Fan, S and Yang, Z

Towards objective human performance measurement for maritime safety: A new psychophysiological data-driven machine learning method

Article

Reliability Engineering and System Safety 233 (2023) 109103

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Towards objective human performance measurement for maritime safety: A new psychophysiological data-driven machine learning method

Shiqi Fan, Zaili Yang *

Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK

A R T I C L E I N F O

Keywords:
Human performance
Human reliability
Human errors
Maritime transport
Maritime education and training
Maritime safety

A B S T R A C T

Human errors significantly contribute to transport accidents. Human performance measurement (HPM) is crucial to ensure human reliability and reduce human errors. However, how to address and reduce the subjective bias introduced by assessors in HPM and seafarer certification remains a key research challenge. This paper aims to develop a new psychophysiological data-driven machine learning method to realize the effective HPM in the maritime sector. It conducts experiments using a functional Near-Infrared Spectroscopy (fNIRS) technology and compares the performance of two groups in a maritime case (i.e. experienced and inexperienced seafarers in terms of different qualifications by certificates), via an Artificial Neural Network (ANN) model. The results have generated insightful implications and new contributions, including (1) the introduction of an objective criterion for assessors to monitor, assess, and support seafarer training and certification for maritime authorities; (2) the quantification of human response under specific missions, which serves as an index for a shipping company to evaluate seafarer reliability; (3) a supportive tool to evaluate human performance in complex emerging systems (e.g. Maritime Autonomous Surface Ship (MASS)) design for ship manufactures and shipbuilders.

1. Introduction

Human factors have been responsible for many transport accidents, across different modes of road, aviation, and maritime transport. For instance, human actions contribute to 60.6% of the investigated marine casualties [1]. Comparatively, the fatalities per accident in the rail, aviation and maritime sectors are higher than in road transport [2,3]. Recently, such low-frequent but high-consequent risky sectors have attracted public interest because of the high uncertainty they are associated with from a safety science perspective. To address such challenges, reliable human performance becomes critical for ensuring safety. Within the context of maritime transport, the development of ship automation stimulates the incorporation of technical and non-technical skills in human performance measurement (HPM) in today's maritime safety studies [4–6]. It is evident that 30.77% of maritime accidents are associated with "poor communication and coordination" and 32.69% "under ineffective supervision and support of the bridge team" [7]. Communication is a significant factor, and teams without effective communication increase the risk of committing errors. An ineffective supervision issue can be seen for lone watchkeeper or working isolated, which exposes hazards derived from workload pressure. Reducing crews onboard and reallocating operators onshore induce new scenarios and challenges for the risk assessment of modern shipping navigation [8]. Although reducing the crew means lower labour costs for ship companies, new risks are observed with technological devices and changing workloads. Improper use of devices and equipment (e.g., BNWAS switched off, alarm system not noticed) contributes to human errors, as evident by the statistic that 37.98% of maritime accidents are associated with it [7]. On the other hand, automation development does not secure a lower workload for seafarers. Over-reliance on Automatic Identification System (AIS) and GPS may result in poor lookout, and the associated factors may contribute to information overload in terms of seafarer cognition. Therefore, an effective tool to realise accurate HPM has become more urgent than even among all the prioritized requirements for maritime safety.

Maritime education and training (MET) serve as a gatekeeper for seafarer recruitment. The regulation on international training standards for seafarers is under the Standards of Training Certification and Watchkeeping (STCW) [9], formed by the International Maritime Organization (IMO) and the United Nations (UN). There are growing concerns about the effectiveness and professionality of the current implemented seafarer certification regime. The STCW emphasises a

* Corresponding author.
E-mail address: z.yang@ljmu.ac.uk (Z. Yang).

https://doi.org/10.1016/j.ress.2023.109103
Received 29 September 2022; Received in revised form 9 January 2023; Accepted 16 January 2023
Available online 18 January 2023

requirement for the qualified training and assessment of seafarers [9]. Simulator training has been widely utilized among industrial operators since the 1990s [10]. It integrates tasks and application context to achieve knowledge transfer [11]. Given this practice, seafarers acquire competencies in the required maritime operations, and assessors/examiners evaluate human performance in the simulation for training purposes. The evaluation can also be supported by the simulator data and parameter calculation [12]. In the current seafarer certification process, an expert is required to evaluate HPM against each competence criterion and, in the meantime, identify all the associated human errors in bridge simulation [13]. However, the practice of having experts undertake the conventional assessment of seafarers in simulation has been criticized due to the possibility of introducing subjective bias. For instance, serial positioning effects, which means people have a better memory of the first and last actions [14]; recognition-primed inferences, which implies that people tend to take actions that are familiar to the assessor and their ability is evaluated by how the assessor would accomplish it [15]. Such bias affects what information is retrieved and perceived. Due to such bias, there is a chance for unqualified trainees to be certified to carry out challenging tasks beyond their actual capacity, endangering maritime safety. It reveals a new research need to improve the validity and reliability of the current maritime HPM methods [16]. In terms of a large number of examinees, limited examination approaches, and insufficient quality of the examiners, the maritime administration has improved subjective training by developing multiple-choice computer-based assessments [16], which pioneers the measures in the maritime HPM domain towards a more objective direction.

In light of human performance in maritime operations, there was evidence of a positive correlation between individual factors and behaviors [17–19]. Various factors, such as mental workload, emotion, and fatigue, affect human behaviors in daily work and decision-making processes [20,13]. To quantify such risk factors, novel and advanced methodologies should be proposed and applied to operator assessment, for which physiological means become a promising solution in recent years. The Electrocardiography (ECG), Electromyography (EMG), Electroencephalography (EEG), skin electrical response, and eye-tracking have been applied to measure the physiological response of human operators [13,21]. However, their associations with human performance and safe navigation have not been thoroughly examined. The relevant studies for driving behaviors were developed in road transport [22,23]. Scant research involves neurophysiological techniques in maritime human reliability analysis (HRA) [20,24]. With the fast development of autonomous ships, the decreasing number of crews onboard and the reallocated responsibility ashore will lead to a fundamental change in MET. Under such circumstances, an effective and reliable assessment method for seafarer qualification is urgently needed to fit and support the fundamental changes introduced by advanced and autonomous vessels. In summary, the demand for seafarer qualification assessment becomes increasingly significant given the limitations of the existing approaches and emerging technologies such as digitalization and autonomation.

This paper aims to develop a new psychophysiological data-driven machine learning method for supporting maritime HPM and facilitating seafarers' qualification evaluation towards an objective perspective. fNIRS has become a popular wearable sensor, a non-invasive brain imaging modality for measuring cortical haemodynamic activity [25] and attracts great interest in risk assessment in some high safety-sensitive sectors [26,27]. Seafarers were divided into two levels of experience based on their STCW qualifications in this study. The experienced group is defined as those who have obtained higher ranked certificates (e.g. MM (master), CM (chief mate), OOW (officer of the watch)), while the inexperienced group consists of the AB (able seaman) and cadets. It develops a new ship-bridge simulator-based scenario to conduct experiments with fNIRS technology and compare the performance of two groups. If the finding reveals the significant differences between experienced group and inexperienced group in terms of their performance indicated by the fNIRS data, while in the meantime, the fNIRS data-driven ANN result of each group is kept a high consistency, we gain the confidence to use fNIRS data and our proposed new methodology to introduce a new objective measure to complement the current purely subjective method in the maritime qualification process in MET. The generated implications and contributions can therefore help shift the paradigm of MET accreditation and qualification/certification assessment in maritime transport. Within the maritime transport context, the practical contributions of this research include (1) incorporation of an objective criterion into the monitoring, assessment, and support of maritime training for maritime authorities; (2) quantification of human response under specific missions, which serves as an index relating to the operator competence of a shipping company; (3) supportive tool to evaluate the human performance in terms of complex system and Maritime Autonomous Surface Ship (MASS) design for ship manufacturers and shipbuilders.

