the Experiment Session . . .	60
3.2	Overall Maximum Facial Temperature T-test Results for the Experimental Sessions . . . . .	61
3.3	Overall Arithmetic Means for the Maximum Facial Temperature . .	62
3.4	Total Correlation between Maximum Facial Temperature and Difficulty Level . . . . .	63
3.5	Total Dispersion Ratio for Maximum Facial Temperature . . . . .	64
3.6	Total Zero-crossing Rate for Maximum Facial Temperature . . . . .	66
3.7	Overall Mean Nasal Temperature T-test Results for Experimental Sessions . . . . .	66
3.8	Overall Arithmetic Means for Mean Nasal Temperature . . . . .	67
3.9	Total Correlation between Mean Nasal Temperature and Difficulty Level . . . . .	68
3.10	Overall Mean Forehead Temperature T-test Results for Experimental Sessions . . . . .	70
3.11	Overall Arithmetic Means for Mean Forehead Temperature . . . . .	70
3.12	Total Correlation between Mean Forehead Temperature and Difficulty Level . . . . .	71
3.13	Overall DEFP T-test Results for Experimental Sessions . . . . .	73
3.14	Overall Arithmetic Means for DEFP . . . . .	74
3.15	Total Correlation between DEFP and Difficulty Level . . . . .	74

---

3.16	Mean H Values for T-test Summary . . . . .	75
3.17	Overall HRV1 T-test Results for Experimental Sessions . . . . .	77
3.18	Overall HRV2 T-test Results for Experimental Sessions . . . . .	77
3.19	Overall TLI1 T-test Results for Experimental Sessions . . . . .	78
3.20	Overall TLI2 T-test Results for Experimental Sessions . . . . .	78
3.21	Overall PDM T-test Results for Experimental Sessions . . . . .	78
4.1	Example rule of single partition complex fuzzy inference model after training . . . . .	94
4.2	Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of single partition complex fuzzy inference model . . . . .	95
4.3	Rule-Base of GT2FM for the First State . . . . .	99
4.4	Linguistic Labels of the Inputs and Output for the First State . . . . .	99
4.5	Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of GT2FM . . . . .	108
4.6	Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of A-GT2-FCM . . . . .	108
4.7	Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of ANFIS . . . . .	109
5.1	Upper and Lower Performance Boundaries . . . . .	130
5.2	Breakdown Percentages (%) for Modelling and Controlling Sessions	131
5.3	Means of Actual Accuracy for Modelling and Controlling Sessions .	131
5.4	Standard Deviations of Actual Accuracy for Modelling and Controlling Sessions . . . . .	131
5.5	Breakdown Percentages (%) for the HMIBC and the E-MBC . . . . .	134
5.6	Means of Actual Accuracy for the HMIBC and the E-MBC . . . . .	134
5.7	Standard Deviations of Actual Accuracy for the HMIBC and the E-MBC . . . . .	134
5.8	Breakdown Percentages (%) for the Task Performance of Single-tasking and Multitasking . . . . .	136

---

5.9	Means of Actual Accuracy for the Task Performance of Single-tasking and Multitasking . . . . .	136
5.10	Standard Deviations of Actual Accuracy for the Task Performance of Single-tasking and Multitasking . . . . .	136
5.11	H Value Summary for T-test between Single-tasking and Multitasking	137
5.12	Breakdown Percentages (%) for the Task Performance of Normal and Fatigue States . . . . .	140
5.13	Means of Actual Accuracy for the Task Performance of Normal and Fatigue States . . . . .	140
5.14	Standard Deviations of Actual Accuracy for the Task Performance of Normal and Fatigue States . . . . .	141
5.15	H Value Summary for T-test between Normal State and Fatigue . . .	141
5.16	Breakdown Percentages (%) for the HMIBC without/with the Fatigue Management . . . . .	145
5.17	Means of Actual Accuracy for the HMIBC without/with the Fatigue Management . . . . .	145
5.18	Standard Deviations of Actual Accuracy for HMIBC without/with the Fatigue Management . . . . .	146
A.1	Overall Maximum Facial Temperature T-test Results for Experimental Session 1 . . . . .	152
A.2	Overall Maximum Facial Temperature T-test Results for Experimental Session 2 . . . . .	152
A.3	Overall Mean Nasal Temperature T-test Results for Experimental Session 1 . . . . .	153
A.4	Overall Mean Nasal Temperature T-test Results for Experimental Session 2 . . . . .	153
A.5	Overall Mean Forehead Temperature T-test Results for Experimental Session 1 . . . . .	153
A.6	Overall Mean Forehead Temperature T-test Results for Experimental Session 2 . . . . .	153

---

A.7 Overall DEFP T-test Results for Experimental Session 1 . . . . .	153
A.8 Overall DEFP T-test Results for Experimental Session 2 . . . . .	154
A.9 Overall HRV1 T-test Results for Experimental Session 1 . . . . .	154
A.10 Overall HRV1 T-test Results for Experimental Session 2 . . . . .	154
A.11 Overall HRV2 T-test Results for Experimental Session 1 . . . . .	154
A.12 Overall HRV2 T-test Results for Experimental Session 2 . . . . .	154
A.13 Overall TLI1 T-test Results for Experimental Session 1 . . . . .	154
A.14 Overall TLI1 T-test Results for Experimental Session 2 . . . . .	155
A.15 Overall TLI2 T-test Results for Experimental Session 1 . . . . .	155
A.16 Overall TLI2 T-test Results for Experimental Session 2 . . . . .	155
A.17 Overall PDM T-test Results for Experimental Session 1 . . . . .	155
A.18 Overall PDM T-test Results for Experimental Session 2 . . . . .	155
B.1 Mean Values of Biomarkers in Single-tasking . . . . .	156
B.2 Mean Values of Biomarkers in Multitasking . . . . .	157
B.3 Standard Deviations of Biomarkers in Single-tasking . . . . .	157
B.4 Standard Deviations of Biomarkers in Multitasking . . . . .	157
C.1 Mean Values of Biomarkers in Normal State . . . . .	158
C.2 Mean Values of Biomarkers in Fatigue State . . . . .	159
C.3 Standard Deviations of Biomarkers in Normal State . . . . .	159
C.4 Standard Deviations of Biomarkers in Fatigue State . . . . .	159



# Nomenclature

A list of the variables and notations used in this thesis is given below. The definitions and conventions set here will be observed throughout the thesis unless otherwise stated. For a list of acronyms, please consult page xvii.

$\epsilon$	Emissivity
$\mu_x$	Mean
$\sigma$	Constant
$\sigma_x$	Standard deviations
$\tau$	Transmittance
$c$	Pearson correlations
$D$	Density matrix
$DT$	Distance tolerance
$E$	State estimation
$e$	Error
$ET$	Error tolerance
$f$	Membership function
$I$	Input vector
$N$	Sample size

$o$	Observation
$P$	Transition matrix
$R$	Fuzzy rule base
$T$	Temperature
$t$	Time
$W$	Radiation
$x$	Model input
$y$	Model prediction

# Acronyms

- A-GT2-FCM** adaptive general type-2 fuzzy c-means modelling. 10, 104, 106, 107, 109, 117, 119, 128, 132, 149
- ANFIS** adaptive neuro-fuzzy inference system. 10, 50, 86, 104, 106, 107, 109, 119
- DEFP** differential energy between philtrum and forehead. 34, 57, 59, 73–75, 77
- DL** difficulty level. 59, 61, 70, 73–75
- E-MBC** energy model-based control. 120, 128, 132, 134, 147
- ECG** electrocardiography. 17, 18, 20, 23, 25, 42–44, 51, 53–55, 77, 87
- EEG** electroencephalography. 15, 17, 23, 25–28, 31, 42–44, 51, 53–55, 77, 87, 122, 147
- EMG** electromyography. 23
- EOG** electroocoulogram. 23, 43, 44
- GT2FM** general type-2 fuzzy model. 98, 105–107, 109, 111, 116–118, 147
- GUI** graphical user interface. 39, 42, 44, 45, 47
- HMI** human-machine interface. 1–3, 5, 6, 9–11, 15, 16, 22, 26–29, 31–42, 49–52, 56, 77, 80, 86, 89, 90, 95, 96, 104–106, 114, 119, 121–126, 128, 130, 131, 134, 141, 144, 146–151
- HMIBC** human-machine interface balance control. 124–126, 128, 131, 132, 134, 139, 142, 145–147

**HRV** heart rate variability. 20, 22, 23, 37, 38, 51–53, 55, 59, 75, 77, 84, 112, 119, 137, 141, 148

**PDM** pupil diameter marker. 31, 32, 45, 51, 52, 60, 75, 77, 114, 119, 148

**ROI** regions of interest. 57

**TLI** task load index. 28, 29, 51–53, 55, 59, 75, 77, 119, 137, 148

**WM** working memory. 28, 53, 121, 122

# Chapter 1

## Introduction

This project aims at producing a new system for the human-machine interface that monitored human operators who undergo psychophysiological loading in order to detect breakdown. This adaptive system was also implemented with a decision-making process as to balance manual/automatic operations depending on the predicted psychophysiological state of the subject at the time. This research was conducted in the Human Performance Laboratory located in the Department of Automatic Control and System Engineering (ACSE) of the University of Sheffield. The real-time experiment and the collection of subjects' data for the analysis and modelling operations followed all the health and safety regulations of the University. The impact of the project findings was directly related to road, work and flight safety whereby recommendations for a new road map on automation and autonomy may be drawn for the future.

### 1.1 Human-Machine Interface

#### 1.1.1 Human-Machine Interface

Human-machine interface (HMI) is naturally introduced with the implementation of automatic systems nowadays, from daily uses of smart phones to international regulation of physical distribution. The combination of a human operator and an automatic system inherited the operational flexibility from manual intervention

while expanded the processing capacity for the information flow. With the development of automatic systems, the communication between the human operators and systems had shifted from basic manual operations to behavioural or even emotional communication. The lack of efficiency within HMI limited such combination to fully utilise the computational ability of advanced automatic systems. The misjudgement of both sides led to the compromised overall performance of operations, the limited data processing of the automatic system and the elevated demands on the human operator.

For the purpose of maintaining the high efficiency of operation state and reducing the threat on the reliability and safety, the communication between the human operator and the automatic system must remain accurate and punctual. It requires a way to assess the human operator's psychophysiological state for performance evaluation and adjust the automation level of the system in response to the breakdown. Therefore, the aims and objectives of this research are:

1. Explore new convenient and efficient human psychophysiological state assessment methods.
2. Create new computational modelling approaches for predicting human psychophysiological state in real-time.
3. Propose new control methods within the interface to ensure the safety and efficiency of the system.

### **1.1.2 Psychophysiological State**

Humans react to the outside stimuli through conditional reflex. A human response is generated from neural communication, executed with the muscular system and regulated with the endocrine. For a complex task, such as operations of automatic systems, it requires a collaboration of multiple higher level activities, including self consciousness, attention, etc. A human action demands an interpretation of core affective for the analyses of situation and feasible resources, a memory recall of previous experience and relevant knowledge for project planning, a motor based on muscular system and regulated with others for the action execution and a self

awareness of emotions and thoughts for memory formation. The monitoring of the psychophysiological state is directly corresponded to human emotions and action whereby predictions of human operator performance in HMI may be made in real time.

### 1.1.3 Psychophysiological State Measurement

The human psychophysiological state measurements utilised in the current researches focused on one or more common used aspects - subjective experience, peripheral physiology, affect modulated startle, central physiology and behaviour. These five aspects indicated part of activities within the central nerve system, cardiovascular system and muscular system. By monitoring these systems, it was able to estimate the attitude of a subject towards a certain task and the energy invested for the performance. As human organs and cells are operated in multi-function style, a single monitoring from a solo aspect cannot differentiate between the activities for voluntary response and involuntary life support. Therefore, it is necessary to assess psychophysiological state from a collective basis.

Electrocardiogram, electroencephalogram, pupil size and facial temperature were selected for the monitoring of subjects' psychophysiological state in this research. Based on previous literature and experiment results, heart rate variability, task load index, pupil diameter maker and facial temperature biomarkers were computed as to indicate task performance related psychophysiological activities at the time.

### 1.1.4 Human-Machine Interface Simulation

A good human-machine interface simulation should be able to introduce workload in laboratory conditions similar to the real world HMI. Apart from the requirement of similarity, it should also be simple to use and reliable to provide distinguishable workload for the reproducibility of experiment. This research selected a mental arithmetic test as the HMI simulation. Compared to the other simulation methods, the mental arithmetic test had significant advantages such as being effective, robust, simple to use and intuitive.

## 1.2 Biomarkers for Psychophysiological State

### 1.2.1 Importance of New Biomarkers

For a full integration of the human operator and the automatic system, one needs to consider the human psychophysiological state and its corresponding behavior in the design of automatic intervention. An effective biomarker is directly linked with task related biological activity and able to reveal the subjects' inner conditions without the use of verbal or behavioral communication. Due to the multi-function of cells and organs, each biomarker inevitably had its own limitation, depending on the method and target of measurement. Existing biomarkers are mostly developed on the use of electrode or other physical attachments. Such measurements suffered lack of accuracy from body contact. Furthermore, it limits the movement of subjects and introduced additional noise to the recordings. It should be noted that the research of human psychophysiological state was at the infancy stage with most theories unverified. The search of new biomarkers was an effective way to increment the understanding on the inner processes whereby recommendations for a more complete conception of psychophysiological state may be drawn.

### 1.2.2 Facial Temperature Biomarkers

It was found that facial temperature, based mainly on vascular system, shared strong connections with several human emotional and behavioral mechanisms. The use of infrared cameras with modern imaging techniques provided a reliable approach to study these relationships in real time. With the experimental results of this research, it was found that the regional temperature of the nasal, forehead and orbital areas showed a consistent correlation with stress related psychophysiological state changes. Compared to the conventional biomarkers, the biomarkers based on facial temperature exhibited similar or even better results in state differentiation, especially for the low workload states which failed to trigger any observable change in conventional biomarkers. It suggested that the facial temperature biomarkers may compensate for the shortage of conventional biomarkers and expand the depth and breadth of human psychophysiological state research.

## 1.3 Modelling and Control of Human-Machine Interface

### 1.3.1 Adaptive Online Modelling

Even with multiple biomarkers monitoring subjects from diverse aspects it was still difficult and sometime even impossible to predict the human task performance from psychophysiological measurements. Previous researches and the use of complex fuzzy inference modelling showed that the information of recordings was time sensitive. It should be noted that human psychophysiological state is unrepeatable. The effectiveness of patterns summarised from previous data rapidly faded with time. Also, each subject's personal knowledge and individual experience were unique and excluding to others. It limited the models and frameworks to be case specific. Thus, a new type-2 fuzzy based modelling approach was introduced to fulfil the requirement of adaptation. The new purposed model consisted of a self-organised structure that could adapt to the time-varying HMI and predict human task performance with high accuracy in real time.

### 1.3.2 Balancing Control

The objective of any endeavour towards the control of a HMI is to correctly balance the engagement of the human operator and the automatic system by considering the operator's capability and capacity. In this way, human effort may be fully integrated into the system without any psychophysiological breakdown due to excessive amount of workload. In this research, a new balancing control method based on adaptive type-2 fuzzy based modelling was investigated in real time. The results proved that such control method was able to maintain the operators' best performance consistently and steadily throughout the entire HMI sessions. In addition, two main origins of the compromised task performance of human operator in the HMI, i.e. fatigue and multitasking, were studied. The impact of these findings may be directly implementable to workload planning, performance improvement and work-rest schedule, whereby recommendations for more advanced controlling method may be drawn.

## 1.4 Thesis Structure

This thesis is organised as follows: Chapter 2 summarises the up-to-date HMI researches from human psychophysiological studies to the modelling and control of HMI. This chapter presents a literature based summary of current psychophysiological state and performance studies. Additionally, the experimental designs for the HMI simulation and the conventional psychophysiological assessment methods are introduced and compared through the findings of previous related researches. Finally, the details of the design configurations which were included in the experiment session are presented.

Chapter 3 presents the new designed facial temperature biomarkers for predicting human psychophysiological state and task performance in the HMI. The experiment results of the ten participants in the HMI simulation are presented for validation.

Chapter 4 starts with a literature based introduction of type-2 fuzzy logic based modelling and its application. It is then followed by a discussion of the limitation of offline models for modelling human psychophysiological state in the HMI with experimental results. Finally, a new type-2 fuzzy-based modelling approach combined type-2 fuzzy sets with state tracking to update the rule base through a Bayesian process is proposed and validated with online experiments.

Chapter 5 proposes a new balancing control for the HMI systems based on the adaptive modelling method in Chapter 4. This newly proposed control approach is validated on mental arithmetic cognitive experiments for the same ten participants and compared with the existing energy model-based control. In addition, the two common origins for human compromised performance in the HMI, fatigue and multitasking, are explored and integrated in the design of control configuration.

Chapter 6 summarises the work and makes recommendations for future research.

This thesis includes work presented in the following papers:

- He C., Mahfouf M., Torres-Salomao L.A. 2018. "Facial Temperature Markers for Mental Stress Assessment in Human-Machine Interface (HMI) Control System". ICINCO 2018.
- He C., Mahfouf M., Torres-Salomao L.A. 2020. "Thermal Imaging for Psychophysiological State Detection in the Human-Machine Interface (HMI) Control System". In: Gusikhin O., Madani K. (eds) Informatics in Control, Automation and Robotics. ICINCO 2018. Lecture Notes in Electrical Engineering, vol 613. Springer, Cham.
- He, C., Mahfouf, M. and Torres-Salomao, L.A., 2021. "An Adaptive General Type-2 Fuzzy Logic Approach for Psychophysiological State Modeling in Real-Time Human-Machine Interfaces". IEEE Transactions on Human-Machine Systems, 51(1):1-11.
- He C., Mahfouf M., Torres-Salomao L.A. "An Adaptive General Type-2 Fuzzy Logic Approach for Real-Time Human-Machine Interface: Self-Organising Control, Multitasking and Fatigue". Under Review. IEEE Transactions on Fuzzy Systems.

## Chapter 2

# Introduction to Human-Machine Interface

This chapter discusses current applications of the human-machine interface as well as their limitations. It also includes a brief introduction of the human psychophysiological state, which is supported by current psychological studies and biological research. Several commonly used psychophysiological state measurement approaches in the relevant research are presented, and the human-machine interface simulation method for this research is selected. Finally, this chapter summarizes the essential expectations for the control frameworks of the human-machine interface.

### 2.1 Human-Machine Interface

Automatic systems were widely implemented in diverse areas from daily life to global regulations from the last decade, such as advanced manufacturing, transportation and clinical medicine. Compared to conventional manual control systems, the combination of an automatic system and a human operator include several significant advantages, including fast reaction to and processing of a large amount of concurrent information [87]. Such a combination may be able to help humans to deal with daily tasks with relatively high efficiency and low error rate. Thanks to this revolutionary intelligent power, automatic systems have commonly been adopted in many important management and operation systems.

There are, however, many barriers yet to be broken for the above combination to fulfil its full potential advantages and reach its theoretical efficiency. The lack of trust and the over-trust with high expectations towards the automatic systems compromised the overall performance of this combination, whereas increases of the operational demands of the human operator was threatening the reliability and the safety of the whole system [87]. Therefore, it is of paramount importance to introduce a mechanism that was able to bridge the communication gap between the automatic system and the so-called 'human in the loop'. Euphemistically speaking, the ideal system should be able to:

1. Estimate the human operator's psychophysiological state and predict their task performance.
2. Generate the corresponding automatic level for implementation and ensure continuous productive operation output.

In order to achieve these requirements, it is necessary to expand current simple direct technical interaction with mental content interaction in HMI [87]. The purpose of such intervention is to maintain the running of HMI in a fast and in an adaptive manner for a sufficiently long period of time regardless of individual differences and singular capabilities.

Humans respond to environmental stimuli through conditioned reflex and subject to personal experience and psycho-physiological state [53]. The HMI consists of communications at three different levels, namely, direct technical, emotional and mental levels. The collaboration at different levels and the diverse dynamic individual states lead, as a result, to a network with high complexity and uncertainty [87]. This requires a shift from traditional modelling approaches, primarily depending on mathematical expressions and physical laws towards data-driven modelling approaches led by powerful pattern recognition and fast data mining operation. These techniques, such as artificial neural networks and fuzzy logic systems, exhibit an extraordinary ability to generate and applied conditional patterns under uncertain environments with a limited amount of data.

For the safety of HMI systems, the critical thinking and reasoning of human operators should always be an intrinsic part of the final decision and cannot be fully replaced by automation (e.g. Boeing 737 MAX crash). The decision of splitting the workload into automatic machines and human operators lead to the human-centred modelling to estimate operator's psych-physiological state. The existing data-driven automatic approaches for psycho-physiological prediction of HMI systems mainly relied on adaptive neuro-fuzzy inference system (ANFIS), type-1 Mamdani fuzzy model [23, 52, 80, 81], proportional integral Mamdani fuzzy model [59–61], type-2 fuzzy model [83, 84] and support vector machines [97, 98]. The majority of the presented models have fixed configurations based on offline training sessions. This has the potential of introducing an intrinsic default in the modelling, for the time-validity of biological data often lasted for a limited period only. As a result, the accuracy of offline models significantly declines with the time, since the human psychophysiological state was highly dynamic. In addition, such model structures restricted the models to those who only share similar patterns to the original training samples, i.e. the established models are not generalising and not robust for a broad implementation.

In order to overcome these shortages, advanced developments of self-organising and adaptive learning were integrated with the existing modelling approaches. For example, in the previous research, Luis proposed the adaptive general type-2 fuzzy c-means modelling (A-GT2-FCM) framework [84] that was able to generate, modify and delete fuzzy rules from continuous observation, so that the framework may be adaptable to the individuals' differences and their psychophysiological changes. However, the nonlinear changes of psychophysiological state and the lack of flexible weight adjustment for inter- and intra-uncertainty compromised the prediction performance of these models. Therefore, these models failed to anticipate the situation changes in advance and required a certain amount of time for reconfiguration, particularly, when the human operator was under some extreme conditions (e.g. psychophysiological breakdown, multitasking and fatigue), or when the performance of such models suffered severe instability and delay, because of the variations in psychophysiological states.

In summary, it is worth noting that current frameworks and models mentioned above may not fulfil one or more of the following essential requirements for HMI systems:

1. **Adaptability:** the ability to reconfigure themselves consistently according to real-time system changes.
2. **Intelligence:** specifically refers to the ability to interpret the system state and modify their inference engine correspondingly.
3. **Robustness:** the ability to handle inter- and intra-uncertainty and being generalised to any human operator.
4. **Being Explicit:** the ability to summarise the learning experiences via easy to understand logical statements.

It is important to address the above mentioned features in the prediction of human operator psychophysiological state. In order to balance the HMI system, a more sophisticated model using easy-to-access psycho-physiological data in addition to existing models and frameworks is required.

## 2.2 Psychophysiological State

Conditioned reflex lay the foundation of human responses to the outside stimuli. The collective network of reflexes presents a general regulation across neural, muscular and endocrine systems. This led to the rise of higher level psychophysiological activities, such as self consciousness, attention, reasoning, memory, cognition, perception and execution [6, 7]. Therefore, the human psychophysiological state may be partially or fully assessed by monitoring the biological processes of some or all organs and systems [53]. Meanwhile, it suggested that the prediction of human performance towards a certain task can also be achieved [67, 95, 96].

### 2.2.1 Core Affective

The psychophysiological state is a temporal segment from core affective, a spontaneous voluntary time-variant collective comprising various transient biological reflexes from both the autonomic arc and the somatic arc [41, 68]. Core affective is the combination of past experience, individual perception and physical evaluation, and consequently generates the primary impression and the basic affective response to any received stimulus [41]. In other words, core affective, which is based on the current physical condition, the scenario analysis and the previous memory, determines feasible responses (from both emotional and behavioural) for any triggered stimuli. Because of such coherence, human responses are regulated and may be interpreted by body biological processes [41].

Present neurological studies locate several brain regions responsible for the interlacement, e.g. amygdala, orbital-frontal cortex, prefrontal cortex, lateral prefrontal cortex, ventral striatum, hypothalamus and anterior cingulate cortex [6, 7, 41, 67, 68, 95, 96]. The positional overlaps within these regions observed with the latest imaging technologies indicate functional fusions of core affect and cognition, intracerebral communication and synchronisation for behavioural coordination. In contrast, modern clinical researches suggest that the formation of psychological disorders, e.g. trait anxiety, is consistently correlated with abnormal neural activity in some specific regions. Therefore, both direct and indirect biological activity measurements over human brain may be applied to the prediction of one's core affective and task performance against certain tasks.

### 2.2.2 Core Affective Assessment

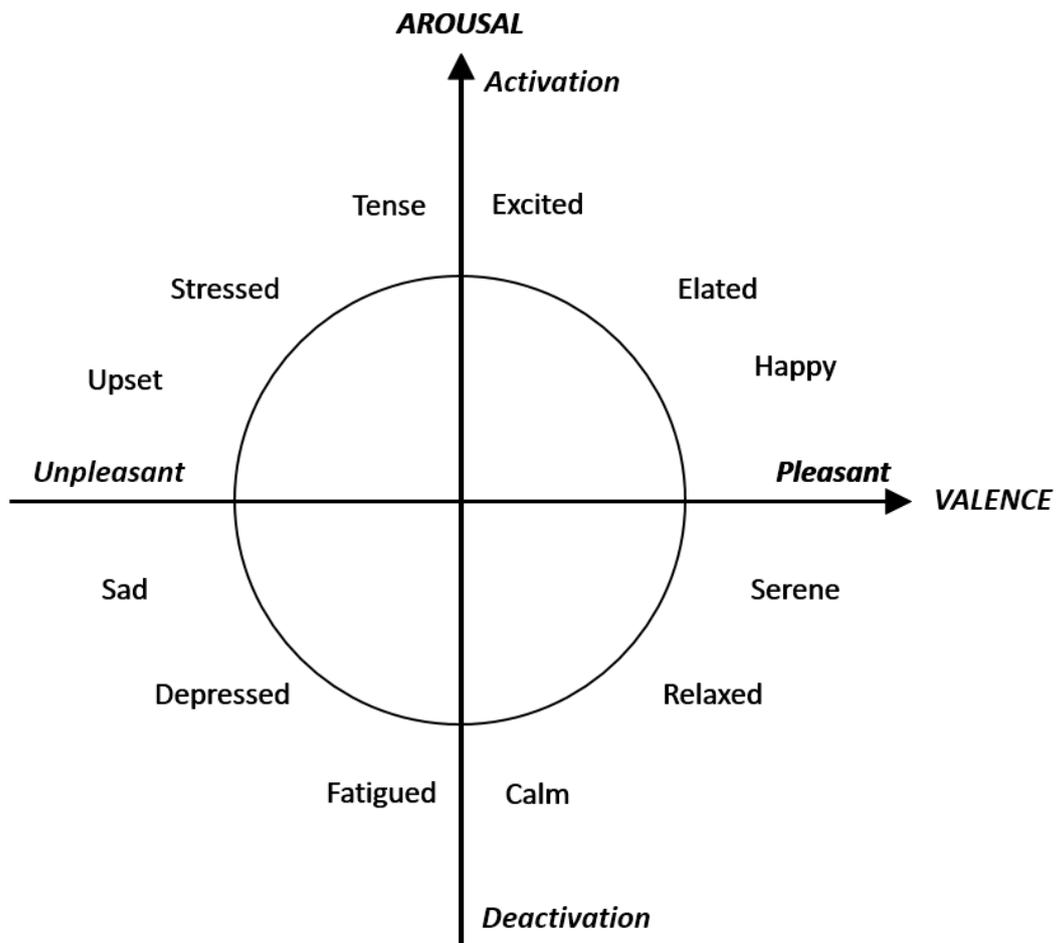
The direct measurement and assessment of human core affective is extremely difficulty and unnecessary, as it is a comprehensive integration and responsible for all inter and intra systems. Therefore, the researches were more focused on the psychophysiological state [3, 16, 53], which is the temporal segment from core affective that has been interpreted by consciousness. Compared to core affective, the psychophysiological state is recognisable and is experienced as some sort of

emotion or feeling that can be recalled by human. Moreover, it is a concise subjective unified description of both body and mind, and therefore, it may be directly expressed and interpreted by others.

There are two types of categorisation system applied in the current research literature - dimensional frameworks and discrete perspectives [3]. Both of them focus on the psychological experiences, for the physiological response varies with individual physical conditions, cultural customs and personal backgrounds. Discrete approaches were developed on the traditional understandings of psychophysiological state, which believe each state as being essentially linked with its own exclusive profile of experience, physiology and behaviour. For example, a psychophysiological state, such as mental stress and its relevant manifestations, is considered as a consequence of intrinsic mechanisms triggered by a specific range of stimulus only [53]. Dimensional approaches, however, consider a psychophysiological state as a collective representations of three fundamental factors: valence, arousal and tendency of approach-avoidance. Specifically, a psychophysiological state may be segregated within these three dimensions and continuously transferred into others as one or more parameter changes [16]. However, it is possible to reconcile two different theories comparatively by distributing various discrete state into different segregated areas of dimensional space, as shown in Figure 2.1.

It is worth noting that the emotional terms used in Figure 2.1 are the representatives of various psychophysiological states. Expressly, emotions prepare humans to respond to environmental stimulus and direct humans to meet the task demands, as they emphasise the evaluation of the human-environment relationship. Emotions are integrated with executive functions for behaviour control and influenced cognition from perception and execution [11, 41, 49, 68]. Meanwhile, the specific behavioural responses and physical conditions are conditionally initiated and motivational dependent. However, as is shown in Figure 2.1, the most psychophysiological states may be decomposed with two basic factors - valence and arousal level, regardless of act or response differences.

Humans communicate with machines through direct technical interactions, emotion based dialogue interactions and mental content interactions [87]. Due



**Figure 2.1:** Combination of Dimensional and Discrete Affective State (data based on [11, 41, 49, 68])

to current technological development and application limitation, HMI is mainly restricted to direct technical interactions, and consequently, compared to human-human interface, human operators' emotions expressed higher intensity and with fewer categories of negative feelings. Future development of HMI emphasise on functional innovative so as to empathetic to operator's emotional needs, physical conditions and mental abilities. Therefore, for the purpose of exploring and integrating the emotional and mental content within the current HMI, it is important to establish a measurement approach for valence and arousal level of human operators in HMI.

### 2.3 Psychophysiological State Measurement

Current research on psychophysiological state measurements focuses mainly on following five aspects - subjective experience, peripheral physiology, affect modulated startle, central physiology and behaviour [3, 8, 16, 24, 45, 53, 66]. The subjective self-report provided the most direct measurement since the psychophysiological state may only be experienced and interpreted by subjects themselves before any communication [3, 16]. Peripheral physiology consists of several involuntary physiological regulations including pupil reaction and circulatory system that are responsible for strengthened sensory representations and control structures [45, 49]. Central physiology monitors the biological changes of psychophysiological state over specific brain regions through electroencephalography (EEG), magnetic resonance imaging (MRI) or other bioimaging technologies, which provide an intrinsic description of task related cognitive process [8, 45]. The variation of the psychophysiological state is consistently correlated with the behavioural pattern change, as the feasible reaction plan was constantly confined to individual time-varying physical conditions. Therefore, the participants' inner states may be revealed via their pupil size, facial temperature or other physical behaviours to a certain degree [24, 66].

Valence indicates the tendency of humans towards a specific stimuli [11, 41, 49, 68], and within HMI it may be seen as the acceptance of operators towards a

given task from the machine side. Because it is a subjective evaluation from the operator himself or herself, the requirement of indirect observation is highly intricate, expensive and sometime even impossible. For example, to monitor from the central physiology prospective, it would be a complete observation of neural activities, such as long term memory recall and symbolic association, and some certain adaptive filtering mechanism maybe needed for irrelevant interference from both human self and environment. Therefore, the most accurate and effective assessment method is a self-report.

Arousal level represents the energy resources of humans invested in the response to a certain stimuli [11, 41, 49, 68], and within HMI it refers to the effort that operators showed to their given task. Different from the valence, the arousal level is able to be assessed with multiple indirect measurement methods beside subjective report. These methods are supported with sufficient physical and medical researches and cover a wide range of human organ systems. The most significant impact may be observed from respiratory system, circulatory system, muscular system and nervous system, and sometime lymphatic system and urinary system if more complex adjustment was required by further reactions. Such indirect assessments and measurements may be completely developed on the online monitoring of biological monitoring. They were capable of providing the most objective and normalised evaluation results without disturbance or interrupt of original HMI.

It is worth noting that, due to the complicity of interface, there is no single measurement or assessment method that may reveal the complete psychophysiological state of human operator in HMI. In practice, the subjective measurement alone may be compromised because of personal experience. Individual self-report entirely depends on the personal knowledge, experience and communication skill. Thus, one's actual psychophysiological state may be occasionally distorted and hard to be understood by others. These variables in between each individual perceptive level also determine an intrinsic issue of unifying or generalising different report. Similarly, there is no singular objective measurement or assessment can reveal the human psychophysiological state accurately alone. Organs and cells are

highly specialised functional units that work collectively. Hence, they may not be segregated for psychophysiological state estimation. The monitoring of a single system may also lead to severe bias, as each organ or system is always under multiple regulations and is responsible for various mechanisms. Therefore, the fusion of several measurements and assessments was necessary and should be incorporated in the design of any model or framework [53]. The combination of different measurements is the only approach to reduced the noise within each measurement method and generate comprehensive estimations from both psychological and physical levels. Yet, the increase of measurement types also leads to the increase of system redundancy and expenses. Therefore, the selection of a limited number of the most efficient biomarkers is required. From previous research, the psychophysiological measurements and assessments applied, most are based on electrocardiography (ECG), EEG, pupil size, heart rate, blood pressure, blood volume, blood volume pulse, respiration, muscle tension, electrodermal activity, galvanic skin response and temperature signals [23, 28, 52, 56, 59–61, 80–84, 97, 98]. These objective measurements are mainly focused on four aspects - peripheral physiology, affect modulated startle, central physiology and behaviour.

### 2.3.1 Peripheral Physiology

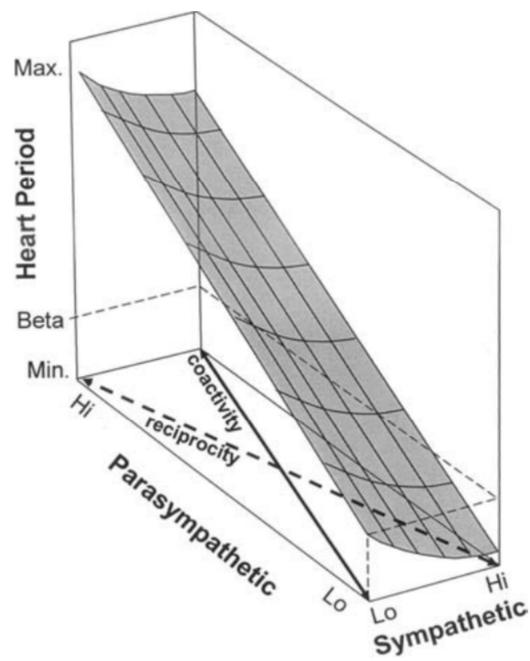
Peripheral physiology studies the electrical activities of nervous system outside the brain and spinal cord. Working as a bridge between the central nervous system and the limbs and organs, peripheral nervous system consists of peripheral nerves, neuromuscular junctions and even some the roots and branches from the cranial nerves and spinal nerves [21]. The part of the peripheral nervous system under the voluntary control of the brain is called somatic nervous system. It transmits signals from the brain to the end organs and created sensory and motion response, which form the foundation of muscle tension and self experience report, whereas, the rest peripheral nervous system creates a self-regulating system called autonomic nervous system. It includes the sympathetic nervous system, parasympathetic nervous system and enteric nervous system that are responsible for maintaining homeostasis homeodynamics. It largely acts on unconscious body

regulations, such as the pupil contraction, heart rate, respiration and perspiration [21]. Current psychophysiological research mainly focuses on the antagonism between the sympathetic nervous system and parasympathetic nervous system, specifically heart rate, for efficiency and convenience. In this project, the ECG is selected for assessing the psychophysiological state from peripheral nervous system.

### 2.3.2 Electrocardiogram

Electrocardiogram (ECG) observes the activities of the heart through the recording of varying electric potential difference. The heart is the core organ of circulatory system, which also includes blood vessels and the blood. The heart, functioned as the primary pump for the blood, has a deterministic impact on the sustainability and versatility of circulatory system [21]. The circulation of blood within the circulatory system starts from the heart and then goes through the lungs, arteries, veins, coronary and portal veins before the return. The whole system is crucial for the transportation of nutrients, oxygen, carbon dioxide, blood cells and hormones, and it also provides support for the immune system and maintained homeostasis through stabilising temperature and pH [21]. This system, with such importance, is directly under the control of somatic nervous system and is highly influenced by cognition process of central nervous system, for the preparation and persecution of responses. Figure 2.2 depicts a simplified model for the relationship between the heart rate and autonomic nervous system. As shown, the heart period and parasympathetic control presents an antagonism effect with the increasing intensity of sympathetic control. Therefore, the observation of heart activities provides a reliable and effective way to access conditions of these key organ systems and assess psychophysiological state [21].

ECG is a non-invasive way of monitoring the heart cardiac cycle by tracing the electrical activity within the heart [13]. Within a normal cardiac cycle, an electrical impulse travels through the sinoatrial node, the atrium, the atrioventricular node and the ventricular septum in a coordinated way. Consequently, the four chambers of heart contract and relax with a recognisable pattern. The neuron electrical sig-



**Figure 2.2:** The Relationship between Heart Rate and Autonomic Nervous System [13]

nals ensure the regulation for two main epochs, diastole and systole, of the heart cycle. Figure 2.3 presents a complete cardiac cycle with three main components - P wave, QRS complex and T wave. The P wave is normally a smooth rounded peak on ECG and corresponded to atrial contraction. The QRS complex consist of the Q wave, R wave and S wave, indicating ventricular contraction. The T wave is a asymmetric peak following the same polarity of the QRS wave, unless some sort of pathology persistent [21]. The most convenient approaches to calculate the heart rate through a ECG recording facilitate the R wave, because of the significant peak inflection. Based on the sequence method, the heart rate (bpm, i.e. beats per minute) may be calculated with dividing the interval distance between two R spikes with a constant value (typically 60,000). A complete heart period (milliseconds) is calculated with the same constant value vice versa [13, 59].

### 2.3.3 Heart Rate Variability

Heart rate variability (heart rate variability (HRV)) is considered to be connected with the respiratory cycle, the blood pressure and the heartbeat fluctuation, which is regulated by the peripheral nervous system and is sometime influenced by the central nervous system [2]. The HRV indicators  $HRV_1$  and  $HRV_2$  in this research, have already been applied in previous experiments and studies at the University of Sheffield [35, 60, 80, 81, 83, 84]. Further research studies proved the efficiency of using HRV in stress assessment [2, 9, 17, 27, 38, 47, 74, 86, 89]. The ECG signal may be divided into three frequency bands, which corresponds to variant biological processes [13]:

1. High frequency band (0.12 Hz–0.4 Hz or 0.15 Hz–0.4 Hz): this frequency band is used for respiratory sinus arrhythmia (RSA), as the heart rate increased and decreased with inspiration and expiration.
2. Medium frequency band (0.05 Hz–0.15 Hz or 0.8 Hz–0.12 Hz): this frequency band is mostly centred around 0.1 Hz and correlated with Mayer waves, brought by periodic arterial blood pressure changes. It is also applied in the measurement of mental workload.

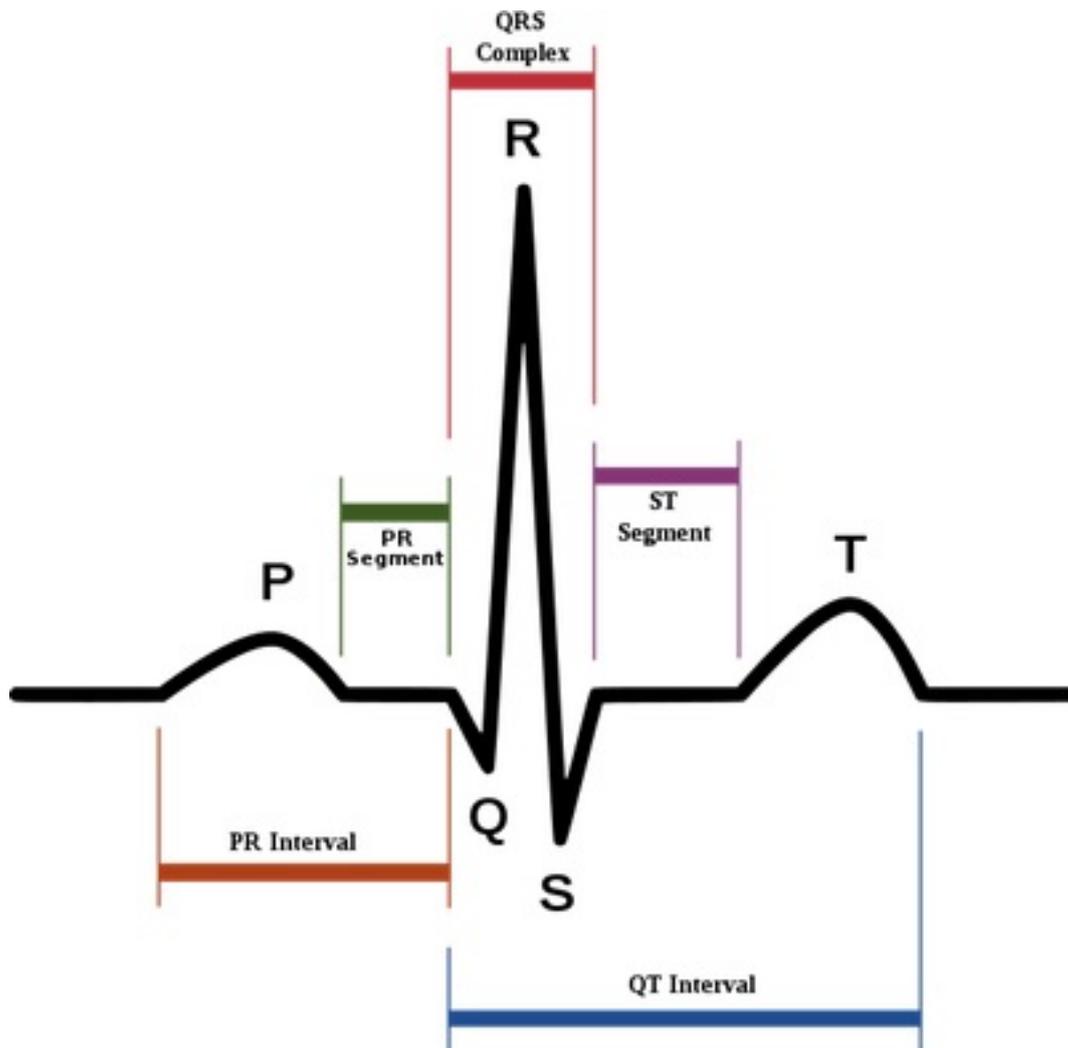


Figure 2.3: Example of ECG Signal [21]

3. Low frequency band (0.003 Hz–0.05 Hz or below 0.003 Hz): this frequency band is mainly related to the physiological control of autonomic nervous system. It has been applied for pathological studies, but it is rarely involved with psychological research [13].

The HRV has a profound correlation with the cardiovascular system. It was found that the change of heartbeat pattern due to any sinus arrhythmia (such as respiratory one) may be precisely represented by the variations in HRV [2]. The following research had further investigated the relationships between the HRV and sinus arrhythmias due to disease, stress, etc [9, 17, 27, 38, 74, 86, 89]. Stress, regardless of due to physical or mental workload, significantly influenced the change of HRV [47]. In some extreme cases such as job loss and death, a prolonged psychophysiological triggered effect may be identified with the HRV [69]. The same report also highlighted the use of HRV as a non-invasive measurement of stress triggered cardiovascular changes for those who lack efficient communication method, e.g. infants. Therefore, HRV has been highly recommended in the research of HMI and applied as an important indicator for human operators' psychophysiological state assessments [60, 61, 80, 81, 84]. However, it is worth noting that HRV responds to sinus arrhythmias regardless of the source of origin. That is, as long as the cardiovascular system changes, the intentional respiration may imitate sympathetic controls that are similar to the stress inducted results. According to the experimental results in [9], the stress related sympathetic control may predominate over the respiration and increase the low frequency component (altering R-R interval). Yet, when in the absence of stress, such low frequency component changes may be still observed with altering respiration alone. Therefore, it is worth noting that, within these HMI experiments, HRV is actually reflecting the workload related cardiovascular variance, prioritising cognitive response and overwriting autonomic regulation.

HRV was integrated in the mental and physical assessments together with other biomarkers in various HMI simulation experiments [35, 60, 80, 81, 83, 84]. Different simulation experiments required different standardisation of physical workload, thus, the ratio between potential mental stress and physical demands

varied. The HRV had succeeded in detecting these mental stress and workload related psychophysiological state changes in Stoop Colour Word testing and mental arithmetic operation [27, 38]. In these experiments, it was found that the HRV may be affected by the breathing control introduced by the verbalisation. Therefore, the combination of other biomarkers to differentiate psychophysiological state from other activities was important for the assessment. Apart from the HRV, task performance, heart rate, arterial blood pressure and hormone monitoring were also commonly used for estimating the psychophysiological state directly and indirectly. Meanwhile, it should be emphasised that the convenience and the reliability of HRV may not be replaced easily with other biomarkers. Compared to the ECG, the measurements based on EEG, electrooculogram (EOG) and electromyography (EMG) suffered more restrictions due to the equipment and operation requirements in a real-time airplane environment [89]. It was suggested that the integration of the HRV and the Bedford scale was able to assess the psychophysiological state real-time for some special conditions, such as during commercial flights.

For this research, the two selected HRV indicators follow the previous experiment work at the University of Sheffield [59–61, 80–84]. The raw ECG signal is filtered with a low-pass filter with a cut-off frequency 5 Hz, and the time-stamps are selected through R peak detection.  $HRV_1$  represented the 0.1 Hz component of the ECG signal, and it is measured by averaging the power spectrum of frequency components from 0.07 Hz–0.14 Hz in a time period of 30 seconds.  $HRV_2$  is the ratio between the standard deviation and the mean value of the heart rate signal in the same time frame. The ECG applied in this research follows the three electrode system from ActiveTwo Biosemi<sup>®</sup>, as shown in Figure 2.4. The LL is located at the end of the xiphoid process slightly towards to the left hand side. The LA is located at the end of the left false rib no. 10. The RA is located below the sternal angle and slightly towards to the right hand side. The three electrodes form a triangle over the heart area.

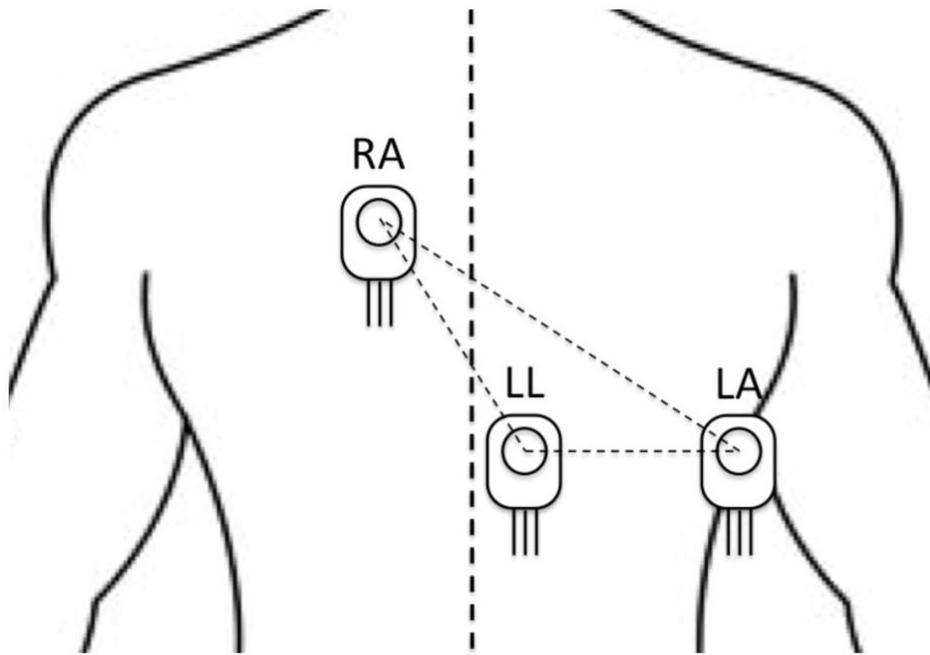


Figure 2.4: ECG Electrodes [1]

### 2.3.4 Central Physiology

Central physiology focuses on the physical activities of central nervous system, which consists of the brain and spinal cord [21]. The whole system controls the most functions of the body and mind. The brain is the center of the central nervous system that is responsible for sensory information processing, voluntary movement, cognition, perception, self-awareness and emotion. The spinal cord provides a highway for the information communication between the brain and the rest of human body. The neurons constitute the most of central nervous system and communicate with each other with electric pulses and neurotransmitters for processing information. Compared to the other human organ systems, the central nervous system processes multiple tasks for variant targets coherently. Due to such importance of central nervous system, it is securely protected with the bones of the skull and spinal column and with the syrinx, a cavity filled with fluid in between for shock absorbance [21].

The complexity of the functional configuration and the enhanced double pro-

tection make the direct measurements and observations of a specific stimuli-triggered response in the central nervous system extremely expensive and complicated. Therefore, current central physiological studies focus on the collective activities at the regional level and are mainly based on non-invasive measurements: EEG and neuroimaging, such as functional magnetic resource imaging (fMRI) and position emission tomography (PET) [8, 11, 32, 41, 45, 49, 51, 53, 54, 68]. Similarly to the ECG, the EEG measures the electrical activities over the skull as the neurons communicate with electrical pulse, whereas the neuroimaging is commonly monitoring the activities of brain regions through the blood flow. Compared to the EEG, and despite the neuroimaging may provide in-depth dimensional perspective on brain activities, it requires expensive and complicated equipment as well as corresponding environment. Meanwhile, EEG is frequently applied in research on the psychophysiological state for its low operational cost.

### 2.3.5 Electroencephalogram

Electroencephalogram (EEG) was initially applied in 1924 by a German psychiatrist, Hans Berger [13]. He identified the electrical activity of neurons at static relaxed awake condition between the forehead and occipital lobe - a rhythmic oscillating wave around 10 Hz. This proved a strong support to link the periodic changes of EEG recordings to different psychophysiological state of humans [13]. Temporal resolution of EEG was remarkable with real-time recordings of millisecond resolution, due to the transient existence of electrical pulse. However, the brain is spherical or elliptical object protected with the skull bone and filled with the syrix, consequently, the volume conductivity of it is not evenly distributed [21]. This inevitably led to the positional distortion of the electromagnetic field recorded by the EEG and low signal to noise ratio. However, this may be solved with improving sampling limitation, i.e. increasing number of electrodes used. Despite these defaults, as a summation of neurons collective activation and deactivation. EEG was able to present clear numerical comparisons between different cerebral cortex regions, e.g. anterior verse posterior or right-hemisphere verse left-hemisphere [13].

According to the frequency analyses, the EEG recordings may be decomposed into five frequency band - delta, theta, alpha, beta and gamma:

- Delta band - It consists of the frequency components of EEG up to 4 Hz. It is mostly identified in adults' slow-wave sleep, associated with deep sleep phase without dreaming. It is also commonly found in children, especially with infants, and together with theta band components slowly diminish with age; whereas, alpha and beta frequency band activities gradually predominate throughout human life [13]. Due to the limitation of age and the inhibition effect, delta band is rarely involved in research of psychophysiological state in working conditions.
- Theta band - It represents the frequency components within the range of 4 Hz–8 Hz on the EEG recordings. It is mainly found in children but also may be observed in adults with emotional stress, such as frustration and disappointment. According to different observation locations, the theta band frequency wave in adults may be correlated with drowsiness or arousal, relaxation or mediation. However, excessive amount of theta frequency band activities for a certain age may indicate abnormal activities, such as decreased vigilance level or impaired information processing. The theta frequency band activity distributes along with the frontal midline being associated with focus and attention, as the anterior cingulate cortex is responsible for the information flow of the paleomammalian cortex [13]. Therefore, it is one of the focused frequency bands in the researches of HMI.
- Alpha band - It includes the frequency components of EEG recordings within the range of 8 Hz–13 Hz. It normally emerges during the state of relaxation or wide wake with eye closed. It is the first observed rhythmic EEG wave by Hans Berger, and it is widely spread over the posterior cortex (i.e. the parietal lobe and the occipital lobe), elevated with dominant side. However, it may also be observed from other regions, such as the temporal lobe and the frontal lobe. Such temporary emergence, under the normal conditions, is mostly associated with the attenuation of physical and mental exertion. The

studies estimates that it is transformed from the theta band rhythmic wave during the puberty, for the frequency for the posterior basic rhythm is commonly lower than 8 Hz [13]. However, since the alpha rhythmic wave rapidly diminished as the eyes opened, it is not engaged within the researches of psychophysiological state in HMI.

- Beta band - It is the collection of the rhythmic waves between the frequency range of 13 Hz–30 Hz in the EEG recordings. It is found to be consistently correspond with cognitive processes like information processing, motor behaviour and attenuation of actions. This frequency band may be further divided into three ranges - the low range (13 Hz–16 Hz), the middle range (16.5 Hz–20 Hz) and the high range (20.5 Hz–28 Hz). The rhythmic wave within the low range emerges with a focus state under relaxed conditions. The middle range rhythm is closely associated with processing sensory information. The frequency component of the high range is normally observed with arousal or anxious state. It is the dominant rhythm distributed symmetrically on the both side of frontal lobe in the alert or concentrated condition. Because of such correlation with vigilance and arousal level, the beta band rhythm is the most focused in the most psychophysiological state studies. However, it is also sensitive and vulnerable to the effect of drug use, pathogens and cortical damage [13].
- Gamma band - It is the summary for the rhythmic waves with frequency over 30 Hz and up to 100 Hz (mostly around 40 Hz) in the EEG recordings. It is observed with cognitive processes associated with valence, imagination and integration of different regions. The oscillation within the gamma band emerges during networking variance simulators for a certain cognitive target [13]. It is usually applied in the researches of regulation for high level cognition or complex motor function.

The method of using EEG recordings is highly depended on the purpose of research, and for the researches of human psychophysiological state the most recognised EEG based indicators are spectral power on a specific frequency range,

event-related potentials and event-related desynchronisation or synchronisation [8, 32, 45, 49, 51, 54]. The spectral energy is mostly aimed for a general estimation of human state for a short period, while the event-related measurements were designed for monitoring consequential response of an instant stimulus. Among all the developed indicators so far, task load index (TLI) is one of frequently used indicators for the psychophysiological researches, which is based on the spectral energy analyses of theta band rhythm and alpha band rhythm [28, 75].

### 2.3.6 Task Load Index

Task load index TLI was initially developed from the EEG recordings by Gevins in 1997 with his team for the research of mental workload [28, 29, 75]. As mentioned earlier, the low frequency wave of the theta band is linked with arousal or stress level as it represents the synchronisation of neurons, and the alpha band rhythmic wave emerges with the state of increased focus and concentration. Based on these two frequency bands, the TLI is able to assess human working memory (WM), which was one's cognitive ability to hold information of a focused event temporarily against continuous disturbances of others [28, 75]. The WM is considered to be important for psychological reasoning, decision making and physical behaviour. In this case, the abstract conceptual WM is mathematically represented by the direct measurement of EEG, which monitored the neural resources involved in the process of solving task. This method ensures a certain level of robustness and stability for the indicator against interference, as the raw EEG recording may be severely influenced by artifacts and voluntary or involuntary moves irrelevant to the researches. In addition, some degree of susceptibility and flexibility introduced by the time-varying neural collective cooperation provides a temporal resolution up to 0.25 seconds, which should sufficed the requirement of most HMI psychophysiological researches [75].

It was found that the decrease of task related WM due to reduced task load, fatigue or psychophysiological breakdown was usually accompanied with the loss of focus and the decrease of performance, e.g. increased reaction time and error rate. This phenomenon always emerged with a degraded theta band activity of

the frontal lobe and elevated alpha band activity of the whole brain [28, 29, 75]. Therefore, in this research, the two selected TLI were  $TLL_1$  and  $TLL_2$ , based on the comparison between the theta band rhythm and the alpha band rhythm, and they were calculated with the following equations:

$$\begin{aligned} TLL_1 &= \frac{P_{\theta, F_z}}{P_{\alpha, P_z}}, \\ TLL_2 &= \frac{P_{\theta, AF_z}}{P_{\alpha, CP_z, PO_z}}, \end{aligned} \quad (2.1)$$

where  $P_{\theta}$  and  $P_{\alpha}$  are the spectrum energy of the theta band rhythmic wave 4 Hz–7.5 Hz and the alpha band rhythmic wave 8 Hz–12.5 Hz. The spectrum energy is calculated by averaging the intensity of rhythmic wave within that specific frequency range over 30 seconds. The electrodes of  $F_z$ ,  $P_z$ ,  $AF_z$  and the combination of  $CP_z$  and  $PO_z$  followed the Biosemi 10/20 system, as is showed in Figure 2.5. Previous research on human performance had confirmed the high efficiency and reliability of using these two TLI indicators for assessing human psychophysiological state in the HMI experiments [60, 80, 81, 83, 84].

### 2.3.7 Affect Modulated Startle

Affect modulated startle measures the magnitude of startle response, as it reveals the valence of human, especially at excited state. The startle response is a general reflex towards to any immediate emerged stimuli that consists of a set of voluntary and involuntary motor actions, such as temporal muscle tension and attention diversion [53]. The whole system functions as a basic protective mechanism for organs and livings creatures against any potential harm. All startle responses are established from current behavioural interruption and vigilance facilitation for a possible threat, which fundamentally depends on the eye movement [53]. Research based on the electromyogram recordings of eye muscle activity had showed that the startle response was effective in assessing human psychophysiological state, particularly for the valence level, as the stimulus triggered approach-avoidance mechanism encouraged or inhibited startle responses, compared to the magnitude of neutral state [53]. However, the muscle contraction and

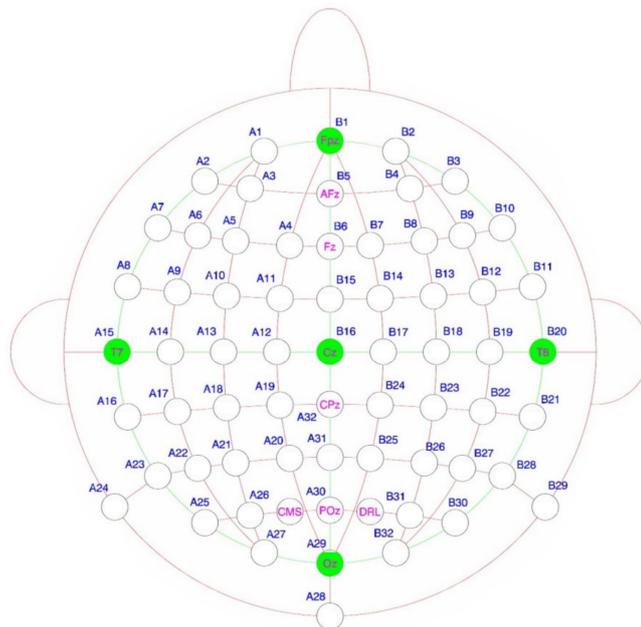


Figure 2.5: EEG Electrodes [1]

relaxation only indicate instant valence state of each individual stimuli, because the intensity of the startle response dropped right after diverting attention. Therefore, the pupil diameter marker (PDM) is applied for measuring human periodic psychophysiological state in the HMI experiment.

### 2.3.8 Pupil Size

The human pupil is the dark circular tissue in the center of the iris of the eye, which allows light to strike the retina through variation of size. Such regulation adjusts the amount of light passing through so that a clear image may be formed on the retina and provide enough details for cognitive information processing. It serves as an adaptive reflex to the environmental stimulus, for the purpose of vigilance control and focus regulation of human. For this reason, the contraction and dilation of pupils are under the control of autonomic nervous system as part of psychophysiological response. The regulation of pupil diameter is controlled by sets of opposite muscles inside the iris, the iris dilator muscles, and guided by the antagonism between the sympathetic and parasympathetic neural branches. Depended on the stimulus and environmental brightness, the pupil diameter may vary from 1.5 mm to 9 mm [21].

The use of pupil size was explored in the research of affective state in the HMI configuration related experiments [42, 57, 58, 72, 82, 91, 97, 98]. In these experiments, the pupil size succeeded in differentiating the stressed and relaxed state of participants during their interfaces with personal computers. The experiment suggested using the combination of the pupil diameter and eye blink rate as the indicator for affective state measurement, such as mental stress in the HMI. Meanwhile, the rate and acceleration of pupil size variation were also experimentally proved to be effective for the psychophysiological state measurement. The efficiency of using pupil size was validated through comparisons with conventional psychophysiological biomarkers in the HMI configuration, including EEG, photoplethysmogram, skin conductivity, blood volume pulse and surface temperature. Additionally, compared to the most traditional psychophysiological biomarkers, the non-intrusive and distanced measurement approach of pupil size minimised

the equipment and environment requirement. Thus, the noise introduced by the psychophysiological state monitoring was reduced during the HMI experiment.

### 2.3.9 Pupil Diameter Marker

The pupil size biomarker used in this project was pupil diameter marker (PDM), which was experimentally tested and applied in the previous HMI researches in University of Sheffield [82–84]. The pupil diameter was recorded in real-time through an eye tracker camera provided by Gazepoint™. Based on the pixel calculation of pupils, in addition to a relative distance measurement and a head movement scale factor, the Gazepoint™ software was able to provide accurate pupil diameters from captured images, via the following equations:

$$\begin{aligned}
 PDM_L &= (LPD) \cdot (LPSF) \cdot (LVF), \\
 PDM_R &= (RPD) \cdot (RPSF) \cdot (RVF), \\
 PDM &= \sum_{i=1}^N \frac{PDM_{L_i} + PDM_{R_i}}{2N},
 \end{aligned} \tag{2.2}$$

where  $PDM_L$  and  $PDM_R$  represent the left and right pupil diameter markers respectively. Similarly,  $LPD$  and  $RPD$  are the left and right pupil diameters in pixels, and  $LPSF$  and  $RPSF$  stand for two pupil scale factors, with  $N$  for the sample number (depend on time frame).  $LVF$  and  $RVF$  are pupil valid flags for left and right pupil respectively, which serve as a filter for blinking and other invalid measurements during the experiment.

### 2.3.10 Behaviour

As mentioned earlier, each psychophysiological state is always involved with a certain set of physical responses, including motor action and organ system variation [11, 41, 49, 68]. Therefore, it should be possible to assess one's psychophysiological state from these behavioural changes, such as vocal characteristics, facial expressions and whole body behaviours [53].

Vocal characteristics infers one's psychophysiological state mainly through voice amplitude and vocal pitch. It was proved that these vocal characteristics

were more connected with arousal level than valence level [53]. However, due to the limitation of current HMI system, verbal communication was rarely applied and therefore vocal characteristics were less focused. Facial expressions, however, have been applied in the researches of assessing human psychophysiological state in the HMI for a while [26, 53, 100]. Facial behaviours appeared to be a reliable and non-intrusive way of reflecting the participants' valence level in real-time experiment. However, the efficiency of facial expression in assessing human psychophysiological state was excessively determined by the coding system, which differentiated variance facial behaviours with muscle assessments. Furthermore, as the core affective is also under the influence of individual experience [41], it may be precarious to assume that facial expression may be a straightforward read-out of a person's psychophysiological state. This intrinsic default of facial behaviour explains most contradictions between the estimation and validation in the experiments. Because of it, facial expression is often applied together with other psychophysiological state measurements, e.g. head movement, eye gaze and blood pressure etc. The whole-body behaviour mostly focuses on communication behaviour expressed bodily. Some studies showed that there may be some distinctive bodily behaviour for some specific emotional state [53]. Yet, the observation of such a behaviour or expression is also influenced by individual difference and might be redundant, as localised electromyogram is able to capture more subtle muscle regulation and provide higher accuracy of psychophysiological prediction while minimised intrusive physical and mental interference [53].

Although the efficiency of behavioural biomarkers are lower than the other type of conventional biomarkers because of individual difference, the non-intrusive approach of behavioural monitoring provided an irreplaceable way of measuring. Nowadays, with the development of thermal imaging, a new method of human psychophysiological state measurement is explored in this research. Apart from the regulation of muscles, the blood also provides support for the physical behaviours and the flow of it may be estimated with human surface heat [21]. Therefore, facial temperature is suggested here to provide an non-intrusive measurement of psychophysiological state.

### 2.3.11 Facial Temperature

The blood flows and the blood vessel sizes are constantly regulated by the hypothalamus and central nervous system, which are also responsible for the psychophysiological state, such as memory coding and emotion processing [16, 39]. Previous studies found that the psychophysiological state may directly impact the human thermoregulation and lead to perceptible temperature changes on the surface skin [12, 14]. The facial behaviour is relatively more sensitive to the variation of psychophysiological state than the rest part of human body. These facial expressions depend on the muscle movements and are supported by the blood, therefore, the above research concentrated on the temperature information captured from the forehead, the periorbital and the nasal regions of the participants. Compared to the rest of face, the soft tissue layer within the forehead region is relatively thinner because of the skull, the temperature over that region was more susceptible to the blood flow in the capillaries, which is regulated by the cardiovascular system [21]. Meanwhile, the blood pressure within the periorbital area followed periodic rhythms associate with the eye ball movement and the muscle activity, for the purpose of switching focused object and regulating vigilance for the surrounding [21].

The temperature of the nasal region indicate a combination of the blood flow from the cardiovascular system and the air flow from the respiratory system, which are all controlled by the autonomic nervous system and are influenced by the central nervous system [21]. Among the rest part of human face, these three regions are the most sensitive to the variation of human psychophysiological state. Moreover, the HMI experiment identified that the indicator based on the spectrum analyses of facial temperature components, the differential energy between philtrum and forehead (DEFP), consistently correspond to the heartbeat rate and the concentration of cortisol in blood [39]. All these studies suggest a significant correlation between the facial temperature and human arousal level.

Therefore, in this project, the psychophysiological biomarkers developed based on the facial temperature are the mean forehead temperature  $\bar{T}_f$ , the maximum facial temperature  $T_{maxf}$ , the mean nasal temperature  $\bar{T}_n$  and the DEFP. The ef-

fectiveness and the efficiency of these facial temperature biomarkers in the HMI experiment are validated in [34, 35]. The thermal recordings are acquired with a FLIR E40bx thermal camera, with a distance of 0.5 meter between the lens and the participant's face. Because the human tissue is heterogeneous, the emissivity is set to be 0.98, following the guideline of emissivity table. The sampling frequency of the thermal camera is set to be 10 Hz and the values of the biomarkers are the averages in a 15-second period window for reducing the effect of auto-calibration. The environment temperature of the lab room is remained around 20 °C and the relative humidity is remained around 50 %.

## 2.4 Human-Machine Interface Simulation

The primary goal of the HMI simulation experiment is to challenge the participants' task solving ability in a similar way to the real world HMI situations. Different from the human-human interfaces, the human participants from the HMI demonstrated lower emotional intensity over less diverse feelings [87], due to lack of emotional content communication in the current HMI. Therefore, the simulation of HMI should be focused on the approaches of introducing adequate adjustable stimuli through workload. Different observable psychophysiological states triggered from these stimuli should resemble human psychophysiological reactions in real-life situations [87].

There were a lot of assessments used in the psychophysiological state researches, such as air traffic control, university test, memory test, cold presser test and coin stacking [23, 28, 56, 97, 98]. However, these tests were considered as being not qualified for the simulation of HMI. Most of them lacked reproducibility or were mixed with human-human interface. In addition, there was no controlled stimulus for variant psychophysiological states, and the performance evaluation was based on subjective experience. These assessments were designed for the experiments aiming at specified transient reactions rather than continuing interface.

The most frequently used HMI simulations in research are the Stroop colour-word interference, the mental arithmetic and the virtual vehicle operation, the

management system simulation [9, 38, 47, 56, 76, 84, 89]. According to the requirement of research, a desirable HMI simulation system should fulfil the following requirements as much as possible:

1. **Difficulty** - the given task should be challenging enough to introduce adjustable psychophysiological workload for observable behavioural changes in body and mind.
2. **Similarity** - the expressed psychophysiological state should cope with existing recognised human states in the real life HMI.
3. **Reliability** - the simulation mechanism should be verified to ensure an explainable, controllable and reproducible experiment.
4. **Simplicity** - the configuration should be clear and direct to minimise the interference of other activities besides the HMI simulation.
5. **Generalisability** - the related task-solving skills should be intuitive and universal to avoid performance difference due to personal experience and ability.

Besides all the requirements above, the desirable HMI simulation should also be efficient in both time and expense. For the time efficiency, it requires a complete simulation process maintain the highest ratio between the interface period and the familiarisation phase. For the expense efficiency, the requirement of the lab room and equipment should be simple and accessible for the most conditions. It is worth noting that the selection of psychophysiological biomarkers should be thoroughly considered. Not only these biomarkers should be convenient to use and available for general cases, but they also should be adaptive to the variance real-time HMI scenarios. Furthermore, a certain degree of robustness should be included to circumvent possible noise within the HMI simulation, and for known and unknown interference existing in the actual HMI.

### 2.4.1 Stroop Colour-Word Interference

The Stroop colour-word interference was created by Stroop and had been commonly applied in psychophysiological research since 1935 [78]. Generally, participants performing the Stroop colour-word interference were required to link the given colours to their corresponding names through related graphic user interface systems on computers. This interference was designed for introducing certain amount of stress on the participants psychologically through limiting the time for each response. A Stroop colour-word interference normally consisted of two phases that were different in presenting the task - non-conflict and conflict. The colour of the letters in a word was congruent with the colour of that words represented in the non-conflict configuration (e.g. the word "green" coloured in green), whereas for the conflict configuration the colour of a word was not consistent with the meaning of that word (e.g. the word "red" coloured in blue). As the experiments proved, compared to the non-conflict configuration with congruous stimuli, the incongruous stimuli of conflict configuration increased reaction time and cognitive error rate [78].

The Stroop colour-word interference has now been applied in many research studies of the human psychophysiological state [27, 38, 56–58, 98]. However, the efficiency of using Stroop coloured-word interference as a reliable HMI simulation was questioned in some studies lately. It was found that the HRV, a traditional and commonly used biomarker for human psychophysiological state mentioned earlier, failed to show significant change during the Stroop colour-word interference [27]. Furthermore, the recordings of biomarkers suggested that there was less observable human psychophysiological state changes during the Stroop colour-word interference, compared to the other HMI simulation such as mental arithmetic experiment [38, 56]. This suggested that the Stroop colour-word interference may not be able to challenge the participants with proper psychophysiological workload and satisfied the requirement for a significant psychophysiological state change. In some other studies, the recorded psychophysiological changes during the Stroop colour-word interference were proved to be more correlated with human natural fundamental regulation for external environmental adaptation and

motor action rather than cognitive activities [9, 76]. According to these experiments, the Stroop colour-word interference could provide contradict results with different environment conditions. Therefore, it should not be considered as a suitable option for the HMI simulation, as it failed to minimise these irrelevant effects.

### 2.4.2 Virtual Vehicle Operation

Under laboratory conditions, the virtual vehicle operation bore the most resemblance to the nowadays HMI systems in the real world, which induced the psychophysiological changes in a similar way through approximately identical interface configurations [26]. The simulated transport vehicles normally included automobiles, airplane and even space shuttle sometime, and the participants were required to operate these virtual vehicle for certain goals in space or time while against some predetermined interference. Virtual vehicle operation was widely applied for the simulation and training of real world HMI systems at present, especially in the human-centered system where human experience and judgment were inherent factors (e.g. captain or pilot).

The virtual vehicle operation had been applied in several research studies of the human psychophysiological state in the HMI, such as in the examples of [26, 36, 47, 89]. For these studies, the virtual vehicle operation was commonly based on automobile driving and the assessment of performance focused on the simulated side slip angle, yawing acceleration and rolling velocity. The human psychophysiological state, such as stress or focus, was normally measured with facial expression, HRV, skin conductivity, respiratory rate, electromyogram and cardiovascular indices. However, the use of virtual vehicle operation in the HMI research studies was severely restricted with its complexity and corresponding high expense, and most importantly, these experimental results lacked reproducibility. It is worth knowing that reproducibility was the only way to guarantee the statistically significant results of a study. Particularly for these research studies comparing the experiment results between the simulation and actual in-car driving, there led to a significant amount of uncertainties within the experiments as many variables were not completely under control, such as road conditions,

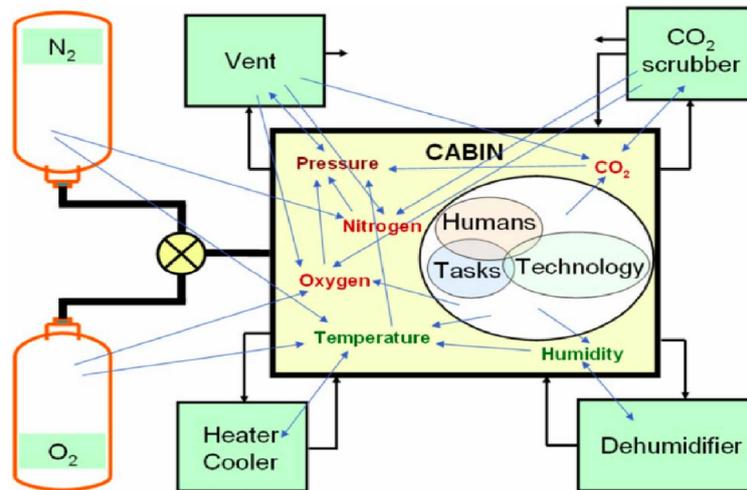
weather conditions and automobile conditions. Meanwhile, the complexity of indoor simulation not only required a relatively large space maintained under controlled conditions, but the expensive equipment also demanded complicated setup that inevitably led to interference to each run. It was experimentally proved that the environmental conditions and maintenance of the system could significantly impact the effectiveness of each simulation [26, 100].

Furthermore, vehicle operation was a common learning process in real life, as the participant could adapt to it through adaptation during each practice and develop individual behaviour from past experience. The performance of the inexperienced participants might be more linked with the proficiency rather than the psychophysiological state, whereas for the experienced ones the experiment might need to last long enough or specially programmed to observe any significant psychophysiological state change. Besides, the personal preference established from daily life also influenced the performance assessment, for people valued differently on different road conditions (e.g. the reduced speed at a curve). The high complexity of the system configuration, the high cost of the equipment and the high uncertainty of the system operation deterred the application of virtual vehicle operation in the human psychophysiological research of the HMI.

### 2.4.3 Management System Simulation

The management system simulation usually requires the participant to maintain a simulated environment within a desired state through adjusting different parameters, i.e. stabilise an unbalanced system simulation with limited inputs. One example of such simulations is the automation enhanced cabin air management systems (aCAMS), which simulates the remote supervision and control of air conditions in a space cabin. This software has been applied in several researches of human psychophysiological state in the HMI [52, 59–61, 80, 81]. A dedicated graphical user interface (GUI) ensures that the whole simulation is done with computers restricted in the lab room and minimised the uncertainty from uncontrolled factors. Meanwhile, such software based configuration guarantees the reproducibility of simulation experiment and the validity of results. Despite of

several advantages, a relatively long phase of familiarisation is necessary and essential for operating aCAMS for it is not intuitive. Similar to the virtual vehicle operation, a certain level of understanding on some key features and functions and a general knowledge about fundamental processes are needed.



**Figure 2.6:** Model of aCAMS [80, 81]

As shown in Figure 2.6, an operator of aCAMS is required to maintain cabin air condition within the breathable range according to the information of interactions. In these management system simulations used in the HMI similar to aCAMS, the participant is normally paired with an automatic controller and also responsible for monitoring the performance of it. The different psychophysiological state may be triggered with the level of automation, such as switching a certain number of subsystems into manual control or tuning different frequency for automatic control intervention. A number of potential failure needed to be anticipated by the operator and the performance is normally assessed based on the time of sustaining simulation within designed range. The experiment results indicated that the change of manual control variables was significantly associated with hu-

man psychophysiological state such as mental entanglement or stress. This sort of induction bore numerous resemblance to the observations from real life human-computer interface, e.g. human operators in modern powder grid regulation.

However, as Figure 2.6 shows, the management system simulation also suffers from the problem of high complexity identical to the virtual vehicle operations. The requirement of proficiency decreased the time efficiency of HMI simulation. Moreover, the specified training for system operation of most management system simulations is not applicable and referable to others, thus increases the costs of the experiments.

#### 2.4.4 Mental Arithmetic Experiment

The mental arithmetic experiment is based on mental arithmetic operation, including elementary addition, subtraction and multiplication. The difficulty of task is fully controllable with the number of operands' digits and the respond time for completing each calculation. Mental arithmetic calculation is intuitive and direct for most people with basic algebra knowledge, and the computational requirement for programming and implementation is relatively low compared to the others. It offers a high efficient way for the HMI simulation, and due to its simplicity the experiment is highly reproducible.

Though the mental arithmetic experiment differed from the previous conventional approaches for the HMI simulation, it has been applied in many human psychophysiological state researches [27, 38, 46, 56, 82–84]. Within these experiments, it was observed that, during the mental arithmetic tests, the task difficulty was significantly linked with the changes in human respiratory system, cognitive workload and cardiovascular system. Compared to the other HMI simulation approaches, the mental arithmetic experiment was selected for the human psychophysiological state research in this project for:

1. Effectiveness: As mention in the previous subsection, the psychophysiological state variances due to the stimuli of the Stroop colour-word interference were less intensive and thus could be interfered with human intrinsic psy-

chophysiological process [9, 38, 56, 76].

2. Robustness: Different from the mental arithmetic experiment, the effectiveness of the virtual vehicle operation was significantly depended on the lab environment and the maintenance [26, 100].
3. Simplicity: Compared to the mental arithmetic test, the efficiency of using the management system simulation was strictly limited with the complex training phase [80, 84].
4. Intuitiveness: The simple knowledge requirement of mental arithmetic operation made it suitable for everyone, regardless of the most inter- and intra-individual differences such as age or gender.

#### 2.4.5 Experimental Configuration

In this research, the mental arithmetic test was selected for the HMI simulation and human psychophysiological state during the experiments was assessed with EEG, ECG, pupil diameter and facial temperatures. The mental arithmetic test was based on a GUI software, which was similar to the one applied in the previous human psychophysiological researches at the University of Sheffield [82–84]. This system had succeeded previously in providing participants controllable workload through a simple and reproducible way. Meanwhile, as mentioned earlier, EEG and ECG were recorded and processed with the hardware and corresponding software from ActiveTwo Biosemi<sup>®</sup>, and a Gazepoint<sup>™</sup> GP3 eye tracker was applied for the pupil diameter measurement. The experiment configuration used for the acquisition of these psychophysiological data followed the previous examples in [52, 60, 61, 80–84]. Facial temperature, as a new psychophysiological measurement approach, was recorded with a FLIR E40bx thermal camera and analysed with Matlab SDKs based system.

The experiment relied on two computers with TCP/IP communication. The first computer was for the participant to perform the mental arithmetic test with a Matlab based GUI system. While, the second computer ran all the software for data acquisition, including the ActiView LabVIEW, Gazepoint<sup>™</sup> Control and

Matlab based thermal imaging system. The exchange of data across each software and between two computers depended on TCP/IP communication. Such hardware setup separated the data acquisition from the participant operation, which minimised intervention to human psychophysiological state in real-time experiments. Furthermore, it reduced the computational requirement and ensured the smoothness of the whole system during running, especially for thermal image processing as thermal imaging live streaming normally took the most of computer drive.

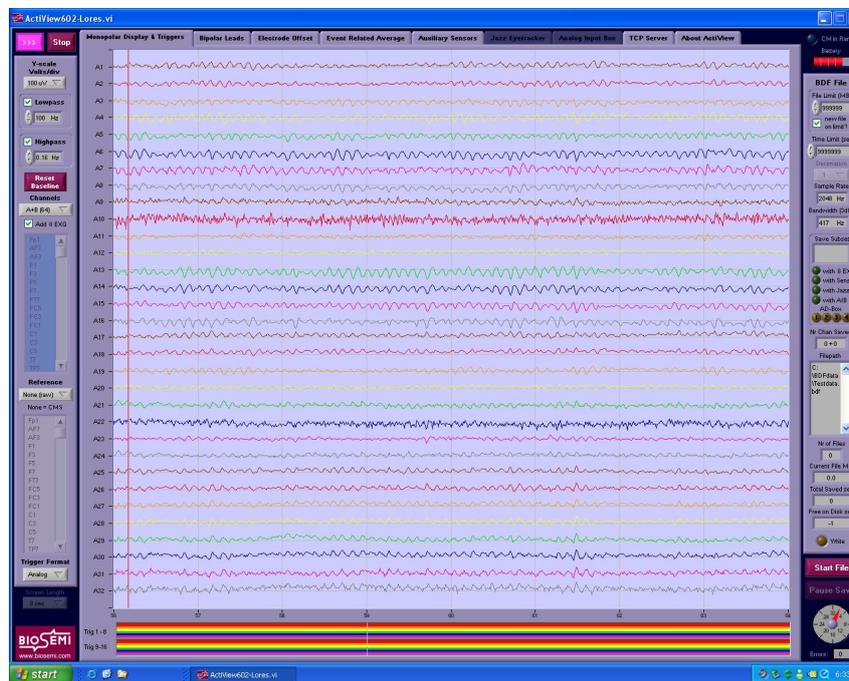
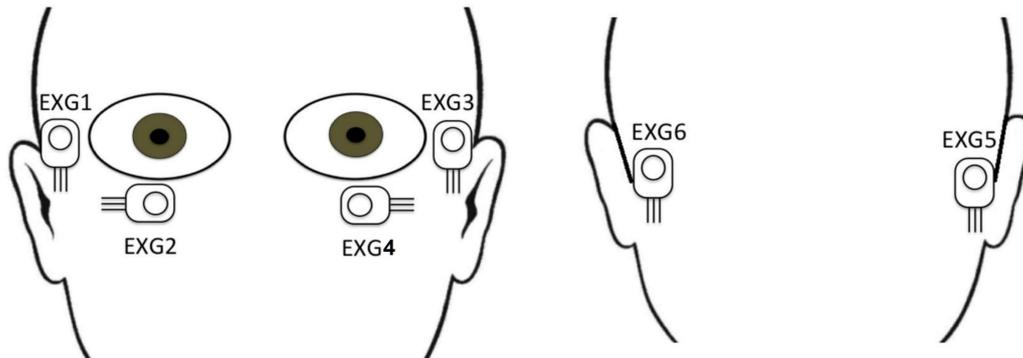


Figure 2.7: Image of ActiView Software by Biosemi [1]

Figure 2.7 shows the ActiView LabView Acquisition software provided by Biosemi<sup>®</sup>. The raw ECG and EEG signals are acquired with a frequency of 2048 Hz, and after the filtering and processing of ActiView software the signals' frequency was reduced to 256 Hz for the convenience of communication. In addition, the EOG of participant is also recorded at the same time for a lateral checkup, as the movement of human eye and eyelid could impact significantly on the recording of EEG signals. Meanwhile, the recordings of EOG may also be applied for future investigations of the relationship between human facial expression and psy-

chophysiological state. Figure 2.8 indicates the positions selected for the EOG measurement in this research. The electrodes *EXG1* and *EXG3* are for the horizontal channels, and the electrodes *EXG2* and *EXG4* form the vertical channels. The reference level consists of signals from the electrodes *EXG5* and *EXG6*, which are placed on the bone behind the ears with less soft tissue.

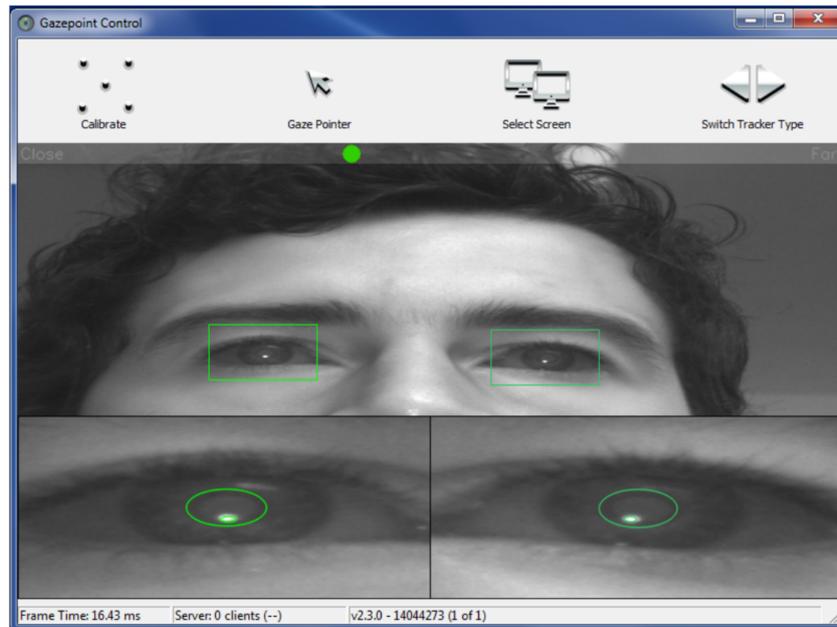


**Figure 2.8:** EOG Electrodes [1]

The Actiview LabView software sends each EEG, ECG and EOG raw data via TCP/IP with 3 bytes sample size. The mental arithmetic test GUI system introduces an extra byte of zero value for computational endianness. These four bytes samples are then unified and transformed into 32-bit format and stored in real-time within the mental arithmetic test GUI system. The data file generated includes the index of time of records (with an interval of 0.0039 s between two adjacent rows), the EEG recordings (at positions of Pz, POz, CPz, Fz and AFz), the ECG recordings (at the positions of LA, RA and LL), the EOG recordings of (at the positions of HR, VR, HL and VL) and the reference (MR and ML).