Furthermore, the theoretical contribution of this paper is to pioneer the use of fNIRS and simulation for HPM in the maritime transport context and compare the performance of seafarers to address the possible subjectivity in the traditional assessment. It develops a new assessment model using a psychophysiological data-driven machine learning approach. The results reveal the potential of the employment of fNIRS data in an assessment framework as a criterion to monitor, assess, and support maritime training.

The structure of the paper is illustrated as follows. Section 2 reviews the literature on operators' training and human performance assessment, as well as the relative HRA research in maritime transportation. Section 3 explains the research methods for the experimental study in maritime scenarios. Then, the results are discussed in Section 4. Section 5 describes the implications of the assessment method and its practical implications. Finally, it is followed by the conclusion in Section 6.

2. Literature review

2.1. Operators' training and performance measurement

Initially aiming at high-risk operations in medical and nuclear sectors, operators' training and performance measurements evaluate the operators' competence so as to improve operation safety. With increased human factor concerns in the transportation field, the relevant literature reveals the significance of operators' HPM in transport safety and uncovers diverse methodologies to assess them. In the case of high-risk operations (e.g. in the nuclear sector), the operators' training is inevitably associated with training simulators that can provide safe and economical working situations [28,29]. The simulator was utilized to investigate human performance based on the parameter calculation and data obtained from regular practical simulator training [12].

The HPM covering technical skills and non-technical skills in the transportation field is quite diverse. It relies on traditional and novel techniques, such as simulators [30], physiological data [31–34], and visual reality (VR) [30]. Meta-analytic research validated the use of cognitive ability, structured interviews, and several personality factors as predictors of training performance [35]. Towards the end of training, HPM is normally introduced as an assurance of the involved trainees' competence and reliability. In this process, different qualifications and certifications are given through examinations by assessors subjectively in the transport sector. It has been argued for a long time about the bias introduced by the assessors. However, few effective solutions have been developed and seen in the current literature, despite the promising direction by the incorporation of psychological methods into the assessment.

Specifically, in the aviation sector, a VR-based training simulator was relatively inexpensive and safe to provide operator training, regardless of the weather conditions [30]. The proposed simulator was utilized for various levels of the trainee and heavy equipment training simulator. To replicate critical transport scenarios, a helicopter simulator was designed for the training purpose of preparing interdisciplinary teams of students for collaborative practice, research, and leadership [36]. In the maritime sector, Jo et al. [37] conducted a survey about maritime cadets' perceptions towards changes in shipping organizations, and the competence of seafarers required in the MASS era using clustering. It was evident that "the traditional seafarers' centric role retainer", the "ship organizational structure domain achiever", and the "new technical competencies builder" were the new competency requirements. In addition, analytical hierarchy analysis (AHP) and principal performance shaping factors (PSFs) were utilized to identify the factors that affect the performance of operators for oil processing and treatment [38]. Regarding train operators, EEG, ECG, electrodermal activity (EDA), photoplethysmography (PPG) and respiration signals were measured to evaluate human performance during train operations [34]. It was evident that the intensity of the beta wave of EEG was higher before and after the train stop due to the increased level of mental arousal and tension. The visual data obtained from a railway overhead gantry equipped with multiple cameras was utilized to estimate safety requirements that can process large imagery of trains and help train operators detect possible malfunction [39]. Dhalmahapatra et al. [40] assessed the developed immersive VR-based safety training simulator that provided training for electric overhead crane (EOT) operators to understand and manage the potential hazards in EOT operations. Moreover, Brandenburger et al. [41] identified train drivers' roles and personnel selection criteria for a highly automated high-speed passenger train. In addition, the transforming from a train driver into a train operator required mental persistence and sustained attention, as well as sharing in tasks of information integration and decision-making [41].

Previous studies illustrated the associated criteria for individual competence (i.e., technical skills, non-technical skills, perceived awareness, and psychophysiological response) are yet well defined, especially in the transportation field. There is a significant research gap in integrating objective measurement with subjective remarking (such as expert knowledge and questionnaires) to support HPM. Wrongly certifying an unqualified operator could lead to the occurrence of catastrophic accidents. Additional assurance beyond the current best practice in HPM in maritime transportation is therefore strongly needed with urgency. One realistic way is to integrate objective HPM using psychophysiological data into the current purely subjective assessment in the maritime sector to generate a hybrid assessment model.

2.2. Human performance in maritime transportation using neurophysiological methods

The assessment of human performance using neurophysiological methods helps recognize task difficulties and evaluate the qualifications of operators under different circumstances. It has been widely applied to evaluate operators' performance on a newly designed system in terms of practical capability [42,43]. As shown in Table 1, there are an increasing number of methods for neurophysiological measurements, such as wearable eye-tracking devices, EEG, and fNIRS, which were integrated with simulators, motion capture devices, heart rate measurement, and augmented reality, to understand operators' behavior and decision-making patterns in various scenarios [20,44–46]. Compared to the other means, the fNIRS has the advantages of higher spatial resolution, flexible restriction on body movement, and relatively low cost. However, such advantages are much more benefited in road transport safety [47,48] and aviation transportation [49,50] than maritime transport. It means that scarce studies are investigated in the maritime domain [51], revealing an applied research gap to fulfil. It is particularly worrisome when the two facts are taken into account, (1) 60.6% of maritime casualties are related to human actions; and (2) MASS develops fast, involving human performance and reliability in a new and largely unknown working environment such as mixed manned and autonomous ships at sea and remote centers onshore [52].

Table 1
Neurophysiological method.

Technique	Application	Advantage	Disadvantage
Eye tracking	Operators' gaze patterns in aviation [53]	Wearable	Limited robust
EEG	Monitor emotion, mental workload and stress in maritime operations [21]	Greater time resolution	Sensitive to body motion and noisy
fNIRS	Applied in real-world scenarios [54] Working memory and cognitive load [25]	Greater spatial resolution, less crosstalk between sites	Sensitive to the light

HPM in maritime transportation is associated with mental workload, fatigue, and situation awareness (SA). Among them, various methods such as the National Aeronautics and Space Administration's Task Load Index (NASA-TLX) [55,56], eye response [45,57], ECG [58], EEG [44, 46], and fNIRS [20] were utilized for the workload assessment. The objective analysis benefited the training and evaluation of seafarers' performance under maritime simulators [44]. Also, time pressure was proved to be significant in affecting operation accuracy and eye movements [59]. In addition, mental fatigue was measured by the inventory or checklist [19,60], EEG [61], and survey [62]. However, fatigue self-assessment was unreliable, so it needed complementation by physiological measurement [63]. Moreover, the SA in the maritime industry was interpreted based on the information obtained from the surrounding environment, including ship, equipment, route, and weather. The SA of operators was proved to be associated with the willingness to take risks [64], which illustrated that the operator performance was influenced by the perceived and comprehended SA. Different from relevant research developed in the aviation and road transport [26,65–69], the seafarers' neurophysiological study reveals more restrains on the experimental design. The seafarers' decision making does not immediately lead to timely results due to ship manoeuvring characteristics. The analysis of their response under routine work is of equal significance as the one in an emergency situation.

Even with the development of MASS, there is increasing attention on the operator assessment and training requirements given new challenges introduced by the remote-control mechanism. Evidence showed that maintaining psycho- and physiological conditions is significant for remote operators [70]. It further proved the critical role of psychological assessment for operators of MASS. In light of emerging complicated situations and visibility constraints of MASS, operators' mental workload has changed, which further influenced the operators' overload threshold [56]. Additionally, a MASS regulatory framework on the competence of remote operators was proposed to redefine the SA by integrating the STCW and Goal-Based Gap Analysis [71].

There are increasing studies utilizing machine learning techniques to propose data-driven methodologies in transportation and reliability fields [7,72,73]. Such methodologies can provide objective risk assessments and decision-making insights in maritime transportation. Regarding ship automation systems, the operational data was measured to investigate ship energy performance and examine the decision to equip vessels with hybrid propulsion [74]. To support the decision support systems, potential ship grounding scenarios and risks were identified and evaluated through a data-driven model using AIS data, nowcast data, and seafloor depth data [75]. Such data-driven methods quantify the performance measurements with minimized subjective bias, which provides a new perspective on maritime safety.