The Gazepoint™ GP3 eye tracker software (see Figure 2.9) is responsible for the acquisition and transmission of pupil size data from the camera. It comes with an embedded TCP/IP communication channel via computational extensible markup language. The images of the human pupil are evaluated with the Gazepoint™ GP3 eye tracker software and are then sent out with a native frequency of 60 Hz (the sampling frequency of the eye tracker camera). The data received and formatted by the mental arithmetic test GUI in real-time includes: time of record (with an interval of 0.0167 s between two adjacent rows), the *x* and *y* coordinates of the left

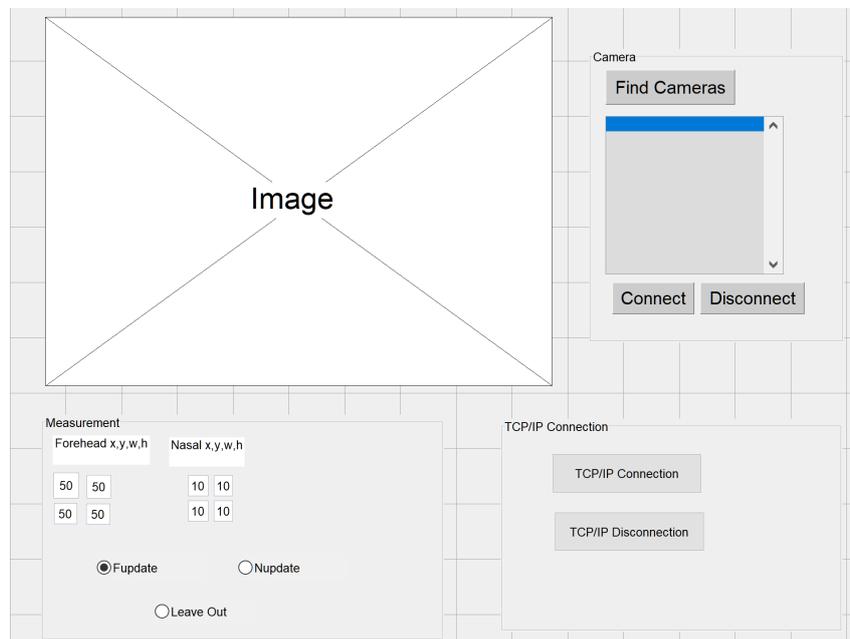
pupil in the image, the left pupil diameter in pixels, the left pupil scaling factor, the left pupil validation flag (binary values of 0 or 1) and followed by the indices for the right pupil in the same order. The mental arithmetic test GUI system calculates the PDM biomarker with the equations mentioned previously.



**Figure 2.9:** GazePoint™ GP3 Eye Tracker Software

The recordings of the FLIR thermal camera are integrated and processed with the self developed Matlab GUI software based on the Matlab support SDKs files from FLIR®. As shown, the software may connect to the thermal camera and project the images in the real-time. The  $x$ ,  $y$ ,  $w$  and  $h$  correspond to the coordinates, width and height of rectangles for measuring the regions of interest. The distortion and displacement due to the head movement may be fixed with updating the position and size of these rectangles during the real-time experiments. The sampling frequency of thermal camera is set to be 10 Hz, and the GUI system stores the raw data locally and sends the processed data with a decimated frequency (1 Hz) to the mental arithmetic test GUI through TCP/IP communication.

The basic configuration of the arithmetic mental test GUI system consists of global variable regulation, TCP/IP receiver and a timer object function to create one-second loops performing the following tasks:

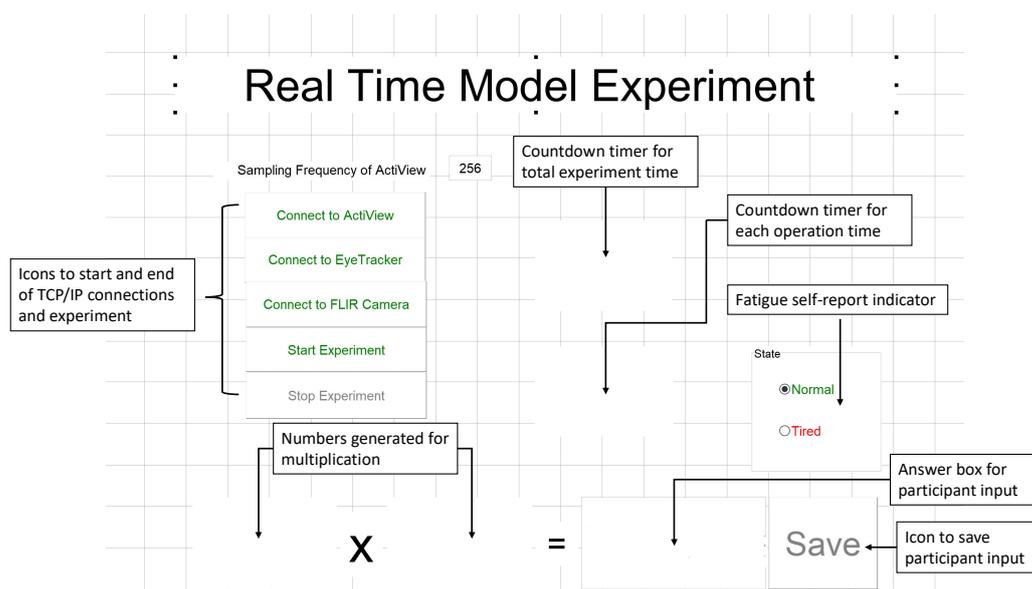


**Figure 2.10:** Thermal Camera Software

1. Open TCP/IP channels and connect to the servers of all biomaker software.
2. Serve as the TCP/IP receiver and store the recording packages in a buffer.
3. Update the timer function based on TCP/IP responses for remaining time of each operation and total experiment.
4. Compile the files for raw data storage.
5. Check the difficulty level of mental arithmetic test and switch the level if required.
6. Generate variance biomarkers via corresponding functions.
7. Evaluate the participant performance with premeditate function.
8. Print out the updated remaining time for participant.
9. Present the new operation task if operation ends because of time or participant's input.
10. Execute function calls for online modelling and control (if requested).

11. Store biomarkers, actual and predicted performance in the model result file, as well as control actions and fatigue index (only for control experiment).

As shown in Figure 2.11, the mental arithmetic test GUI software provides an access for the participant to operate multiplication task with limited experiment time and reaction time. It also ensures the operation of experiment was fully under the control of participant as they may stop at anytime. The difficulty of multiplication may be expressed with the number of operands' digits and the countdown timer of remaining operation time presented in the boxes. In the modelling experiments, each experiment follows arranged phases with different difficulty level for the evaluation of model prediction. While in the controlling experiments, the difficulty level during each experiment follows the guidance from the controller to cooperate with human operator psychophysiological state. Also, the fatigue function was also implemented for further adjustments if requested.



**Figure 2.11:** Mental Arithmetic Test GUI

The mental arithmetic test GUI applied in this research provides the task that requires the participant to complete the multiplication of two numbers (include

both positive and negative integrals). The digits of the random generated numbers and the time scale for answering each question constitute different difficulty levels of mental arithmetic tasks. The numbers are generated from the Matlab stochastic function, and the four difficulty levels defined in this research are shown in Table 2.1. The total operation time for one mental arithmetic test lasted for twelve minutes. For the modelling experiments there were four experiment phases with equal time period corresponding to four difficulty levels, whereas for the controlling experiments the difficulty level varied in relate to the changing of operators' psychophysiological state.

Difficulty Level	Level 1	Level 2	Level 3	Level 4
Response Time (sec.)	10	5	10	5
Digit of the First Numbers	1	1	1	1
Digits of the Second Number	1	1	2	2
Value in the Plots	0.25	0.5	0.75	1

**Table 2.1:** Mental Arithmetic Difficulty Level

The whole mental arithmetic experiment received the ethical approval and followed the health and safety restrictions of University of Sheffield. All the psychophysiological measurements applied were independent from the direct electrical current and ran on batteries with limited voltages. For hygiene, the electrodes were carefully washed and maintained both before and after each experiment, and alcohol free wipes were applied at human-electrode contact area before and after experiment. The participants were informed with the aims and objectives of the research as well as the complete experiment detail and rights during the experiment. All data collected from the research remained anonymous and strictly under the regulation of the University data protocols. The participants were asked to avoid consuming any food or drink at least one hour prior to the experiment and abstain from taking any medicine, caffeine or alcohol drink one day before to avoid any psychophysiological alteration. They were also advised to maintain relaxed both mentally and physically before the experiment. During the experiment, they were told to avoid clenching jaws or massive movements in head and leg to reduce the contamination within the recordings. Talking and specking were

also not allowed during the experiment unless specifically for the personal request or for the purpose of research.

## 2.5 Control of the Human-Machine Interface

Currently, HMI constitutes the foundation of current world, ranging from day-to-day individual mobile communication to global logistics. New automatic algorithms and advanced computational systems allow the HMI systems to be implemented in many aspects of society and initiated revolutionary changes in diverse areas. These advanced computational systems provide the opportunity to process large information in a relatively short period of time for the human operators in various complex fields, such as remote surgery and industrial networks [87]. Though automatic algorithm have largely relieved manual work for basic redundant tasks and compensated human performance, human intervention is still necessary for high level communication and general regulation. Therefore, certain approaches and methods to regulate level of automation remain important for the further application and development of the HMI systems.

In the past few decades, the requirement of human operators in the HMI systems had significantly shifted from working on constant predictable laborious tasks to irregular complex immediate issues. This indicated that as automatic systems in the HMI developed, the work hour of human operator decreased while the work intensity increased. Thus, for such combination of an automatic system with a human operator, the problem often rose from asymmetries between continuously varying effort required for task-solving and relatively fixed process capacity of human operators. System failures and performance breakdowns were commonly correlated with the effort required exceeded human capability. While the insufficient workload resulted in attentional shift and vigilant degradation, and finally led to compromised process capacity. These phenomena were considered to be rooted in the prejudices of distrust or over-trust towards the automation in the HMI. They had influenced the overall performance of HMI and sometimes threaten the safety of whole system [87]. Since constantly assessing task difficulty

and determining automatic level were inessential most of the time and demanded extra effort from human operators, a more acceptable and practical approach was to balance human-machine effort level according to human operator psychophysiological states.

For the purpose of balancing human-machine effort level, it is of great importance to measure and predict human psychophysiological state accurately and immediately in the real-time HMI. Current existing models and frameworks for predicting human psychophysiological state in the HMI systems, e.g. ANFIS, Mamdani-type fuzzy model and interval type-2 fuzzy model, were mostly developed from data-driven approaches [28, 39, 52, 61, 84]. They approximately failed to satisfied the primary demand of balancing human-machine effort level: to assess human psychophysiological state promptly and estimate the trend of operator performance for tuning the automation correspondingly. The lack of investigation for potential origins and possible causes of human operators' compromised the task performance created a disastrous flaw for these existing control approaches. Such passive control approaches without prediction of future human behaviour may not prevent operator mental breakdown or other scenarios and prepare the automatic system adapt to the situation in advance. Therefore, an ideal control method to balancing human-machine effort level in the HMI systems should meet the following fundamental requirements:

1. Reliability: the control configuration should be established from solid assessments of human psychophysiological state in real-time.
2. Adaptability: the control algorithm should be able to tolerate the estimation error and was robust to the interference within the HMI.
3. Perception: the control action should be developed from an accurate understanding of the change of human psychophysiological states and human cognitive performances.
4. Decisiveness: the control intervention should remain punctual and effective to ensure the overall efficiency and safety of HMI systems.

It was of great interest to address these requirements in the control of HMI system. Therefore, with the aim of balancing human-machine effort level, a new computational framework combining state tracking technique based on existing human psychophysiological state estimation algorithm is needed.

## 2.6 Summary

This section summarised current development of human-machine interface research from human psychophysiological state assessments to the modelling and control within the interface. After comparing the different human psychophysiological state measurements, the HRV based on ECG, the TLI based on EEG and PDM were recommended as the most suitable biomarkers for the HMI. Meanwhile, the up-to-date thermal imaging technology allowed future promising biomarkers for this research to be developed on human facial temperature. These findings provided a theoretical foundation for the modelling and control of HMI and some guidance for the design of new facial temperature related biomarkers as well.

## Chapter 3

# Facial Thermal Imaging for Psychophysiological State Detection

Psychophysiological state prediction is of great importance to the HMI as far as both safety and reliability are concerned. In this chapter, the use of facial temperature changes for predicting psychophysiological state and task performance has been investigated. The effectiveness of using facial temperature with the thermal camera to estimate the human psychophysiological state has been validated with the statistical results from a carefully designed HMI experiments with ten (10) healthy subjects. The new facial temperature biomarkers have exhibited a similar or even better ability to differentiate various psychophysiological state in comparison with the traditional biomarkers (e.g. HRV, TLI and PDM). The mean nasal temperature and the differential energy between philtrum and forehead (DEFP) have been shown to be more sensitive to the psychophysiological state changes comparing to the conventional biomarkers. The maximum facial temperature and the mean forehead temperature have also shown clear correlations with psychophysiological state and task performance.

## 3.1 Background

### 3.1.1 Psychophysiological State Detection

The human operator's performance on a certain task is dependent on his or her attention span, cognition, perception and execution, which all develop from the basic conditional reflex [6, 7]. Therefore, monitoring the activities of some specific neurons and subsystems they regulate proved to be a valid approach to assess one's psychophysiological state [6, 7, 41]. In the area of human-machine interface research, the assessment of the human operator's psychophysiological state usually combines peripheral physiology, startle response, central physiology and behaviour. The frequently used measurements are ECG, EEG, pupil size, blood pressure, blood volume, blood volume pulse, respiration, muscle tension, electrodermal activity, galvanic skin and temperature signals [23, 52, 61, 81, 82, 84, 97].

HRV from ECG and TLI from EEG are the most common and recommended psychophysiological state biomarkers. HRV consistently corresponds to the cardio-respiratory system, which is sensitive to the changes of psychophysiological stress [9, 48]. The aim of TLI is to calculate one's work memory (WM), which constituted one's ability to maintain the focus on one specific task regardless of the surrounding interference [28, 75]. However, EEG and ECG measurements are normally involved with using the electrodes to record voltage differences across the skin. Such a requirement limits the movement and the range of movement of the human operator and disturbed his or her mental state as well. Meanwhile, the measurements here remains sensitive to the noise introduced by defective skin-electrode connections and surrounding electromagnetic fields. The high complexity of EEG and ECG measurements restricts the efficiency and the safety of applying HRV and TLI in real-world situations. Therefore, it was important to design and integrate new psychophysiological state biomarkers with the existing system to overcome these constraints.

### 3.1.2 Facial Temperature

It was found that the thermal homeostasis, the emotion and the fight or flight response all share the same area of the brain - hypothalamus [16]. Apart from the hypothalamus, the parasympathetic and sympathetic nervous system also involves in the psychophysiological induced thermal regulation [39]. The control of blood flow and the contraction and dilation of blood vessels were consistently correlated with the psychophysiological state. These connections constituted the foundation of human conditioned response, and as a result, these regulations led to the observable temperature changes at the level of surface skin [12, 14]. Due to the periodic regulation of blood circulation, most temperature regulations demonstrated circular periods around 10 seconds, and others varied from seconds to minutes [40, 73].

Infrared cameras have hitherto provided a reliable means of documenting the facial temperature in real time without body contact, and they usually have fewer requirements for the work environment compared with EEG and ECG. The temperature of the target is measured by the recorded infrared radiation, the emissivity, the transmittance and the surrounding temperature [85]. The total radiation  $W$  captured by the camera is as follows:

$$W = W_{obj} + W_{ref} + W_{atm}, \quad (3.1)$$

where  $W_{obj}$ ,  $W_{ref}$  and  $W_{atm}$  are the infrared radial emission from the target object, surrounding environment reflection and atmosphere, and they are calculated via the following equations:

$$\begin{aligned} W_{obj} &= \epsilon_{obj} \cdot \tau_{atm} \cdot \sigma \cdot (T_{obj})^4, \\ W_{ref} &= (1 - \epsilon_{obj}) \cdot \tau_{atm} \cdot \sigma \cdot (T_{ref})^4, \\ W_{atm} &= (1 - \tau_{atm}) \cdot \sigma \cdot (T_{atm})^4, \end{aligned} \quad (3.2)$$

where  $T_{obj}$ ,  $T_{ref}$  and  $T_{atm}$  are the temperatures of the target object, surrounding environment and atmosphere,  $\sigma$  is a constant.  $\epsilon_{obj}$  is the emissivity of the object that is depended on the material, and  $\tau_{atm}$  is the transmittance of the atmosphere

which is determined by the distance and the relative humidity. Infrared cameras provide a non-contact measurement that is free from the psychophysiological changes for connecting the electrodes in EEG and ECG. A comparison between areas and a tracking of changing thermal patterns can be obtained with the real-time two-dimensional thermal image from the camera [85].

The facial temperature recorded by infrared cameras has been shown to be a potential valid reflection of the human mental state nowadays. Previous studies have also shown that temperature readings from the forehead, the periorbital and the nasal regions are closely correlated to psychophysiological state [62, 63, 65]. The thin soft tissue layer of the forehead make the observation of temperature change more convenient than the other areas, while the high blood pressure around the eye orbit formed the maximum facial temperature point in the periorbital region. The temperature of the nasal regions represent the regulation of cardio-respiratory system, as it is under the control of both blood flow and air flow. Also, it was proven that the energy spectrum from the frequency analysis of the facial temperature was consistently correlated with the heartbeat rate and cortisol level [39]. Thus, the biomarkers based on the facial temperature readings from the camera had a better chance of providing accurate psychophysiological state estimation instantaneously, and also retained adequate distance from the subject as compared to the HRV and TLI.

## **3.2 Experimental Setup**

### **3.2.1 Participants**

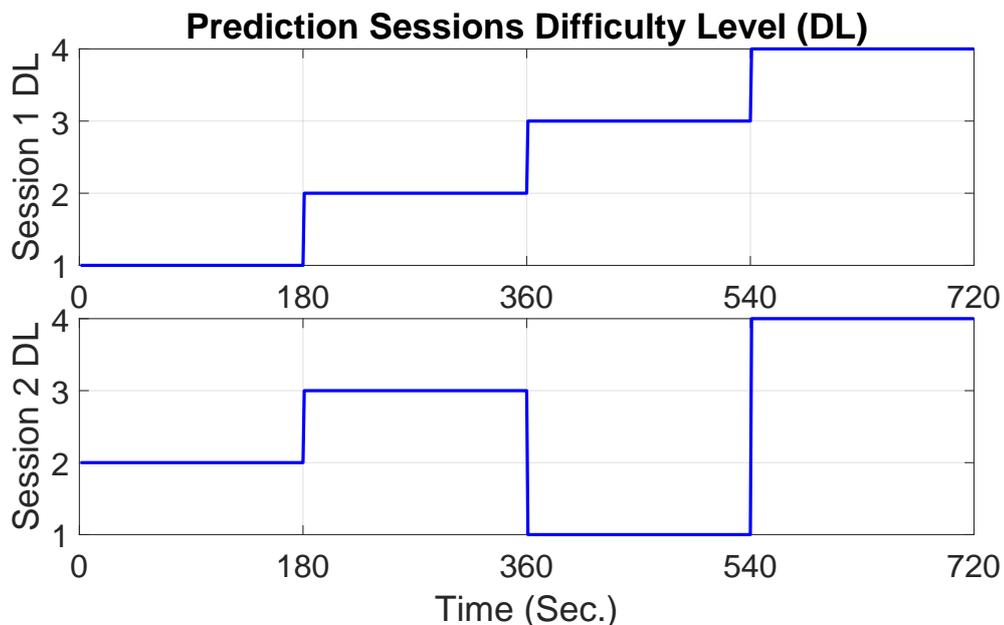
The participants of this experiment included a total of ten (10) healthy research students, selected from the Department of Automatic Control and Systems Engineering at the University of Sheffield (UK). The selection of participants covered both genders from different races and countries to avoid any bias. They were between 22 to 30 years old, with an average age of 25. In addition, all the participants were required to abstain from taking any medicine, caffeine or alcohol drink at least two hours before the experiment to avoid abnormal psychophysiological

reaction. They were informed with the experiment procedures and their rights during the experiment, and all the experiments were conducted under the Health and Safety regulations of the University of Sheffield.

### 3.2.2 Simulation Experiments

The mental arithmetic experiment was selected as the HMI simulation. It was based on the Matlab GUI app developed by a previous PhD student Luis A. Torres-Salomao in the University of Sheffield [82, 84].

The entire experiment for each student lasted around 40 minutes, including two 12-minute mental arithmetic test sessions and one 12-minute comparison session as a control group in the interval, with 2-minute breaks in between the sessions. There were a total of four difficulty levels and each difficulty lasted 180 seconds that should comply with the time requirement for a full transition of human psychophysiological state. As shown in Figure 3.1, the order of difficulty levels varied between two mental arithmetic test sessions for eliminating the inference of human adaptation, where the participants subjectively controlled their task performances and psychophysiological states based on their previous experiences.



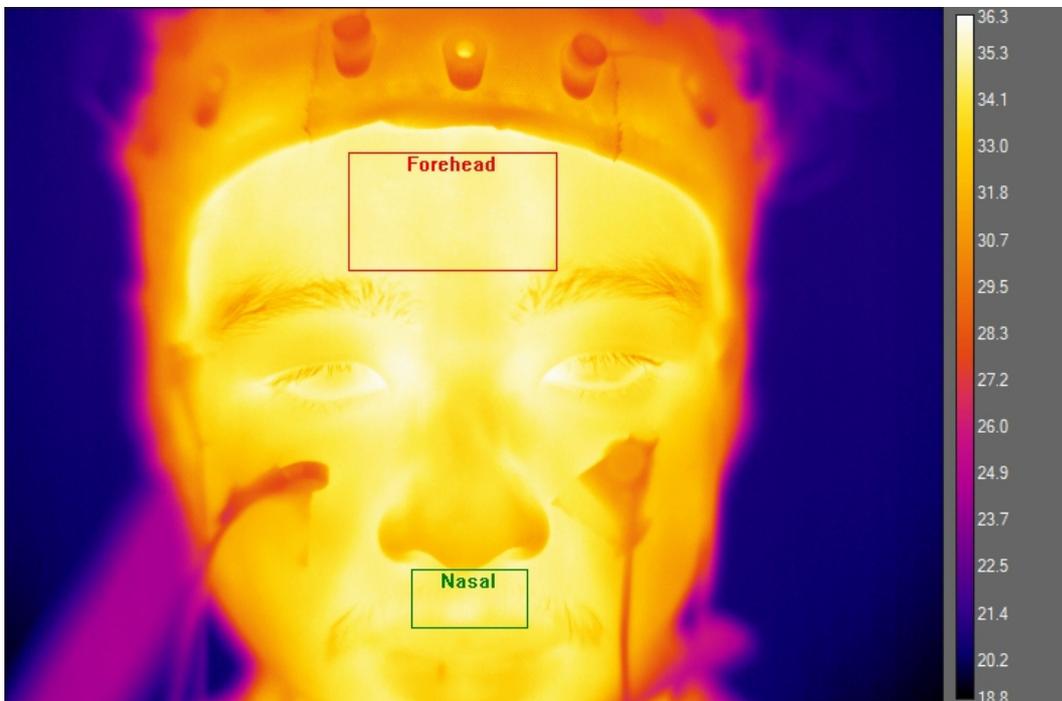
**Figure 3.1:** Difficulty levels for facial temperature experiment session 1 (progressive increase) and session 2 (random)

### 3.2.3 Temperature Data Acquisition

The facial temperature was recorded with a FLIR E40bx thermal camera that is positioned horizontally 0.5 meter away from the participant's face. The emissivity of the human tissue was 0.98, and the sampling frequency of the camera was set to be 10 Hz. The laboratory room provided a steady environment temperature and a relative humidity of around 20.0 °C and 50 % respectively.

## 3.3 Developed Facial Temperature Biomarkers

The thermal imaging sequences were analysed in MATLAB® to extract the temperature magnitude and frequency changes in the regions of interest (ROI), which were the periorbital area, the nasal area and the forehead (see Figure 3.2). The developed biomarkers based on these were the maximum facial temperature (around periorbital), the mean nasal temperature, the mean forehead temperature and the differential energy between philtrum and forehead (DEFP).



**Figure 3.2:** Regions of interest: periorbital, nasal and forehead

The maximum facial temperature  $T_{maxf}$  is captured by the point measurement

of the camera. The calculation is as follows:

$$T_{maxf} = \left\{ \max(T_{(i,j)}) \mid \forall i \in L, \forall j \in W \right\}, \quad (3.3)$$

where  $T_{(i,j)}$  represents the temperature value recorded at pixel  $(i,j)$ , and  $L$  and  $W$  are the numbers of the columns and rows of pixels in a frame.

The mean nasal temperature  $\bar{T}_n$  is captured by the rectangle measurement of the camera. The calculation is as follows:

$$\bar{T}_n = \left\{ \frac{1}{XY} \sum T_{(i,j)} \mid \forall i \in X, \forall j \in Y \right\}, \quad (3.4)$$

where  $T_{(i,j)}$  represents the temperature value recorded at pixel  $(i,j)$  from the selected rectangle region over the nasal area, and  $X$  and  $Y$  are the numbers of the columns and rows of pixels in the region.

The mean forehead temperature  $\bar{T}_f$  is captured by the rectangle measurement of the camera. The calculation is as follows:

$$\bar{T}_f = \left\{ \frac{1}{MN} \sum T_{(i,j)} \mid \forall i \in M, \forall j \in N \right\}, \quad (3.5)$$

where  $T_{(i,j)}$  represents the temperature value recorded at pixel  $(i,j)$  from the selected rectangle region over the forehead, and  $M$  and  $N$  are the numbers of the columns and rows of pixels in the region.

The differential energy between philtrum and forehead (DEFP) is obtained by the differential energy from the Fourier transform analysis of the temperatures between the philtrum and forehead. After filtering the raw data with a bandpass filter (0.4 Hz–4 Hz representing 24–240 beats per minute), the Fast Fourier Transform (FFT) is performed with a window size of 10 seconds to extract the energy spectrum within that frequency range. The final value is then generated from the subtraction between the maximum energy value of the philtrum temperature and the mean energy value of the forehead temperature. The whole process is summarised via the following equations (the Discrete Fourier Transform DFT to represent the Fast Fourier Transform FFT for simplicity):

$$\begin{aligned}\hat{S}_{Tp,j}[k] &= \left\{ \sum_{n=0}^{N-1} \bar{T}_p[2-n-j]e^{-i\frac{2\pi}{N}kn} \mid \forall k \in [4, 40], N = 100 \right\}, \\ \hat{S}_{Tf,j}[k] &= \left\{ \sum_{n=0}^{N-1} \bar{T}_f[2-n-j]e^{-i\frac{2\pi}{N}kn} \mid \forall k \in [4, 40], N = 100 \right\}, \\ DEFP_j &= \max \hat{S}_{Tp,j} - \bar{\hat{S}}_{Tf,j}\end{aligned}\quad (3.6)$$

where  $\bar{T}_n$  and  $\bar{T}_f$  represents the mean philtrum temperature and the mean forehead temperature respectively.  $N$  is the sample size of the selected window.

### 3.4 Evaluation of Facial Temperature Biomarkers

In this research, the efficiencies of different biomarkers are determined by their abilities to measure human operators' psychophysiological state in real-time. As are already mentioned in the previous chapter, in the normal state, there exist synchronised connections between human operators' psychophysiological state and the task they are dealing with. Therefore, in reverse, the difficulty level (DL) and the task performance capable of representing psychophysiological state in some degree. However, the stability of human operators' task performance are compromised by the disturbances from the random mental arithmetic tasks and the mathematical abilities of different individuals, which were irrelevant to the actual psychophysiological state. Conversely, the DL was associated with a prolonged 3-minute subsection with similar mental arithmetic tasks, thus, it was able to represent a general human operator's psychophysiological state within that corresponding period. Accordingly, the following evaluations of facial temperature biomarkers,  $T_{maxf}$ ,  $\bar{T}_n$ ,  $\bar{T}_f$  and DEFP, were focused on the analysis of correlations between the indicators and the DL.

#### 3.4.1 Maximum Facial Temperature

##### T-test H Value

In this research, the two-sample T-test is applied to compare the efficiency of using facial temperature as suitable psychophysiological indicators with HRV, TLI and

PDM. The test statistic is given as follows:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2(\frac{1}{N_1} + \frac{1}{N_2})}}, \quad (3.7)$$

where  $t$  is a quantile with  $N_1 + N_2 - 2$  degrees of freedom which is an indicator ratio between the samples and noises, and the pooled sample variance  $s^2$  was measured as follows:

$$s^2 = \frac{\sum_{i=1}^{N_1} (x_i - \bar{x}_1)^2 + \sum_{j=1}^{N_2} (x_j - \bar{x}_2)^2}{N_1 + N_2 - 2}, \quad (3.8)$$

in which the two  $\bar{x}$  were the means of two compared sequences, and  $N_1$  and  $N_2$  represented the sample sizes, see Table 3.1 for details. The null hypothesis of the test was that there is no statistically significant difference between the two samples. The  $t$  value is checked with a critical value at the probability level of 0.05 to determine the acceptance or rejection of the null hypothesis.

	Phase 1	Phase 2	Phase 3	Phase 4
Sample Sizes	150	180	180	180

**Table 3.1:** Sample Sizes of Different Phase within the Experiment Session

According to the two-sample T-test, each indicator is tested for its ability to distinguish different psychophysiological state. For  $H = 0$ , we found no significant difference observed within a 5% confidence level. For  $H = 1$ , there were significant differences between the data from two adjacent subsection with different difficulty levels. Original H value of the two-sample T-test is a binary indicator for showing whether difference can be found between two compared sequences. By summering the H values over the different participants, it can be found whether the change of human psychophysiological state had triggered any change in the recordings of biomarkers during the experiments. For this research, the H value was considered as the primary criterion for the assessments of biomarkers, because it was the most direct and clear estimation of how sensitive a biomarker was to a change in psychophysiological state. A biomarker with the higher H value required less

effort to extract psychophysiological related information from its records.

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	0.9500			
DL 3	1.0000	1.0000		
DL 4	1.0000	1.0000	0.8500	

**Table 3.2:** Overall Maximum Facial Temperature T-test Results for the Experimental Sessions

The overall average H value of the maximum facial temperature for all the ten volunteers was 0.9667. This suggested that the maximum facial temperature was sensitive to the operators’ psychophysiological state changes. As shown in Table 3.2, the biomarker was able to differentiate the different psychophysiological states of the participants with medium workloads. For some individuals, the low workloads of DL 1 & 2 were not high enough to trigger observable psychophysiological changes between these two difficulty levels, thus, the performance of the biomarker was compromised. Meanwhile, some participants met their mathematical limitation on mental arithmetic with DL 3 and failed to handle the more challenging DL 4 properly. This led to no actual psychophysiological change exist and to be observed when the difficulty level increased from 3 to 4. To support the finding of the H value, the following sections presented a further investigation on how the some features of the maximum facial temperature indices were consistently correlated with psychophysiological state.

**Arithmetic Mean**

The arithmetic mean was one of the most observable statistical differences for the temperature indicators. Table 3.3 summarises the overall arithmetic means of maximum facial temperature for all the participants in the incremental and random difficulty level sessions. As shown in this table, the maximum facial temperature was generally increased as the task difficulty level increased, regardless the order of the difficulty levels. It is worth noting that the mean value of difficulty level 4 was lower than the mean value of difficulty level 3 in the random sessions. This was due to the fatigue and breakdown of the participants, considering it was the

last subsection for an hour-long experiment (plus the setup time). Therefore, the maximum facial temperature are able to directly indicate human psychophysiological state.

Mean value (°C)	DL 1	DL 2	DL 3	DL 4
Incremental session	36.1602	36.1568	36.1775	36.1802
Random session	36.3570	36.4011	36.4083	36.3328

**Table 3.3:** Overall Arithmetic Means for the Maximum Facial Temperature

The overall mean value of the maximum facial temperature for all the ten volunteers was 36.2713 °C. Changes in temperatures for different sessions were mostly around 0.4397 °C, and around 0.0634 °C for any two adjacent subsection. In addition, compared to the period of the control sessions, both the mean and the standard deviation of the maximum facial temperature for the experimental sessions showed some increases, e.g. from 35.7329 and 0.4108 °C to 36.2713 and 0.4894 °C.

### Correlation

Correlation checked the dependence between the biomarkers and the operators' psychophysiological states. In this research, the Pearson product-moment correlation coefficients were measured via the following equation:

$$\rho_{xo} = \frac{\sum_{i=1}^N (x_i - \bar{x})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (o_i - \bar{o})^2}} \quad (3.9)$$

where  $x$  is the recordings of the biomarkers and  $o$  is the observation (difficulty level).  $N$  is the total sample size and was set to be 692 for a session.  $\bar{x}$  and  $\bar{o}$  are the mean values of the biomarkers and the observation. Table 3.4 summarises the correlation between the maximum facial temperature and the difficulty level for the ten participants.

In the experiment with ten (10) participants, the maximum facial temperatures for five subjects were consistently correlated with their psychophysiological state

Participant	Incremental session	Random session
1	0.7304	0.4788
2	-0.4925	0.1148
3	-0.2055	0.2654
4	0.5436	0.6491
5	-0.0560	0.0046
6	0.5066	0.2670
7	-0.7933	-0.1562
8	-0.0994	-0.4141
9	-0.1684	0.5163
10	-0.3923	0.5736

**Table 3.4:** Total Correlation between Maximum Facial Temperature and Difficulty Level

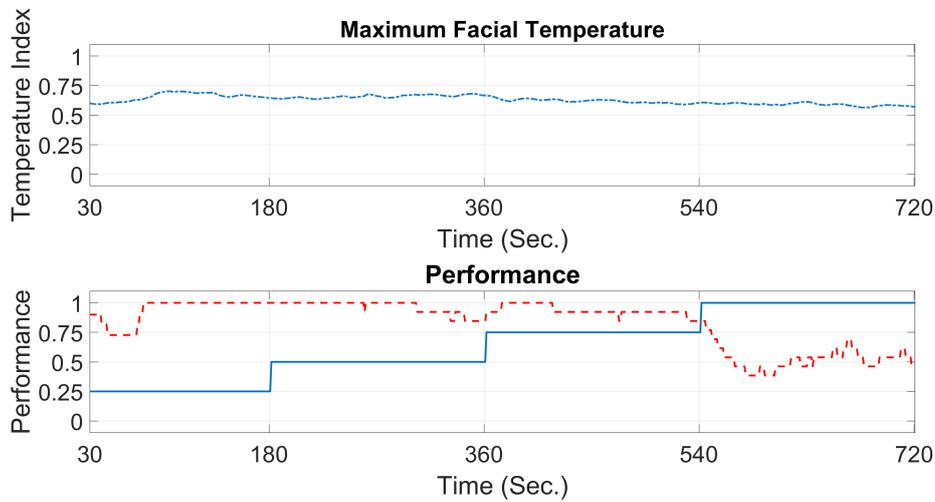
for both increasing and random difficulty order (see Table 3.4). Among the five participants, two showed negative correlations between the temperature and the accuracy, while the other three demonstrated positive correlations. Therefore, the psychophysiological state induced change of the maximum facial temperature was predominantly affected by the individual differences. Meanwhile, the absolute value of correlation coefficient between the maximum facial temperature and the task accuracy performance were relatively high in some cases, e.g. Figure 3.3. This suggested that the maximum facial temperature showed a great potential of representing an online indicator.

### Dispersion Ratio

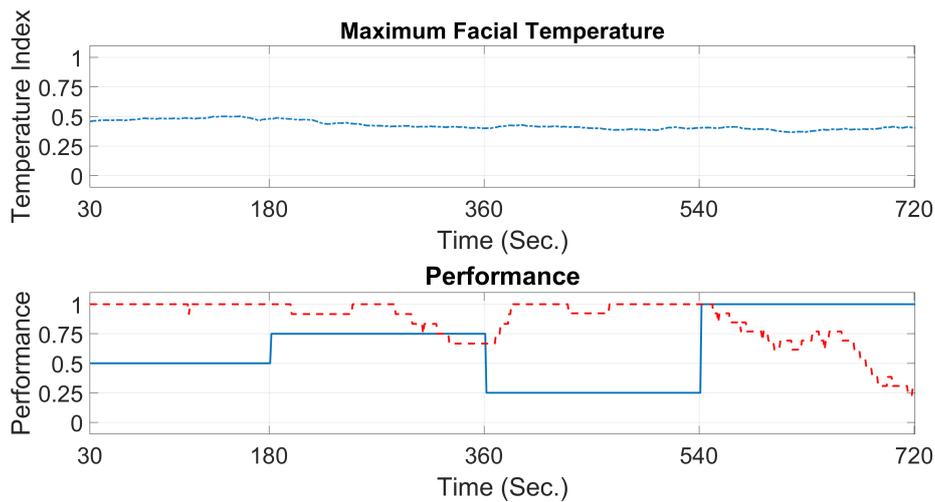
The dispersion ratio measured the density of data distribution, which is the ratio between the variances  $\sigma_x$  and the means  $\mu_x$ , as the following equation shows:

$$D = \sigma_x^2 / \mu_x. \quad (3.10)$$

Table 3.5 shows the total dispersion ratio of the maximum facial temperature for the ten participant. The table suggests that the dispersion ratios of the rest states are different from the experimental sessions. This was due to the change of thermal regulation based on the circular system. However, there was no significant



(a) Normalised maximum facial temperature (-.-), accuracy (- -) and difficulty level (-) plots for participant 1. Session 1 with elevated difficulty levels



(b) Normalised maximum facial temperature (-.-), accuracy (- -) and difficulty level (-) plots for participant 1. Session 3 with random difficulty levels

**Figure 3.3:** Maximum Facial Temperature

Dispersion ratio	Incremental session	Random session
rest states	0.6665	0.6174
DL 1	0.5769	0.6391
DL 2	0.6656	0.6305
DL 3	0.7135	0.5730
DL 4	0.6376	0.5956

**Table 3.5:** Total Dispersion Ratio for Maximum Facial Temperature

general pattern that can be observed between the difficulty level and dispersion ratio. Nevertheless, the differences of dispersion ratio between different difficulty level were significant. This may contribute to the change of stress level or arousal of the psychophysiological state in response to different requirements of the task. Therefore, the dispersion ratio of the maximum facial temperature was also indicative of the human psychophysiological state.

### Zero-Crossing Rate

The zero-crossing rate measures the rate of changes within a data recording. It checks the stability and frequency component within the data. In this experiment, the zeros set to be the mean values  $\bar{x}$  of each phase and the rate was measured through the following equations:

$$zcr = \frac{1}{N-1} \sum_{i=1}^{N-1} 1_{R_{<0}}((x_i - \bar{x})(x_{i-1} - \bar{x})), \quad (3.11)$$

where  $N$  is the sample size of  $x$  in each phase (see Table 3.1) and  $1_{R_{<0}}((x_i - \bar{x})(x_{i-1} - \bar{x}))$  is an indication function defined as follows:

$$1_{R_{<0}}((x_i - \bar{x})(x_{i-1} - \bar{x})) = \begin{cases} 1, & \text{if } (x_i - \bar{x})(x_{i-1} - \bar{x}) < 0 \\ 0, & \text{if } (x_i - \bar{x})(x_{i-1} - \bar{x}) \geq 0 \end{cases} \quad (3.12)$$

Table 3.6 summarises the total zero-crossing rate for maximum facial temperature of the participants in the mental arithmetic experiments. Similarly to the dispersion ratio, the pattern between the difficulty level and zero-crossing rate was not obvious in general. However, the differences of the zero-crossing rate for the maximum facial temperature between the rest states and the experimental sessions were significant. This may also be the results of the regulation of blood flow due to the changes in psychophysiological states. Still, the zero-crossing rate suggests that there exist some consistent connection between the maximum facial temperature and the participant's psychophysiological state.

Zero-crossing rate	Incremental session	Random session
rest states	0.1378	0.1313
DL 1	0.1026	0.1263
DL 2	0.1161	0.1171
DL 3	0.1115	0.1099
DL 4	0.1257	0.0925

**Table 3.6:** Total Zero-crossing Rate for Maximum Facial Temperature

### 3.4.2 Mean Nasal Temperature

Following the results of maximum facial temperature analyses, the statistical analyses of mean nasal temperature were focused on the T-test H value, arithmetic mean and correlation.

#### T-test H Value

The overall average H value for the mean nasal temperature encompassing all the ten participants in the experiment was 0.9500. This indicated that the mean nasal temperature was also sensitive to the psychophysiological changes.

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	0.9500			
DL 3	1.0000	0.9500		
DL 4	0.9500	0.9500	0.9500	

**Table 3.7:** Overall Mean Nasal Temperature T-test Results for Experimental Sessions

As shown in Table 3.7, the mean nasal temperature demonstrated an excellent ability to differentiate various psychophysiological states. Compared to the performance of the maximum facial temperature, the mean nasal temperature had a more balanced performance throughout the entire experiment. However, the lower H value suggested that the mean nasal temperature shared a weaker connection to the operators' psychophysiological state. It was found, in some cases, that the moustache of male participants disturbed the air flow in the nasal area. This may be one of the reasons for the compromised overall average H value of the mean nasal temperature indicator. To further investigated the relationships between the

psychophysiological state and the mean nasal temperature, the arithmetic mean and the correlation analyses were also provided.

**Arithmetic Mean**

Table 3.8 shows the overall arithmetic means of the mean nasal temperature for all the participants in the incremental and random difficulty level sessions. As shown in the table, the mean nasal temperature was slightly reduced as the task difficulty level increased in the both sessions. The mean values of difficulty level 4 were significantly lower than the mean values of difficulty level 1 in the both two sessions. This was due to the increased speed of air flow at the nasal area when the participants’ raised their arousal or stress level for more challenging tasks. The consistent performance of the mean nasal temperature in the both two sessions suggested that it was able to represent the human psychophysiological state precisely.