Currently, the best practice of HPM in the marine industry is to use subjective judgements, which is probably biased for the performance evaluation of seafarers. Previous research in maritime transportation has yet to incorporate any neurophysiological methodology into seafarers' performance assessment. Furthermore, there is not a well-established methodology for psychological measurement in HPM;

moreover, the question as to how to quantify and convert the psychological criterion for seafarer qualification remains unanswered in the maritime sector. Therefore, the development of a hybrid approach to predict seafarers' qualifications will open a new direction for improving HPM in the sector. An fNIRS-based HPM during routine work can significantly improve the current practice in MET. Within this context, it can not only fulfil the research gap that is largely left open in the existing literature but also tackle new human-related risks due to the growth of the use of digitalised technologies (e.g. MASS) in maritime operations.

2.3. New contributions

Based on a new data-driven machine learning approach, this study uses an fNIRS technology to collect seafarers' psychological data and an artificial neural network (ANN) to realize the prediction and classification of seafarers with different levels of qualifications. More specifically, the theoretical contributions include:

(1) Holistic use of fNIRS and maritime simulation to derive objective data that can be used to model HPM.
(2) Development of a hybrid assessment model using haemoglobin data and ANN.
(3) Pioneering the use of machine learning methods (e.g. ANN) to deal with psychophysiological data and predict seafarers' experience in the maritime area.
(4) Real case analysis through a hybrid approach to classifying seafarers of different qualifications.

From a managerial perspective, maritime administrative/authorities can use the findings of this work to define an objective criterion/value that can be used to supplement subjective assessment methods and evaluate a seafarer's competence more rationally. It will aid in avoiding/reducing errors in the certification process of seafarer qualification. When the result of the traditional subjective assessment is different from it of the new method, an in-depth assessment should be launched to reassure the qualification is well reflected by the actual competence of the tested seafarer.

3. Methods

In this section, the study conducted experiments and compared the performance of two groups (i.e., experienced and inexperienced seafarers with different levels of qualifications). Two groups of seafarers (e.g., experienced and inexperienced) of 40 participants (20 in each group) were invited to undertake an assessment of ship anti-collision in a ship bridge simulation. This study collected the psychophysiological data of maritime operators under two different levels of workloads, i.e. distraction and no distraction situations. The "no distraction" situation was developed in several different timeframes, including a 5 min baseline, a 20 min watchkeeping period with observation and attention maintained till the target vessel was spotted, a 10 min decision-making phase till the evasive manoeuvre was performed. On the other hand, the "distraction" situation was developed based on the same timeframes, where an additional task was required to report the position of the own ship.

The objective HPM was conducted using a psychophysiological data-driven machine learning method. Specifically, the analysis consisted of fNIRS data and ship distance data. Among them, raw fNIRS data was pre-processed and trained using ANN model. The results of the model presented the predicted participant's experience. The results revealed the possibility of using psychophysiological data to predict the experience. Then the distance between ships was compared and analyzed to explain human behaviors within the two distinguished groups.

3.1. Participants

The participants were recruited from the Nautical Institute UK, as shown in Table 2. There were 20 experienced seafarers, with an average age of 44.6 years old. Their average employment time was 213.4 months. The 20 inexperienced participants were, on average, 25 years old, with 27.2 months of service time as seafarers. Regarding the criteria of participant selection, all participants require not to have a history of head injury or currently taking medication for anxiety and high blood pressure [20].

3.2. Experiment protocol

There were two levels of workload situations developed for the experiment. The experiment protocol is illustrated in Fig. 1. After a 5 min baseline period, the participants were requested to keep a lookout in the bridge simulator. Such a watchkeeping period ended when participants spotted the target vessel (they pressed the button to inform investigator). On average, the duration of the watchkeeping period was about 20 min. With reference to the baseline time, the watchkeeping was divided into four sections (w1, w2, w3, w4), with each section for about 5 min. Then, it came into a decision-making period. The participants needed to observe the position and action of target vessels and then evaluate whether and when to make a manoeuvre. When they altered the course of the own ship (it was recorded in the simulator), the exercise ended. However, to reflect the features at the early (after spotting the vessel) and later (before manoeuvring) stages, the decision-making process was divided into two sections (d1, d2), with each for about 5 min. The fNIRS data were collected through the above procedures (w1-d2).

The fNIRS data were measured by Nirsport 88 continuous wave fNIRS device that consists of 8 sources and 8 detectors that emit near-infrared light. It was collected to reveal the brain activities of participants in different sections of the experiment, which reveals objective HPM using psychophysiological data. The deoxygenated haemoglobin (HbR) and HbO levels were obtained from the montage designed by NIRSite, as shown in Fig. 2. This montage covered the prefrontal cortex area that explained brain functions of working memories and decision-making. There were 7 sources and 7 detectors, resulting in a total of 15 channels of HbO and HbR. The specific montage was divided into three sub-areas: left dorsolateral prefrontal cortex (DLPFC) (channel 1–5); central DLPFC (channel 6–10); right DLPFC (channel 11–15).

In addition, care was taken to avoid hair from the eyebrows or side of the head interfering with detectors and sources. The staff in the control room next to the simulator recorded the target spotted time with the corresponding distance (distance 1), and the course changed time with the corresponding distance (distance 2).

3.3. Data analysis

3.3.1. fNIRS raw data analysis

Raw fNIRS data (15 channels) was pre-processed. The Interpolate function was used to fill the data in each channel where there was detector saturation. Then the data quality function was applied to check and identify any "poor quality" channels in which the signal was too weak. After removing discontinuities and spike artefacts, a low-pass filter was applied to reduce high-frequency instrument noise and

Table 2
Two groups of participants.

Group	Age (year)	Employment (month)	Gender	STCW qualification
Experienced	44.6	213.4	20 males	MM, CM, OOW
Inexperienced	25	27.2	18 males, 2 females	AB, cadet

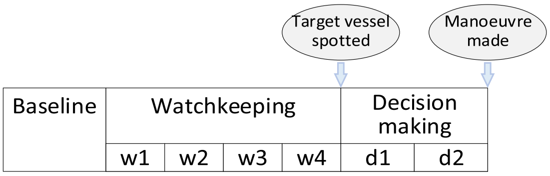

Fig. 1. Experiment protocol.

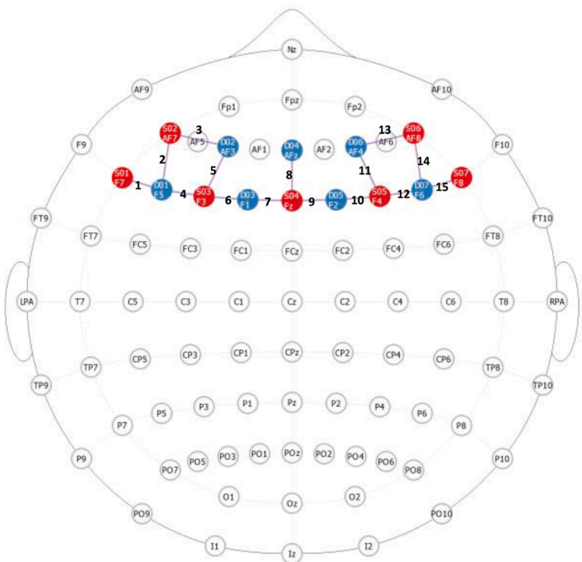

Fig. 2. fNIRS montage - where redpoint refers to 'Source', blue one refers to 'Detector', and purple lines refer to channels (source: Authors).

physiological noise such as fast cardiac oscillations. The pre-processed data was imported for haemodynamic state calculation using the modified Beer-Lambert law [76]. It reveals changes in HbO, HbR and total haemoglobin (Hb). All fNIRS results were reported in micromoles (μM).

The analysis was conducted to investigate how onboard tasks influenced the neurophysiological activation of experienced and inexperienced seafarers concerning collision avoidance. Moreover, it determined differences between left, central, and right DLPFC activities. As far as the data analysis is concerned, there was a transformation of the data called Correlation-Based Signal Improvement (CBSI) that forces HbO and HbR to be negatively correlated and controls for head movement, which was developed by Cui et al. [77]. As HbR is transformed into the inverse of HbO after this point, only HbO data were used in the subsequent analyzes.

The results of fNIRS data analysis demonstrated significant brain activities in different sections for different groups. This statistical analysis helps extract elements as input data for the experience prediction model.

3.3.2. Artificial neural network modelling

In the current literature, basic statistical methods have been used for the fNIRS data analysis. Few studies utilize machine learning methods to facilitate data analysis and processing. ANN is derived from biological neural networks with multiple neurons. It has a single input and output but may also have none, one or many hidden layers. Among an orientated network, the layers of parallel processing elements are neurons. Each layer is connected to the backward layer by interconnection strengths or weights. A perceptron network with one or more hidden layers is called a multilayer perceptron network, which is widely implemented in the ANN [78]. The ANN has the advantage of offering numerical models on relationships between complex nonlinear data but does not require any prior assumption [79]. Also, it can naturally reduce the effect of the data noise. Compared with other machine learning methods such as Support Vector Machine (SVM), ANN has the disadvantage of overfitting [80]. However, it can be mitigated by integrating with other optimization approaches such as the genetic algorithm (GA) [81]. The SVM often outperforms ANN in terms of small data due to an improved parameter selection [79].

This study uses the multilayer perceptron network model of ANN to predict seafarers' experience. The ratio of training data to testing data was selected as 7:3, which aided in yielding the highest prediction accuracy in the literature [82]. The filtered channels of fNIRS which revealed a significant main effect on the experience were selected as neurons to conduct the ANN modelling. The participants' experience was the dependant variable, and a multilayer perception network method was utilized. The ANN results will reveal the predicted participant's experience, so as to enable the objective HPM using psychophysiological data-driven approach.

4. Results

This study conducts objective HPM using a psychophysiological data-driven machine learning method, supported by human behavior analysis. The HPM analysis consists of two parts: (1) ANN prediction model using fNIRS data; (2) human behavior comparison using ship distance data. On the one hand, the raw fNIRS data is pre-processed, followed by ANOVA analysis. The significant ROI during the significant period is selected as the input of the ANN prediction model to predict seafarer experience. On the other hand, the ship distance data extracted from the simulator is used as a support to analyze human behaviors, revealing the manoeuvre differences between the two groups. The data analysis flow is shown in Fig. 3.

4.1. Neuron definition using fNIRS data

There are 40 participants' fNIRS data for the analysis. The fNIRS data are divided into three Regions of Interest (ROI) corresponding to the left-lateral, medial and right-lateral areas of prefrontal cortex. After data pre-processing, HbO data are averaged for each task period, i.e. four periods of watchkeeping and two periods of decision-making. HbO data for each ROI are subjected to a 2 (experienced/inexperienced) x 2

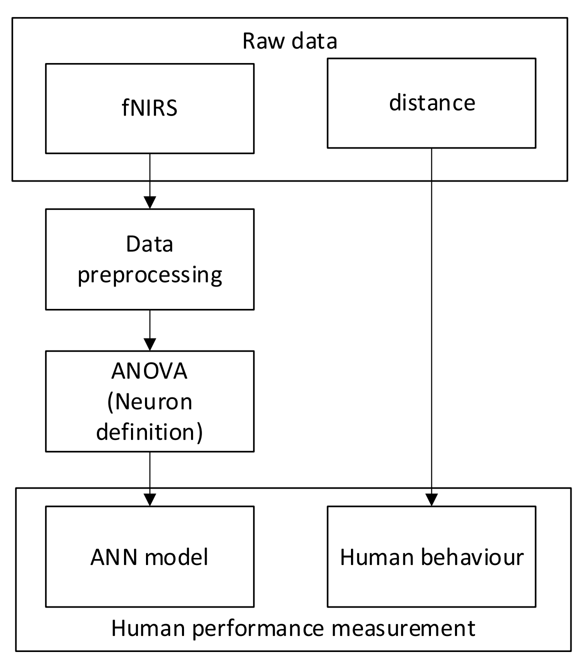

Fig. 3. Data analysis flow.

(distraction/no-distraction) x 6 (task period) ANOVA. This statistical analysis helps define neurons as the input of the ANN model.

Analysis of left-lateral and medial ROI fails to reveal any statistically significant main effects or interactions. However, analysis of HbO data from the right-lateral ROI reveals a significant main effect for task period [$F(5,30) = 3.76$, $p=.02$, $\eta_p^2=0.4$], as well as significant interactions between experience x task period [$F(5,30) = 2.30$, $p=.05$, $\eta_p^2=0.27$]. Post-hoc testing indicates that average HbO at the right-lateral ROI is significantly lower during W3 and W4 than all other periods ($p < .05$); this effect is illustrated in Fig. 4 [20]. At the beginning of wathkeeping phase (w1), the right-lateral ROI's HbO levels reflect the normal state of seafarers with wahckeeping task. At the later phase of watchkeeping (w3 and w4), it shows a decline of HbO, which indicates the boredom of seafarers in these periods. From d1 (decision 1) period, seafarers' HbO levels have increased after they spot the target ship. Then, they are supposed to think about whether to take actions for collision avoidance while observing the target ship and making the decision, which makes the HbO level continuously increase at d2 (decision 2) period.

The results illustrate the value of investigating seafarers' specific areas of brain activities in a task. In the ANN model, the HbO level of particular brain areas with significant differences can be extracted as data input to predict the model. Therefore, the right-lateral ROI is selected for the subsequent step modelling. Regarding MET assessments, the changes in HbO show the possibility of utilising psychophysiological data to predict the professional levels of trainees. It predicts seafarers' qualification levels under different scenarios, complementing the subjective evaluation by assessors. In addition, when the traditional HPM by assessors is inconsistent with the result predicted by fNIRS, extra measures should be taken to ensure seafarers reach the proper level of competence.

The interaction effect between experience x task period is also explored using t-tests, as seen in Fig. 5. These tests reveal that the average HbO is higher for the experienced participants at the right-lateral ROI, but only during the fourth period of watchkeeping (w4) when the ship is spotted [$t(36)=2.78$, $p < .01$]. It means that HbO of experienced group is significantly different from the HbO of inexperienced group at the right-lateral ROI during w4. Such fNIRS data in channels 11–15 can be used to distinguish experienced and inexperienced groups.

The results demonstrate the significant differences among two groups of seafarers in HbO levels during the fourth period of watchkeeping (w4). In the ANN model, the HbO level during w4 can be extracted as the input data to predict the model. It shows that different tasks or stages of a task are related to varying levels of HbO of humans. With the development of MASS, the new workspace will introduce new duties and new competence requirements. However, current knowledge and assessment criteria for assessors to evaluate MASS operators are limited at the early stage of MASS. Given challenges in a remote-control mechanism, psychological assessment for operators of MASS plays a critical role. The results of fNIRS data analysis provide a perspective to observe and evaluate human performance in designed scenarios for advanced ships. It remedies the insufficient expert knowledge in judging operator qualifications in MASS environments. In this way, utilising fNIRS data in a specific period to predict the experience of operators can serve as a reliable validation and evaluation method for HPM in everyday working scenes onboard or onshore.

Therefore, the HbO at the right-lateral ROI is extracted for neurons used as the input of the proposed ANN prediction model. In other words, the fNIRS data of channels 11–15 during w4 period are utilized as neurons for further analysis. A detailed statistical analysis of significant differences in results can be found in Fan et al. [20].