Mean value (°C)	DL 1	DL 2	DL 3	DL 4
Incremental session	33.9163	34.0144	33.9394	33.7517
Random session	33.9331	34.1069	33.9776	33.8240

**Table 3.8:** Overall Arithmetic Means for Mean Nasal Temperature

The general mean value of the mean nasal temperature for all ten (10) test subjects was 33.9293 °C. The change of temperature over the different sessions ranges from 0.1527 up to 1.8739 °C with a mean value of 0.7758 °C, and from 0.0021 to 0.8933 °C with a mean value of 0.2287 °C for any two nearby sub-sessions. Meanwhile, the standard deviation of the mean nasal temperature remained around 1.0285 °C for all the participants across the entire experiment. It suggested that the mean nasal temperature was able to be a relatively stable indicator for predicting the psychophysiological state.

**Correlation**

Table 3.9 summarises the correlation between the mean nasal temperature and the difficulty level of all the ten participants for the entire experiment. As shown in

Participant	Incremental session	Random session
1	0.6804	0.5682
2	-0.4725	-0.1972
3	-0.3730	0.4792
4	0.8899	0.5022
5	-0.0239	-0.3436
6	0.7615	-0.3980
7	0.6955	0.3215
8	0.2656	0.6778
9	0.4606	0.6736
10	-0.2865	0.4220

**Table 3.9:** Total Correlation between Mean Nasal Temperature and Difficulty Level

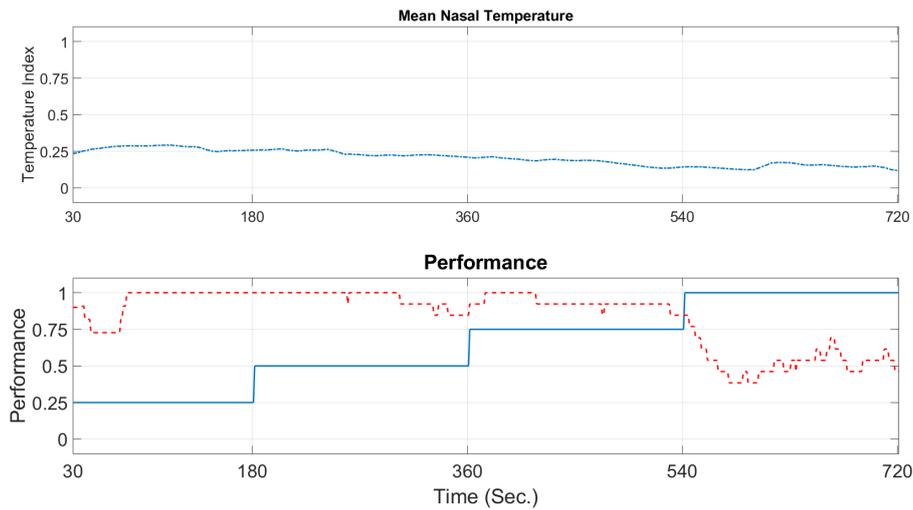
this table, there were seven participants (1, 2, 4, 5 & 7-9) presented consistent correlations between their mean nasal temperatures and psychophysiological states regardless of the difficulty order. Within these seven subjects, five showed negative correlations between their mean nasal temperature and task accuracy, while the other two provided positive correlations. This suggested that the changes of the mean nasal temperature were also consistently correlated to the change of the subjects' psychophysiological performance, which supported the findings from previous research mentioned in the background section [62, 63, 65], see Figure 3.4 for example.

### 3.4.3 Mean Forehead Temperature

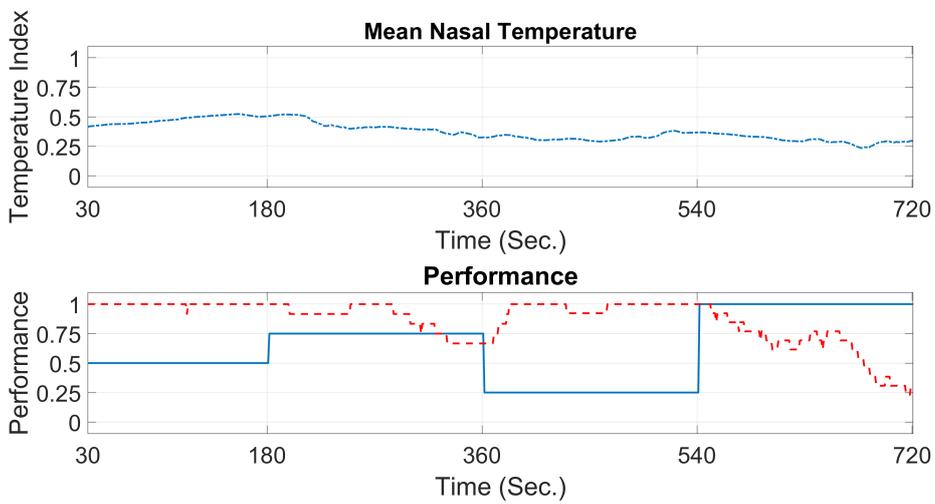
Similar to the mean nasal temperature analyses, the statistical analyses of mean forehead temperature were focused on the T-test H value, arithmetic mean and correlation.

#### T-test H Value

The overall average H value for the mean forehead temperature for all the ten participants in the entire experiment was 0.9333. Compared to the maximum facial temperature and the mean nasal temperature, the mean forehead temperature was slightly less sensitive to the operators' psychophysiological changes.



(a) Normalised mean nasal temperature (-.-), accuracy (- -) and difficulty level (-) plots for participant 1. Session 1 with elevated difficulty levels



(b) Normalised mean nasal temperature (-.-), accuracy (- -) and difficulty level (-) plots for participant 1. Session 3 with random difficulty levels

**Figure 3.4:** Mean Nasal Temperature

Table 3.10 summarises the H value for the mean forehead temperature in each difficulty levels. As shown in this table, the performance of the biomarker was slightly compromised in the comparisons of DL 1 and DL 2, DL 1 and DL 3. This was mainly because as compared to the other facial areas the forehead had a relatively thinner soft tissue layer. The reduced blood flow of that area made the temperature more susceptible to the environment disturbances, especially when the participants were under lower arousal states or stress levels and the temperature changes were minor. Nevertheless, the overall H value analysis of mean forehead temperature suggested that it was still a credible biomarker for assessing human psychophysiological state.

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	0.9000			
DL 3	0.8500	0.9500		
DL 4	0.9500	1.0000	0.9500	

**Table 3.10:** Overall Mean Forehead Temperature T-test Results for Experimental Sessions

#### Arithmetic Mean

Mean value (°C)	DL 1	DL 2	DL 3	DL 4
Incremental session	34.3702	34.4102	34.4096	34.2615
Random session	34.4604	34.4890	34.5064	34.4035

**Table 3.11:** Overall Arithmetic Means for Mean Forehead Temperature

Table 3.11 presents the overall arithmetic means of the mean forehead temperature for all the participants in the incremental and random difficulty level sessions. It is worth noting that despite the different difficulty level orders, the changes of the mean forehead temperature followed the same pattern: it increased with the difficulty levels at first and then decreased with the difficulty levels. This suggested there were two significantly different behaviors for the mean forehead temperature in the lower and higher arousal states (stress levels). It was found that when participants were under the low difficulty levels, the temperature was

predominately determined by the blood flow, which increased with the stress and raised the temperature of forehead. However, once the temperature was raised after a certain threshold (individual depend), the perspiration gradually took the control of forehead temperature regulation. As a result, the mean forehead temperature presented a negative correlation with the difficulty level when the participants were under higher stress.

The overall mean value of the mean forehead temperatures across all the ten participants was 34.4130 °C. The mean temperature difference between the sessions ranged from 0.1090 to 1.2535 °C with a mean value of 0.4802 °C, and were mostly around 0.1285 °C between any two neighbouring subsection. Different from the mean nasal temperature, the standard deviation of mean forehead temperature were increased from 0.8405 to 1.3546 °C in the incremental sessions, and from 0.9957 to 1.0689 °C in the random sessions. This also indicated the pronounced nonlinearity within the mean forehead temperature increased with human stress level, as it was regulated by both blood flow and perspiration.

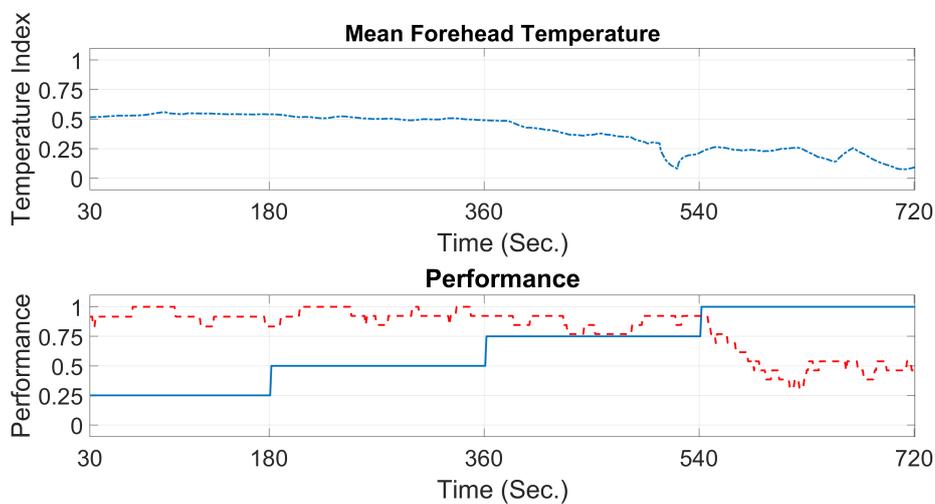
**Correlation**

Correlation (%) Participant	Incremental session	Random session
1	0.5083	0.2948
2	-0.5341	0.4979
3	-0.4672	-0.2347
4	0.1116	0.0486
5	0.7872	0.0304
6	0.5434	0.3464
7	0.7581	0.7197
8	-0.3537	0.0800
9	-0.3508	0.2770
10	-0.7146	0.1366

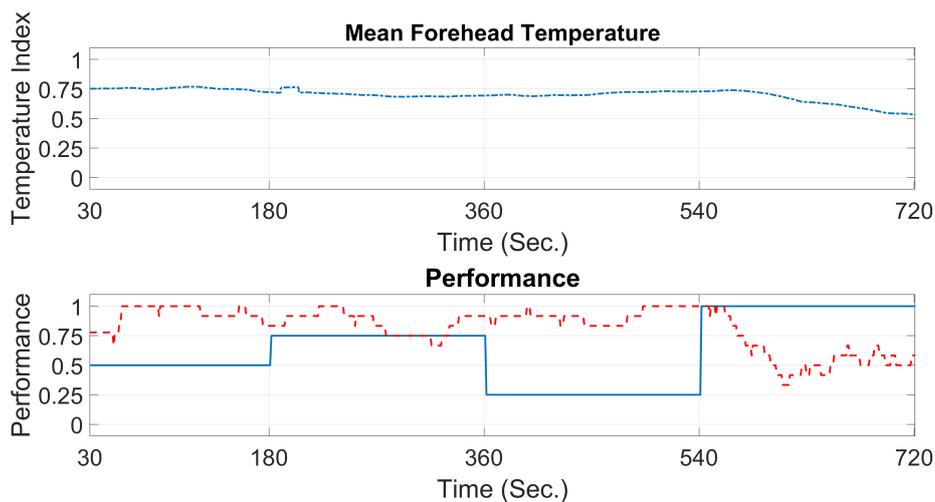
**Table 3.12:** Total Correlation between Mean Forehead Temperature and Difficulty Level

Table 3.12 shows the correlation between the mean forehead temperature and the difficulty level of all the ten participants for the whole experiment. There were six participants showed consistent correlations between the difficulty level and the

mean forehead temperature in both two different difficulty orders. Meanwhile, there were eight participants presented consistent correlations between their task performance and the mean forehead temperature in both sessions. An example was provided in Figure 3.5. Within these eight participants, half of them showed positive correlations, while the rest showed negative correlations. This indicated that the changes of the mean forehead temperature were able to assess the human psychophysiological state directly in most cases.



(a) Normalised mean forehead temperature (-.-), accuracy (- -) and difficulty level (-) plots for participant 7. Session 1 with elevated difficulty levels



(b) Normalised mean forehead temperature (-.-), accuracy (- -) and difficulty level (-) plots for participant 7. Session 3 with random difficulty levels

**Figure 3.5:** Mean Forehead Temperature

### 3.4.4 Differential Energy between Philtrum and Forehead

Similarly to the previous biomarkers, the statistical analyses of differential energy between philtrum and forehead (DEFP) included the T-test H value, arithmetic mean and correlation.

#### T-test H Value

The overall average H value for the DEFP of all the ten participants in the whole experiment was 0.9667. Compared to the previous biomarkers, the DEFP remained relatively high sensitivity in the most situations.

As shown in Table 3.13, the H value of DEFP was only compromised in the session of DL 4. As already stated, the DEFP was correlated with the heartbeat rate and the cortisol level. It was mainly under the influence of the human hormone regulations, and the changes of it was affected by the typical hormone regulation effects, e.g. slow but lasting. Both experimental sessions were ended with the DL 4, when the cortisol level of participants could already reached the peak value and the similar psychophysiological effect might remained for a long period. Therefore, compared to the previous subsections, the DEFP in DL 4 led to slightly less significant statistical changes.

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	1.0000			
DL 3	1.0000	1.0000		
DL 4	0.9000	0.9500	0.9500	

**Table 3.13:** Overall DEFP T-test Results for Experimental Sessions

#### Arithmetic Mean

Table 3.14 provides the overall arithmetic means of DEFP for all the ten participant in the whole experiment. In the both sessions, the mean value of DEFP decreased when the difficulty level increased from 1 to 3. As DEFP was closely linked to the cortisol level, this inverse ratio between the difficulty level and DEFP clearly indicates the positive correlation between the participant stress level and the task

difficulty level. However, compared to the neural regulation, the hormone regulation suffered a relatively larger hysteresis. As a consequence, the mean value of DEFP in DL 2 of the random session was abnormally higher than those in DL 1 and 3. Also, it is worth noting that the overall mean value of DEFP sudden increased in DL 4 in both sessions. This was mainly because of the sudden increases in DEFP for these participants who experienced performance breakdown in the DL 4.

Mean value (°C)	DL 1	DL 2	DL 3	DL 4
Incremental session	0.1383	0.0587	0.0513	0.1476
Random session	0.1529	0.1922	0.1157	0.2139

**Table 3.14:** Overall Arithmetic Means for DEFP

The mean DEFP for the ten subjects was 0.1265 during the experiment sessions. Compared to the control sessions, the DEFP values increased up to around 10 %. DEFP values changed around 0.1117 for crossing the two neighboring subsection and 0.1163 for crossing different sessions. Compared to the other biomarkers, the DEFP was relatively more sensitive to the participant psychophysiological or task performance breakdown. This suggested though DEFP suffered more severe hysteresis than other biomarkers, it was able to provide a more direct indication of human psychophysiological state, especially for the stress level.

### Correlation

Correlation (%) Participant	Incremental session	Random session
1	0.1129	-0.3704
2	-0.1458	0.0284
3	0.1003	0.2257
4	-0.6103	-0.4327
5	0.0626	-0.1361
6	-0.4007	0.1812
7	-0.1391	-0.2961
8	0.5303	0.1472
9	0.5021	-0.0739
10	-0.1086	0.0749

**Table 3.15:** Total Correlation between DEFP and Difficulty Level

Table 3.15 summarises the correlation between the DEFP and the difficulty level for all the participants across the whole experiment. As shown in this table, there were only four participants (3, 4, 7 & 8) showed consistent correlations between their DEFP and the difficulty levels in both sessions. Meanwhile, among these four participants, there were only two participants (4 & 7) showed consistent correlations between their DEFP and the task performance. This was mainly due to the breakdown in the subsection of difficulty level 4, see Figure 3.6 for example. In the first pair of plots in Figure 3.6, the DEFP showed a clear increase at time around 530s, while the accuracy of the participant continuously decreased for the next 60 seconds. Meanwhile, in the second pair of plots, the increases of DEFP were somewhat following the compromised performance in the subsections of DL 2 and 4. Therefore, these observations prove that DEFP is an efficient indicator for assessing the stress level relevant to the psychophysiological state.

### 3.4.5 Comparison of Facial Temperature Biomarkers with Conventional Biomarkers

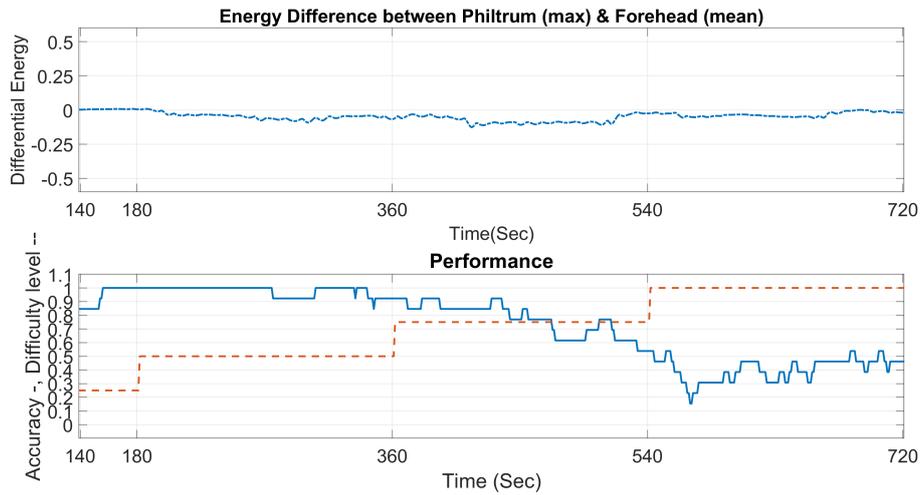
HRV and TLI mentioned previously were selected as the biomarkers for the psychophysiological state estimation for years, and PDM was validated as an effective biomarker for the psychophysiological state assessment [9, 28, 48, 75, 82, 84].

The details about the test results are summarised in Table 3.16.

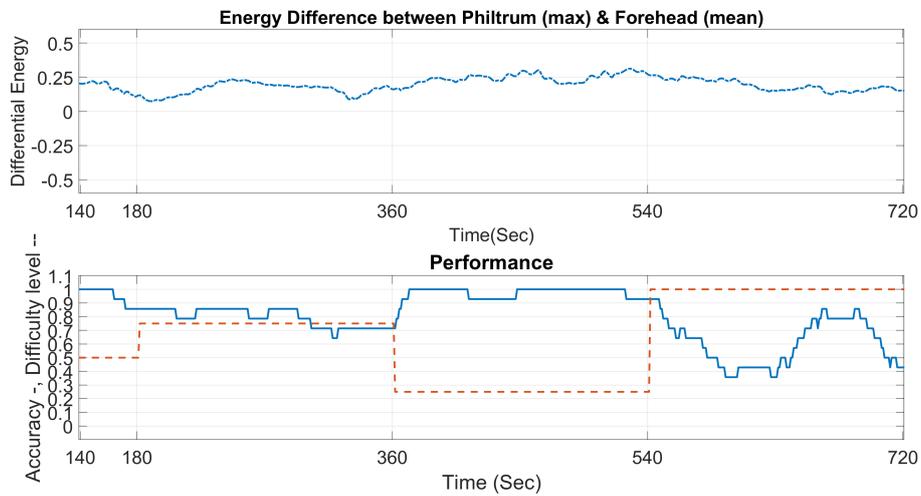
Biomarkers	Incremental session	Random session	Mean
$\bar{T}_n$	0.9667	0.9667	0.9667
$\bar{T}_f$	0.9333	0.9667	0.9500
$T_{maxf}$	0.9667	0.9000	0.9333
DEFP	0.9667	0.9667	0.9667
HRV1	0.9167	0.9333	0.9250
HRV2	1.0000	1.0000	1.0000
TLI1	0.7333	0.8833	0.8083
TLI2	0.8667	0.9333	0.9083
PDM	0.9500	0.8833	0.9167

**Table 3.16:** Mean H Values for T-test Summary

The H values of the temperature markers were all above 0.9 and higher than the scores of the most conventional biomarkers. Therefore, in general, the new



(a) Differential Energy (-.-), accuracy (- -) and difficulty level (-) plots for participant 4. Session 1 with elevated difficulty levels



(b) Differential Energy (-.-), accuracy (- -) and difficulty level (-) plots for participant 4. Session 3 with random difficulty levels

**Figure 3.6:** Differential Energy between Philtrum and Forehead

markers developed from the data relating to the facial temperature exhibited a similar or even a better ability to assess the subjects' psychophysiological state changes in comparison with the traditional biomarkers. This is mainly because the biomarkers based on facial temperature were more sensitive to the minor changes in the psychophysiological state during low work load periods, especially for the maximum facial temperature, the mean nasal temperature and DEFP, as shown in Tables 3.17, 3.18 and Appendix A.

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	0.9000			
DL 3	1.0000	0.9000		
DL 4	1.0000	0.9000	0.8000	

**Table 3.17:** Overall HRV1 T-test Results for Experimental Sessions

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	1.0000			
DL 3	1.0000	1.0000		
DL 4	1.0000	1.0000	1.0000	

**Table 3.18:** Overall HRV2 T-test Results for Experimental Sessions

As already mentioned, the conventional biomarkers based on EEG and ECG had some major limitations - noise from bad electrode-skin connections, the interference of other non HMI relevant activities and the complex equipment setup. These limitations had significantly restricted the efficiency of using conventional biomarkers for assessing human psychophysiological state, see examples in Table 3.17, 3.19 and 3.20. Compared with the facial temperature biomarkers, the statistical data of HRV and TLI showed less diverse characteristics to differentiate each psychophysiological state. However, in Table 3.18 the overall H values for HRV2 were all equal to 1, it is worth knowing that HRV2 is the ratio between the standard deviation and the mean of heart rate. HRV2 was more sensitive to noise which were due to bad electrode-skin connections and muscle movements. As shown in Table 3.21, the efficiency of using PDM was also limited for the same reason - interference of other non HMI relevant activities, i.e. movements of eye

and head. Therefore, a combination of both the new facial temperature biomarkers and the conventional biomarkers should be the optimal solution for assessing the human psychophysiological state. The facial temperature biomarkers were able to compensate for the low accuracy of conventional biomarkers and increase the reliability of the overall observations.

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	0.8500			
DL 3	0.7500	0.8500		
DL 4	0.9000	0.8000	0.7000	

**Table 3.19:** Overall TLI1 T-test Results for Experimental Sessions

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	0.9500			
DL 3	0.8500	0.9500		
DL 4	1.0000	0.9500	0.7500	

**Table 3.20:** Overall TLI2 T-test Results for Experimental Sessions

H value	DL 1	DL 2	DL 3	DL 4
DL 1				
DL 2	0.8500			
DL 3	0.9000	0.9000		
DL 4	0.9500	0.9500	0.9500	

**Table 3.21:** Overall PDM T-test Results for Experimental Sessions

### 3.5 Discussion

The effectiveness and efficiency of using the facial temperature to estimate the subjects' psychophysiological state changes was validated via the experiments, yet in practice these biomarkers were still limited by two major problems: auto-calibration of the camera and the subjects' head movement.

Auto-calibration was an intrinsic design of the camera to deal with the problem of thermal drift in the data recording. The thermal drift was related to the abnormal temperature shifts in the recordings, and it was introduced by the changing

temperature of the camera itself. The process was automatically programmed to calibrate the camera depending on the temperature change of the camera. During the one or two seconds of auto-calibration, the camera measured the temperature within itself rather than the target. As a consequence, the recording was disturbed by the sudden fluctuations that needed to be manually removed during data processing. The lack of actual data in those periods of time affected the ability of the facial temperature indicators to reflect the psychophysiological state at those precise moments. In addition, auto-calibration also significantly interrupted the frequency analysis of the temperature signal. The effect of the calibration was limited by switching on the camera at least ten minutes before each experiment. However, this method only reduced the number of calibration cycle for a 12-minute recording instead of eliminating them completely. Better cameras may well provide more effective solutions to this problem in the future.

Both FLIR ResearchIR and MATLAB<sup>®</sup> were only able to support fixed spatial windows for data extraction. However, the subject's head movements were unavoidable for any long periods of observation. Therefore, spatial windows with fixed positions and fixed shapes were not capable of handling the displacement and distortion caused by these movements. Therefore, due to the failure of tracking the regions of interest, the biomarkers based on the data were not able to faithfully represent the actual temperature changes in those areas consistently, and thus their efficiency was compromised. Unfortunately, in contrast to the object tracking of the normal RGB images, the thermal image lacked enough shape contrast for the normal tracking algorithm to follow. Since there was little research on thermal image tracking and barely any actual algorithm, the participants were advised to be conservative with their head movements, which in this case may be considered as a source of disturbance to the psychophysiological state in human-machine interface, e.g. multitasking.

The tracking of the region of interest seems to represent the toughest challenge among all the other mentioned challenges. However, a new tracking algorithm based on particle-filter might be a useful solution to this problem [20, 22, 50, 90, 101, 102]. The algorithm, built on the Matte algorithm that was based on the pixel

dependence, is able to deal with nonlinear motion within the predict-update cycle in a simple way. Despite the limitations of the current thermal imaging technique, the facial temperature was proved to be a reliable tool for psychophysiological state measurement in HMI.

### **3.6 Summary**

In summary, the experimental results of the facial temperature validated the effectiveness and efficiency of using thermal imaging for psychophysiological state estimation. Such a method proposed a more reliable marker for assessing the psychophysiological state of the operator. Furthermore, the combination of the facial temperature and other well-known biomarkers significantly increased the robustness of the system and the precision of the prediction, as the facial temperature measurement required no body contact and was more sensitive to the changes within the low mental stress states. These findings provided solid support for the modelling of the human operators' psychophysiological state during the human-machine interface in the next chapter.

## Chapter 4

# Adaptive General Type-2 Fuzzy Modelling for Psychophysiological State Prediction

In this chapter, a new type-2 fuzzy-based modelling approach is proposed to assess the human operators' psycho-physiological states for both safety and reliability of the human-machine interface systems. Such a new modelling technique combined type-2 fuzzy sets with state-tracking to update the fuzzy rule-base through a Bayesian process. These new configurations led to an adaptive, robust and explicable computational framework that may be utilised to identify the dynamic (i.e. real-time) features without prior training of implementation. Validated on mental arithmetic cognitive real-time experiments with the psycho-physiological data for ten (10) participants, the proposed framework outperformed other paradigms (e.g. an adaptive neuro-fuzzy inference system and an adaptive general type-2 fuzzy c-means modelling approach) in terms of disturbance rejection and adaptive learning capabilities. Furthermore, good accuracy and performance were obtained via the proposed framework when compared to other models that have been presented in the related literature. In addition, the

new framework may be considered to be a promising development in human-machine interface systems where it may be utilised to (i) estimate the human operator performance in real-time (ii) develop advanced control mechanisms for human-machine interface, (iii) investigate the origins of the compromised human operator performance and (iv) identify and remedy psycho-physiological breakdown at their early stages.

## 4.1 Background

### 4.1.1 Fuzzy Logic Development

The original fuzzy logic was found by Prof. Lotfi Zadeh in 1973 and the first fuzzy sets were proposed in his seminal paper in 1965 [55]. In 1974 and 1975, Mamdani and Assilian made a major breakthrough by implementing a type-1 fuzzy rule-based control to a nonlinear system. Other applications of type-1 fuzzy systems also began to appear, such as in the control of Sendai city subway system and a water treatment system in Japan. Zadeh also introduced the type-2 fuzzy sets as an extension of the type-1 fuzzy sets, which were later developed by Mizumoto, Tanaka, Nieminen, Dubois and Prade [55]. In 1998 and 2001, Karnik and Mendel extended the works of the type-2 fuzzy logic systems by providing two practical algorithms for computation.

### 4.1.2 Fuzzy Logic

Different from the conventional Boolean logic which is based on “true or false” (1 or 0), fuzzy logic computes on “degree of truth”. Fuzzy logic considers 0 and 1 as the two extreme cases of truth and includes various states of truth in between them. In this way, two different forms of knowledge, objective knowledge (e.g. mathematical representations) and subjective knowledge (e.g. linguistic information), were able to be coordinated and synergised within an unified system. The idea was initially introduced by Zadeh in 1965 in the shape of type-1 fuzzy sets [55]. The left plot in Figure 4.1 shows an example of a type-1 fuzzy set. When only the numbers of 0.3,  $x$ , 0.5 and 0.8 are considered in the  $X$  domain, the type-1

fuzzy set may be written as  $\{0/0.3, 0.5/x, 1/0.5, 0/0.8\}$ , where  $0.5/x$  represented that the number  $x$  had a membership degree of 0.5. While, for a crisp set, the membership degree for each element was only either 0 or 1. In addition, fuzzy logic is also able to include the secondary “degree of truth” which describes the reliability of the initial “degree of truth”. This normally forms a type-2 fuzzy set, within which the membership degrees of one or more elements are described by a range values rather than single values. As shown in the second plot of Figure 4.1, the membership degree of the number  $x$  becomes 0.3 to 0.7, with equal (for interval type-2 fuzzy sets) or unequal (for general type-2 fuzzy sets) intensity for each membership degree value.

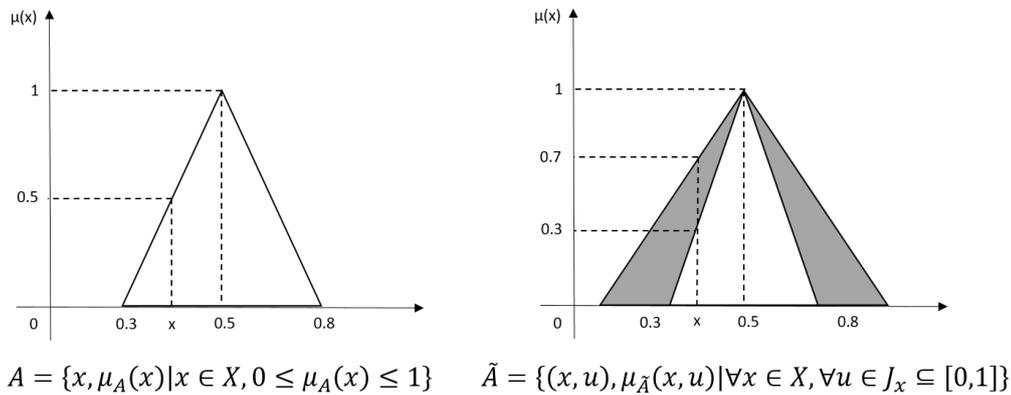
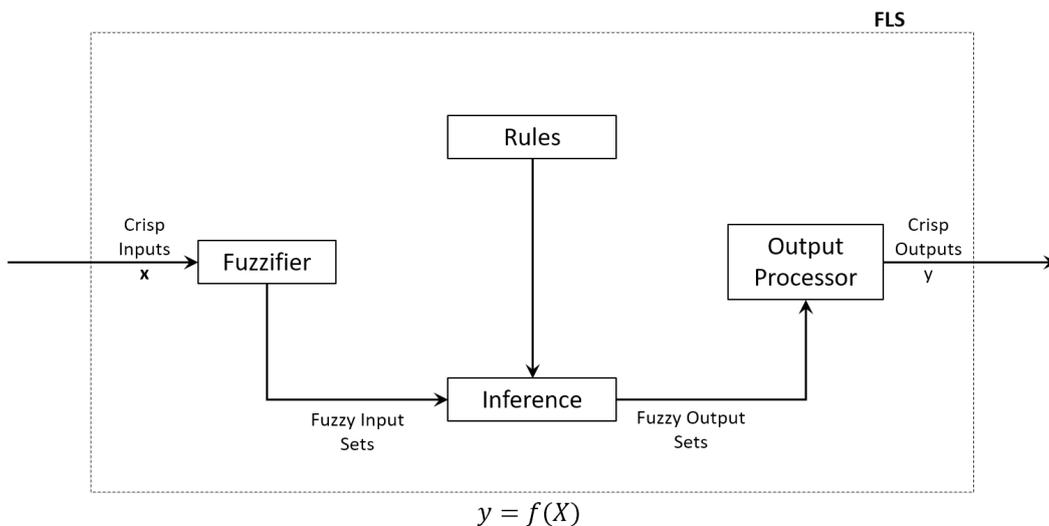


Figure 4.1: Diagram of type-1 fuzzy sets and type-2 fuzzy sets

### 4.1.3 Fuzzy System

A typical fuzzy logic system consists of a fuzzifier, a rule-base, an inference and an output processing [55]. The fuzzifier transfers the crisp inputs into fuzzy input sets. For example, in this research, the readings of biomarkers are transferred into linguistic variables (e.g. ‘Low’ or ‘High’). The rule-base contains a series of rules that can be a summary from previous knowledge or experience. The rules are structured as “IF-THEN” statements, in which the IF part is the antecedent and the THEN part is the consequent. Within the inference, the fuzzy inputs determine which rules are related and are assigned the weights for the corresponding rules, and the final fuzzy output sets are calculated based on the weights and the rules.

For example, when the reading of HRV1 is 0.75, the rules of “if the value of HRV1 is ‘Medium’, then the arousal level is ‘Medium’”, and “if the value of HRV1 is ‘High’, then the arousal level is ‘High’”, are assigned with weights 0.5 and 0.5 correspondingly, and the fuzzy output is something similar to “the arousal level is ‘Medium High’”. In this research, the developed system is based on type-2 fuzzy logic sets, which means that the rules themselves are fuzzy and are described with their own membership functions and weights of corresponding rules are represented with ranges. Each value within the ranges is with different intensity (general type-2 fuzzy set). Therefore, the output processing of the system includes a type-reducer prior to the defuzzifier that transferred the type-1 fuzzy sets to crisp outputs.



**Figure 4.2:** Example of a simple fuzzy logic system

Compared with the conventional systems such as neural network, fuzzy systems share unique advantages as follows:

1. Flexibility: nonlinearity is well described with representations of “degrees of truth”.
2. Simplicity: “IF-THEN” structure unify linear and nonlinear patterns without hard-coding for complex problems.
3. Intuitiveness: the transparent rule-bases are transparent as well as inter-

pretable to human for justifying the system results.

Fuzzy logic deals with uncertainties intrinsically and can handle small sizes of data and sparsity in data better than other machine learning paradigms. Such advantages are even remarkably enhanced for type-2 fuzzy logic systems. However, the complicated base structures of type-2 fuzzy logic systems also lead to additional mathematical inference engines. These may result in negative impacts as follows and limit the application of such systems:

1. Convergence: there may not exist a numerical closed-form solution existed for the problem with current mathematical approaches.
2. Intensity: the available approach may be computationally taxing and time-consuming.
3. Interpretability: the extra layer of “degree of truth” may add difficulty in understanding and the rule-bases lack of transparency.

With regards to these problematic features, several approaches were proposed to restrict the issues while retaining the most advantages [55]. One of the most significant approaches was to create systems based on interval type-2 fuzzy sets, where the simplification of the secondary membership degree provides systems with an opportunity for real-time simulations. In this research, the type-2 fuzzy logic modelling is based on interval type-2 fuzzy sets in the inference process for the purpose of fast computation. However, the element-specified secondary membership degrees were reintroduced during the defuzzification to ensure that the same flexibility and accuracy of the results can be achieved as general type-2 fuzzy logic systems.

It is worth noting that type-2 fuzzy logic systems represent an ongoing research area for system modelling and control with many new approaches addressing the mathematical problems of inference and defuzzification. Detailed explanations and summaries on this specific topic may be found in [55].

#### 4.1.4 Type-2 Fuzzy Logic Modelling

The psychophysiological state of a subject in HMI consists of both inner consciousness and outer behaviour. Therefore, it would be tricky to find the associated conventional mathematical model representations. Existing models and frameworks for predicting the psychophysiological state include ANFIS, Mamdani-type fuzzy model, proportional integral Mamdani fuzzy model, type-2 fuzzy model and support vector machines (SVMs) [23, 28, 39, 52, 61, 81, 84, 97]. ANFIS models often suffer from over-fitting problems because of lack of adaptation. SVM models require to specify kernel functions for individuals to optimise the feature extraction operation. Mamdani-type fuzzy models are transparent and are more efficient to describe the subjective part of the state. Type-2 fuzzy models can achieve the best prediction results for they are able to handle uncertainty with less data requirement and are usually less prone to over-fitting.

Different subjects respond to the same stimuli differently. For example, the same level of psychophysiological pressure may contribute to one's efficiency in completing a given task but compromise others. Such uncertainty requires that the general model be tolerant to individual differences and to recognise the correct correlation between the psychophysiological state and task performance. The systems monitored by the biomarkers were also responsible for other psychophysiological regulations beside HMI. In addition, the unconscious movements of the participants may introduce noise and data loss for the recordings. Therefore, the model must have some level of robustness. The ideal model needs to be implemented for the online real-time process. This requires the model to be adaptive to continuous observations and to self adjust its structures and parameters as the participant's psychophysiological states evolve with time.

Fuzzy logic models are capable of handling a large amount of uncertainty, which is supported with a wide range of previous research studies in the literature [23, 28, 39, 52, 59–61, 80, 81, 83, 84, 97, 98]. A type-2 fuzzy model is capable of dealing with the heuristic or linguistic uncertainty within the system. Additionally, it is also able to tolerant to random uncertainties which limit current predictive approaches. The systems based on type-2 fuzzy sets are effective in

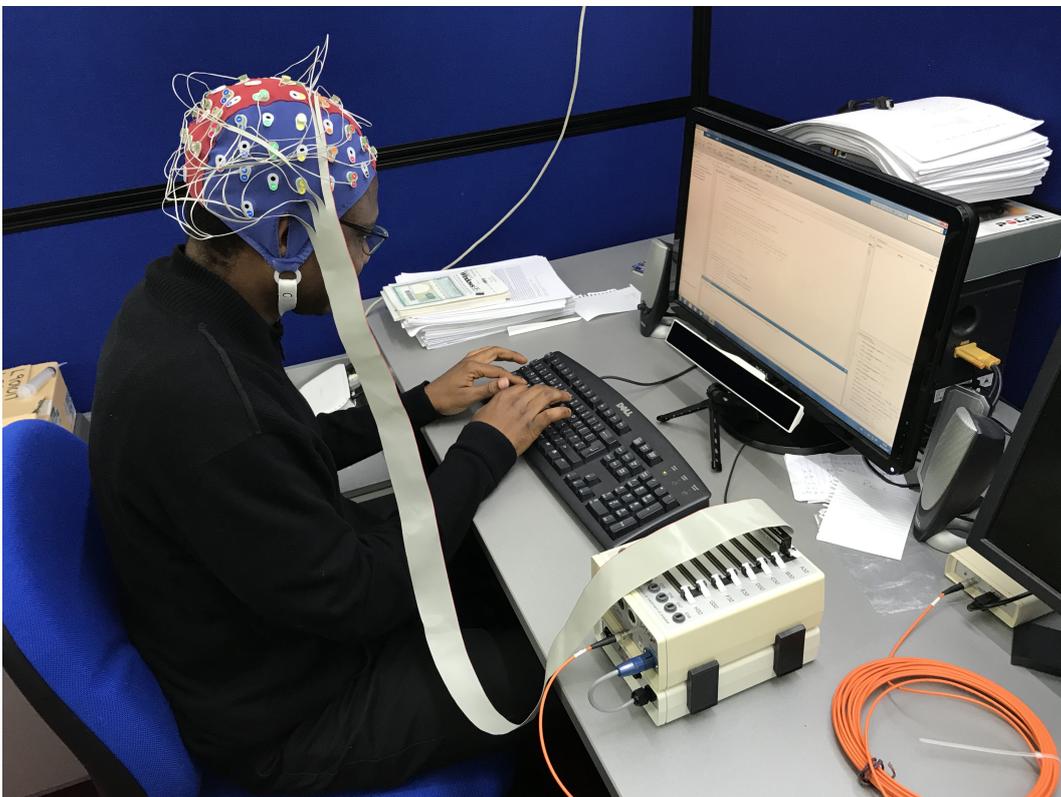
the circumstances where there are the uncertainties in both the fuzzy rule and the measurement [55]. As already mentioned, the human psychological responses to the same stimulus are individual-dependent. Hence, in this research, type-2 fuzzy logic systems are selected to create a generalising model that may be easily interpreted by the human.

## **4.2 Experimental Setup**

### **4.2.1 Data Acquisition**

The data relating to the selected psychophysiological biomarkers for this experiment were collected from four major measurements - EEG, ECG, pupil sizes and facial temperatures. EEG and ECG were recorded by the Biosemi<sup>®</sup> ActiveTwo system with a sampling frequency of 2048 Hz. EEG signals were collected from the 32-channel system according to the standard Biosemi 10/20 layout. ECG signals were collected from the triangle 3-lead system covering the heart area. The initial filtering and reconstruction of signals were processed by the Biosemi<sup>®</sup> ActiView software. Pupil movements were monitored with a Gazepoint<sup>™</sup> eye-tracking camera, and the sizes were calculated by the Gazepoint<sup>™</sup> software. Facial temperatures were based on the infrared imaging recordings from a FLIR E40bx thermal imaging camera at a sampling frequency of 10 Hz. The methods of data recordings followed the experimental frameworks in [23, 35, 81–84]. Figure 4.3 shows a participant during the experiment in the research.

The mental arithmetic applied in this research was based on the Matlab<sup>®</sup> GUI app similar to the one used in [35, 82–84]. The participants in this research were ten (10) healthy students from the University of Sheffield ageing from 22 to 30. The selected subjects included both genders from different countries and backgrounds. The participants were advised to abstain from taking any medicines, coffee or alcohol at least two hours before the experiment to avoid any bias in task performance and psychophysiological measurements.



**Figure 4.3:** A Picture of a participant during the HMI experiment

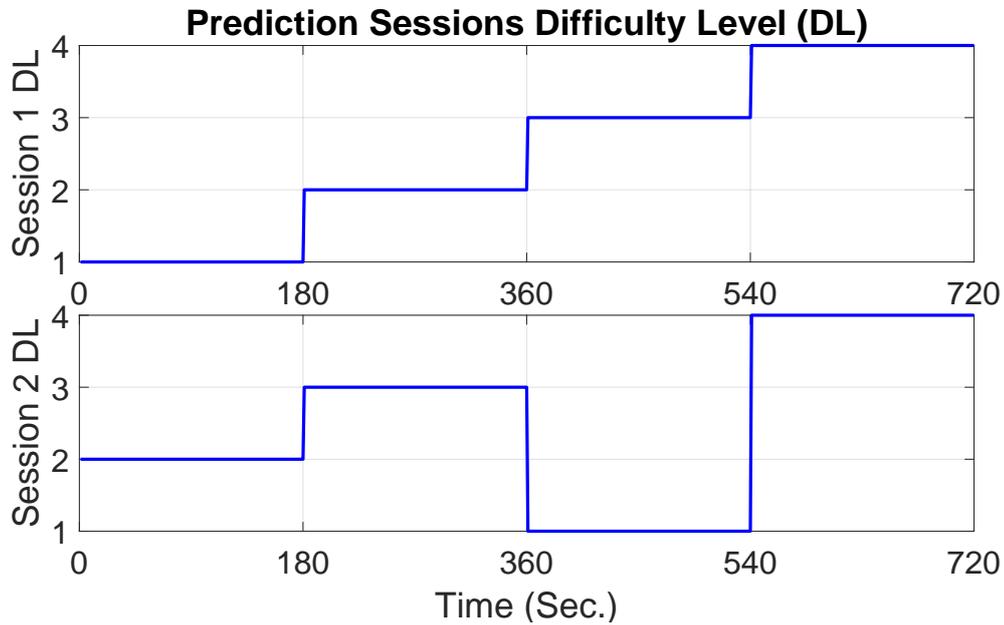
### 4.2.2 Modelling Experiment Configuration

For the modelling configuration, the psychophysiological data were recorded from the participants during the premeditated HMI simulation sessions. The prediction of subjects' task performances was generated by the computational framework of the adaptive general type-2 fuzzy model based on the recordings in real time.