4.2. Human performance measurement

HPM in this paper consists of two parts. The first part classifies the seafarer experience using fNIRS data through the ANN model, and the second part analyzes the human behavior among two groups using ship distance data. The study utilizes a multilayer perceptron network with one hidden layer to generate the ANN model. The input of the ANN model is 36,529 pieces of fNIRS data, each with 5 columns (channels 11–15) reflecting the HbO of the right-lateral ROI; the output is the predicted "experience". This numerical model reveals relationships between fNIRS data of 5 channels and the participant's experience. The results of the first part (ANN model) show that it is possible to extract the fNIRS data to predict the seafarer experience, so as to judge whether the seafarer is qualified to be an experienced professional. The second part shows the differences in the distance when ship was spotted and the distance when the manoeuvre was made. The results of two parts show that using fNIRS and distance data helps objectively predict the experience level of a seafarer and evaluate human behaviors.

4.2.1. Model prediction and human performance for a non-distraction situation

A non-distraction situation is with the timeframe of baseline, watchkeeping, and decision making. The fNIRS data of the right ROI are extracted as neurons because they are significantly higher during the

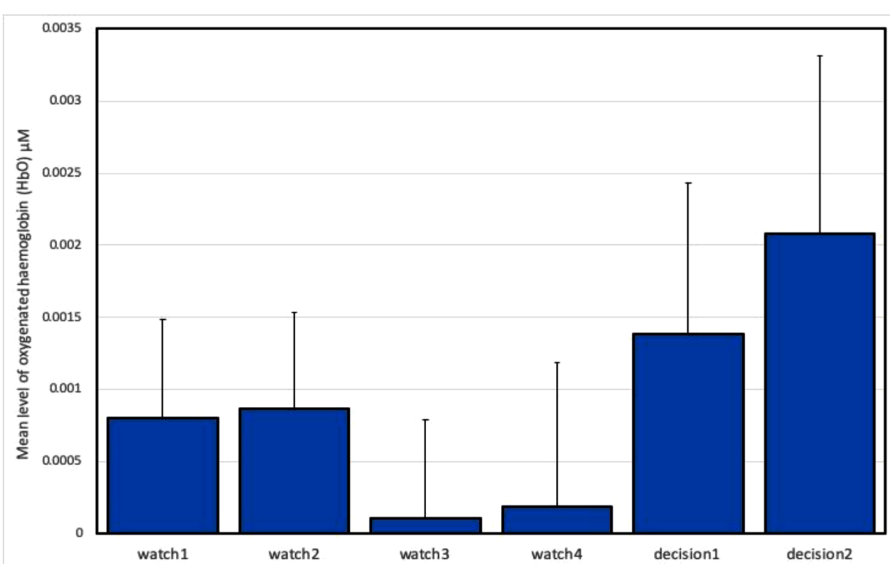

Fig. 4. Mean HbO and standard error during all Task Periods for Right-Lateral ROI [20].

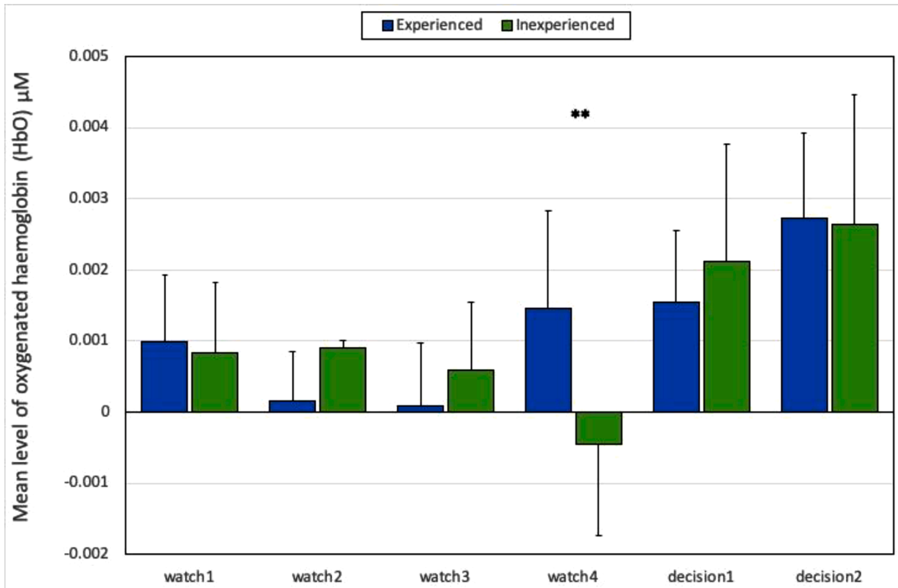

Fig. 5. Mean HbO and standard error in Right-Lateral ROI for Task Period x Experience Interaction. ** = significant difference at $p < .01$.

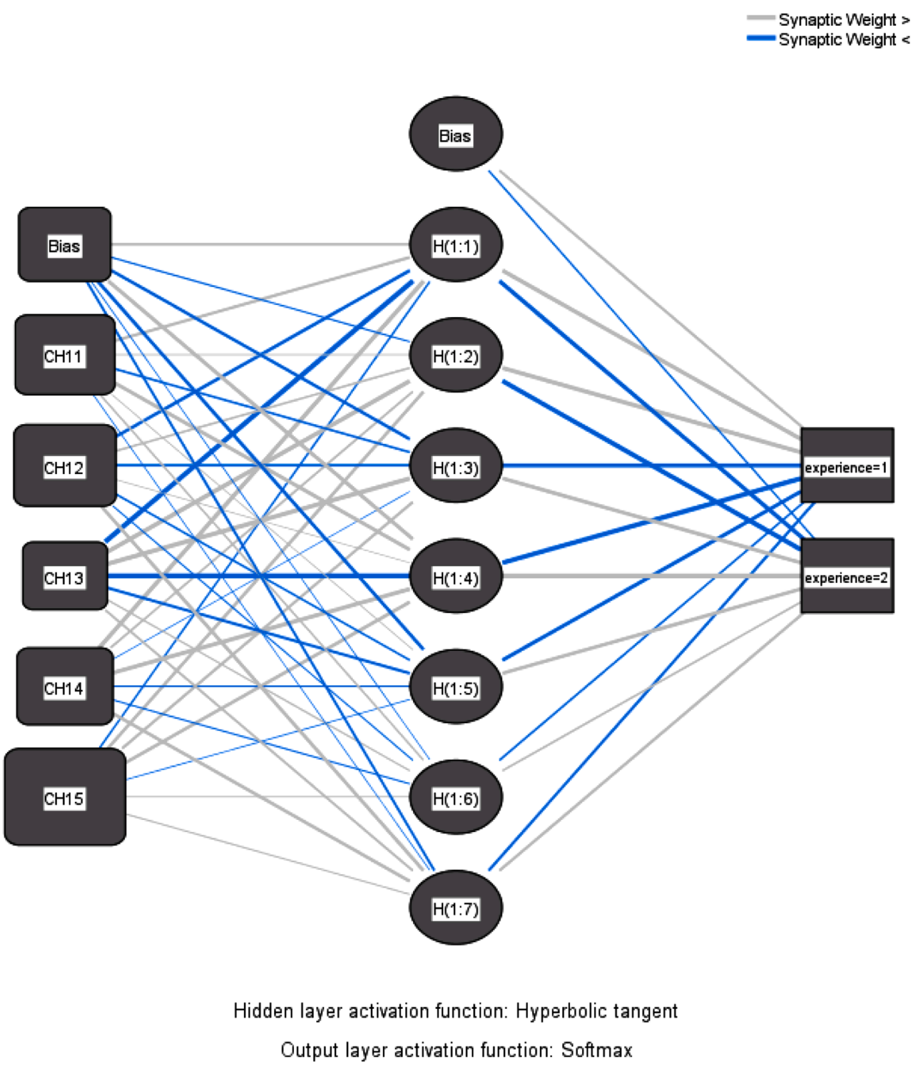

Fig. 6. Network diagram for the non-distraction situation.

fourth period of watchkeeping (w4) when the ship was spotted. In the ANN model, 70% of data is randomly selected as training data, while 30% is used as testing data. As a result, the network diagram for the non-distraction situation is illustrated in Fig. 6. There is one hidden layer for the ANN.