The whole prediction experiment for one subject lasted approximately 30 minutes, including two 12-minute mental arithmetic test sessions and a 5-minute break in the interval. The participants were required to complete a two-number multiplication within a certain amount of time in the mental arithmetic test. In each test session, there were four 3-minute phases with different difficulty levels. The first difficulty level required the subjects to answer the multiplication questions of two random one-digit numbers within ten seconds. Compared with the first difficulty level the second difficulty level only provided five seconds for each question. The third and fourth levels followed the same answering time pattern as the first and second levels except for the switching of the questions to the multiplication of a one-digit number and a two-digit number both randomly generated. The order of the difficulty levels was different between two mental arithmetic sessions for checking the adaptiveness of the model (see Figure 4.4).

## 4.3 Offline Modelling

The previous research studies had showed that the conventional data driven modelling approaches were insufficient for predicting human operator's psychophysiological state [52, 60, 61, 80, 81]. The uncertainty of HMI was considered as being the major cause of limited performance of these models. Compared with the training group, the performance of these offline models in the testing group constantly dropped beyond the expectations. It suggested that a considerable amount of uncertainty was not included during the modelling. Based on the previous psychophysiological literature, it was known that the human performance of a certain task was a collective results of multiple biological processes and instant environment factors. Such stochastic system significantly complicated demands for the



**Figure 4.4:** Difficulty level for prediction experiment session 1 (progressive increase) and session 2 (random)

model structure and base function. In addition, considering the timing, human factor and other environment factors, each experiment is irreproducible. Therefore, it is worth noting that the feasible data for modelling was always insufficient to describe the complete dynamic of the HMI. For modelling, the uncertainty within the data may be solved with complex model and advanced optimisation, while the uncertainty outside the data required the model to be adaptive in real time. It is worth noting that the complexity and adaptability of a model conflicts at a given time and computing power. For the purpose of balancing the optimisation and adaptation, complex fuzzy inference modelling was applied to investigate the main source of uncertainty within the HMI. Complex fuzzy logic inference employed a simplified type-2 fuzzy modelling approach that ensured high computational efficiency with enhanced nonlinear pattern recognition. The test results should be able to identify whether the compromised prediction performance of conventional offline models and frameworks came from limited complexity bound by existing modelling approaches or from the limited information presented in the feasible laboratory data.

### 4.3.1 Complex Fuzzy Sets and Logic

Complex fuzzy sets were initially proposed by Daniel Ramot etc. in 2002 [70, 71], and their corresponding logic in modelling was implemented in 2005 by Scott Disk [18]. Compared with the traditional fuzzy membership functions, the range based on complex fuzzy sets extend the membership functions from  $[0, 1]$  to an unit circle on the complex plane. Thus, with the expression of membership in a set of complex numbers, the complex fuzzy sets provide a secondary dimension for the uncertainty. In contrast to a traditional fuzzy set, the primary membership in the value range  $[0, 1]$  is retained as the amplitude or the real part of the membership grade, while the secondary membership is represented with the phase or the imaginary part of the membership grade. The following equations represent a complex fuzzy set  $S$ , defined on a universe of discourse  $U$  with Euler's formula:

$$S = \{(x, \mu_s(x)) | x \in U\}, \quad (4.1)$$

$$\mu_s(x) = r_s e^{jw_s(x)},$$

where, for any element  $x$  within the set, its corresponding membership function  $\mu_s(x)$  is characterised by the magnitude  $r_s$  and the phase  $w_s$ . By definition, the real number  $r_s$  represents the primary membership within the range  $[0, 1]$ , while the secondary membership is determined by the angular value  $w_s$  and  $j^2 = -1$ .

### 4.3.2 Single Partition Complex Fuzzy Inference Modelling

Most current models and frameworks developed using complex fuzzy logic have achieved a high accuracy and a reliable performance, especially when dealing with periodic data [15, 79, 92–94]. Compared to the general type-2 fuzzy logic systems, complex fuzzy logic system retains the secondary information of uncertainty through imaginary dimension. Thus, the original surface represented by the general type-2 fuzzy logic sets in the 3 dimensional space is simplified to be a trajectory. Therefore, complex fuzzy logic systems benefit from the simplicity of using complex numbers in inference while being able to handle the intra and inter uncertainty better. A complex fuzzy logic based system is commonly combined

within neural networks to form models and frameworks that are able to fulfill the following requirements:

1. Comprehensive - be able to handle both numerical data and linguistic data
2. Intuitive - generate simple fuzzy rules and model structure for interpretation
3. Efficient - ensure rule inferences in parallel

In this research, a single partition complex fuzzy inference modelling is applied. This approach generates a single feature partition for each rule that consists of the antecedents in the form of type-1 fuzzy Gaussian membership function and the consequences with complex fuzzy singleton membership functions. This structure guarantees transparency and interpretability of the relationship between the phase variances of the consequences and the feature partitions of the antecedents. By definition, the interaction within different partition may be represented by the straightforward vector interference. Such arrangement avoids the linear growth of rule-based size with additional features with the grid partition method. Consequently, a clustering algorithm for generating the rule-base of the model may be simplified, and the computational requirement of model implementation is reduced.

For each partition in a feature of antecedents in the single partition complex fuzzy inference model, the Gaussian membership function is as follows:

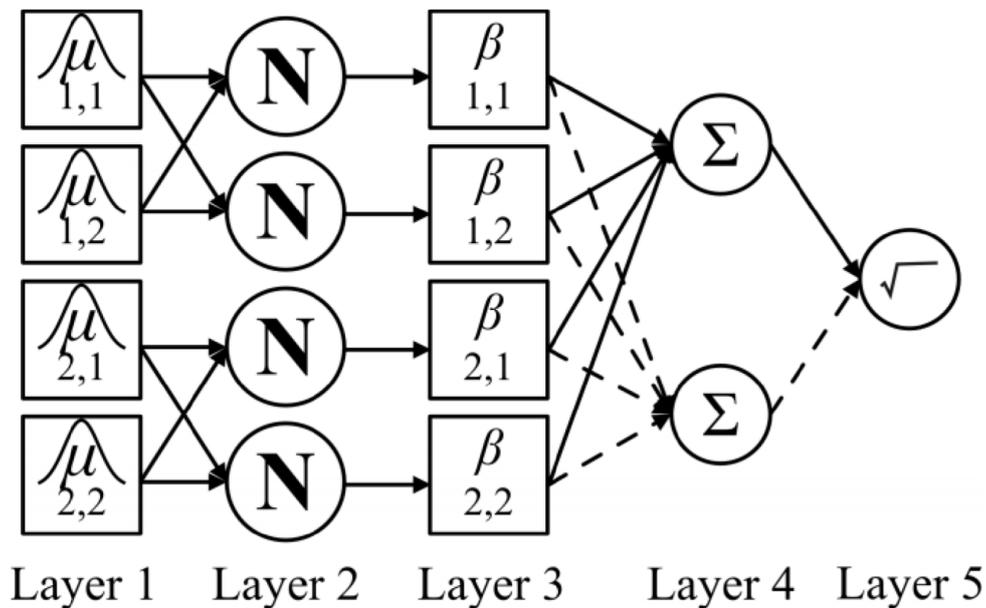
$$\mu_s(x) = e^{-(x-c)^2/2\sigma^2}, \quad (4.2)$$

where  $c$  and  $\sigma$  are the centre and the standard deviations of corresponding partition respectively. For each consequences, the complex singleton membership function is as follows:

$$\begin{aligned} \beta_s(x) &= r_s e^{jw_s(x)}, \\ \beta_s^{Re}(x) &= r_s \cos(w_s(x)), \\ \beta_s^{Im}(x) &= jr_s \sin(w_s(x)), \end{aligned} \quad (4.3)$$

where  $r_s$  is the magnitude of corresponding consequence and  $w_s(x)$  is the phase determined by the inferences between input  $x$  and partitions.  $\beta_s^{Re}(x)$  represents the real part of the vector for final result computation, while the imaginary part  $\beta_s^{Im}(x)$  is stored for additional improvement on the transparency of the model.

The single partition complex fuzzy inference model implemented in this research is based on the Mamdani type fuzzy rule-base with singleton defuzzification. Figure 4.5 presents the structure of models that resembles the radial basis function neural network. A group of membership degrees is generated through the first fuzzification layer with every partition membership function of each feature. The second layer provides the normalisation for the membership degrees. The third layer performs the algebraic product as the implication operation. The normalised membership degree is multiplied with the consequence complex singleton membership function. In the fourth layer, the complex products of the third layer are separated and summed up in the real and imaginary group respectively. The last layer calculates the magnitude of the complex number formed from the fourth layer as the final output.



**Figure 4.5:** The structure of single partition complex fuzzy inference model

Each feature is to be partitioned into 4 partitions and is labelled with corre-

sponding Gaussian membership functions mentioned above, with the intersection set at the 0.5 membership value. The initial magnitudes of the consequence are assigned with the partial correlation coefficients from analysing the training data. The initial phases of the consequence are obtained by partitioning  $2\pi$  evenly. The optimisation of the real number parameters  $c$ ,  $\sigma$  and  $w$  are based on Levenberg-Marquardt optimisation method and the iteration is set to be 1000 times. The established models are trained with the recording of the first session and are tested on the data of the second session. In this research, the input vector of the model was  $I(t) = [HRV_1(t), HRV_2(t), TLI_1(t), TLI_2(t), PDM(t), \bar{T}_n(t), \bar{T}_f(t), T_{maxf}(t), DL(t)]$  and the corresponding output was performance prediction. The models were trained and tested individually according to different subjects. A model was trained with the data from session 1 (incremental difficulty level) of one single participant and then was tested with the data from session 2 (randomised difficulty level) of the same participant. After the training, the final model consisted of 36 rules. Table 4.1 shows an example rule from a final single partition complex fuzzy inference model based on Gaussian membership function.

**Table 4.1:** Example rule of single partition complex fuzzy inference model after training

	$HRV_1(t)$	$HRV_2(t)$	$TLI_1(t)$	$TLI_2(t)$	$PDM(t)$
Weight	-1.205	-1.461	0.619	3.224	0.445
Angle	-0.542	5.128	5.133	5.926	3.119
Sigma	0.100	0.158	0.010	0.654	0.039
Center	0.053	0.983	0.972	2.134	0.970
	$\bar{T}_n(t)$	$\bar{T}_f(t)$	$T_{maxf}(t)$	$DL(t)$	Performance
Weight	0.249	0.228	-1.791	0.573	-1.204
Angle	3.821	5.014	4.311	4.874	-0.542
Sigma	0.052	0.013	0.512	0.097	0.100
Center	0.901	0.752	1.846	0.923	0.053

### 4.3.3 Evaluation of the Single Partition Complex Fuzzy Inference Model

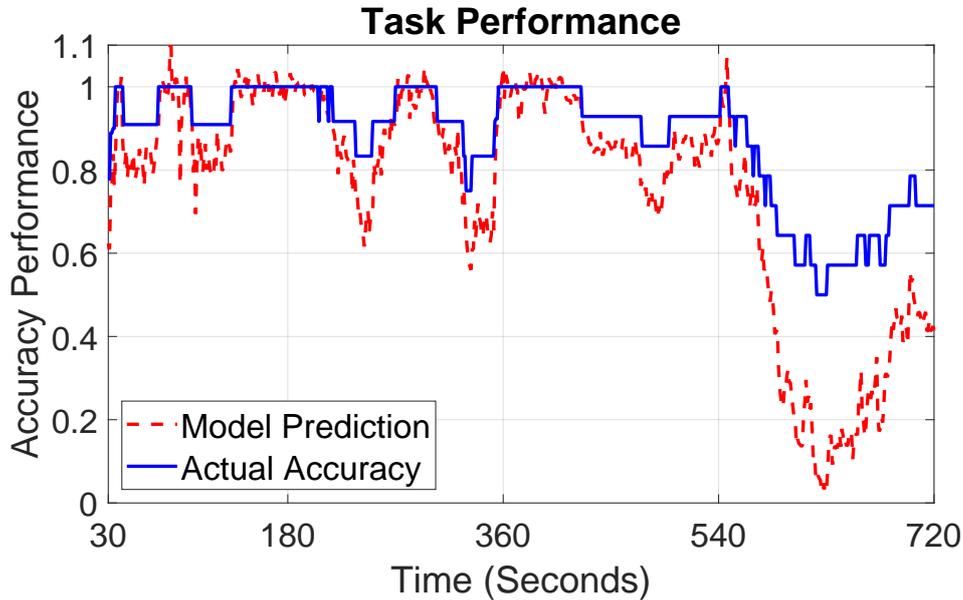
Assessments of the single partition complex fuzzy model were based on the Pearson Correlation and the Root Mean Squared Error. The details of the computation are presented in the next section. Table 4.2 summarises the model performance

**Table 4.2:** Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of single partition complex fuzzy inference model

Participant	Correlation (%)		Root Mean Squared Error (%)	
	Training	Testing	Training	Testing
01	95.823	63.874	9.994	32.301
02	98.764	41.122	18.666	39.560
03	99.217	55.451	13.943	28.610
04	99.565	47.223	6.522	27.145
05	98.121	7.444	10.640	38.767
06	99.360	52.118	7.196	24.992
07	98.859	74.056	13.778	35.881
08	98.744	59.939	18.384	33.754
08	98.363	61.440	18.085	38.443
10	99.033	72.110	12.446	34.157
Mean	98.551	51.192	12.877	33.080

for all ten participants in the HMI simulation experiment. The model for each participant is trained with the data of incremental difficulty session and is tested with randomised difficulty sessions. As shown in this table, the performance of a single partition complex fuzzy model varied significantly between the training data and the testing data. The correlation between the predictions and actual performance dropped from 98.55% to 51.19%, while the root mean squared error rose from 12.87% to 33.08%. These results matched the performances of similar offline models and frameworks applied in this field previously [83, 84].

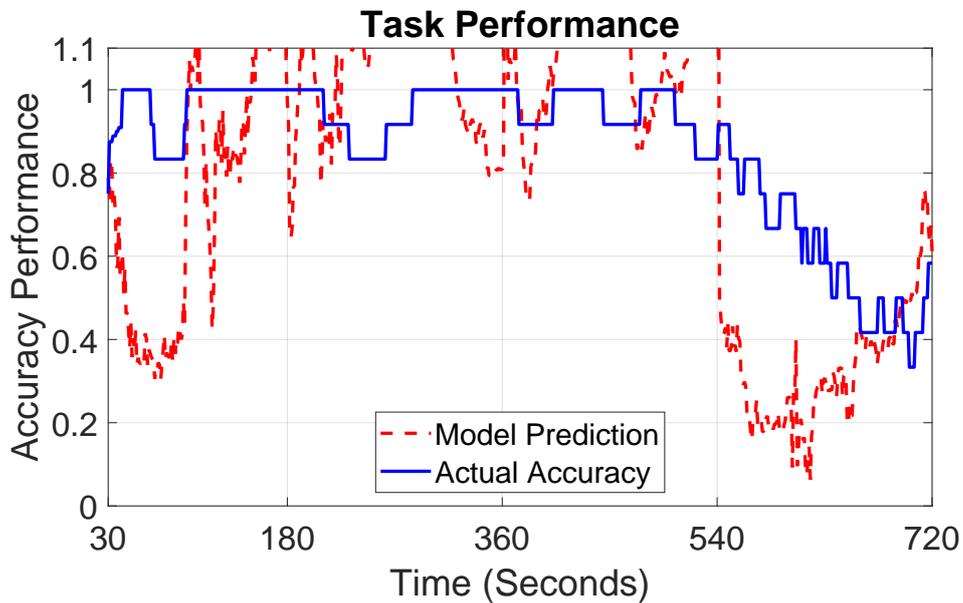
Figures 4.6 and 4.7 presents the comparisons of the single partition complex fuzzy model prediction and the actual performance of the participant 08 in both sessions. It may be seen in Figure 4.6 that the model prediction roughly matched with the actual accuracy performance. It is reasonable to assume that the model may be further improved with more iterations of optimisation or more advanced optimisation techniques. However, the Root Mean Square Standardised Error of models in the training was around 0.863 which less than 1. Suffice to say that for the training session the variability considered by the single partition complex fuzzy model closely matched the variability of data. This suggested that the the nonlinearity within the HMI was not the main reason for the compromised performance of offline models and frameworks.



**Figure 4.6:** Accuracy performance from the participant 08 and from the single partition complex fuzzy inference model prediction in training session

As shown in Figure 4.7, the model prediction failed to represent the actual performance based on the recordings of testing group. It is worth noting that the the Root Mean Square Standardised Error of the models increased from around 0.863 in the training sessions to around 2.043 in the testing sessions. The single partition complex fuzzy models implemented are offline models with the fixed parameters. Therefore, it may be found that the variability of the data changes between the training sessions and testing sessions. These changes may due to the changes of difficulty level order, environmental factors and participants' psychophysiological state. The model performance of the testing group clearly shows that these time-varying changes reshaped the participants' behaviour in the HMI the cannot be simply neglected. Meanwhile, it also indicates that the patterns presented in the training data and the testing data lacked resemblance.

Similar offline model failures may be found in [83]. These offline type-2 fuzzy logic based modelling results had clearly showed that the type-2 fuzzy sets were capable of representing the nonlinearity relationship between the biomarkers' recordings and the participants' performance in the HMI. However, these patterns varied with time because of the uncontrollable human psychophysiological states,

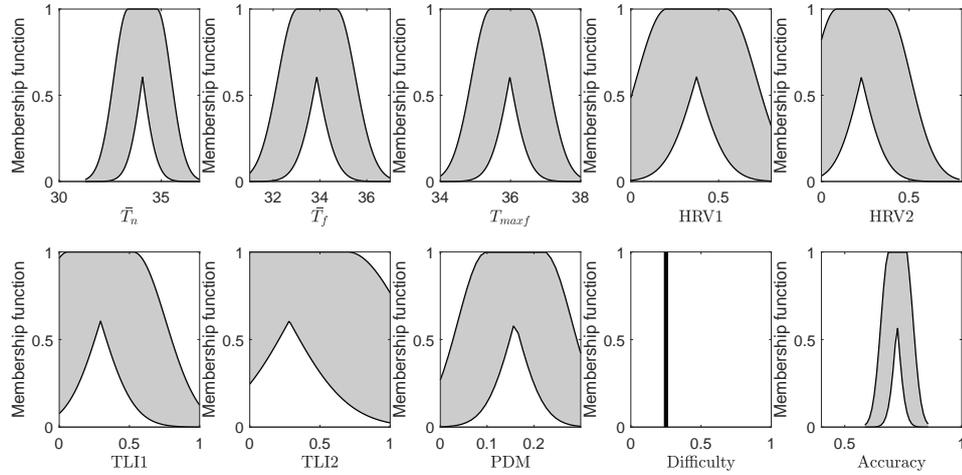


**Figure 4.7:** Accuracy performance from the participant 08 and from the single partition complex fuzzy inference model prediction in testing session

e.g. valence level, and environment factors, such as timing. Therefore, in order to achieve the acceptable accuracy in the prediction, it was important to create online models that can be adaptive to these time-varying states and can ultimately provide predictions based on updated patterns.

#### 4.4 Adaptive General Type-2 Fuzzy Modelling

An adaptive general type-2 fuzzy framework was selected and devised to exploit the advantages of type-2 fuzzy logic to handle the intra and inter uncertainty while achieving fast adaptive learning with Bayesian theory. The model first predicted the trend of the subject's future performance based on the latest biomarker recordings. The type-2 fuzzy framework used the psychophysiological data as the input and applied the centre-of-sets (COS) type-reduction based method to generate the initial prediction of each fired fuzzy rules. The final prediction combined all the predictions with the performance trend. When the final prediction was varying from the actual observed value, the selected fuzzy rule from the rule-base was updated with the observation recursively via Bayesian theory [43].



**Figure 4.8:** Example rule of GT2FM for the first state

One example fuzzy rule of the adaptive general type-2 fuzzy model (GT2FM) is illustrated in Figure 4.8, where the shaded area represented the footprint of the first degree uncertainty, and its sample corresponding linguistic form reads as follows:

**Rule 1:** *IF  $\bar{T}_n$  is large and  $\bar{T}_f$  is medium and  $T_{maxf}$  is small and  $HRV_1$  is small and  $HRV_2$  is large and  $TLI_1$  is large and  $TLI_2$  is small and  $PDM$  is large, THEN the task accuracy is small.*

The primary degree of uncertainty of the fuzzy rule remains fixed during the inference, whereas the secondary degree of uncertainty varies with the current state estimation. Table 4.3 summarises two sample fuzzy rules of the adaptive general type-2 fuzzy model, which describes the relationships between the inputs and output under the condition of the same difficulty level. The linguistic labels applied in the fuzzy rules are illustrated in Table 4.4.

The following steps explain the inference mechanism leading to an output:

1. Calculate the latest transition matrix  $P$  for each state; in a period of time, different task performance ranges indicate different  $K$  states of the subject (the states describe the performance in descending order). The entries in the first row of the transition matrix  $P$  represents the probability for the

**Table 4.3:** Rule-Base of GT2FM for the First State

Inputs					
	$\bar{T}_n$	$\bar{T}_f$	$T_{maxf}$	$HRV_1$	$HRV_2$
Rule 1	Large	Medium	Small	Small	Large
Rule 2	Medium	Large	Medium	Large	Small

Inputs				Output	
	$TLL_1$	$TLL_2$	$PDM$	$DL$	$TaskAccuracy$
Rule 1	Large	Small	Large	0.25	Small
Rule 2	Medium	Medium	Small	0.25	Large

**Table 4.4:** Linguistic Labels of the Inputs and Output for the First State

Linguistic Labels	$\bar{T}_n$	$\bar{T}_f$	$T_{maxf}$	$HRV_1$	$HRV_2$
Small	<33.5	<33.9	<36.0	<0.37	<0.18
Medium	33.5-34.1	33.9-34.1	36.0-36.1	0.37-0.53	0.18-0.23
Large	>34.1	>34.1	>36.1	>0.53	>0.23

Linguistic Labels	$TLL_1$	$TLL_2$	$PDM$	$TaskAccuracy$
Small	<0.18	<0.28	<0.14	<0.72
Medium	0.18-0.29	0.28-0.30	0.14-0.16	0.72-0.97
Large	>0.30	>0.30	>0.16	>0.97

adjacent two states 0 & 2 and the same state 1 following the first state 1 (task performance measurement is continuous), and similarly for the remaining rows:

$$P = \begin{bmatrix} P_{(1,0)} & P_{(1,1)} & P_{(1,2)} \\ P_{(2,1)} & P_{(2,2)} & P_{(2,3)} \\ \vdots & & \\ P_{(K,K-1)} & P_{(K,K)} & P_{(K,K+1)} \end{bmatrix}, \quad (4.4)$$

$$P_{(S_n=i, S_{n+1}=j)} = P_{(S_{n+1}=j|S_n=i)} \cdot P_{(S_n=i)}, \quad (4.5)$$

where  $P_{(i,j)}$  is estimated via Equation 4.5 in a certain amount of time, except  $P_{(1,0)} = P_{(K,K+1)} = 0$ .  $P_{(S_n=i)}$  represent the total probability of initial state  $i$ , and  $P_{(S_{n+1}=j|S_n=i)}$  represent the probability of state  $j$  given the initial state  $i$ .

2. Perform the state estimation for the current time  $t$  with the transition matrix  $P$ ; The following gives the expectation of each state to be presented at the time  $t$ :

$$E = \begin{bmatrix} E_1 & E_2 & \dots & E_K \end{bmatrix}, \quad (4.6)$$

$$E_t = E_{t-1} \cdot P_{t-1}, \quad (4.7)$$

where  $E_{(i,j)}$  denotes the expectation of the state  $j$  at the time  $i$ .

3. Compute the upper and lower membership functions  $\bar{F}$  &  $\underline{F}$  for firing the fuzzy rules; The rule-base  $R$  consists of the  $4K$  fuzzy rules describing each state with  $M$  different difficulty levels:

$$R = \begin{bmatrix} R_{(1,1)} & \cdots & R_{(1,K)} \\ \vdots & \ddots & \\ R_{(M,1)} & & R_{(M,K)} \end{bmatrix}, \quad (4.8)$$

where  $R_{(i,j)}$  represents the fuzzy rule describing the state  $j$  at the difficulty level  $i$ . The input values of each rule are range values with means  $\mu_x$  and standard deviations  $\sigma_x$ , representing the measurement uncertainty and individual difference. The firing of the fuzzy rules involves the  $K$  fuzzy rules covering every state of the participants for the difficulty levels that they currently experience. The inference between the input values and one firing fuzzy rule depends on the Gaussian functions, and for simplification, the functions has the same standard deviation value from that rule. The following summarises the inference processes to find the upper and lower membership functions  $\bar{f}$  &  $\underline{f}$  for one single input  $x$  and one firing fuzzy rule:

- (a) if  $x < \mu_x - \sigma_x$ , then find  $m$  and  $n$  that satisfy:

$$\begin{aligned} f(m|x, \sigma_x^2) &= f(m|\mu_x + \sigma_x, \sigma_x^2), \\ f(n|x, \sigma_x^2) &= f(n|\mu_x - \sigma_x, \sigma_x^2), \end{aligned} \quad (4.9)$$

where:

$$f(\bar{x}|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\bar{x}-\mu)^2}{2\sigma^2}}, \quad (4.10)$$

which gives:

$$\begin{aligned} \underline{f} &= \frac{f(m|x, \sigma_x^2)}{f(x|x, \sigma_x^2)}, \\ \bar{f} &= \frac{f(n|x, \sigma_x^2)}{f(x|x, \sigma_x^2)}. \end{aligned} \quad (4.11)$$

(b) if  $\mu_x - \sigma_x \leq x \leq \mu_x + \sigma_x$ , then find  $l$  that satisfy:

$$f(l|x, \sigma_x^2) = f(l|\mu_x, \sigma_x^2), \quad (4.12)$$

which gives:

$$\begin{aligned} \underline{f} &= \frac{f(l|x, \sigma_x^2)}{f(x|x, \sigma_x^2)}, \\ \bar{f} &= 1; \end{aligned} \quad (4.13)$$

(c) if  $\mu_x + \sigma_x < a$ , then find  $m$  and  $n$  that satisfy:

$$\begin{aligned} f(m|x, \sigma_x^2) &= f(m|\mu_x - \sigma_x, \sigma_x^2), \\ f(n|x, \sigma_x^2) &= f(n|\mu_x + \sigma_x, \sigma_x^2), \end{aligned} \quad (4.14)$$

which gives:

$$\begin{aligned} \underline{f} &= \frac{f(m|x, \sigma_x^2)}{f(x|x, \sigma_x^2)}, \\ \bar{f} &= \frac{f(n|x, \sigma_x^2)}{f(x|x, \sigma_x^2)}. \end{aligned} \quad (4.15)$$

This way, the final membership functions for one firing fuzzy rules  $\bar{F}$  &  $\underline{F}$  are:

$$\begin{aligned} \underline{F} &= \left\{ \max(\underline{f}_i) | \forall i \in L \right\}, \\ \bar{F} &= \left\{ \max(\bar{f}_i) | \forall i \in L \right\}, \end{aligned} \quad (4.16)$$

where  $L$  represents the number of the inputs.

4. Find the initial prediction of each firing fuzzy rule with the transition matrix  $P$  and the membership functions  $F$ . Similarly to the input values, each fuzzy rule has a range for the output values  $\forall y \in [\underline{y}, \bar{y}]$ . The following summarises the type-reduction processes to find the prediction value  $y_k$  for one firing fuzzy rule  $k$  with the transition matrix  $P$  and membership function  $F$  (sort  $\underline{y}$  &  $\bar{y}$  in ascending order):

(a) if  $P_{(k,k-1)} < P_{(k,k+1)}$  or for the fuzzy rule representing the last state

$P_{(k,k-1)} < P_{(k,k)}$ , then

$$y_k = \frac{\sum_{n=1}^k \bar{F}^n \cdot \underline{y}^n + \sum_{n=k+1}^K \underline{F}^n \cdot \underline{y}^n}{\sum_{n=1}^k \bar{F}^n + \sum_{n=k+1}^K \underline{F}^n}, \quad (4.17)$$

(b) if  $P_{(k,k+1)} < P_{(k,k-1)}$  or for the fuzzy rule representing the first state

$P_{(k,k+1)} < P_{(k,k)}$ , then

$$y_k = \frac{\sum_{n=1}^{k-1} \underline{F}^n \cdot \bar{y}^n + \sum_{n=k}^K \bar{F}^n \cdot \bar{y}^n}{\sum_{n=1}^{k-1} \underline{F}^n + \sum_{n=k}^K \bar{F}^n}, \quad (4.18)$$

(c) if  $P_{(k,k-1)} = P_{(k,k+1)}$  or for the fuzzy rule representing any middle state

$P_{(k,k)} > \max(P_{(k,k-1)}, P_{(k,k+1)})$  or for both fuzzy rules representing two

end states  $P_{(k,k)} = P_{(k,k-1)} + P_{(k,k+1)}$ , then

$$\begin{aligned} y_{(k,l)} &= \frac{\sum_{n=1}^k \bar{F}^n \cdot \underline{y}^n + \sum_{n=k+1}^K \underline{F}^n \cdot \bar{y}^n}{\sum_{n=1}^k \bar{F}^n + \sum_{n=k+1}^K \underline{F}^n}, \\ y_{(k,h)} &= \frac{\sum_{n=1}^k \underline{F}^n \cdot \bar{y}^n + \sum_{n=k+1}^K \bar{F}^n \cdot \underline{y}^n}{\sum_{n=1}^k \underline{F}^n + \sum_{n=k+1}^K \bar{F}^n}, \\ y_k &= \frac{y_{(k,l)} + y_{(k,h)}}{2}. \end{aligned} \quad (4.19)$$

The main idea of the type-reduction algorithm is to keep the prediction consistently corresponding to the tendency measured from the state tracking. Take Equation (4.18) for example,  $y_k$  is a maximised prediction. Since the probability of switching to a better state dominates, the likelihood for  $\bar{y}^n$  from a back state decreases, and for  $\bar{y}^n$  from a front state increases. For  $n < k$ , the  $y_k$  calculation uses the lower membership weights; for  $n \geq k$ , the  $y_k$  calculation uses the upper membership weights. This algorithm ensures that the prediction is maximised by the transition matrix.

5. Generate the final prediction  $\hat{y}_t$  from the state estimations  $E_t$  and the initial predictions  $y_{(t,k)}$ ; The following gives the final prediction of the model at the time  $t$ :

$$Y_t = \left[ y_{(t,1)} \quad y_{(t,2)} \quad \cdots \quad y_{(t,K)} \right], \quad (4.20)$$

$$\hat{y}_t = Y_t \cdot E_t^T, \quad (4.21)$$

where  $Y_t$  donates the set of all individual predictions from every firing fuzzy rule at the time  $t$ .

The adaptive general type-2 fuzzy modelling algorithm used two fuzzy membership sets for the prediction computation. The primary memberships represent the individual difference and measurement uncertainty. The type-reduction of the primary membership weights is based on the participant state tracking, which forms the secondary membership sets of the model. This membership is computed with the comparison between the input vector to the selected fuzzy rules, and the prediction is generated based on the latest state information. The modelling algorithm utilises a simplified inference to combine the statistical estimation and the fuzzy logic mechanism. Thus, it takes into account the data uncertainty and aligns this uncertainty with the forecast without a computationally expensive type-reduction algorithm that limits the use of general type-2 fuzzy logic sets [55] nowadays. In addition, the intra-uncertainty is integrated with the framework by a simplified learning algorithm. Intra-uncertainty developed with time and gradually reduces the reliability of the model. Therefore, an adaptive learning algorithm based on Bayesian theory [43] is implemented for updating the rule-base.

The adaptive learning algorithm follows the following steps:

1. Calculate the prediction error and check it is within the maximum error tolerance  $ET_{max}$ . The adaptive learning algorithm is only applied if the error between the prediction  $\hat{y}_{t-1}$  and the observation  $o_{t-1}$  at the time  $t-1$  exceeds the limitation:

$$\|\hat{y}_{t-1} - o_{t-1}\| > ET_{max} \quad (4.22)$$

2. When the learning algorithm is needed, update the selected fuzzy rule with the observation  $o_{t-1}$  using the Bayesian theory [43]. The selected fuzzy rule has the same subject state under the same difficulty level as the observation  $o_{t-1}$ , with the means  $\mu_{t-1}$  and the standard deviations  $\sigma_{t-1}$  for the inputs and output. The observation  $o_{t-1}$  is described with the Gaussian functions

via the means  $\mu_{o,t-1}$  and the standard deviations  $\sigma_{o,t-1}$ . Since the prior (rule) and the posterior (observation) are of the same type (both Gaussian), the posterior mean and the posterior standard derivation of the conjugate prior for the normal distribution are calculated as follows:

$$\begin{aligned} E(\mu_t|\mu_{o,t-1}) &= \frac{\sigma_{t-1}^2 \cdot \mu_{t-1} + \sigma_{o,t-1}^2 \cdot \mu_{o,t-1}}{\sigma_{t-1}^2 + \sigma_{o,t-1}^2}, \\ \text{Var}(\sigma_t|\sigma_{o,t-1}) &= \frac{\sigma_{t-1}^2 \cdot \sigma_{o,t-1}^2}{\sigma_{t-1}^2 + \sigma_{o,t-1}^2}, \end{aligned} \quad (4.23)$$

where  $\sigma_{o,t-1}$  are equal to the initial values of standard deviations of the fuzzy rule, considering that the individual difference and the measurement uncertainty are time-independent.

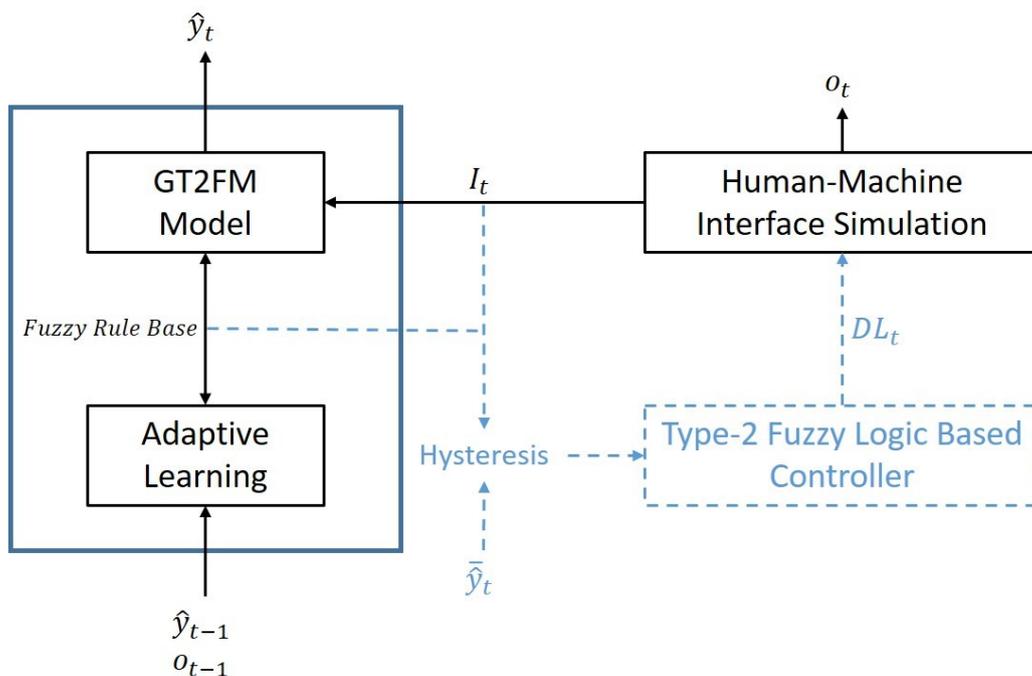
3. Calculate the distance between the new fuzzy rule and the observation and checked it with the maximum distance tolerance  $DT_{max}$ :
  - (a) if  $\|\mu_t - \mu_{o,t-1}\| > DT_{max}$ , replace  $\mu_{t-1}$  &  $\sigma_{t-1}$  with  $\mu_t$  &  $\sigma_t$  and repeat the Bayesian update;
  - (b) else if  $\|\mu_t - \mu_{o,t-1}\| \leq DT_{max}$ , stop learning algorithm and replace the old fuzzy rule with the new one.

## 4.5 Evaluation of Adaptive General Type-2 Fuzzy Modelling

This section focuses on the prediction results of the adaptive general type-2 fuzzy framework mentioned in the previous section. The fuzzy model is implemented in the HMI mental arithmetic experiments for online real-time experiments. This section also includes the prediction results of a generalised offline ANFIS model and the real-time A-GT2-FCM model based on the same experimental data for comparisons. The evaluations and summaries of the experiment results in the following should be able to demonstrate the performances of the system.

### 4.5.1 Model Configuration

The adaptive GT2FM is built using the computational frameworks of MATLAB<sup>®</sup>. The HMI mental arithmetic experiment includes four different difficulty levels  $M = 4$ . In this research, the input vector for the system was  $I(t) = [HRV_1(t), HRV_2(t), TLL_1(t), TLL_2(t), PDM(t), \bar{T}_n(t), \bar{T}_f(t), T_{maxf}(t), DL(t)]$ . The corresponding output of the HMI simulation system is the Actual Accuracy  $o(t)$ , which is also recorded as being the Observation for the prediction at time  $t + 1$ . The prediction output of the model is the Predicted Accuracy  $\hat{y}(t)$ . Figure 4.9 shows the diagram of the GT2FM that is designed for the whole HMI simulation.



**Figure 4.9:** Diagram of the GT2FM for the HMI simulation experiment (controller and other sections in - - will be the subject of the next chapter, chapter 5)

The GT2FM model starts to generate the prediction 30 seconds after the experiment begins. The computational framework consists of eight (8) fuzzy rules in total, which leads to two (2) fuzzy rules per difficulty level and divides the operator task performance into two states. The Root Mean Square Standardised Error  $e_{rmsse}$  is introduced to evaluate the complexity of the model and to compare

it with the real HMI simulation results, as follows:

$$e_{rmsse} = \sqrt{\frac{\sum_{i=1}^n [(\hat{y}_i - o_i) / \hat{\sigma}_o]^2}{n}}, \quad (4.24)$$

where  $\hat{\sigma}_o$  is the standard error of the observations and  $n$  is the length of the data. The mean error for ten participants is  $e_{RMSE} = 0.1376 < 1$ , which indicated the variability of the model should suffice the HMI prediction and there is no need to divide any extra state further for introducing more fuzzy rules.

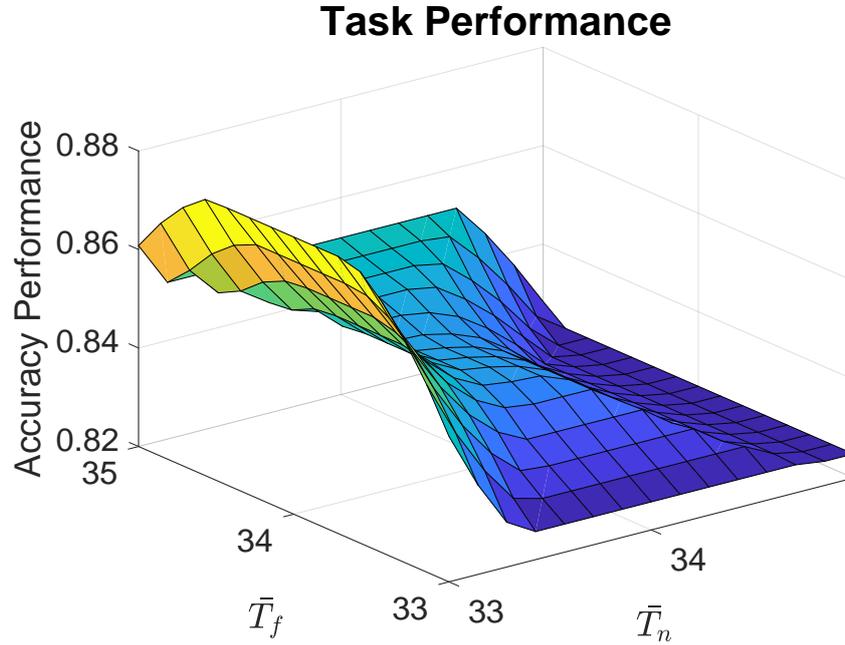
The generalised offline ANFIS model was constructed and trained with the Matlab® build-in functions *genfis* and *anfis*. The model implemented fuzzy c-means (FCM) to each participant, dividing 5 clusters for each input. The ANFIS model was trained with all the experimental data from the first session and then validated with the individual data from the second session for each participant. The real-time A-GT2-FCM model was the same one applied as in the previous research [84]. The error tolerance  $ET_{max}$  of the A-GT2-FCM model was set to be 0.01, the same as in the GT2FM model.

#### 4.5.2 Modelling Results Comparison

As described in the previous section, one complete experiment consisted of two HMI mental arithmetic test sessions for real-time modelling. The participant took the first session with incremental difficulty levels and then the following second session with randomised difficulty levels, with a ten-minute interval.

The model framework started with an initial rule-base from the generalised results in the previous experiment [35], with the first state estimation matrix  $E_1 = [1 \ 0]$  and the first transition matrix  $P = [1 \ 0; 0 \ 1]$ . Figure 4.10 shows the membership function between the task performance and two facial temperature readings  $\bar{T}_n$  and  $\bar{T}_f$ , which was based on the two initial fuzzy rule for describing the first state. The adaptive learning algorithm thereupon calculated the new state estimation and elicited the new individual dependent fuzzy rules according to the psychophysiological recordings and the observations.

The Pearson Correlations  $c$  and the Root Mean Squared Error  $e_{rmse}$  were in-



**Figure 4.10:** Membership for the performance function  $f_A(\bar{T}_n, \bar{T}_f)$

roduced to assess the prediction of the adaptive general type-2 fuzzy model. Table 4.5 shows the correlations and the errors between the observations and the predictions for each participant in each session. The calculations of these indices for  $n$  samples were via the following equations:

$$c_{\hat{y},o} = \frac{1}{n-1} \sum_{i=1}^n \left( \frac{\hat{y}_i - \mu_{\hat{y}}}{\sigma_{\hat{y}}} \right) \left( \frac{o_i - \mu_o}{\sigma_o} \right), \quad (4.25)$$

$$e_{rmse} = 100 \cdot \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - o_i)^2}{n}},$$

where  $\mu_{\hat{y}}$  and  $\sigma_{\hat{y}}$  are the mean and standard deviation of the prediction  $\hat{y}$ , respectively, and  $\mu_o$  and  $\sigma_o$  are the mean and standard deviation of the observation  $o$ . The sampling rate for the model is 1 Hz so the total samples for one session is  $n = 690$ .