The ANN model prediction error for the training dataset is 4.8%, while for the testing dataset, it is 5%. The area under the ROC curve (AUC) is widely used to represent the effectiveness of ANN's accuracy in prediction and classification [83]. The higher AUC, the better the test is. From the curve in Fig. 7, the AUC is 0.991, indicating the validity of the model and its satisfactory prediction results. In the light of all 5 channels of fNIRS data, channel 15 is the most important independent variable (with an importance of 0.309) for the operator's qualification prediction. It proves that the fNIRS data is distinguished between the experienced and inexperienced groups, in the non-distraction situation. The results show that the proposed ANN model effectively predicts the seafarer experience using fNIRS data, with an accuracy of 95%, in the non-distraction situation. As this study is the first to use psychological data and a machine learning approach to predict seafarers' qualifications, it will be difficult to undertake a like-to-like benchmark. Nevertheless, the accuracy of ANN model prediction reflects the reliability of results compared to the realistic seafarer experience. Therefore, the new method can help, as an additional assurance, detect and reduce subjective bias in maritime HPM.

On the other hand, human behavior analysis is conducted using ship distance data. With regards to "Distance 1″, where the own ship spotted the target ship, the experienced seafarers performed with a distance of 4.94 nm while inexperienced seafarers at 4.57 nm. In the light of "Distance 2″, where the own ship altered the course to avoid the collision, experienced seafarers (3.41 nm) outperformed the inexperienced group (1.90 nm) by manoeuvring at a greater distance, as seen in Fig. 8. It can be seen that, in the non-distraction situation, experienced seafarers spotted the vessel and made manoeuvres earlier than inexperienced seafarers did.

4.2.2. Model prediction and human performance for a distraction situation

A distraction situation is within the same timeframe as the non-distraction situation but with reporting missions at specific intervals. The fNIRS data of the right ROI during the fourth period of watchkeeping (w4) is distracted as neurons for the model. As a result, the network diagram for the distraction situation is illustrated in Fig. 9. There is one hidden layer for the ANN.

The model prediction error for the training data set is 3.9%, while the prediction error for the testing dataset is 4.3%. From the ROC curve in Fig. 10, the area under the curve is 0.989, which indicates a satisfactory prediction result. Among all 5 channels of fNIRS data, channel 14 is the most important independent variable (with an importance of 0.293) for the operator's qualification prediction. In addition, it proved that the fNIRS data is distinguished between the experienced and inexperienced groups, in the distraction situation. The results show that the proposed ANN model effectively predicts the seafarer experience using fNIRS data, with an accuracy of 95.7% in the distraction situation.

Similarly, human behavior analysis is conducted using ship distance data. Regarding "Distance 1″, where the own ship spotted the target ship, experienced seafarers performed with 4.60 nm while inexperienced seafarers performed with 4.45 nm. In the light of "Distance 2″, where the own ship altered the course to avoid the collision, experienced seafarers, on average, manoeuvred at 2.10 nm while the inexperienced group did at 1.80 nm, as seen in Fig. 11. In a distraction situation, experienced seafarers spotted the vessel and altered the course earlier than the inexperienced seafarers. Due to the distracting tasks, the difference between the experienced and inexperienced group is not as significant as the results in Section 4.2.1 (the non-distraction situation).

5. Discussions and implications

The findings from the ANN model prove the acceptable prediction accuracy for the seafarer experience guidance, as shown in Figs. 6 and 9. Given such circumstances, human performance can be assessed based on the newly proposed approach with the support of the use of fNIRS. It is obvious that based on this new approach, seafarers' performance in daily work (e.g., watchkeeping) and even their qualifications can be quantified and evaluated objectively to supplement the currently used subjective assessment.

The results from fNIRS can help examiners to avoid their subjective bias in seafarers' qualification examinations. Collecting fNIRS seafarers' data during daily duty onboard makes it possible to predict seafarers' qualification levels under different scenarios. If the traditional HPM by experts in one exam is inconsistent with the one predicted by fNIRS, extra measures should be taken to ensure they reach the proper level of competence. It will shift the paradigm of the current seafarer qualification and certification mechanism. With the support of fNIRS, the seafarer's qualification can be predicted from their training process, which will significantly complement the current subjective evaluation of operator performance using expert knowledge and self-assessment questionnaires. Furthermore, the conventional assessment of seafarers against the data influcing their behavior such as time and distance when seafarers take actions reveals limited statistical characteristics for various qualification levels of seafarers. This objective behavior data could not be proved to be scientifically robust to indicate seafarers' qualifications in both non-distraction and distraction situations. However, the new approach by the hybrid of an ANN model and an fNIRS technology proves the applicability of such HPM and shows a satisfactory prediction accuracy rate. Therefore, based on the new findings from this experiment, HPM using psychophysiological aid in a maritime case opens a new paradigm for MET.

In light of the seafarer/operator training in the maritime domain, the proposed objective assessment method serves as a criterion for expert judgement and supports maritime training. Specifically, HPM using psychophysiological data records the brain signals during seafarer training. The proposed ANN model using such psychophysiological data can then be used to predict seafarer experience with an accuracy of 95%. Based on records and analysis results from fNIRS data, it provides a promising new solution towards a more rational assessment of whether a seafarer meets the criteria of a standard qualification through psychophysiological measurement. One significant challenge for onboard training (OBT) is the deficiency of administration efforts to monitor OBT activities [84]. It was evident that the issue of certificates of competence for seafarers is based on the OBT Record Book verified by maritime

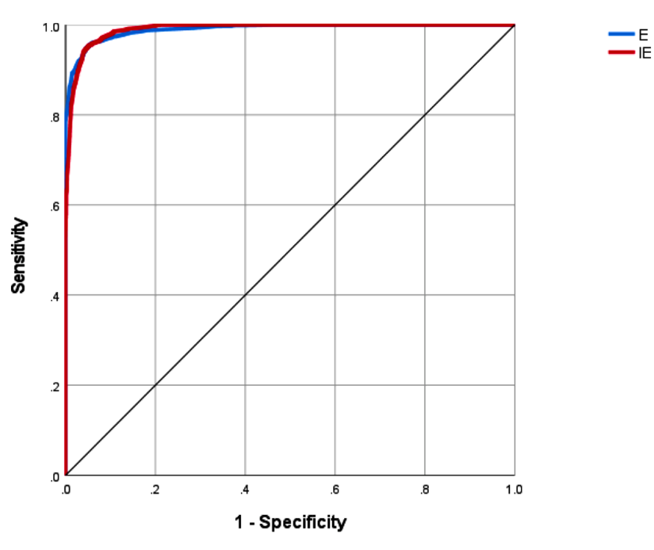

Fig. 7. ROC curve for non-distraction group.

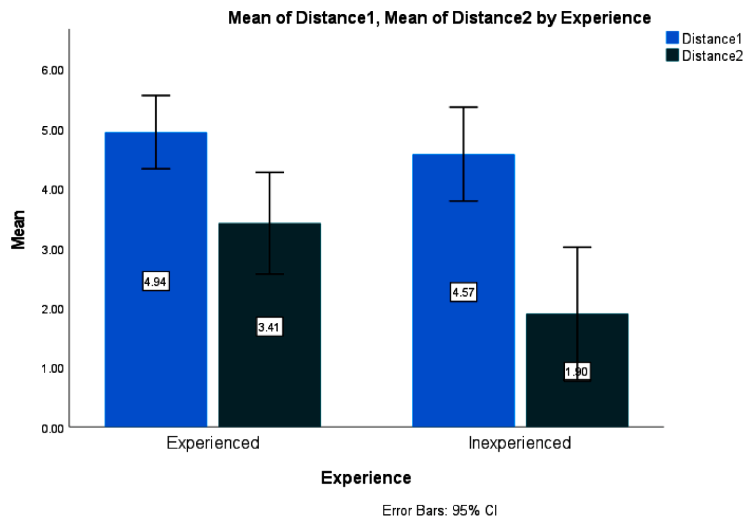

Fig. 8. Means of distances for seafarers in non-distraction situation.