Tables 4.5, 4.6 and 4.7 summarises the prediction results for all the participants from the real-time online modelling of GT2FM and A-GT2-FCM and the off-line generalised ANFIS. From Table 4.5, it can be seen that the mean correlations and the mean errors of the GT2FM remained consistent. Compared to the A-GT2-FCM (Table 4.6) and the ANFIS models (Table 4.7), the prediction results of GT2FM

**Table 4.5:** Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of GT2FM

Participant	Correlation (%)		Root Mean Squared Error (%)	
	Session 1	Session 2	Session 1	Session 2
01	98.755	99.558	3.628	2.430
02	98.331	98.857	1.725	1.476
03	99.520	99.424	2.265	2.237
04	99.609	99.402	2.205	1.982
05	98.193	98.693	4.109	3.635
06	99.504	99.376	2.504	2.773
07	98.916	99.077	2.081	2.208
08	98.627	99.128	2.276	2.392
08	98.291	98.960	2.365	1.972
10	98.678	99.224	2.192	1.706
Mean	98.840	99.171	2.535	2.281

**Table 4.6:** Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of A-GT2-FCM

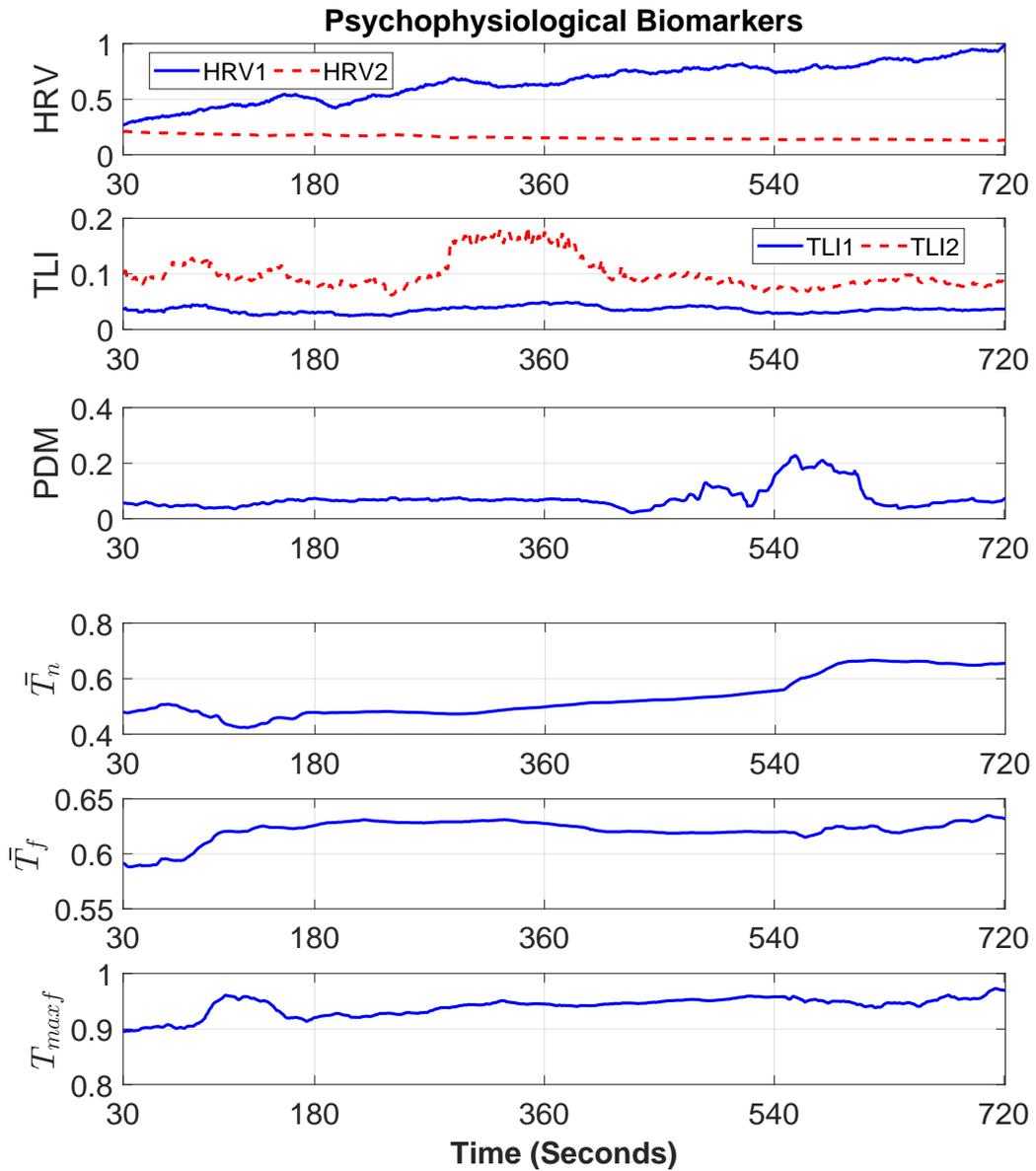
Participant	Correlation (%)		Root Mean Squared Error (%)	
	Session 1	Session 2	Session 1	Session 2
01	97.060	98.156	5.536	4.977
02	95.433	97.344	2.931	2.375
03	98.211	97.705	4.957	4.541
04	98.675	97.537	4.201	4.159
05	95.669	96.768	6.505	5.995
06	98.800	97.953	3.965	5.086
07	97.046	98.501	3.583	2.877
08	96.457	97.190	3.710	4.398
08	96.691	97.256	3.397	3.387
10	97.391	98.199	3.178	2.758
Mean	97.142	97.663	4.198	4.058

**Table 4.7:** Correlations and Root Mean Squared Errors (RMSE) for Real Accuracy versus Predicted Accuracy of ANFIS

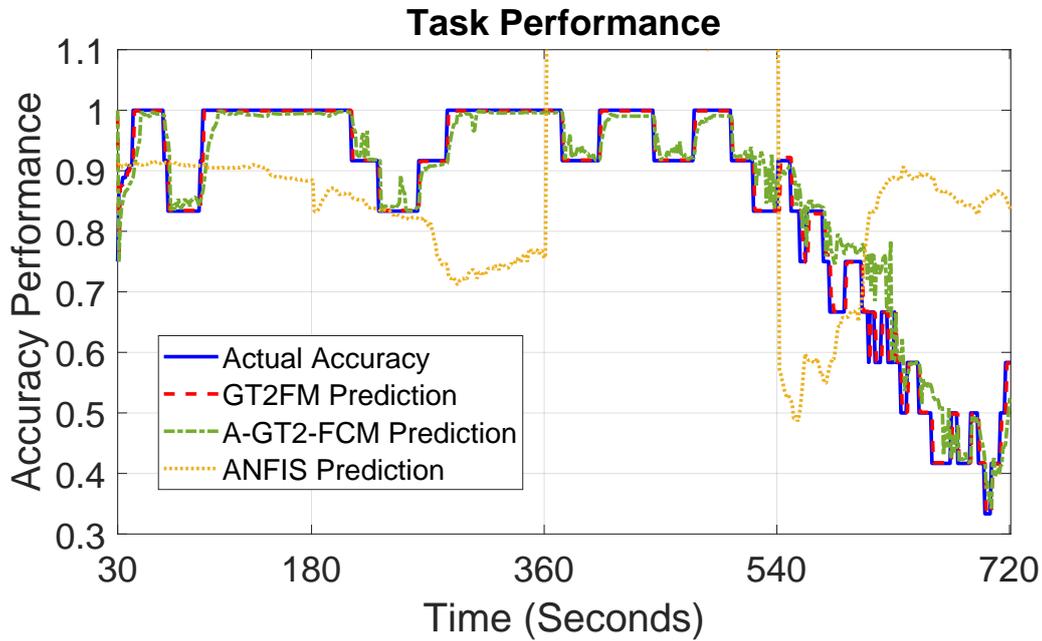
Participant	Correlation (%)		Root Mean Squared Error (%)	
	Session 1	Session 2	Session 1	Session 2
01	80.631	72.272	13.740	18.590
02	70.028	57.243	6.828	52.290
03	84.880	59.187	12.280	22.800
04	96.763	73.328	6.223	47.430
05	80.741	72.895	12.800	20.460
06	90.991	-32.545	10.800	82.180
07	75.955	91.371	9.326	8.505
08	75.979	22.890	8.987	23.640
08	84.388	79.680	7.793	15.350
10	93.127	56.762	4.961	37.030
Mean	83.348	61.813	9.374	32.828

achieved the highest correlations and the lowest error rates across all the participants. Overall, the GT2FM model performed well and consistently throughout the entire experiment. Based on the prediction outcome, the GT2FM model presented an excellent predictive ability due to the forecast prediction of the participant’s state. The learning algorithm was capable of fast individual features extraction without any prior knowledge or specific training. To further evaluate the ability of the model, Figures 4.11, 4.12, 4.13 and 4.14 show the detailed time sequences of the prediction and the psychophysiological biomarkers for participant 08 real time experiment results of both sessions.

From the task performance plot (Figure 4.12), compared to the A-GT2-FCM and ANFIS model, the predictions of the GT2FM matched with the actual performance of the participant the most. It could be observed how fast the adaptive general type-2 fuzzy model adjusted itself at the beginning of the experiment. The plot also showed the GT2FM model capability of handling high-frequency state change from the last phase. The psychophysiological inputs for all the models are presented in Figure 4.11. It is worth noting that all the psychophysiological biomarkers were suffering from a certain degrees of delay in representing participant inner state. However, the GT2FM model was still able to maintain the delay within 1 to 2 seconds throughout all participants in this research, despite of the



**Figure 4.11:** Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM,  $\bar{T}_n$ ,  $\bar{T}_f$ ,  $T_{maxf}$ ) for the participant 08 in session 2



**Figure 4.12:** Accuracy performance from the participant 08 and from the adaptive general type-2 fuzzy model prediction in session 2

intra-parameter variations. The merging between the existing fuzzy rules and the observations kept the rule-base simple and up-to-date. It ensured the efficiency and effectiveness of the prediction inference process.

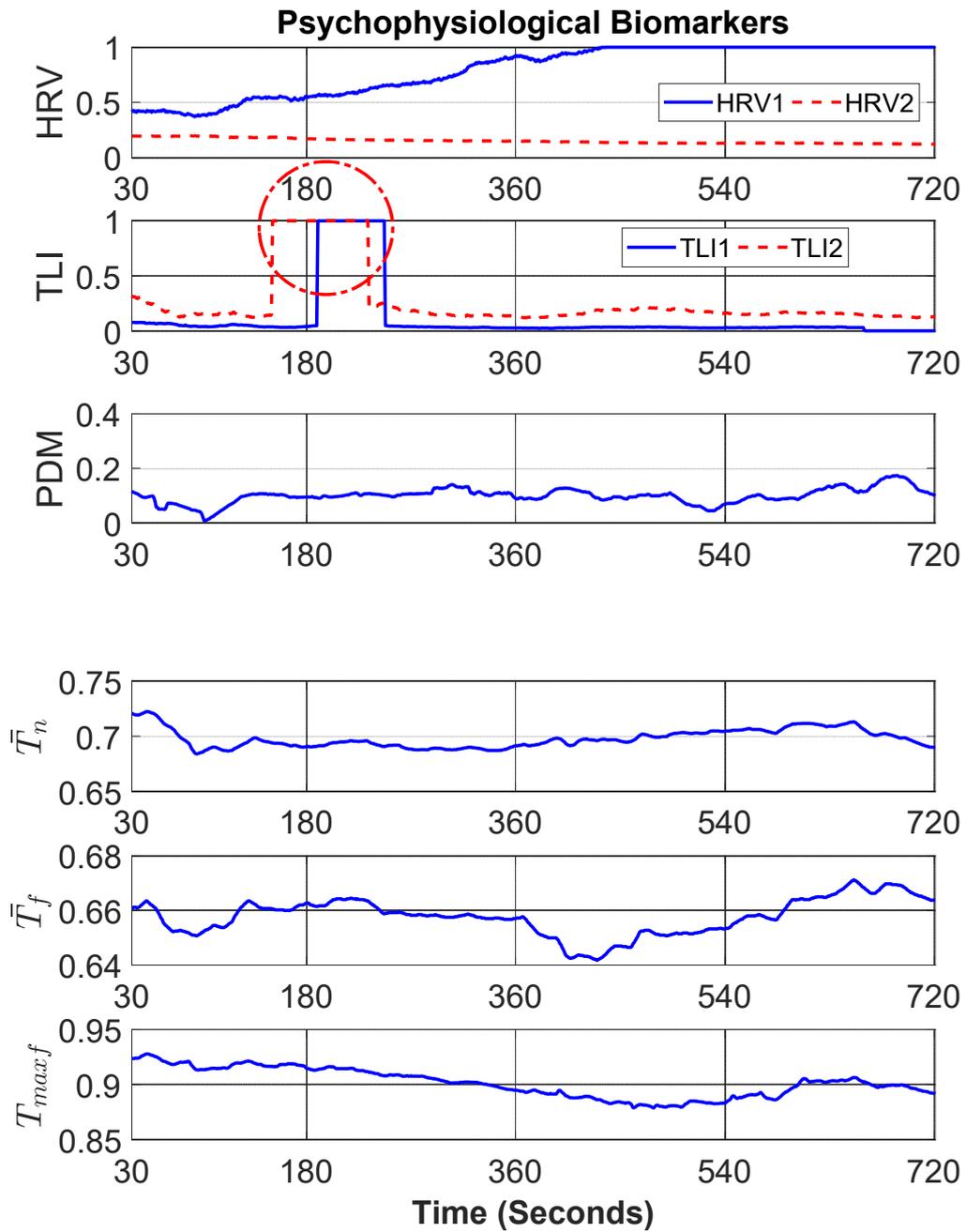
The psychophysiological biomaker readings ( $HRV_1$ ,  $HRV_2$ ,  $TLL_1$ ,  $TLL_2$ ,  $PDM$ ,  $\bar{T}_n$ ,  $\bar{T}_f$ ,  $T_{maxf}$  and  $DL$ ) in all figures of this section were all normalised for the purpose of illustration only. The adaptive general type-2 fuzzy model did not require any type of normalisation to operate.

### 4.5.3 Adaptive Learning of General Type-2 Fuzzy Model

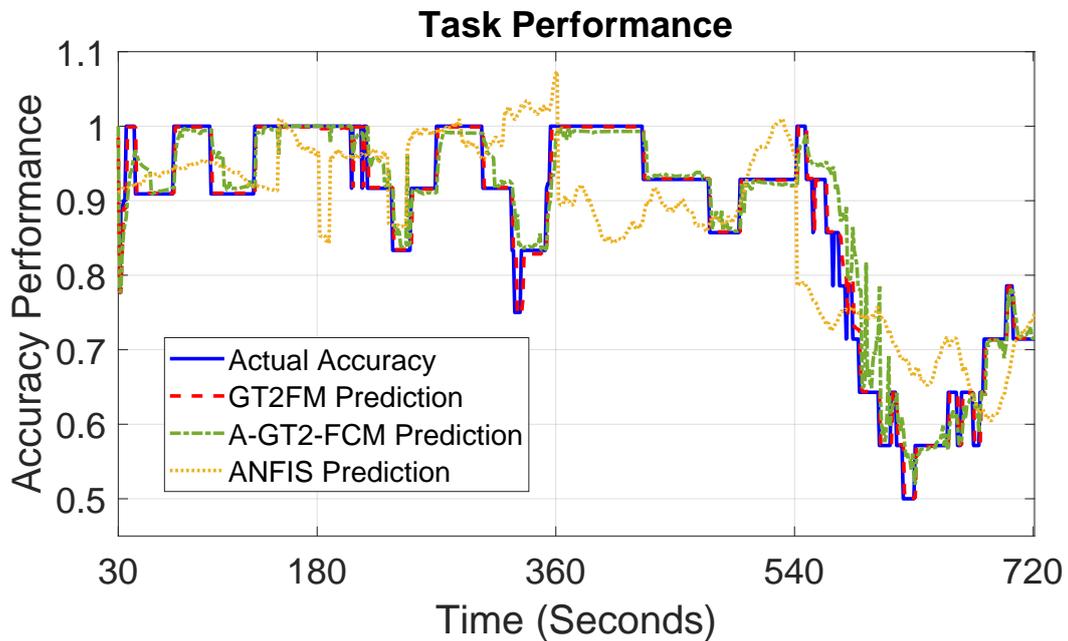
Table 4.5, Figures 4.12 and 4.14 illustrate the adaptive learning of the adaptive general type-2 fuzzy model via comparisons between the predictions and the observation. As already stated above, the GT2FM algorithm combined the inter- and intra-uncertainty within the type-2 fuzzy sets. The state tracking algorithm finalised the prediction according to the trend estimation and the probability. The learning algorithm kept the fuzzy rule configuration consistent with the current situation. In this research, the adaptive learning of GT2FM may be interpreted as

follows:

1. The model was able to self-organise in real-time. From the task performance plots of Figures 4.12 and 4.14, it may be observed that the model quickly adjusted itself at the beginning of the experiment and when the participant's task performance became unstable. The psychophysiological indices of the participants varied with multiple factors besides the task load. Thus, the psychophysiological recordings demonstrated clear different patterns even for the same person with the same difficulty level (e.g., the psychophysiological recordings plots of Figures 4.11 and 4.13). However, similar performances from all participants suggested that the learning and self-organising abilities of the model were sufficient for this intra-uncertainty.
2. The model was generalising for every participant. The model did not require individual-based calibration or off-line training for the operation. The initial rule-base, the first state estimation matrix and the first transition matrix were universal for all the participants. It is worth noting that the initial rule-base was based on sample mathematical estimations from previous inputs and output data. The initial statistical means and deviations only influenced the speed of the convergence (personalise) rather than itself. Despite the inter/intra participant uncertainty, the model had succeeded in extracting these uncertainties and transferring them into recognisable patterns. For example, comparing the psychophysiological biomarker recordings in Figures 4.13 and 4.15, there were significant differences between these biomarker values even when participants 03 and 08 were under the same experimental conditions. It could also be observed that participant 08 showed higher values than participant 03 in all the facial temperature indicators  $\bar{T}_n$ ,  $\bar{T}_f$  and  $T_{maxf}$ . HRV1 provided another evident inter-subject variation. Compared with the ones from participant 03 these HRV1 values of participant 08 were doubled. These may explain the performance differences within the participants based on the work memory theory. Inter-differences of the predicted accuracy at the beginning of the prediction in the task performance



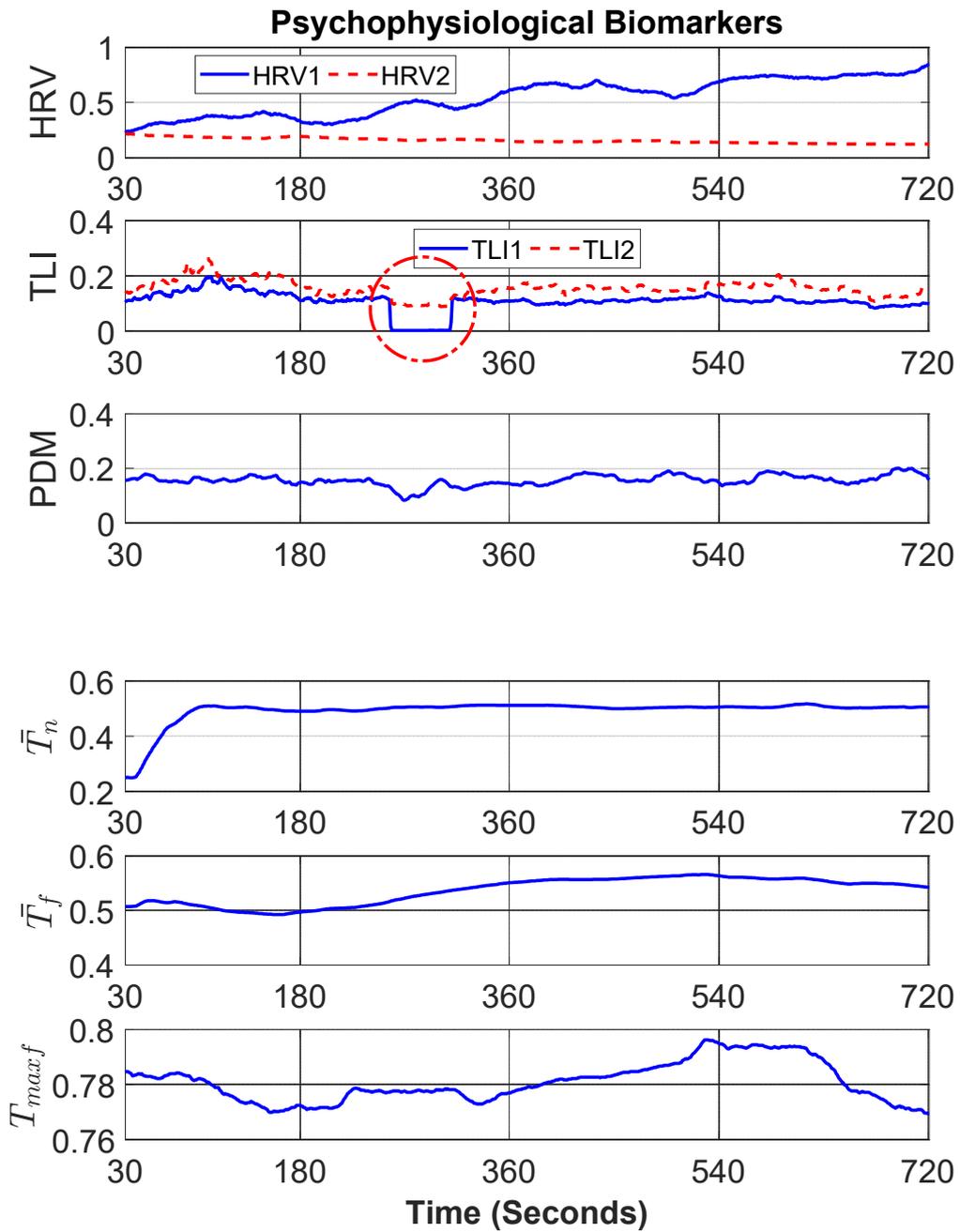
**Figure 4.13:** Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM,  $\bar{T}_n$ ,  $\bar{T}_f$ ,  $T_{maxf}$ ) for the participant 08 in session 1



**Figure 4.14:** Accuracy performance from the participant 08 and from the adaptive general type-2 fuzzy model prediction in session 1

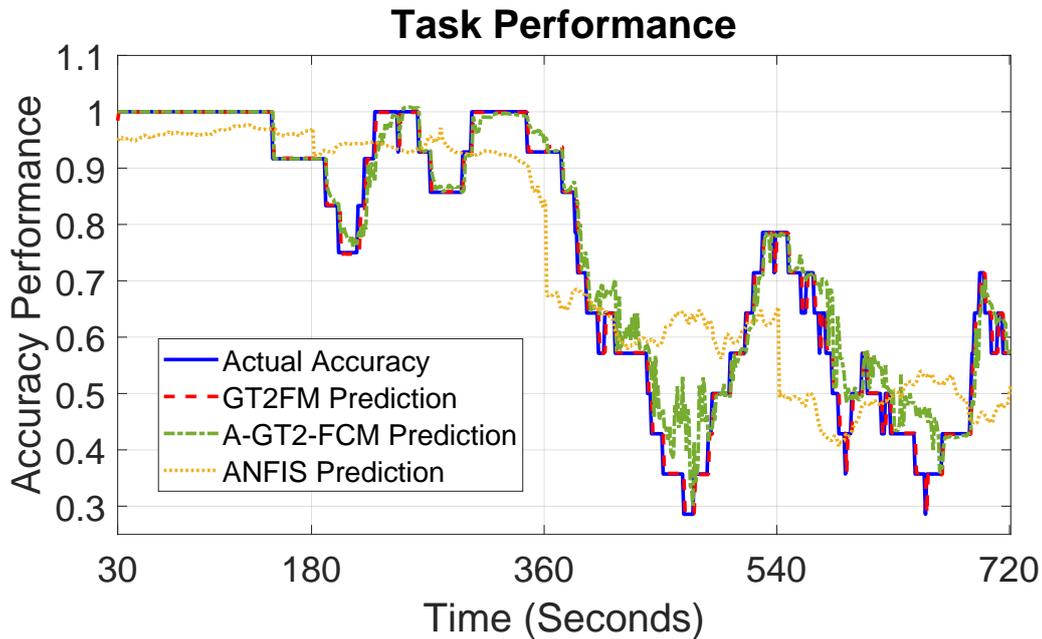
plot (Figures 4.14 and 4.16) were also obvious. This suggested that the initial fuzzy rule-base described participant 03 more precisely than for participant 08.

3. The model was able to manage temporal information loss and noise. In the real-time HMI simulation experiment, the information losses occasionally happened because of sudden disconnections between the electrodes and the subject. Meanwhile, the noise was introduced with task-irrelevant events such as unconscious movement. The red circles in Figures 4.13 and 4.15 indicated some cases of such information loss. The noise within the recordings may be conspicuous and may be persistent throughout the whole session. One extreme example may be found in the psychophysiological recording plots for PDM in Figure 4.13, the value drops nearly to 0 during the first phase, which is clearly impossible for human pupils and could only be a result of eye tracker's misinterpretation for something else. Still, the model had managed to maintain the high accuracy during these periods from the task performance plots of Figures 4.14 and 4.16. The combination of different



**Figure 4.15:** Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM,  $\bar{T}_n$ ,  $\bar{T}_f$ ,  $T_{maxf}$ ) for the participant 03 in session 1

biomarkers provided the model with the ability to quickly switch the lead biomarkers it depended on and maintained the consistency of the model prediction. In some other cases where all the facial temperature biomarkers  $\bar{T}_n$ ,  $\bar{T}_f$  and  $T_{maxf}$  were taken out from the model input vector for the purpose of testing, the time lag between the model prediction and the observation was still steadily remained within three seconds.



**Figure 4.16:** Accuracy performance from the participant 03 and from the adaptive general type-2 fuzzy model prediction in session 1

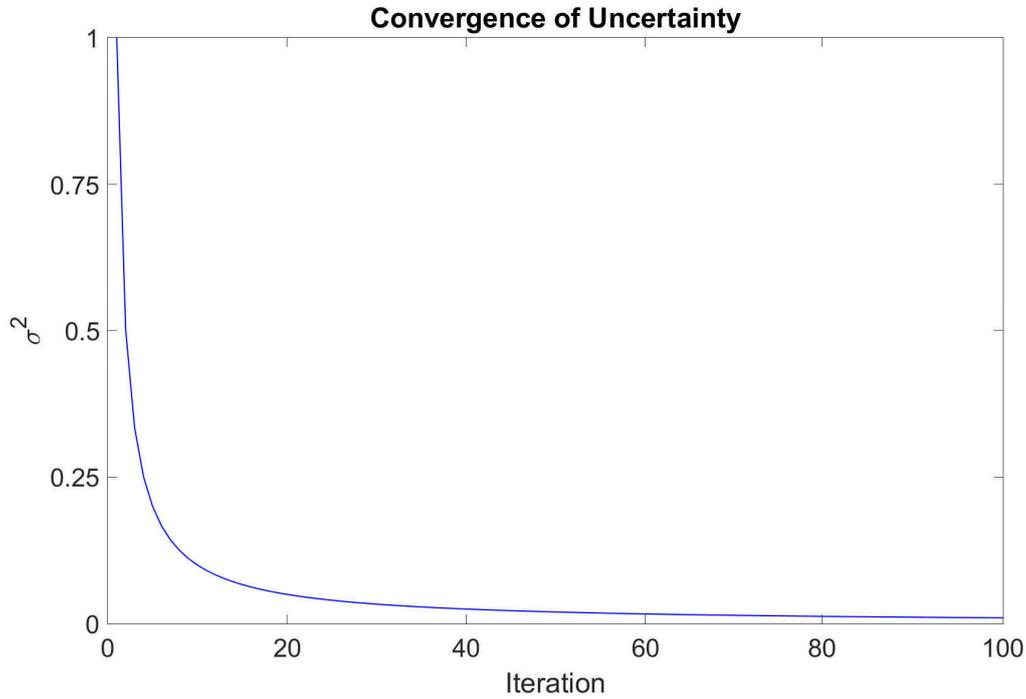
#### 4.5.4 Computation of General Type-2 Fuzzy Model

Compared with the conventional type-2 fuzzy modelling systems, GT2FM also demonstrated several excellent computational advantages in addition to its advanced adaptive features. The unified secondary uncertainty of GT2FM ensured the existence of numerical closed-form solutions and simplified the type-reduce process. The adaptive learning algorithm based on Bayesian theory had significantly contribution to the implementation of GT2FM in online experiments. The details of these computational advantages can be summarised as follows:

1. The fuzzy sets of GT2FM were based on the interval type-2 fuzzy sets, where

all the secondary uncertainty was equal. The simplification avoided the definition, computation and storage of extra uncertainty from the secondary dimension. In fact, in this research, it was also hard and unnecessary to link such uncertainty to any actual physical factor. The use of interval type-2 fuzzy sets reduced the redundancy of the modelling system and made the calculation of the numerical close-form solutions feasible and straightforward. As a result, it guaranteed a fast inference between the input sets and the fired fuzzy rules. However, compared to the type-1 fuzzy sets, the primary uncertainty of GT2FM was enhanced (e.g. Figure 4.8). Such uncertainty was applied on both the input sets and the fired rules (the same standard deviations  $\sigma_x$ ). Meanwhile, the state estimation involved in the type-reduce process combined the final prediction with tendency measurement. Compared to the conventional general type-2 fuzzy modelling systems, this configuration could achieve equal or higher flexibility and accuracy while maintaining lower computational requirement.

2. The learning algorithm based on the Bayesian theory contributed to the computation of GT2FM in three major aspects. First, compared to other updating method (e.g. A-GT2-FCM), the Bayesian updating was straightforward and represented a relatively fast way to generate a close-form solution. In the experiment, it demonstrated a better convergence and increased the overall performance of model computation. Second, it restricted the size of fuzzy rule-base by continuously integrating the new observations with the closest existing rules. This ensured the complexity of the inference process and the storage of the rule-base remained simple and consistent throughout the experiment. Third, the algorithm allowed the uncertainty of the rules continuously converge with the updated observations. Figure 4.17 shows an example of how the uncertainty value converged with the iterations of the adaptive learning process, as the fuzzy rule-base representation transferred from the general to the specific individual. It is reasonable to believe that with each Bayesian updating, the difficulty of computation was remarkably reduced and the prediction outcome was significantly improved.



**Figure 4.17:** Example of uncertainty convergence with the learning algorithm

3. In this research, the model was implemented on Matlab<sup>®</sup> GUI system. The redundancy of the overall system had severe influence of model computation, such as the delay caused by the equipment-software desynchronisation and the inefficient build-in function. Real-time online image processing required a massive amount of computational resource which led to the limited memories shared by other experiment equipment software and caused desynchronisation and delay. Meanwhile, the redundancy of Matlab<sup>®</sup> build-in functions occupied extra computational resources, caused system error with preassigned threshold and slowed down the computation. It is worth noting that the experiment results in this project were not able to present the full capacity of GT2FM. Therefore, it was reasonable to expect better performance outcome of GT2FM with a clean and clear coding structure on a more specified program language and a better computational environment.

## 4.6 Summary

This section focused on the prediction of the human operator psycho-physiological state in the HMI system. The mental arithmetic was selected as the simulation of HMI systems for 10 participants. In addition to the previous psycho-physiological biomarkers TLI, HRV and PDM, new facial temperature indicators were introduced and integrated with others for assessing the operators' psycho-physiological state.

A new modelling approach named adaptive general type-2 fuzzy modelling was proposed to predict human psycho-physiological state based on real-time experiments. Such a new modelling approach integrated system uncertainty with type-2 fuzzy sets and state tracking with defuzzification to estimate human psycho-physiological state. The model prediction results were compared to participant-specific ANFIS and A-GT2-FCM, and it was found that the proposed model outperformed the other models presented in the related literature. The design of adaptive learning algorithm based on Bayesian theory proved its ability to extract patterns from observations in real-time. With the estimation and classification of psycho-physiological state, high accuracy and reasonable correlation were achieved, including during the breakdown periods.

In summary, the results of the simulation experiment provided an evaluation for applying adaptive general type-2 fuzzy modelling to systems similar to HMI. Furthermore, it created the foundation of more advanced control mechanisms for HMI systems and can be applied for the exploration of the origins of human operator compromised performance in the future. The next chapter will introduce a new type-2 fuzzy logic-based control algorithm based on this adaptive general type-2 fuzzy modelling algorithm to balance the human effort in the human-machine interface.

## Chapter 5

# Adaptive General Type-2 Fuzzy Controller for Balancing Human-machine Interface

This chapter presents a new type-2 fuzzy logic-based control algorithm for balancing the human-machine interface systems via adjusting the automation level of the machine correspondingly to the estimation of the human operator psychophysiological state. The newly proposed control approach combines the state estimation of the human operator with the type-2 fuzzy sets to ensure equilibrium between the task requirements (i.e. difficulty level) and the human operator feasible effort (i.e. psychophysiological states). Validated on mental arithmetic cognitive experiments for ten (10) participants and compared with the existing energy model-based control (E-MBC), the results revealed the impacts of the multitask and fatigue on human operator performance. A selection of indicators are presented for the detection of these two human compromised task performance related psychophysiological states. In addition, the new control approach with fatigue management showed the best performance with the highest consistency and stability throughout the experiment. These findings allowed for path-opening to the prevention of the human operator psychophysiological breakdown and for the further development of other human-machine interface systems.

## 5.1 Background

In this research, two major origins of human compromised task performance, multitasking and fatigue, were investigated for developing a more advanced control strategies for the HMI systems.

### 5.1.1 Multitasking

The term task may be referred to a cognitive or behavioural aim that requires some certain corresponding responses to achieve [44]. Undoubtedly, such a broad definition may result in some inconsistency under some special circumstances, e.g. hierarchical tasks and multistep tasks. However, from the prospective of neural science, a task may be explained with the synchronisation activities of the human brain. In this way, multitasking may be expressed as a goal-directed, instructed or self-instructed behaviour for maintaining information from multiple items over a short period in the WM [5]. A typical cognitive process of task stimulated response consists of task state tracking, feasible resource evaluation and outcome prediction. Consequently, in addition to dual concurrent task that may share the same information and pathways, simultaneous regulations such as serial task switching (interruptions and resumptions) existed in the multitasking. Both dual-tasking and task-switching can be identified to having a significant influence on the human task performance from three aspects, namely, cognitive structure, flexibility and plasticity, and this increased performance costs in the individual tasks [44]. The research of the stimulus onset asynchrony, when it is not long enough for sequential behaviour, in dual task paradigms show the more temporal overlap of task processes the worse human task performance may be observed. Meanwhile, as compared to the single-task, task-switching between different goals requires switching estimation and the coordination of at least two task sets, including cognitive process and behavioural response. The switch costs of task-switching is for changing the task set regardless of the interference of other activities. While the coordination costs represent the preparation and maintenance of multiple task sets coherently. It is typical to find that task performance is worse in task switching

than in the repetitive work [44]. The influence of multitasking on human task performance is, therefore, of great importance for controllers to balance the HMI systems.

This may be explained by the synchronisation of neurons and neural networks that are responsible for promoting the neural communication and memory processing [25]. The human cognition study suggests that the compromised performance is related to the processing of information stream and memory in the cortex [88]. As already mentioned, WM indicates one's ability to maintain the task related information against interference from other irrelevant stimuli. It is based on various EEG frequency components that are generated by the collaboration of transient inter neural electromagnetic pulses. Subtle synchronous coordination laid the foundation of the multi sensory modality combination. Neural populations selected by encoded sensory stimuli modalities formed rapid and transient connections through functional enhanced synchronisation [88]. Previous research papers and books found that such neural activities may lead to observable oscillation patterns in the gamma waves (25 to 100 Hz) on the human EEG over the prefrontal cortex [4]. The simultaneous maintenance of increasing items in WM is accomplished by continuously regulated cross-frequency coupling. Hence, the increasing memory load from multitasking compromises the task performance significantly [4]. Meanwhile, attention is related to the EEG alpha rhythm and gamma rhythm. Evidence showed, at a given cortical site, the alpha rhythm regulated the information processed by that area and the gamma rhythm adjusts attention for different sensory modalities [88]. Therefore, attention control and WM for completing tasks are competing for a common cortical circuit substrate. Additionally, multitasking also requires the reconfiguration of prefrontal cortex tuning profiles to map different task contexts constantly. The transform of network impacts the coding space for decision making and might lead to a complete behavioural choice drift in order to cope with other task context [77]. These findings also prove the possibility of assessing human multitasking with measurable psychophysiological indicators, such as EEG.

### 5.1.2 Fatigue

The term fatigue describes the psychophysiological state of experiencing tiredness and is commonly associated with a compromised performance from both the cognitive and physical levels [30]. For the physiological prospective, fatigue may lead to strength reduction, exercise capacity loss, effort level increase and muscle power decrease. For the psychological prospective, fatigue may refer to an overwhelming feeling of tiredness and impaired cognitive performance, such as a diminished short-term memory, a reduced vigilance and a degraded communication skills [30]. Previous statistical research papers and books have shown a high correlation between the human fatigue level and the human error rate [19, 30]. Most HMI nowadays focused on cognitive tasks more than physiological tasks. Therefore, this project researched mainly the acute mental fatigue that may be defined as an emotion involved with tiredness or exhaustion and often associated with reluctant behaviour towards the current task, including reduced commitment and impaired performance [10, 64]. Compared with multitasking, the fatigue had a more significant impact on human task performance - cognitive slowing and cognitive lapses that prolonged the reaction time for the task and increased the probability of breakdowns in physical performance [31, 37]. For the safety and efficiency of the HMI systems, the control strategies should be capable of detecting such fatigue of human operators and respond adequately to avoid any breakdown.

Although mental fatigue is a well known psychophysiological state, there is no unified definition nor straightforward measurement [30]. As a brain-derived emotion, fatigue is mainly assessed by the subjective self-report and the reaction time [10, 30, 31, 99]. There are two observable traits in task performance may be linked with mental fatigue - cognitive slowing (slowdown in short reaction time tasks) and cognitive lapses (psychophysiological breakdowns lasting from seconds to hours) [31]. Current studies have established mathematical models explaining the relationships between fatigue, sleep deprivation and working hours [31, 37, 99]. These research studies focused on fatigue generated from overloaded physical or mental tasks and sleep deprivation based on the studies of long time experiment. However, some partial equations from these existing model

may be applied to the modelling of acute mental fatigue in HMI. It was found that, under a constant workload, mental fatigue level approximated a circadian self-sustaining oscillator [99]. From these psychophysiological research works, it may be proved that fatigue was able to be estimated with the regulation of human body homeostasis.

However, it was proved that such acute mental fatigue could be restrained or relieve with certain methods [10]. Based on the research of fundamental goal directed behaviour, a human is only motivated for a certain task when its benefits overcame its effort costs. Therefore, for a prolonged period of time, fatigue rose when the accumulated effort invested in eventually outweighed the potential benefits. Previous records showed the minimised requirement of effort level may significantly improve experience of solving tasks. Therefore, acute mental fatigue may be regarded as an adaptive strategy to constraining effort on high reward goals and increasing overall energy efficiency. It revealed that such a fatigue may be overrode with increased task reward or relieved with decreased effort requirement temporarily. However, it must be aware that the suppression of acute mental fatigue for a prolonged time came at a price, as it elevated psychophysiological stress level and contributed to burnout and long term health disorder. Meanwhile, the individual difference was also important, considering that the evaluation of efforts and benefits different differed from person to person.

## 5.2 Experimental Setup

### 5.2.1 Human-Machine Interface Simulation

The mental arithmetic experiment had been selected for this research because of its effectiveness, simplicity and intuitiveness [33]. For the consistency, the experiment environment and the ten (10) selected participants were the same as in previous modelling part of this study, except for the difficulty level of the task, where it was adjusted by the human-machine interface balance control (HMIBC) controller rather than predetermined. The simulation system took records of heart rate variability ( $HRV_1$  and  $HRV_2$ ), task load index ( $TLL_1$  and  $TLL_2$ ), pupil sizes ( $PDM$ )

and facial temperature biomarkers ( $\bar{T}_n$ ,  $\bar{T}_f$  and  $T_{maxf}$ ) from the human operators in real-time. Based on the psycho-physiological monitoring, the adaptive generate type-2 fuzzy model provided task performance prediction, psychophysiological state estimation and the rule-base for the controller to determine the necessity of the appropriate intervention to meet the expectation of the HMI system.

### 5.2.2 Controlling Experiment Configuration

For the control configuration, the HMIBC controller was implemented in order to adjust the task difficulty level according to the participant psychophysiological state prediction based on the adaptive general type-2 fuzzy model in real-time. The psychophysiological recordings of the human operator under fatigue and multitasking conditions were collected concurrently.

The entire control experiment consisted of three 12-minute mental arithmetic test sessions and two 5-minute breaks in between for each participant. The participants went through four 30-second phases with all four difficulty levels without any control intervention to familiarise themselves with the system in the beginning. The controller then adjusted the difficulty level according to the model predictions and the subjective fatigue report.

A fatigue indicator was integrated with the software for studying the human operator psychophysiological fatigue in the HMI and fatigue management in the first and the last mental arithmetic test sessions. The indicator relied on the self-report of the participants during the experiments and marks the period when the subjects experienced fatigue-related feelings, e.g. tired, bored or anxious. In the last mental arithmetic test session, the controller combined the fatigue indicator to improve the control efficiency and created a more intelligent system based on participant subjective self-evaluations. The second mental arithmetic test session focused on the multitasking of human operators. In addition to the normal tasks, the participants were required to maintain a casual conversation throughout the whole session (secondary task).

### 5.3 Human-Machine Interface Balance Control

A control action depends on the inference between the input vector  $I_t$  and the predicted output  $\hat{y}_t$ , the state estimations  $E_t$ , the fuzzy rule-base  $R$  from the adaptive general type-2 fuzzy model.

The aim of the human-machine interface balance control (HMIBC) is to balance the HMI system. Therefore, the controller must be able to maintain the consistent overall performance against the time while maximising the subjects' participation. The idea of balance control is to assign the participant with the suitable task load according to the mapping between the psychophysiological states and the difficulty levels. In this way, the task performance remains at a suitable range for each difficulty level.

The detailed inference process for a control prediction is as follows:

1. Check whether a control action is necessary. A control action is only introduced when:

$$\begin{aligned}\bar{y}_t &= \frac{\sum_{i=t-N+1}^t \hat{y}_i}{N}, \\ \bar{y}_t &\notin [y_{(i,l)}, y_{(i,u)}],\end{aligned}\tag{5.1}$$

where  $y_{(i,l)}$  and  $y_{(i,u)}$  represent the lower and the upper performance boundaries of the current difficulty level  $i$ , and  $N$  stood for the length of selected performance window.

2. Compute the densities between selected fuzzy rules and input vector  $I_t$  using Gaussian membership function. The selected fuzzy rules are as follows:

$$R = \begin{bmatrix} R_{(1,m)} & R_{(1,n)} \\ R_{(2,m)} & R_{(2,n)} \\ \vdots & \vdots \\ R_{(M,m)} & R_{(M,n)} \end{bmatrix},\tag{5.2}$$

where the predicted output is between the two states  $y_m \geq \hat{y}_t > y_n$  (if  $\hat{y}_t \geq y_1$ , then select all the fuzzy rules describing the 1st and the 2nd state; correspondingly, the  $K-1$  and  $K$  for  $y_K \geq \hat{y}_t$ ). The density  $D_{(i,j)}$  between the

input vector  $I_t$  to each selected fuzzy rule  $\mu_{(i,j)}$  &  $\sigma_{(i,j)}$  were:

$$D_{(i,j)} = e^{-\frac{(I_t - \mu_{(i,j)})^2}{2\sigma_{(i,j)}^2}}. \quad (5.3)$$

3. Decide on the type of control action and the density matrix  $D_t$ :

(a) If  $\hat{y}_t < y_{(i,l)}$ , then the current difficulty level  $DL_t$  should be decreased.

Correspondingly, The density matrix  $D$  is as follows:

$$D = \begin{bmatrix} D_{(1,m)} & D_{(1,n)} \\ \vdots & \vdots \\ D_{(DL_t-1,m)} & D_{(DL_t-1,n)} \end{bmatrix}; \quad (5.4)$$

(b) If  $\hat{y}_t > y_{(i,u)}$ , then the current difficulty level  $DL_t$  should be increased.

Correspondingly, The density matrix  $D$  is as follows:

$$D = \begin{bmatrix} D_{(DL_t+1,m)} & D_{(DL_t+1,n)} \\ \vdots & \vdots \\ D_{(M,m)} & D_{(M,n)} \end{bmatrix}. \quad (5.5)$$

4. Generate the new difficulty level  $DL_{t+1}$  from the state estimation  $E_t$ . The control strategy depends on the comparison between two state estimations  $E_{(t,m)}$  &  $E_{(t,n)}$ :

(a) If  $E_{(t,m)} \geq E_{(t,n)}$ , then the participants' performance tends to increase.

The suitable new difficulty level  $DL_{t+1}$  should be generated from  $m$  states and is close to the current input vector to achieve the most possible improved performance, as follows:

$$DL_{t+1} = a, \text{ where} \quad (5.6)$$

$$D_{(a,m)} = \left\{ \max(D_{(i,m)}) \mid \forall D_{(i,m)} \in D \right\};$$

(b) If  $E_{(t,m)} < E_{(t,n)}$ , then the participants' performance tends to decrease.

The suitable new difficulty level  $DL_{t+1}$  should be generated from  $n$

states and is far from the current input vector to avoid the most possible declined performance, as follows:

$$DL_{t+1} = b, \text{ where} \quad (5.7)$$

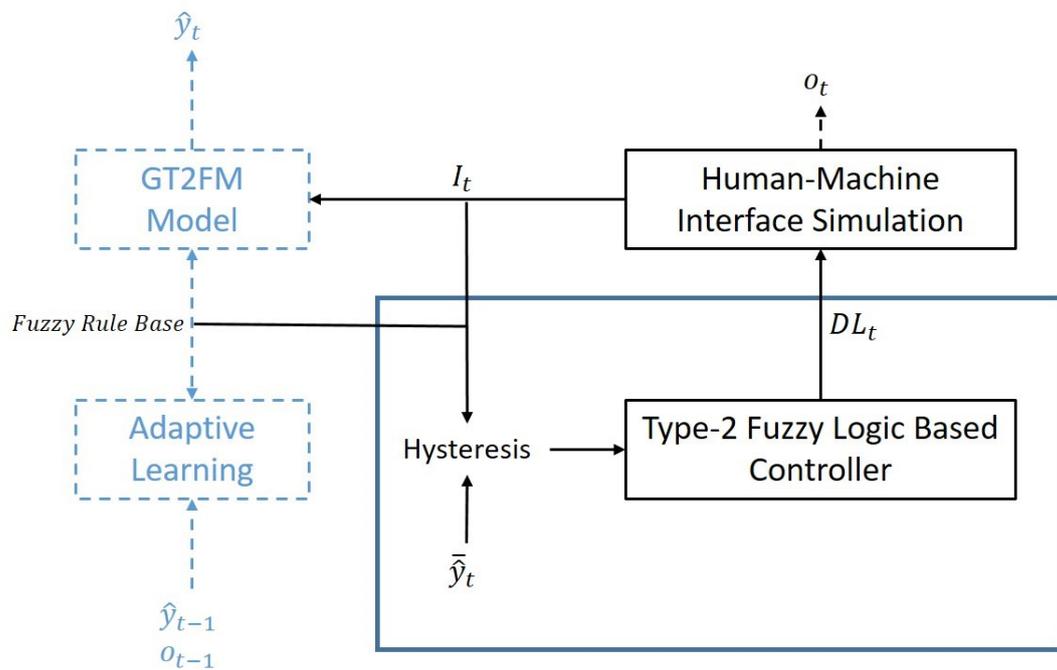
$$D_{(b,n)} = \left\{ \min(D_{(i,n)}) \mid \forall D_{(i,n)} \in D \right\}.$$

## 5.4 Evaluation of Impaired Task Performance and Controller

This section focuses on the experimental results of the HMIBC controller based on adaptive general type-2 fuzzy framework mentioned in the previous section, as well as the analyses of the human operator fatigue and multitasking states in the HMI system. The controller is integrated within the HMI mental arithmetic experiments for online real-time controls, and the results are compared with the HMI simulation sessions without any controller. Additionally, this chapter also provides the results of E-MBC based on A-GT2-FCM framework for comparison. It is worth mentioning that any real-time experiment session is impossible to replicate, even for the same subject under the same conditions. The evaluations and summaries of the comparison results, however, should still be helpful to show some different characteristics among the control-free sessions and two types of control methods.

### 5.4.1 Controller Configuration

The HMIBC controller based on the adaptive general type-2 fuzzy framework started working from 140 seconds after the experiment begins. Figure 5.1 shows the diagram of the controller integrated within the HMI simulation. The hysteresis of the control action was 10 seconds for the participant to adjust to the new difficulty level and the performance window was  $N = 10$ . This ensured the participant had adequate time to adapt to any new difficulty level applied. Table 5.1 summarises the performance boundaries for each difficulty level in this experiment. An intervention of the task difficulty level from the controller depended on the input vector  $I(t) = [HRV_1(t), HRV_2(t), TLI_1(t), TLI_2(t), PDM(t), \bar{T}_n(t), \bar{T}_f(t), T_{maxf}(t)]$ , the state estimation matrix  $E_t = [e_a \ e_b]$  and the rule-base  $R$  of



**Figure 5.1:** Diagram of the controller for the HMI simulation experiment (GT2FM and other sections in - - is the subject of the previous chapter, chapter 4)

eight type-2 fuzzy set based rules from the modelling framework as mentioned in the previous chapter.

**Table 5.1:** Upper and Lower Performance Boundaries

Difficulty Level	1	2	3	4
Accuracy (%)	$\leq 0.975$	$0.875 \sim 0.925$	$0.825 \sim 0.875$	$\geq 0.775$

### 5.4.2 Controller Evaluation

Similarly to the modelling experiment in the previous chapter, the control experiment started with a general rule-base from the previous thermal temperature experiment and the first state estimation matrix  $E_1 = [1 \ 0]$ . The assessment of control efficiency considered two aspects - the consistency and stability of the human operator task performance. The stability represents the controllers' ability to avoid the abnormal human operator psychophysiological state (e.g. breakdown) and restricts the consequent impact on the HMI system. In this experiment, it was quantified with breakdown percentages, which measured the occupation of the breakdown period (when the actual accuracy was less than 0.667 - the participant failed on more than four consecutive questions) over the whole time for each difficulty levels and the entire session. The consistency required maintaining the HMI overall performance at a reasonable level regardless of the changes in the human operator psychophysiological state and task difficulty level. For this research, this consistency was assessed by the statistical means and the standard deviations of the actual accuracy for each difficulty levels over the whole session.

Tables 5.2, 5.3 and 5.4 summarised the breakdown percentages, means and standard deviations of two modelling sessions and one controlling session for ten participants. Compared with the two modelling sessions with fixed difficulty level, the controlling session showed more stability and consistency in the task performance. From Tables 5.2, it can be seen that the overall breakdown percentage of the controlling session was the lowest among the three sessions for an accuracy trade-off between the easy and hard difficulty levels. It revealed that the controller succeeded in restricting the duration and influence of the abnormal human psy-

chophysiological state by adjusting the task difficulty to suit the human operator. Based on Table 5.3, the controller was able to help the participants to maintain a balanced task performance with relatively high accuracy throughout the HMI simulation with their various psychophysiological states. This avoided the inequality between the needs for human effort and the operator’s capacity. Furthermore, the task performances of the human operators, in general, were more stable with the controller as shown in Table 5.4, especially during the high difficulty level periods. Figure 5.2 provided an example for comparing the task performances of the same participant with and without the HMIBC. Compared with the control free session, the accuracy of the human operator remained at a reasonable level without any sudden decrease following psychophysiological breakdown.

**Table 5.2:** Breakdown Percentages (%) for Modelling and Controlling Sessions

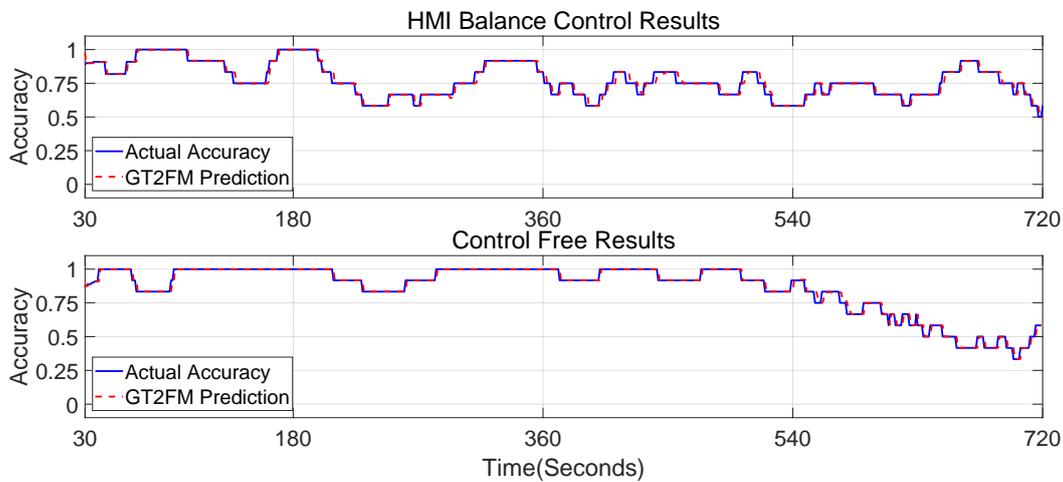
Sessions	Difficulty Levels				Total
	1	2	3	4	
Incremental Difficulty	0.261	0.444	12.725	52.501	17.111
Random Difficulty	2.565	0.074	12.176	49.897	16.800
Controlled Difficulty	2.258	6.345	6.289	20.800	10.271

**Table 5.3:** Means of Actual Accuracy for Modelling and Controlling Sessions

Sessions	Difficulty Levels				Total
	1	2	3	4	
Incremental Difficulty	0.961	0.942	0.795	0.587	0.816
Random Difficulty	0.948	0.956	0.815	0.605	0.826
Controlled Difficulty	0.862	0.873	0.890	0.861	0.855

**Table 5.4:** Standard Deviations of Actual Accuracy for Modelling and Controlling Sessions

Sessions	Difficulty Levels				Total
	1	2	3	4	
Incremental Difficulty	0.059	0.066	0.085	0.126	0.181
Random Difficulty	0.077	0.052	0.091	0.163	0.183
Controlled Difficulty	0.131	0.053	0.091	0.108	0.138



**Figure 5.2:** Real-time experiment results for the participant 08 with and without the HMIBC

### 5.4.3 Controller Comparison

Figure 5.3 compares the real-time experiment results for the HMIBC controller and the E-MBC controller. From the second set of plots in this figure it may be observed that the control actions from the HMIBC controller were more flexible and adaptive to the human operator psychophysiological changes than these from the E-MBC controller, despite the two controllers having the same hysteresis for generating new control actions. In contrast, the E-MBC controller increased the difficulty level when a clear performance breakdown was presented 540 seconds after the experiment started, which significantly damaged the task performance in the first set plots of the same figure. It is also worth noting that the lack of flexibility in the control actions of the E-MBC restricted the human operator's ability to perform high demanding tasks, as the time for the level four task was remarkably less in the experiment with the E-MBC. It should also be mentioned that the prediction accuracy of the modelling framework is of great importance for the efficiency of the control actions. In this respect, the E-MBC controller was compromised, for it is based on the A-GT2-FCM framework [84] with less accurate psychophysiological prediction (as shown in the previous chapter, chapter 4).

To further illustrate the HMIBC control's capabilities, additional statistical comparisons were provided in Tables 5.5, 5.6 and 5.7. From Table 5.5, and com-

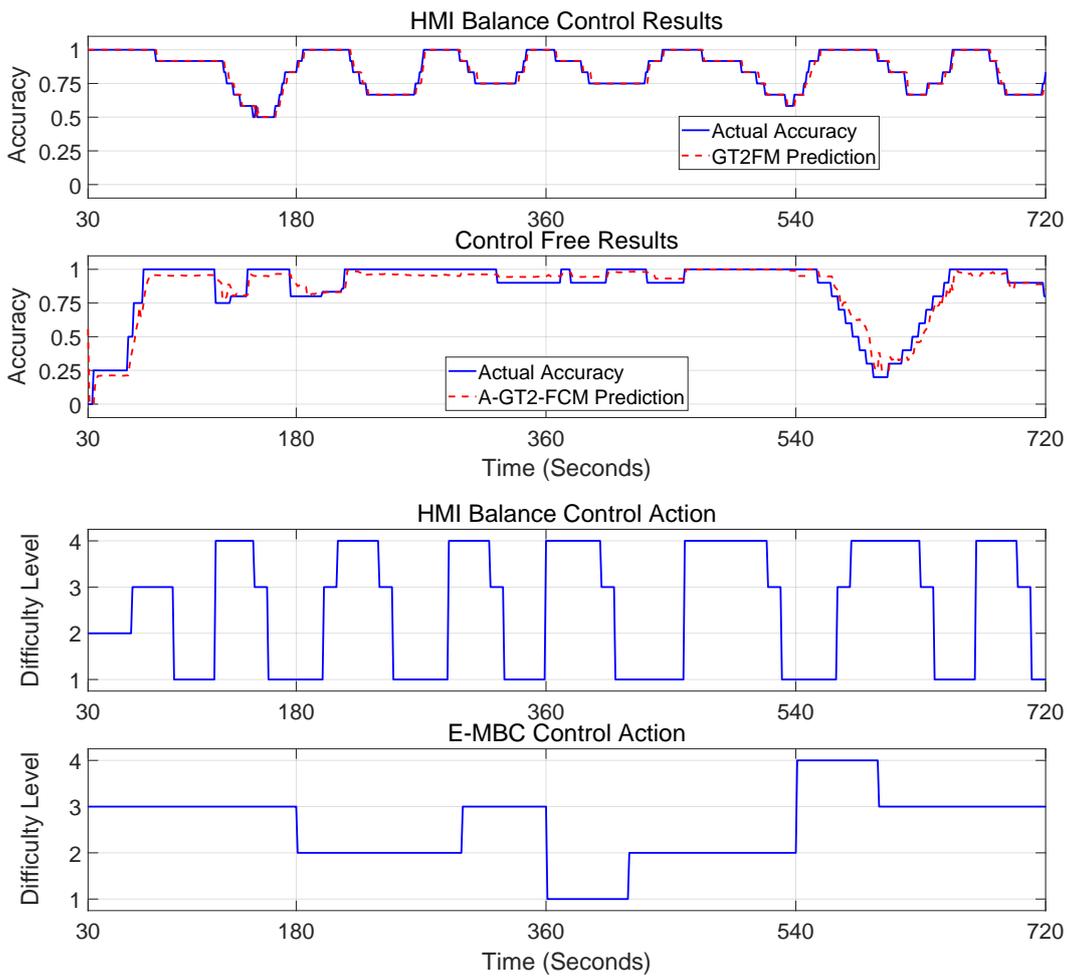


Figure 5.3: Real-time experiment results for the HMIBC and the E-MBC

pared with the E-MBC, the HMIBC significantly reduced the likelihood of breakdown, especially for difficulty levels 3 and 4. Apart from minimising the likelihood of breakdown, the HMIBC ensured consistency of human operators' performance throughout different difficulty levels with limited persistent disturbances in Tables 5.6 and 5.7. In summary, the HMIBC provided a more reliable and efficient approach for balancing the task requirement and human operator psychophysiological state in HMI systems.

**Table 5.5:** Breakdown Percentages (%) for the HMIBC and the E-MBC

Sessions	Difficulty Levels				Total
	1	2	3	4	
HMIBC	5.170	0.000	8.333	3.732	4.916
E-MBC	0.002	0.001	18.621	32.209	11.728

**Table 5.6:** Means of Actual Accuracy for the HMIBC and the E-MBC

Sessions	Difficulty Levels				Total
	1	2	3	4	
HMIBC	0.857	1.000	0.801	0.853	0.853
E-MBC	0.939	0.963	0.808	0.700	0.864

**Table 5.7:** Standard Deviations of Actual Accuracy for the HMIBC and the E-MBC

Sessions	Difficulty Levels				Total
	1	2	3	4	
HMIBC	0.140	0.000	0.164	0.125	0.141
E-MBC	0.049	0.067	0.269	0.272	0.225

#### 5.4.4 Multitasking

As one of the main contributions to the human operator compromised performance or even breakdown, multitasking in the HMI systems was investigated. The requirement of the participates to fulfil both a casual conversation and the mental arithmetic test concomitantly was significantly impacted on their task performance in the HMI simulation. Compared with the values from the original single-tasking sessions in Tables 5.8, 5.9 and 5.10, the total breakdown percentage

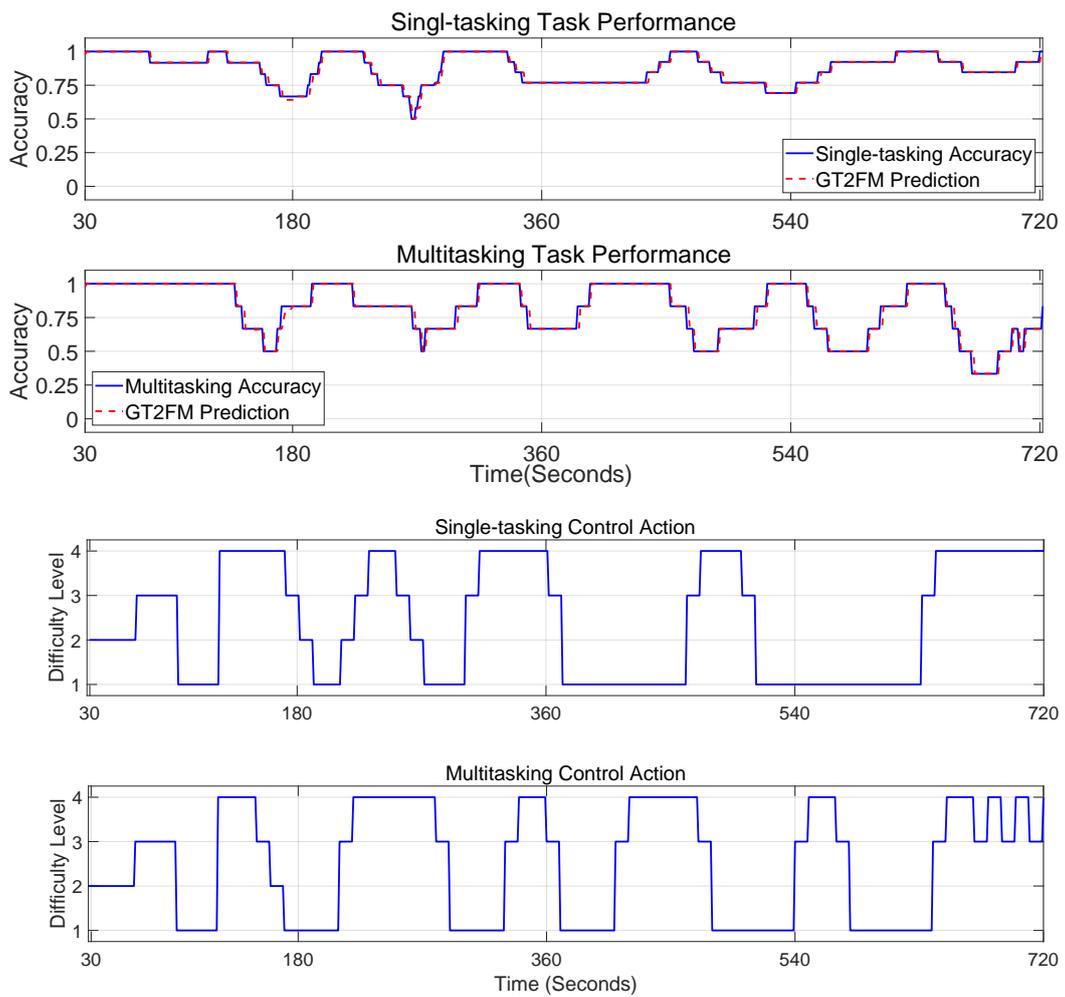


Figure 5.4: Real-time experiment results for the participant 10 in Single-tasking and Multitasking

of the multitasking sessions rose to 11.54 %, whereas the total mean and standard deviation of accuracy performance fell to 0.839 and 0.133. The increases of breakdown likelihood in difficulty levels 1, 2 and 3 and the decreases of the task performance for difficulty levels 1 and 3 were a significant influence on the overall compromised performance. Figure 5.4 provides the real-time experiment results for participant 10 in the sessions with the single-tasking and multitasking. It showed that, compared with the single-tasking session, the task performance of the human operator in multitasking was less stable with the increased likelihood of breakdown.

**Table 5.8:** Breakdown Percentages (%) for the Task Performance of Single-tasking and Multitasking

Sessions	Difficulty Levels				Total
	1	2	3	4	
Single-tasking	2.252	6.344	6.287	20.805	10.279
Multitasking	6.171	7.140	14.765	12.414	11.548

**Table 5.9:** Means of Actual Accuracy for the Task Performance of Single-tasking and Multitasking

Sessions	Difficulty Levels				Total
	1	2	3	4	
Single-tasking	0.862	0.873	0.890	0.861	0.855
Multitasking	0.843	0.884	0.823	0.862	0.839

**Table 5.10:** Standard Deviations of Actual Accuracy for the Task Performance of Single-tasking and Multitasking

Sessions	Difficulty Levels				Total
	1	2	3	4	
Single-tasking	0.131	0.053	0.091	0.108	0.138
Multitasking	0.117	0.079	0.130	0.107	0.133

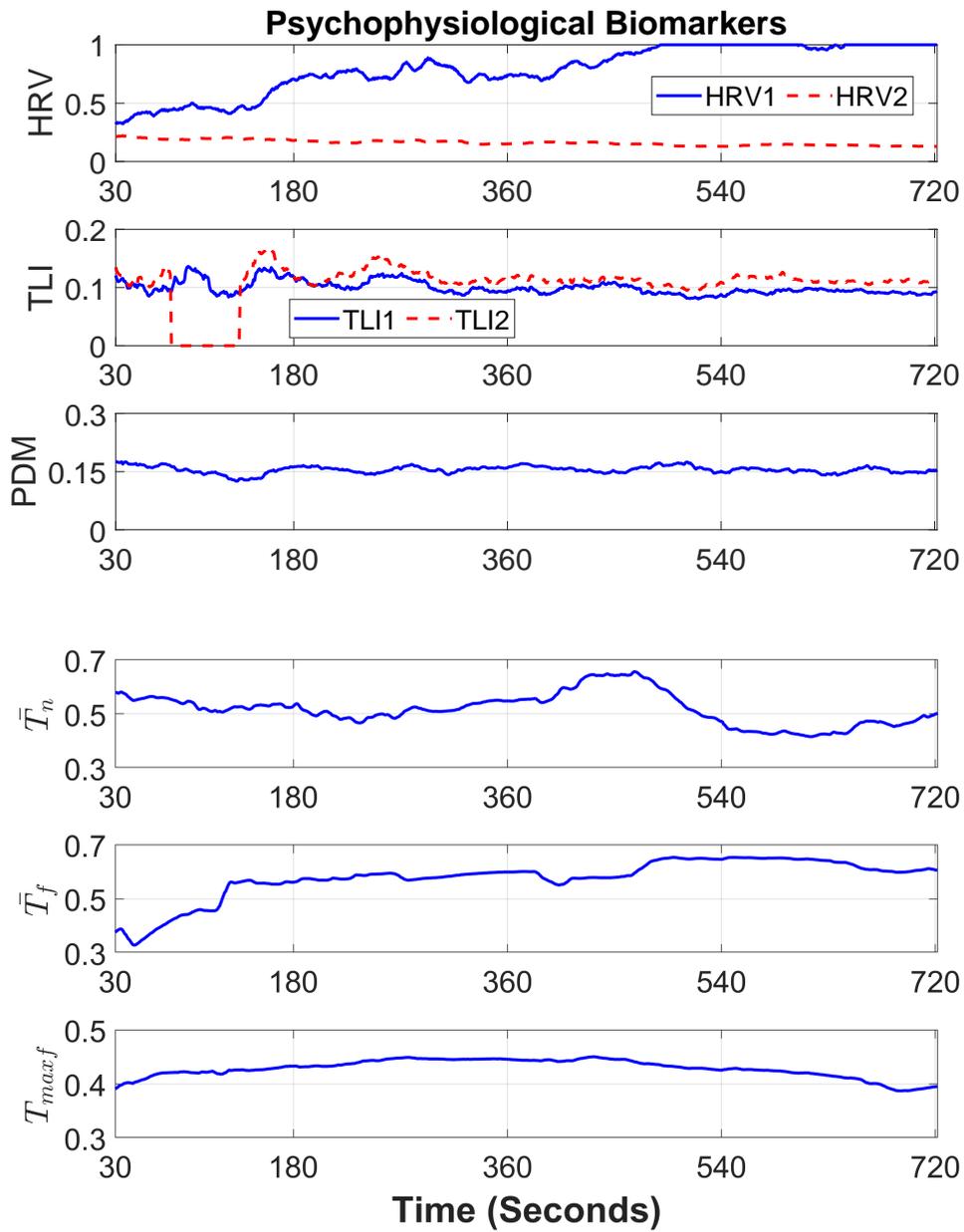
Despite the fact that multitasking was a subjective behaviour, the psychophysiological changes may still be observed and was measurable for the feature extraction of system modelling. A two-sample T-test was applied in this research to compare the differences of psychophysiological biomarker indices between the

normal and multitasking sessions. According to the principle of the two-sample T-test, a significant difference in marker readings provided a  $H$  value of 1, while  $H = 0$  if no difference can be distinguished at a 5 % confidence level. Table 5.11 summarises the  $H$  values for all the psychophysiological biomarkers generated by the same ten participants. It showed that the mean nasal temperature ( $\bar{T}_n$ ), the task load index 1 (TLI1), the mean forehead temperature ( $\bar{T}_f$ ) and the heart rate variable 1 (HRV1) demonstrated significantly different patterns when the participants were engaged in multitasking.

**Table 5.11:** H Value Summary for T-test between Single-tasking and Multitasking

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	0.909	0.727	0.909	1.000	0.884
$\bar{T}_f$	0.909	0.727	0.818	0.900	0.837
$T_{maxf}$	0.909	0.636	0.909	0.800	0.814
HRV1	0.909	0.455	0.800	1.000	0.786
HRV2	1.000	0.727	0.818	0.800	0.837
TLI1	0.818	0.818	0.818	1.000	0.861
TLI2	0.727	0.727	0.727	0.900	0.767
PDM	0.818	0.909	0.818	0.700	0.814

For the  $\bar{T}_n$  and TLI1, observable differences were able to be found in the mean and standard deviation of accuracy performance for each difficulty level and the entire sessions. From the level 1 to 4, the  $\bar{T}_n$  remained between 29 and 30.00 °C with a standard deviation of 2.310 °C in the normal session and rose to 33.080 °C with a reduced deviation of 0.206 °C. This variation may be due to the change of the respiratory system regulation in multitasking. The mean value and the standard deviation of TLI1 decreased from 0.103 to 0.088 and from 0.035 to 0.16 throughout all difficulty levels when switched to multitasking. These suggested that the multitasking suppressed the overall working memory and restrained the activation level of the participant. Similarly to the  $\bar{T}_n$  and TLI1,  $\bar{T}_f$  and HRV1 presented distinguishable differences in their standard deviations between the normal working condition and multitasking, where their values being reduced from 0.760 to 0.140 °C and from 0.212 to 0.038. Such decreases may also be regarded as an indication of the reduced excitement of the participants in multitasking. Figures 5.5



**Figure 5.5:** Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM,  $\bar{T}_n$ ,  $\bar{T}_f$ ,  $T_{maxf}$ ) for the participant 02 in multitasking session

and 5.6 show the psychophysiological biomarkers and task performance of the participant 02 during the multitasking session with HMIBC controller. It may be found that the changes of psychophysiological biomarkers were not entirely synchronised with the change of the participant performance. For example, the task performance met a decrease around 540s, where there was barely any change in the recordings of most biomarkers. These findings corresponded to the psychological theory mentioned in [4, 88] that multiple tasks fought for the limited brain resource, such as attention and working memory, and as a result an additional effort was required for the regulation of these tasks (See Appendix B for the detailed summary).

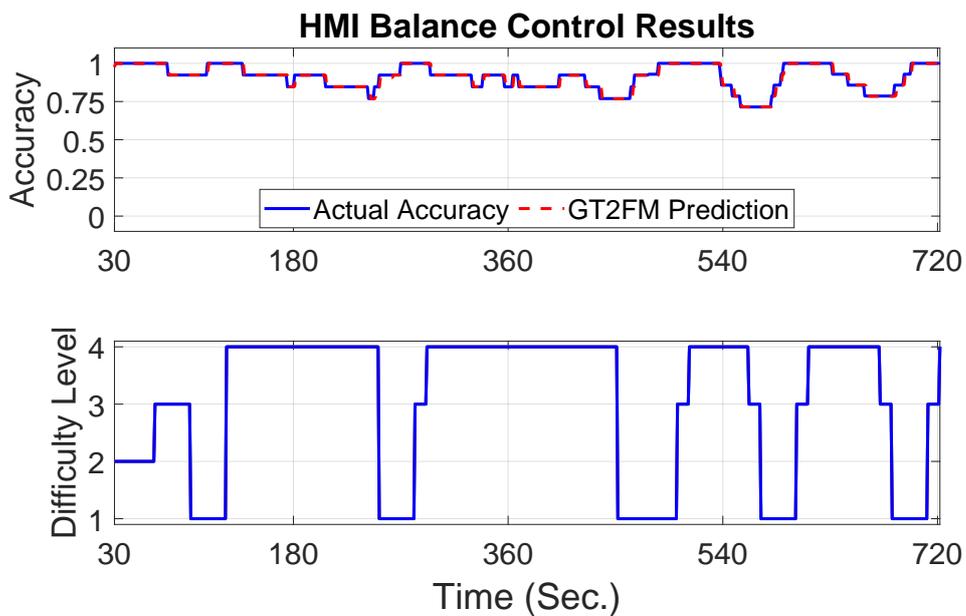
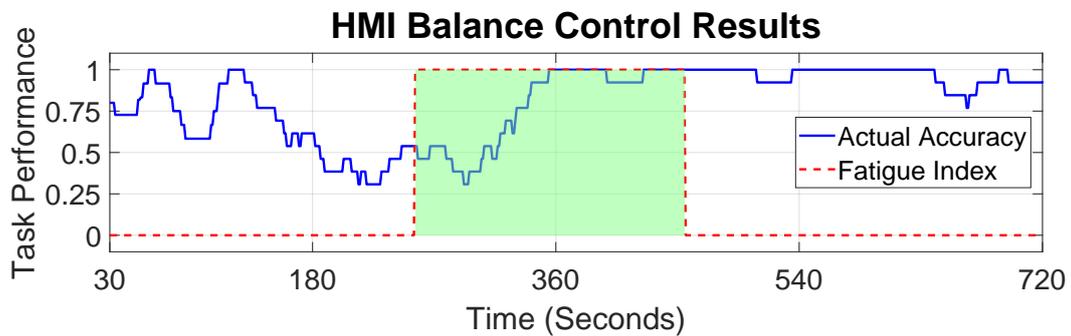


Figure 5.6: Accuracy performance and HMIBC control action from the participant 02 in multitasking session

### 5.4.5 Fatigue

As already stated, fatigue is a subjective cognition of the human psychophysiological conditions that cannot be objectively quantified or qualified. Different from multitasking, fatigue analysis depends on the subjective self-reports from the human operators. Figure 5.7 presents an example of the extraction of the fatigue state



**Figure 5.7:** Real-time experiment results for the participant 03

during this session. The shaded area covered the time when the human operator experiences fatigue-related feelings such as frustration and triteness. From Tables 5.12, 5.13 and 5.14, the statistical results of the ten participant showed, for the normal and the fatigue states, their mean breakdown percentages as being 4.75 % and 11.89 % respectively, and their corresponding means and standard deviations of accuracy performances being 0.868 & 0.133 and 0.735 & 0.157. These figures suggest similar patterns as in other research studies [30, 31, 89, 99] that fatigue is able to compromise human task performance significantly and eventually leads to psychophysiological breakdown.

**Table 5.12:** Breakdown Percentages (%) for the Task Performance of Normal and Fatigue States

Sessions	Difficulty Levels				Total
	1	2	3	4	
Normal	2.657	5.815	7.124	11.917	4.573
Fatigue	1.899	85.008	25.008	50.056	11.895

**Table 5.13:** Means of Actual Accuracy for the Task Performance of Normal and Fatigue States

Sessions	Difficulty Levels				Total
	1	2	3	4	
Normal	0.899	0.877	0.887	0.809	0.868
Fatigue	0.863	0.515	0.790	0.589	0.735

Different from multitasking, fatigue is difficult to identify solely from the psychophysiological biomarkers. Similarly, the two-sample T-test was applied from

**Table 5.14:** Standard Deviations of Actual Accuracy for the Task Performance of Normal and Fatigue States

Sessions	Difficulty Levels				Total
	1	2	3	4	
Normal	0.097	0.107	0.108	0.126	0.133
Fatigue	0.070	0.034	0.070	0.106	0.157

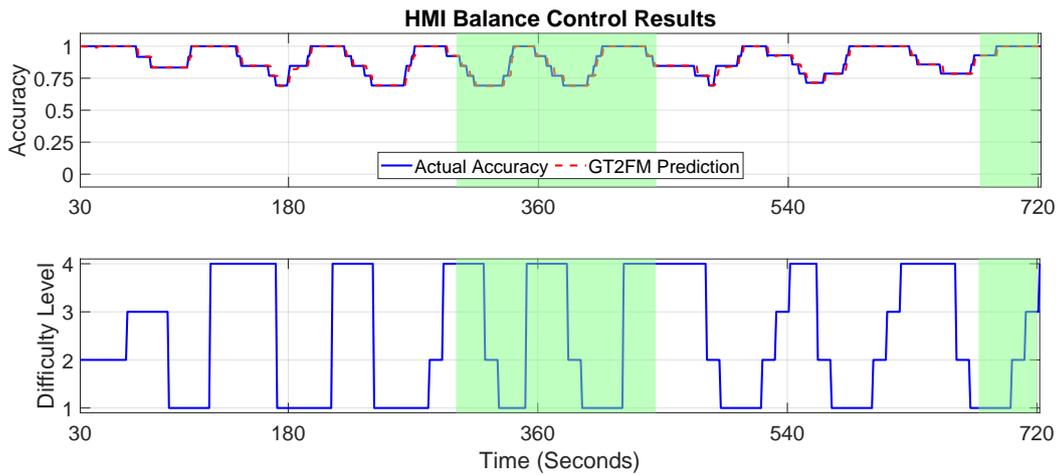
ten-sample experiment to compare these biomarkers' effectiveness and efficiency to differentiate fatigue from the normal psychophysiological state, see Table 5.15. The most correlated indicators appeared to be the heart rate variable 1 (*HRV1*), the mean forehead temperature ( $\bar{T}_f$ ) and the maximum facial temperature ( $T_{maxf}$ ). The relatively low H values for all the biomarkers indicate that fatigue was leaning more towards a psychological state rather than a physiological state, and combining the subjective methods with existing indirect measurements is of great importance for the detection and assessment of the human operator fatigue in the HMI systems.

**Table 5.15:** H Value Summary for T-test between Normal State and Fatigue

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	0.875	0.667	0.750	0.625	0.733
$\bar{T}_f$	0.875	1.000	0.750	0.625	0.800
$T_{maxf}$	0.875	0.833	0.750	0.750	0.800
HRV1	0.750	1.000	0.875	0.875	0.867
HRV2	0.500	0.833	0.875	0.375	0.633
TLI1	0.500	0.667	0.375	0.750	0.567
TLI2	0.500	0.833	0.625	0.625	0.633
PDM	0.750	0.667	0.750	0.625	0.700

Significant differences may be found for HRV1, where the total mean values were 0.460 and 0.623 for the normal state and the fatigue state respectively. The mean value for the fatigue state increased from 0.517 to 0.627 with a reduced standard deviation across the difficulty levels 1 to 4, whereas the mean value for the normal state was around 0.46 with a relatively larger standard deviation. Compared with the normal psychophysiological state,  $\bar{T}_f$  for the fatigue state rose around 1 °C and was maintained at 33.47 °C throughout all difficulty levels.

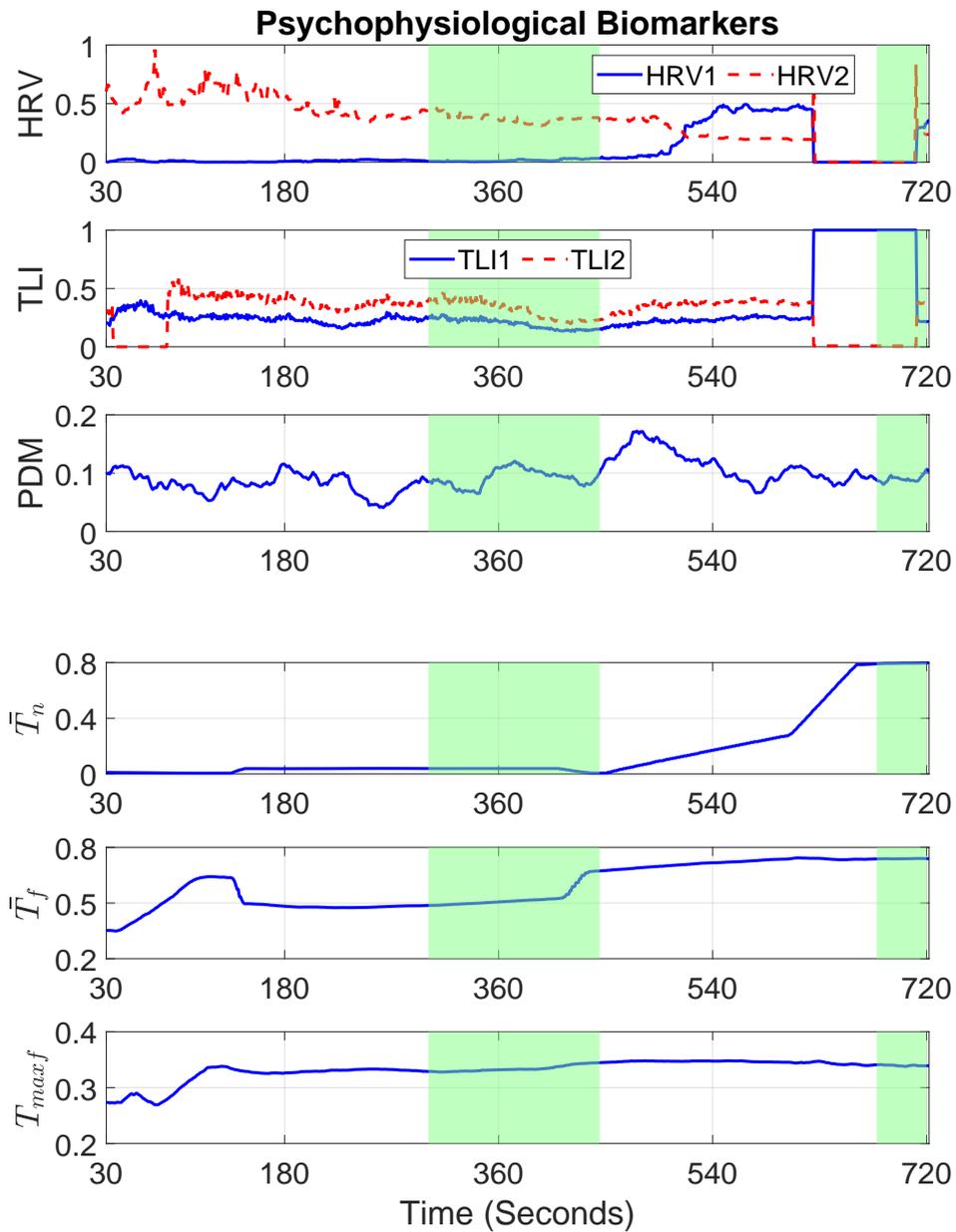
The total standard deviation of the  $\bar{T}_f$  showed an obvious decrease from 0.706 to 0.120 °C when the participant entered the fatigue state. The significant difference for the  $T_{maxf}$  was that the general standard deviations for the normal and the fatigue states were 0.146 °C and 0.025 °C. Figures 5.9 and 5.8 show the psychophysiological biomarkers and task performance of the participant 06 during the multitasking session with HMIBC controller. It may be found that some of the psychophysiological biomarkers were sensitive to the fatigue state. For example, there was a rapid increase for  $\bar{T}_f$  during the fatigue period around 360s. The three indicators above were consistently correlated with the cardiovascular system, and the lack of fluctuation might suggest that the participant lacked engagement for the task (See Appendix C for a more detailed summary).