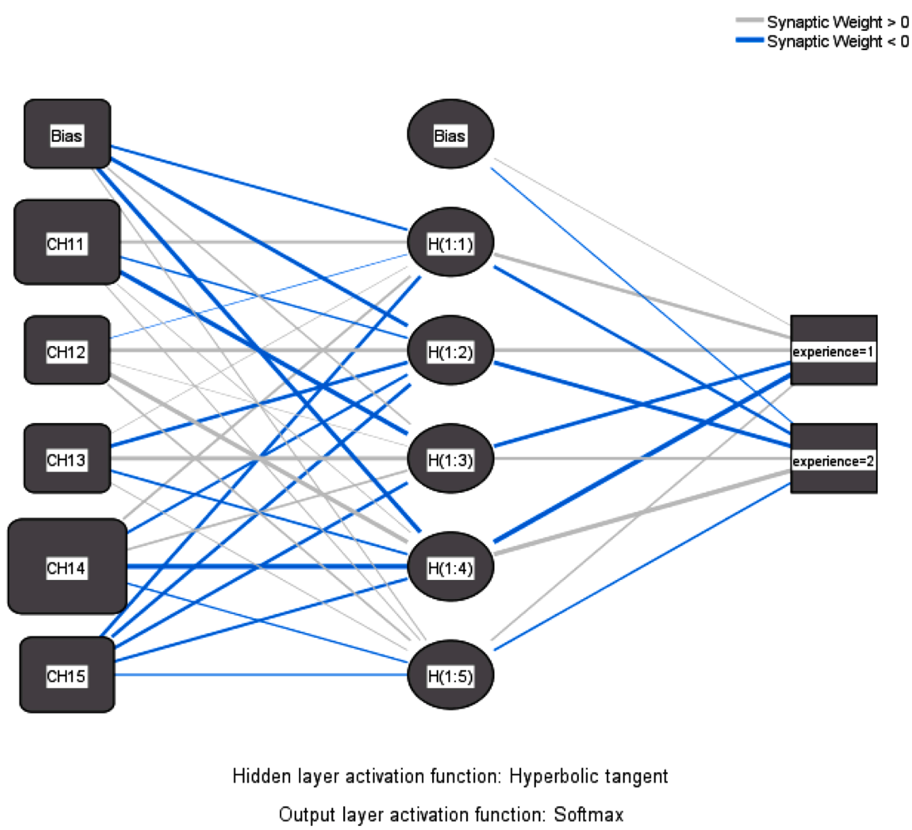

Fig. 9. Network diagram for the distraction situation.

authorities' officials. However, the effectiveness of the method through the OBT Record Book is arguable, possibly causing the misclassification of a seafarer with the competence at the board line of two neighboring classes/ranks. Using the proposed approach, it serves as a new reference to aid the currently established OBT to classify a seafarer with the required competence more rationally. It will be beneficial when the results from the newly proposed fNIRS-ANN approach and OBT conflict. Therefore, it has many insights to further explore the proposed method and see how it can complement the operator experience assessment in practice. Psychophysiological evaluation technologies, such as EEG [44] and fNIRS, can support the HPM given objective data collection. In this way, the results indicate recommendations on the "qualified" or "unqualified" responses of seafarers in maritime training. The approach can be tailored to make it applicable in other areas of similarity, such as offshore and other transport modes. Instead of using the established HRA models [85,38], this assessment conquers data limitation and expert knowledge bias research gaps. In this regard, maritime authorities can take it as an alternative way to supervise, monitor, and assess the operator training process.

Moreover, fNIRS-based HPM supports the measurement and quantification of human response under specific missions, serving as an indicator of qualified seafarers. It shows the possibility for the shipping company to evaluate employees given specific scenarios, integrated with the simulator training. Specifically, the training dataset can be collected

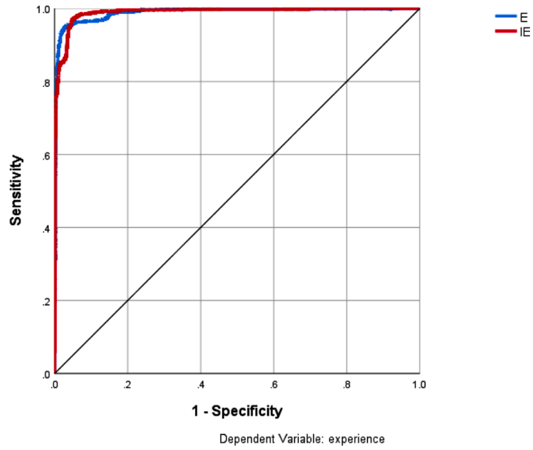

Fig. 10. ROC curve for distraction group.

from the qualified employees given specific scenarios or tasks, and the test data is from employees to be evaluated. The proposed ANN model distinguishes skilled employees by analyzing their brain activity signals. The experienced and qualified seafarers will be selected to commence the nautical tasks. In contrast, unqualified and inexperienced seafarers are expected to have additional training courses or transfer to other tasks. In summary, the criteria for "qualified" employees can be produced based on previous traditional checklists and the fNIRS index for qualified seafarers. Further, the efficient evaluation of employees relieves the pressure on ship companies to reduce the crew size as far as ship automation is concerned. It provides the guideline to optimise human resource allocation. In this way, qualified employees are elected to ensure safe shipping, which will compensate for the above deficiency of maritime training.

The proposed fNIRS-based measurement illustrates the possibility of predicting operators' qualification from daily maritime duty (watchkeeping) rather than emergency responses. It explains the brain activity differences between experienced and inexperienced operators in the face of common working scenarios. Therefore, the assessment methodology could act as a predictor of operators' reliability levels under various circumstances. The development of MASS requires new perspectives on human factors and seafarers' qualifications, given advanced technologies and new scenarios in shipping [37,86]. Besides the emergency response, human performance under typical working scenarios requires more attention for the MASS. There are various assumptions and plans for the autonomous ship design within remote operators, which urgently need a reliable validation and evaluation method for mental workload in such everyday working scenes. Given the new design of autonomous ships, the proposed methodology will serve as a tool to test human performance onboard or onshore. Under the new deployment for autonomous ships, ship manufactures and shipbuilders could utilize this supportive tool to evaluate human performance in designing complex systems so the design can be assured as safe and human-centric as possible. The results are expected to establish the guideline for the ergonomic design of both traditional ships and the MASS. In addition, the proposed methodology shows insights into HPM in other high-risk sectors, such as the autonomous industry [87]. The automation development in other transport sectors can also improve human training and education accreditation, which is beyond the maritime case. Subjective assessment of human performance in transport will be significantly enhanced by introducing novel methodology and technology using the new approach in this study as a foundation. It will help evaluate the operator's qualification and predict the human performance failures in a wide range of different transport sectors and applications.