**Figure 5.8:** Accuracy performance and HMIBC control action from the participant 06 in fatigue session (shade green area represented self-report fatigue)

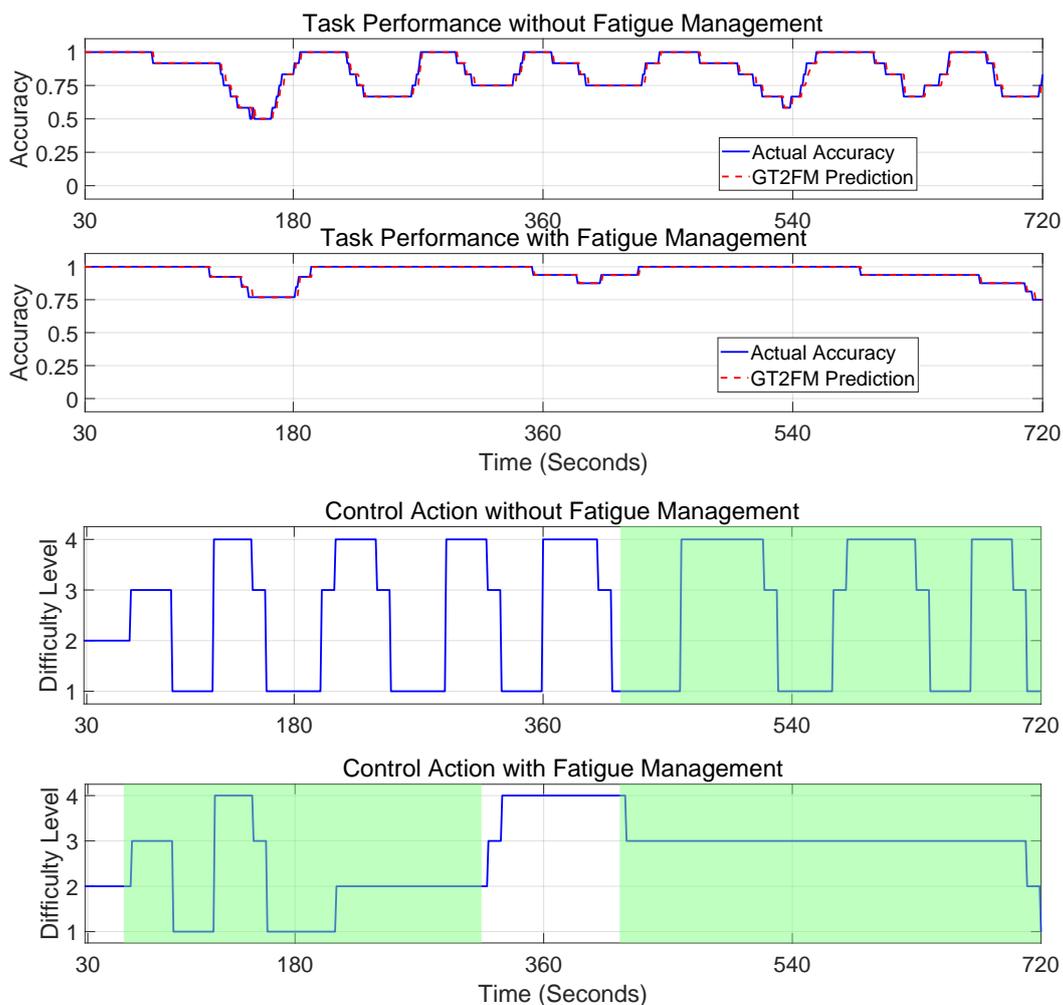
#### 5.4.6 Fatigue Management

Based on the theoretical study [10], fatigue was a psychological state related to the effort-reward equilibrium. For each task, the cognitive system analysed the reward from completing the task and the effort required. Fatigue rose when the effort overcame the reward and encouraged the organism to seek lower reward targets and/or lower effort strategies. In this case, it compromised the task performance of the human operator and eventually caused psychophysiological break-



**Figure 5.9:** Psychophysiological biomarker recordings (HRV1, HRV2, TLI1, TLI2, PDM,  $\bar{T}_n$ ,  $\bar{T}_f$ ,  $T_{maxf}$ ) for the participant 06 in fatigue session (shade green area represented self-report fatigue)

down. The mathematical model in another research [99] illustrated that within a certain period (such as an hour) human effort could be regarded as a self-sustaining oscillation. These findings suggested two strategies for avoiding the breakdown and reversing fatigue to a certain degree - increasing the potential rewards and/or decreasing the task demand. For this research, the potential reward of completing the task was mainly from the satisfaction of answering questions correctly, whereas the task demand was fundamentally established based on the difficulty level. Since the psychological satisfaction was subjective and individual-dependant, the reduction of the task difficulty level during fatigue state was integrated for more intelligent control of the HMI system.



**Figure 5.10:** Real-time HMIBC experiment results for the participant 01 without vs. with the fatigue management

Figure 5.10 presents the real-time experiment results for participant 01 in the sessions with and without fatigue management in the HMIBC. It may be observed from the second set of plots in Figure 5.10 that the controller with fatigue management adjusted the task difficulty level during the periods of fatigue state (the shaded areas) according to the fatigue index. In the first plot of the same figure, without fatigue management integrated in the controller, the task performance nearly decreased to the breakdown point (0.6) around 540 seconds. Compared with the task performance without fatigue management, the task performance with fatigue management showed fewer oscillations and remained at a relatively higher level (see the second plot of Figure 5.10). It is also worth noting that in the session with fatigue management the participant maintained a higher accuracy in difficulty level 4 for a longer time in a single trial. This indicated that fatigue management was not only constructive for the human operator to perform stably and accurately, but it also had allowed for path-opening to advanced human performance in complex situations.

**Table 5.16:** Breakdown Percentages (%) for the HMIBC without/with the Fatigue Management

Sessions	Difficulty Levels				Total
	1	2	3	4	
Original Control	2.251	6.345	6.280	20.800	10.272
With Fatigue	3.454	6.089	3.115	9.012	4.152

**Table 5.17:** Means of Actual Accuracy for the HMIBC without/with the Fatigue Management

Sessions	Difficulty Levels				Total
	1	2	3	4	
Original Control	0.862	0.873	0.890	0.861	0.855
With Fatigue	0.894	0.850	0.851	0.880	0.893

**Table 5.18:** Standard Deviations of Actual Accuracy for HMIBC without/with the Fatigue Management

Sessions	Difficulty Levels				Total
	1	2	3	4	
Original Control	0.131	0.053	0.091	0.108	0.138
With Fatigue	0.107	0.081	0.096	0.070	0.111

Furthermore, Tables 5.16, 5.17 and 5.18 compared the breakdown percentages, means and standard deviations of the original HMIBC session and the HMIBC with the fatigue management session for ten participants. Compared with the task performance in the original HMIBC session, the task performance with fatigue management had a higher accuracy and stability, and the probability of the breakdown was significantly reduced. In Table 5.16, the breakdown percentages of the fatigue management session were significantly reduced for the difficulty levels 3 & 4, which also led to a decrease in the overall percentage. This suggested that fatigue management was an efficient strategy to avoid the psychophysiological breakdown of human operators in the HMI system. According to the results in Table 5.17, fatigue management had ensured the human operators' performance well especially in the difficulty level 4. This suggested that fatigue management was an effective way of inspiring improved human performance, which was essential for solving even more demanding tasks. Additionally, the overall task performance of the human operators was more stable and more consistent with fatigue management as shown in Table 5.18.

## 5.5 Summary

This chapter presented a new type-2 fuzzy logic-based control algorithm for the HMI systems based on the human operator psychophysiological state. The HMI systems were simulated with the mental arithmetic cognitive experiment for 10 participants. In addition to the comparison of the various control methods presented previously, the two main causes of human operator compromised task performance, i.e. multitasking and fatigue, were also investigated.

The new HMIBC approach adjusted the task difficulty level to respond to

the varying human operator psychophysiological state. It combined the state estimation of the adaptive general type-2 fuzzy modelling and the human state-performance correlation described by the type-2 fuzzy logic sets. Compared with the results of the E-MBC, the ones of new HMIBC showed better performances with higher consistency and stability with all the participants.

The experiment results also identified the impacts of the multitasking and fatigue on human operator task performance. In general, multitasking detection mainly depended on the temperature changes in the nasal area and the working memory from the EEG signal, whereas fatigue detection required a combination of these with the subjective self-report. Finally, a new HMIBC integrated with fatigue management achieved the best performance as compared with the other controllers. It suggested the importance of integrating existing psychological findings with balancing the HMI systems.

The combination of existing psychophysiological findings on human task performance and the HMIBC algorithm guaranteed the equilibrium between the task requirement and the human operator feasible effort in the HMI systems. It explored the full potential of the HMI systems without the psychophysiological breakdown of the human operator. This ensured a further possible application of the HMIBC and GT2FM framework in the fields that share similar characteristics with the HMI systems.

## Chapter 6

# Conclusion and Future Work

### 6.1 Conclusion

This research thesis started by introducing the current human psychophysiological studies. It was found that human psychophysiological state played a core role in the task performance of human operators in the HMI. A summary of the existing biomarkers for human psychophysiological state assessment was provided in addition to commonly used HMI simulation configurations from previous research studies. These findings contributed to the experiment design of this research and suggested a human-centred modelling and control for the HMI systems.

Based on the discussions in relation to the facial temperature in the previous studies, this research developed new biomarkers, facial temperature biomarkers, to assess the human psychophysiological state in real-time HMI. With the same experiment configuration across ten participants, the newly developed biomarkers performed equally or even better in differentiating different psychophysiological state than the conventional biomarkers TLI, HRV and PDM, especially in the range of low workload state. The experiment results suggested that the mean nasal temperature and the differential energy between philtrum and forehead were more sensitive to the human psychophysiological state changes in comparison to the conventional biomarkers. The maximum facial temperature and the mean forehead temperature demonstrated clear correlations with the operators' state and task performance in the real-time HMI. These findings proved the efficiency and

effectiveness of using facial temperature biomarkers for the research of the HMI.

The results of participant specific complex fuzzy inference modelling showed that the time-varying pattern was the major issue for current offline models and frameworks. Therefore, an adaptive type-2 fuzzy-based modelling approach was proposed and applied to assess human operators' psychophysiological states for both safety and reliability of the HMI systems. The new modelling approach was compared with offline ANFIS models and online A-GT2-FCM frameworks under the same experimental conditions, and its prediction accuracy exceeded that of previously mentioned approaches for all participants. In terms of both the disturbance rejection and the learning capabilities, the prediction of proposed models achieved the best performance compared to other models that have been presented in the relevant literature. The design of state tracking to update the rule base through a Bayesian process allowed the model to incorporate time-varying patterns and individual difference across participants in real-time. The implementation of type-2 fuzzy rules was able to take into account the inter and intra participant uncertainty. The computational frameworks based on these new features were able to adapt to the dynamic changes within the HMI while maintaining interpret-ability and robustness. These new configurations successfully lead to an adaptive, robust and transparent computational framework that can be utilised to identify dynamic (i.e., real-time) features without prior training. It was believed that this new modelling approach would be a promising development in human-machine interface systems and relevant research studies. The simplicity of the designed type-2 fuzzy logic inference makes it open to new and exciting development of advanced modelling and control mechanisms for many other similar challenges in real world situations where the human-centred design is required.

In addition, this research introduced a new balancing control approach based on the adaptive type-2 fuzzy-based modelling to maintain the equilibrium between the human operators and the automation, along with the exploration of compromised task performance. Validated with the same ten participants in mental arithmetic cognitive experiments, this new innovative control outperformed the energy model-based control presented in the previous research, in terms of break-

down avoidance and human engagement. Meanwhile, the relationships between the operators' performance and abnormal psychophysiological states, fatigue and multitasking, were revealed in the HMI experiments. Furthermore, a selection of biomarkers was agreed upon for the detection of these two compromised performance related states in the HMI. In addition, the new control approach combined with fatigue management demonstrated the best performance with the highest consistency and stability throughout the experiment for all ten participants. These findings open paths for the identification and remedy of the human operator psychophysiological breakdown in the early stage and the further exploration of human psychophysiology in similar human-centred operations.

## 6.2 Future Recommendations

The importance of adaptive modelling and control for the HMI has rapidly increased with the incremental implementation of automatic systems in the real world. Future work should be particularly focused on the psychophysiological state of human operators. As the information flow showed an exponential growth with the development of advanced software and hardware, the operational demands may exceed the maximum capacity of the human operators and endanger the safety and reliability of the whole system. Looking from this perspective, it was of great importance to fully understand the machines behind human task performance, e.g. decision making and action. Compared to the other relevant researches, this research study successfully combined the latest psychophysiological theories and clinical findings with the modelling and control. The use of facial temperature biomarkers and the tracking of human psychophysiological state should open new research directions. However, these were just a mere use of current discoveries and there is a lot more to be explored, such as the facial expressions which have accumulated a lot of attention in human psychophysiological research studies in the field of anthropology [26, 53, 100]. Meanwhile, the advanced imaging technique was able to measure the pulse and calculate heartbeat based on skin image. The development of these non-invasive or contact assessment tools

from imaging analyses may significantly reduce the interference introduced by the psychophysiological measurement.

This research explored the impact of fatigue and multitasking on human task performance and selected possible biomarkers for the identification of these states. Yet, the integration of these findings in the modelling and control of the HMI still remains at its infancy. More studies of neuroscience and psychology are still required to support a clear glimpse of the actual human cognitive process and to create a generalised, transparent and accurate model structure. Further investigation on the fatigue and multitasking should be able to contribute to more intelligent controlling methods with higher level functions such as workload planning and work/rest schedules. The findings of fatigue control indicated the existence of the elasticity within the human capacity. Future research study which focuses on these pattern may be directly related to system safety and human well-being.

It is worth noting that current research study of HMI was limited to the HMI in a relatively short time period. Neither the length of time nor the intensity of workload may be completely matched with the HMI systems as applied in the real world. A human may adapt to the HMI in many different ways and the long term psychophysiological effect is still yet to be discovered. It is reasonable to believe that the documentation of human operator psychophysiological state and the intelligent adaptation of model and control over a long period of time should be one of the next focus points of any future HMI research study.

## Appendix A

# Summary of Two Sample T Test for Facial Temperature Experiment

H value	subsection 1	subsection 2	subsection 3	subsection 4
subsection 1				
subsection 2	0.9000			
subsection 3	1.0000	1.0000		
subsection 4	1.0000	1.0000	0.9000	

**Table A.1:** Overall Maximum Facial Temperature T-test Results for Experimental Session 1

H value	subsection 2	subsection 3	subsection 1	subsection 4
subsection 2				
subsection 3	1.0000			
subsection 1	1.0000	1.0000		
subsection 4	1.0000	1.0000	0.8000	

**Table A.2:** Overall Maximum Facial Temperature T-test Results for Experimental Session 2

**Chapter A. Summary of Two Sample T Test for Facial Temperature Experiment 1**

H value	subsection 1	subsection 2	subsection 3	subsection 4
Subsection 1				
Subsection 2	1.0000			
Subsection 3	1.0000	0.9000		
Subsection 4	0.9000	0.9000	0.9000	

**Table A.3:** Overall Mean Nasal Temperature T-test Results for Experimental Session 1

H value	Subsection 2	Subsection 3	Subsection 1	Subsection 4
Subsection 2				
Subsection 3	0.9000			
Subsection 1	1.0000	1.0000		
Subsection 4	1.0000	0.9000	1.0000	

**Table A.4:** Overall Mean Nasal Temperature T-test Results for Experimental Session 2

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.9000			
Subsection 3	1.0000	1.0000		
Subsection 4	1.0000	1.0000	0.900	

**Table A.5:** Overall Mean Forehead Temperature T-test Results for Experimental Session 1

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.9000			
Subsection 3	0.7000	0.9000		
Subsection 4	0.9000	1.0000	1.0000	

**Table A.6:** Overall Mean Forehead Temperature T-test Results for Experimental Session 2

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	1.0000			
Subsection 3	1.0000	1.0000		
Subsection 4	0.9000	0.9000	1.0000	

**Table A.7:** Overall DEFP T-test Results for Experimental Session 1

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	1.0000			
Subsection 3	1.0000	1.0000		
Subsection 4	0.9000	1.0000	0.9000	

**Table A.8:** Overall DEFP T-test Results for Experimental Session 2

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.9000			
Subsection 3	1.0000	0.9000		
Subsection 4	1.0000	0.9000	0.8000	

**Table A.9:** Overall HRV1 T-test Results for Experimental Session 1

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.9000			
Subsection 3	1.0000	0.9000		
Subsection 4	1.0000	1.0000	0.8000	

**Table A.10:** Overall HRV1 T-test Results for Experimental Session 2

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	1.0000			
Subsection 3	1.0000	1.0000		
Subsection 4	1.0000	1.0000	1.0000	

**Table A.11:** Overall HRV2 T-test Results for Experimental Session 1

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	1.0000			
Subsection 3	1.0000	1.0000		
Subsection 4	1.0000	1.0000	1.0000	

**Table A.12:** Overall HRV2 T-test Results for Experimental Session 2

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.7000			
Subsection 3	0.8000	0.8000		
Subsection 4	0.8000	0.8000	0.5000	

**Table A.13:** Overall TLI1 T-test Results for Experimental Session 1

**Chapter A. Summary of Two Sample T Test for Facial Temperature Experiment 15**

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	1.0000			
Subsection 3	0.7000	0.9000		
Subsection 4	1.0000	0.8000	0.9000	

**Table A.14:** Overall TLI1 T-test Results for Experimental Session 2

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.9000			
Subsection 3	0.9000	0.9000		
Subsection 4	1.0000	0.9000	0.7000	

**Table A.15:** Overall TLI2 T-test Results for Experimental Session 1

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	1.0000			
Subsection 3	0.8000	1.0000		
Subsection 4	1.0000	1.0000	0.8000	

**Table A.16:** Overall TLI2 T-test Results for Experimental Session 2

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.9000			
Subsection 3	0.8000	1.0000		
Subsection 4	1.0000	1.0000	1.0000	

**Table A.17:** Overall PDM T-test Results for Experimental Session 1

H value	Subsection 1	Subsection 2	Subsection 3	Subsection 4
Subsection 1				
Subsection 2	0.8000			
Subsection 3	1.0000	0.8000		
Subsection 4	0.9000	0.9000	0.9000	

**Table A.18:** Overall PDM T-test Results for Experimental Session 2

## Appendix B

# Summary of Means and Standard Deviations of the Psychophysiological Biomarkers for Multitasking

**Table B.1:** Mean Values of Biomarkers in Single-tasking

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	30.181	29.935	30.928	29.472	30.129
$\bar{T}_f$	33.320	32.981	33.228	33.220	33.187
$T_{maxf}$	36.026	35.989	35.969	35.946	35.982
HRV1	0.649	0.447	0.584	0.599	0.570
HRV2	0.236	0.231	0.218	0.244	0.232
TLI1	0.089	0.111	0.099	0.112	0.102
TLI2	0.095	0.127	0.112	0.121	0.114
PDM	0.102	0.123	0.128	0.122	0.119

**Table B.2:** Mean Values of Biomarkers in Multitasking

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	33.200	32.893	33.060	33.174	33.080
$\bar{T}_f$	33.830	33.502	33.659	33.723	33.677
$T_{maxf}$	35.940	35.951	35.930	35.933	35.939
HRV1	0.731	0.434	0.579	0.621	0.591
HRV2	0.196	0.277	0.259	0.205	0.235
TLI1	0.078	0.097	0.093	0.084	0.088
TLI2	0.088	0.106	0.112	0.098	0.101
PDM	0.134	0.116	0.115	0.119	0.121

**Table B.3:** Standard Deviations of Biomarkers in Single-tasking

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	2.399	2.506	2.386	1.945	2.309
$\bar{T}_f$	0.677	0.925	0.812	0.626	0.760
$T_{maxf}$	0.166	0.165	0.136	0.115	0.146
HRV1	0.232	0.187	0.217	0.212	0.212
HRV2	0.093	0.063	0.083	0.080	0.080
TLI1	0.034	0.048	0.028	0.031	0.035
TLI2	0.040	0.060	0.040	0.040	0.045
PDM	0.047	0.049	0.041	0.043	0.045

**Table B.4:** Standard Deviations of Biomarkers in Multitasking

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	0.283	0.171	0.174	0.196	0.206
$\bar{T}_f$	0.198	0.122	0.119	0.121	0.140
$T_{maxf}$	0.049	0.035	0.030	0.038	0.038
HRV1	0.117	0.059	0.098	0.095	0.092
HRV2	0.031	0.028	0.027	0.037	0.030
TLI1	0.020	0.010	0.011	0.022	0.015
TLI2	0.021	0.008	0.008	0.026	0.016
PDM	0.046	0.022	0.016	0.027	0.028

## Appendix C

# Summary of Means and Standard Deviations of the Psychophysiological Biomarkers for Fatigue

Table C.1: Mean Values of Biomarkers in Normal State

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	29.249	28.590	30.052	29.519	29.336
$\bar{T}_f$	33.110	32.583	32.724	33.139	32.893
$T_{maxf}$	36.032	36.030	35.954	35.953	35.993
HRV1	0.550	0.358	0.408	0.521	0.460
HRV2	0.252	0.258	0.269	0.264	0.261
TLI1	0.105	0.119	0.112	0.119	0.114
TLI2	0.107	0.141	0.133	0.124	0.126
PDM	0.109	0.122	0.124	0.125	0.120

**Table C.2:** Mean Values of Biomarkers in Fatigue State

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	29.849	32.115	31.796	29.191	30.646
$\bar{T}_f$	33.463	33.585	33.496	33.355	33.467
$T_{maxf}$	35.965	35.819	35.842	35.911	35.889
HRV1	0.517	0.651	0.706	0.627	0.623
HRV2	0.292	0.180	0.201	0.212	0.224
TLI1	0.082	0.082	0.087	0.097	0.087
TLI2	0.077	0.035	0.077	0.094	0.073
PDM	0.137	0.138	0.113	0.171	0.140

**Table C.3:** Standard Deviations of Biomarkers in Normal State

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	1.777	1.502	2.146	1.920	1.829
$\bar{T}_f$	0.658	0.716	0.853	0.612	0.706
$T_{maxf}$	0.159	0.138	0.160	0.127	0.146
HRV1	0.180	0.116	0.162	0.184	0.161
HRV2	0.082	0.055	0.093	0.088	0.079
TLI1	0.030	0.044	0.033	0.021	0.032
TLI2	0.037	0.059	0.046	0.034	0.044
PDM	0.042	0.053	0.053	0.043	0.047

**Table C.4:** Standard Deviations of Biomarkers in Fatigue State

Difficulty Level	1	2	3	4	Mean
$\bar{T}_n$	0.878	0.028	0.862	0.013	0.473
$\bar{T}_f$	0.186	0.018	0.163	0.087	0.120
$T_{maxf}$	0.055	0.004	0.031	0.003	0.024
HRV1	0.123	0.028	0.116	0.020	0.075
HRV2	0.068	0.003	0.053	0.008	0.035
TLI1	0.013	0.003	0.008	0.005	0.007
TLI2	0.009	0.002	0.003	0.004	0.004
PDM	0.020	0.010	0.009	0.014	0.013

# References

- [1] Headcaps. <https://www.biosemi.com/headcap.htm>. Accessed: 2020-02-20.
- [2] S. Akselrod, D. Gordon, F.A. Ubel, D.C. Shannon, A.C. Barger, and R.J. Cohen. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *science*, pages 220–222, 1981.
- [3] M. Asgari, G. Kiss, J. Van Santen, I. Shafran, and X. Song. Automatic measurement of affective valence and arousal in speech. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 965–969. IEEE, 2014.
- [4] N. Axmacher, M.M. Henseler, O. Jensen, I. Weinreich, C.E. Elger, and J. Fell. Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, 107(7):3228–3233, 2010.
- [5] A. Baddeley. Working memory. *Science*, 255(5044):556–559, 1992.
- [6] L.F. Barrett. Solving the emotion paradox: Categorization and the experience of emotion. *Personality and social psychology review*, 10(1):20–46, 2006.
- [7] L.F. Barrett, B. Mesquita, K.N. Ochsner, and J.J. Gross. The experience of emotion. *Annu. Rev. Psychol.*, 58:373–403, 2007.
- [8] M.Y. Bekkedal, J. Rossi, and J. Panksepp. Human brain eeg indices of emotions: delineating responses to affective vocalizations by measuring frontal theta event-related synchronization. *Neuroscience & Biobehavioral Reviews*, 35(9):1959–1970, 2011.

- [9] L. Bernardi, J. Wdowczyk-Szulc, C. Valenti, S. Castoldi, C. Passino, G. Spadacini, and P. Sleight. Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. *Journal of the American College of Cardiology*, 35(6):1462–1469, 2000.
- [10] M.A. Boksem and M. Tops. Mental fatigue: costs and benefits. *Brain research reviews*, 59(1):125–139, 2008.
- [11] M.M. Bradley, M. Codispoti, D. Sabatinelli, and P.J. Lang. Emotion and motivation ii: sex differences in picture processing. *Emotion*, 1(3):300, 2001.
- [12] G.L. Brengelmann. Body surface temperature: Manifestation of complex anatomy and physiology of the cutaneous vasculature. In *Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE*, volume 3, pages 1927–1930. IEEE, 2000.
- [13] J.T. Cacioppo, L.G. Tassinary, and G. Berntson. *Handbook of psychophysiology*. Cambridge University Press, 2007.
- [14] N. Charkoudian. Skin blood flow in adult human thermoregulation: how it works, when it does not, and why. In *Mayo Clinic Proceedings*, volume 78, pages 603–612. Elsevier, 2003.
- [15] Z. Chen, S. Aghakhani, J. Man, and S. Dick. Ancfis: A neurofuzzy architecture employing complex fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 19(2):305–322, 2010.
- [16] J. Clay-Warner and D.T. Robinson. Infrared thermography as a measure of emotion response. *Emotion Review*, 7(2):157–162, 2015.
- [17] R. Colombo, G. Mazzuero, F. Soffiantino, M. Ardizzoia, and G. Minuco. A comprehensive pc solution to heart rate variability analysis in mental stress. In *Computers in Cardiology 1989, Proceedings.*, pages 475–478. IEEE, 1989.
- [18] S. Dick. Toward complex fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 13(3):405–414, 2005.

- [19] D.F. Dinges. An overview of sleepiness and accidents. *Journal of sleep research*, 4:4–14, 1995.
- [20] J. Dowdall, I.T. Pavlidis, and P. Tsiamyrtzis. Coalitional tracking. *Computer Vision and Image Understanding*, 106(2):205–219, 2007.
- [21] R. Drake, A.W. Vogl, and A.W. Mitchell. *Gray's Anatomy for Students E-Book*. Elsevier Health Sciences, 2009.
- [22] H. Driessen and Y. Boers. Map estimation in particle filter tracking. 2008.
- [23] E. El-Samahy, M. Mahfouf, L. Torres-Salomao, and J. Anzures-Marin. A new computer control system for mental stress management using fuzzy logic. In *Evolving and Adaptive Intelligent Systems (EAIS), 2015 IEEE International Conference on*, pages 1–7. IEEE, 2015.
- [24] W. Ellermeier and W. Westphal. Gender differences in pain ratings and pupil reactions to painful pressure stimuli. *Pain*, 61(3):435–439, 1995.
- [25] P. Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, 9(10):474–480, 2005.
- [26] H. Gao, A. Yüce, and J.P. Thiran. Detecting emotional stress from facial expressions for driving safety. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 5961–5965. IEEE, 2014.
- [27] A. Garde, B. Laursen, A. Jørgensen, and B. Jensen. Effects of mental and physical demands on heart rate variability during computer work. *European journal of applied physiology*, 87(4-5):456–461, 2002.
- [28] A. Gevins and M.E. Smith. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, 4(1-2):113–131, 2003.
- [29] A. Gevins, M.E. Smith, L. McEvoy, and D. Yu. High-resolution eeg mapping of cortical activation related to working memory: effects of task difficulty,

- type of processing, and practice. *Cerebral cortex (New York, NY: 1991)*, 7(4): 374–385, 1997.
- [30] C.D. Griffith and S. Mahadevan. Inclusion of fatigue effects in human reliability analysis. *Reliability Engineering & System Safety*, 96(11):1437–1447, 2011.
- [31] G. Gunzelmann, M.D. Byrne, K.A. Gluck, and L.R. Moore Jr. Using computational cognitive modeling to predict dual-task performance with sleep deprivation. *Human Factors*, 51(2):251–260, 2009.
- [32] E. Harmon-Jones, P.A. Gable, and C.K. Peterson. The role of asymmetric frontal cortical activity in emotion-related phenomena: A review and update. *Biological psychology*, 84(3):451–462, 2010.
- [33] C. He, M. Mahfouf, and L. A. Torres-Salomao. An adaptive general type-2 fuzzy logic approach for psychophysiological state modeling in real-time human-machine interfaces. *IEEE Transactions on Human-Machine Systems*, 51(1):1–11, 2021.
- [34] C. He, M. Mahfouf, and L.A. Torres-Salomao. Thermal imaging for psychophysiological state detection in the human-machine interface (hmi) control system. In *International Conference on Informatics in Control, Automation and Robotics*, pages 214–229. Springer, 2018.
- [35] C. He, M. Mahfouf, and L.A. Torres-Salomao. Facial temperature markers for mental stress assessment in human-machine interface (hmi) control system. In *ICINCO*, 2018.
- [36] J.A. Healey and R.W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on intelligent transportation systems*, 6(2):156–166, 2005.
- [37] H. Heuer, T. Kleinsorge, W. Klein, and O. Kohlisch. Total sleep deprivation increases the costs of shifting between simple cognitive tasks. *Acta Psychologica*, 117(1):29–64, 2004.

- [38] N. Hjortskov, D. Rissén, A.K. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard. The effect of mental stress on heart rate variability and blood pressure during computer work. *European journal of applied physiology*, 92(1-2):84–89, 2004.
- [39] K. Hong and S. Hong. Real-time stress assessment using thermal imaging. *The Visual Computer*, 32(11):1369–1377, 2016.
- [40] Y. Houdas and E. Ring. *Human body temperature: its measurement and regulation*. Springer Science & Business Media, 2013.
- [41] M. Inzlicht, B.D. Bartholow, and J.B. Hirsh. Emotional foundations of cognitive control. *Trends in cognitive sciences*, 19(3):126–132, 2015.
- [42] D. Kahneman. *Thinking, fast and slow*. Macmillan, 2011.
- [43] M.C. Kennedy and A. O’Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.
- [44] I. Koch, E. Poljac, H. Müller, and A. Kiesel. Cognitive structure, flexibility, and plasticity in human multitasking—an integrative review of dual-task and task-switching research. *Psychological bulletin*, 144(6):557, 2018.
- [45] S. Koelstra, C. Muhl, M. Soleymani, J.S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1):18–31, 2012.
- [46] W.J. Kop, D.S. Krantz, R.H. Howell, M.A. Ferguson, V. Papademetriou, D. Lu, J.J. Popma, J.F. Quigley, M. Vernalis, and J.S. Gottdiener. Effects of mental stress on coronary epicardial vasomotion and flow velocity in coronary artery disease: relationship with hemodynamic stress responses. *Journal of the American College of Cardiology*, 37(5):1359–1366, 2001.
- [47] Y. Kuriyagawa and I. Kageyama. A modeling of heart rate variability to estimate mental work load. In *Systems, Man, and Cybernetics, 1999. IEEE SMC’99*

- Conference Proceedings. 1999 IEEE International Conference on*, volume 2, pages 294–299. IEEE, 1999.
- [48] Y. Kuriyagawa and I. Kageyama. A modeling of heart rate variability to estimate mental work load. In *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on*, volume 2, pages 294–299. IEEE, 1999.
- [49] P.J. Lang and M.M. Bradley. Emotion and the motivational brain. *Biological psychology*, 84(3):437–450, 2010.
- [50] A. Levin, D. Lischinski, and Y. Weiss. A closed-form solution to natural image matting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):228–242, 2008.
- [51] C. Lithari, C. Frantzidis, C. Papadelis, A.B. Vivas, M. Klados, C. Kourtidou-Papadeli, C. Pappas, A. Ioannides, and P. Bamidis. Are females more responsive to emotional stimuli? a neurophysiological study across arousal and valence dimensions. *Brain topography*, 23(1):27–40, 2010.
- [52] M. Mahfouf, J. Zhang, D.A. Linkens, A. Nassef, P. Nickel, G.R.J. Hockey, and A.C. Roberts. Adaptive fuzzy approaches to modelling operator functional states in a human-machine process control system. In *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*, pages 1–6. IEEE, 2007.
- [53] I.B. Mauss and M.D. Robinson. Measures of emotion: A review. *Cognition and emotion*, 23(2):209–237, 2009.
- [54] J. McGlone. Sex differences in human brain asymmetry: A critical survey. *Behavioral and Brain Sciences*, 3(2):215–227, 1980.
- [55] J.M. Mendel. *Uncertain rule-based fuzzy logic systems: introduction and new directions*. Prentice Hall PTR Upper Saddle River, 2001.
- [56] E.S. Mezzacappa, R.M. Kelsey, E.S. Katkin, and R.P. Sloan. Vagal rebound and recovery from psychological stress. *Psychosomatic medicine*, 63(4):650–657, 2001.

- [57] F. Mokhayeri, M. Akbarzadeh-T, and S. Toosizadeh. Mental stress detection using physiological signals based on soft computing techniques. In *Biomedical Engineering (ICBME), 2011 18th Iranian Conference of*, pages 232–237. IEEE, 2011.
- [58] F. Mokhayeri and M. Akbarzadeh-T. Mental stress detection based on soft computing techniques. In *Bioinformatics and Biomedicine (BIBM), 2011 IEEE International Conference on*, pages 430–433. IEEE, 2011.
- [59] A. Nassef, M. Mahfouf, D. Linkens, E. Elsamahy, A. Roberts, P. Nickel, G. Hockey, and G. Panoutsos. The assessment of heart rate variability (hrv) and task load index (tli) as physiological markers for physical stress. In *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany*, pages 146–149. Springer, 2009.
- [60] A. Nassef, C.H. Ting, M. Mahfouf, D.A. Linkens, P. Nickel, G.R.J. Hockey, and A.C. Roberts. A new framework for real-time adaptive fuzzy monitoring and control for humans under psychophysiological stress. In *BIOSIGNALS (2)*, pages 320–325, 2008.
- [61] A. Nassef, M. Mahfouf, C.H. Ting, E. El-Samahy, D.A. Linkens, and M.A. Denai. Hybrid physiological modeling of subjects undergoing cyclic physical loading. In *Biosignals*, pages 252–257, 2010.
- [62] B. Nhan and T. Chau. Infrared thermal imaging as a physiological access pathway: a study of the baseline characteristics of facial skin temperatures. *Physiological measurement*, 30(4):N23, 2009.
- [63] B.R. Nhan and T. Chau. Classifying affective states using thermal infrared imaging of the human face. *IEEE Transactions on Biomedical Engineering*, 57(4):979–987, 2010.
- [64] T.D.O. Noakes. Fatigue is a brain-derived emotion that regulates the exercise behavior to ensure the protection of whole body homeostasis. *Frontiers in physiology*, 3:82, 2012.

- [65] A. Nozawa and M. Tacano. Correlation analysis on alpha attenuation and nasal skin temperature. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(01):P01007, 2009.
- [66] T. Partala and V. Surakka. Pupil size variation as an indication of affective processing. *International journal of human-computer studies*, 59(1):185–198, 2003.
- [67] L. Pessoa. On the relationship between emotion and cognition. *Nature reviews. Neuroscience*, 9(2):148, 2008.
- [68] L. Pessoa. How do emotion and motivation direct executive control? *Trends in cognitive sciences*, 13(4):160–166, 2009.
- [69] S.W. Porges. Cardiac vagal tone: a physiological index of stress. *Neuroscience & Biobehavioral Reviews*, 19(2):225–233, 1995.
- [70] D. Ramot, R. Milo, M. Friedman, and A. Kandel. Complex fuzzy sets. *IEEE Transactions on Fuzzy Systems*, 10(2):171–186, 2002.
- [71] D. Ramot, M. Friedman, G. Langholz, and A. Kandel. Complex fuzzy logic. *IEEE Transactions on Fuzzy Systems*, 11(4):450–461, 2003.
- [72] P. Ren, A. Barreto, Y. Gao, and M. Adjouadi. Affective assessment of computer users based on processing the pupil diameter signal. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 2594–2597. IEEE, 2011.
- [73] J. Ring and R. Dear. Temperature transients: a model for heat diffusion through the skin, thermoreceptor response and thermal sensation. *Indoor Air*, 1(4):448–456, 1991.
- [74] R. Sloan, P. Shapiro, E. Bagiella, S. Boni, M. Paik, J. Bigger, R. Steinman, and J. Gorman. Effect of mental stress throughout the day on cardiac autonomic control. *Biological psychology*, 37(2):89–99, 1994.

- [75] M.E. Smith, A. Gevins, H. Brown, A. Karnik, and R. Du. Monitoring task loading with multivariate eeg measures during complex forms of human-computer interaction. *Human Factors*, 43(3):366–380, 2001.
- [76] R.M. Stern, W.J. Ray, and K.S. Quigley. *Psychophysiological recording*. Oxford University Press, USA, 2001.
- [77] M.G. Stokes, M. Kusunoki, N. Sigala, H. Nili, D. Gaffan, and J. Duncan. Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, 78(2): 364–375, 2013.
- [78] J.R. Stroop. Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6):643, 1935.
- [79] D.E. Tamir, N.D. Rishe, and A. Kandel. Complex fuzzy sets and complex fuzzy logic an overview of theory and applications. In *Fifty years of fuzzy logic and its applications*, pages 661–681. Springer, 2015.
- [80] C.H. Ting, M. Mahfouf, D.A. Linkens, A. Nassef, P. Nickel, A.C. Roberts, M.H. Roberts, and G.R.J. Hockey. Real-time adaptive automation for performance enhancement of operators in a human-machine system. In *Control and Automation, 2008 16th Mediterranean Conference on*, pages 552–557. IEEE, 2008.
- [81] C.H. Ting, M. Mahfouf, A. Nassef, D.A. Linkens, G. Panoutsos, P. Nickel, A.C. Roberts, and G.R.J. Hockey. Real-time adaptive automation system based on identification of operator functional state in simulated process control operations. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(2):251–262, 2010.
- [82] L. Torres-Salomao, M. Mahfouf, and E. El-Samahy. Pupil diameter size marker for incremental mental stress detection. In *E-health Networking, Application & Services (HealthCom), 2015 17th International Conference on*, pages 286–291. IEEE, 2015.

- [83] L.A. Torres-Salomao, M. Mahfouf, and O. Obajemu. Interval type-2 fuzzy logic adaptive modelling for human operators undergoing mental stress. *IFAC Proceedings Volumes*, 47(3):9880–9885, 2014.
- [84] L.A. Torres-Salomao, M. Mahfouf, E. El-Samahy, and C.H. Ting. Psychophysiological based real-time adaptive general type 2 fuzzy modeling and self-organizing control of operator’s performance undertaking a cognitive task. *IEEE Trans. Fuzzy Systems*, 25(1):43–57, 2017.
- [85] R. Usamentiaga, P. Venegas, J. Guerediaga, L. Vega, J. Molleda, and F.G. Bulnes. Infrared thermography for temperature measurement and non-destructive testing. *Sensors*, 14(7):12305–12348, 2014.
- [86] A. Vincent, F.I. Craik, and J.J. Furedy. Relations among memory performance, mental workload and cardiovascular responses. *International Journal of Psychophysiology*, 23(3):181–198, 1996.
- [87] S. Walter, C. Wendt, J. Böhnke, S. Crawcour, J.W. Tan, A. Chan, K. Limbrecht, S. Gruss, and H.C. Traue. Similarities and differences of emotions in human–machine and human–human interactions: what kind of emotions are relevant for future companion systems? *Ergonomics*, 57(3):374–386, 2014.
- [88] X.J. Wang. Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, 90(3):1195–1268, 2010.
- [89] D.W. Watson. Physiological correlates of heart rate variability (hrv) and the subjective assessment of workload and fatigue in-flight crew: a practical study. In *Human Interfaces in Control Rooms, Cockpits and Command Centres, 2001. People in Control. The Second International Conference on (IEE Conf. Publ. No. 481)*, pages 159–163. IET, 2001.
- [90] H.Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. 2012.
- [91] K. Yamanaka and M. Kawakami. Convenient evaluation of mental stress

- with pupil diameter. *International journal of occupational safety and ergonomics*, 15(4):447–450, 2009.
- [92] O. Yazdanbakhsh and S. Dick. Time-series forecasting via complex fuzzy logic. In *Frontiers of Higher Order Fuzzy Sets*, pages 147–165. Springer, 2015.
- [93] O. Yazdanbakhsh and S. Dick. Forecasting of multivariate time series via complex fuzzy logic. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47(8):2160–2171, 2017.
- [94] O. Yazdanbakhsh, A. Krahn, and S. Dick. Predicting solar power output using complex fuzzy logic. In *2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS)*, pages 1243–1248. IEEE, 2013.
- [95] R. Zeelenberg, E.J. Wagenmakers, and M. Rotteveel. The impact of emotion on perception: Bias or enhanced processing? *Psychological Science*, 17(4): 287–291, 2006.
- [96] R. Zeelenberg, E.J. Wagenmakers, and M. Rotteveel. The impact of emotion on perception: Bias or enhanced processing? *Psychological Science*, 17(4): 287–291, 2006.
- [97] J. Zhai and A. Barreto. Stress detection in computer users based on digital signal processing of noninvasive physiological variables. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 1355–1358. IEEE, 2006.
- [98] J. Zhai, A.B. Barreto, C. Chin, and C. Li. Realization of stress detection using psychophysiological signals for improvement of human-computer interactions. In *SoutheastCon, 2005. Proceedings. IEEE*, pages 415–420. IEEE, 2005.
- [99] K. Zhang and X. Li. Human-robot team coordination that considers human fatigue. *International Journal of Advanced Robotic Systems*, 11(6):91, 2014.

- 
- [100] R. Zheng, S. Yamabe, K. Nakano, and Y. Suda. Biosignal analysis to assess mental stress in automatic driving of trucks: Palmar perspiration and masseter electromyography. *Sensors*, 15(3):5136–5150, 2015.
- [101] Y. Zhou, P. Tsiamyrtzis, and I.T. Pavlidis. Tissue tracking in thermophysiological imagery through spatio-temporal smoothing. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 1092–1099. Springer, 2009.
- [102] Y. Zhou, P. Tsiamyrtzis, P. Lindner, I. Timofeyev, and I. Pavlidis. Spatiotemporal smoothing as a basis for facial tissue tracking in thermal imaging. *IEEE Transactions on Biomedical Engineering*, 60(5):1280–1289, 2013.