6. Conclusion

This study proposes an fNIRS-based HPM method to quantify the

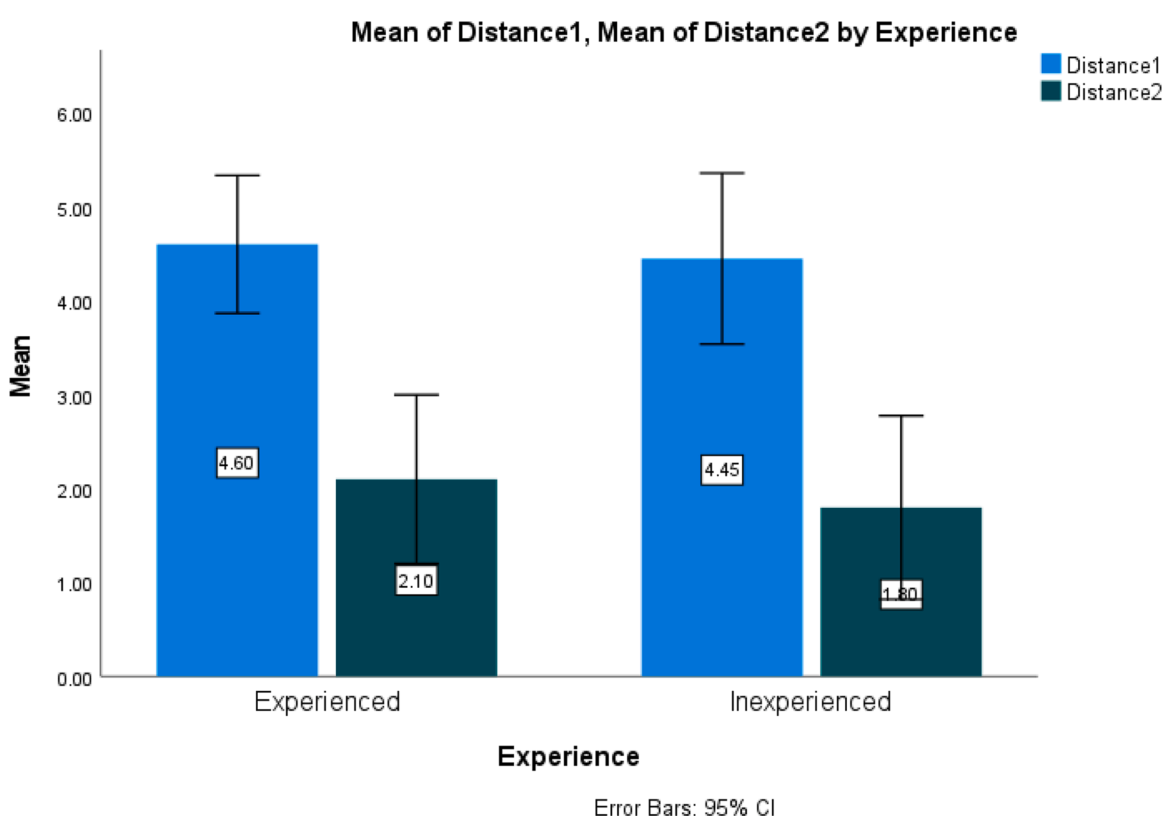

Fig. 11. Means of distances for seafarers in distraction situation.

psychophysiological activities and predict the operators' qualifications. 40 participants were recruited to conduct a new experimental maritime transport study, measured by the fNIRS technology. Two situations, distraction and non-distraction, were deployed for 20 participants. In each situation, there were 10 experienced seafarers and 10 inexperienced ones. Then, the statistical method was used to identify significant brain areas (right prefrontal cortex) and task period (late watchkeeping period) for the operators based on the fNIRS data. The right prefrontal cortex of operators in the w4 task period was selected as neurons for the ANN modelling. The machine learning algorithm was applied to the classification of seafarers' experiences. Lastly, the prediction rates of operators' experience for the distraction and non-distraction groups were 95.7% and 95.0%, respectively. It proves the applicability of fNIRS for operator experience assessment under daily work scenarios.

To conclude, the contributions of this paper lie in the use of fNIRS and maritime simulation in a holistic way to predict seafarers' qualifications and hence reduce the subjective bias introduced by the assessors in the certification process. The proposed objective assessment model utilizes an ANN algorithm to classify the seafarers' experience using haemoglobin data. The result obtained from the performance of 40 participants in ship bridge simulation reveals insights into the HPM framework using psychophysiological data, which benefits and supports MET. Nevertheless, there are limitations to this study. Developing the training dataset requires a sufficient sample size in each scenario. It will be costly when other researchers and industries utilize the proposed method and fNIRS technology to conduct HPM. Participants were constrained in their seats and not allowed to walk in the bridge simulation room. Although reducing movement artefacts of data collection, this constraint is inconsistent with seafarers' daily work scenes. Future studies will be undertaken to address the limitation by applying wireless equipment and exploring algorithms to reduce relevant artefacts induced by the walking and movement of seafarers. In addition, the ship bridge team consist of a group of seafarers instead of a sole worker. There are interactions between team members and the hierarchy in the existing bridge team. This study only considered individual perception and decision-making. In future work, teamwork should be addressed in human performance measurement.

CRediT authorship contribution statement

Shiqi Fan: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Project administration. **Zaili Yang:** Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Zaili Yang reports financial support was provided by European Research Council. Zaili Yang reports financial support was provided by EU Framework Programme for Research and Innovation.

Data availability

The authors do not have permission to share data.

Acknowledgments

This work is financially supported by the EU H2020 ERC Consolidator Grant program (TRUST Grant No. 864724) and EU H2020 Marie Curie RISE ENHANCE project (MSCA-RISE 823904).

References

[1] EMSA 2021. Annual overview of marine casualties and incidents 2021.
[2] ERA 2016. Railway safety performance in the European Union Brussels, Union (2016).
[3] Papadimitriou E, Schneider C, Tello JA, Damen W, Vrouenraets ML, Broeke AT. Transport safety and human factors in the era of automation: what can transport modes learn from each other? Accid Anal Prev 2020;144:105656.
[4] Kim DH. Human factors influencing the ship operator's perceived risk in the last moment of collision encounter. Reliab Eng Syst Saf 2020;203:107078.
[5] Sotiralis P, Ventikos NP, Hamann R, Golyshev P, Teixeira AP. Incorporation of human factors into ship collision risk models focusing on human centred design aspects. Reliab Eng Syst Saf 2016;156:210–27.
[6] Yildiz S, Ugurlu O, Wang J, Loughney S. Application of the HFACS-PV approach for identification of human and organizational factors (HOFs) influencing marine accidents. Reliab Eng Syst Saf 2021;208:107395.
[7] Fan S, Blanco-Davis E, Yang Z, Zhang J, Yan X. Incorporation of human factors into maritime accident analysis using a data-driven Bayesian network. Reliab Eng Syst Saf 2020;203:107070.
[8] Liu C, Chu X, Wu W, Li S, He Z, Zheng M, Zhou H, Li Z. Human–machine cooperation research for navigation of maritime autonomous surface ships: a review and consideration. Ocean Eng 2022;246:110555.
[9] Stcw I. International convention on standards of training, certification and watchkeeping for seafarers, (STCW) 1978, as amended in 1995/2010. London, UK: International Maritime Organisation; 2011.
[10] Marcano L, Haugen FA, Sannerud R, Komulainen T. Review of simulator training practices for industrial operators: how can individual simulator training be enabled? Saf Sci 2019;115:414–24.
[11] H. Spetalen & R. Sannerud 2015. Experiences with the use of simulation as transfer strategy. https://hdl.handle.net/20.500.12199/440 Online accessed on 18 September. 2022.
[12] Petkov G, Todorov V, Takov T, Petrov V, Stoychev K, Vladimirov V, Chukov I. Safety investigation of team performance in accidents. J Hazard Mater 2004;111:97–104.
[13] Fan S, Zhang J, Blanco-Davis E, Yang Z, Wang J, Yan X. Effects of seafarers' emotion on human performance using bridge simulation. Ocean Eng 2018;170:111–9.
[14] Murdock Jr BB. The serial position effect of free recall. J Exp Psychol 1962;64:482.
[15] Wilson TD, Brekke N. Mental contamination and mental correction: unwanted influences on judgments and evaluations. Psychol Bull 1994;116:117.
[16] Ernstsen J, Nazir S. Performance assessment in full-scale simulators - A case of maritime pilotage operations. Saf Sci 2020;129:104775.
[17] Akhtar MJ, Utne IB. Human fatigue's effect on the risk of maritime groundings - a Bayesian network modeling approach. Saf Sci 2014;62:427–40.
[18] Akhtar MJ, Utne IB. Common patterns in aggregated accident analysis charts from human fatigue-related groundings and collisions at sea. Marit Policy Manag 2015;42:186–206.
[19] Besikci EB, Tavacioglu L, Arslan O. The subjective measurement of seafarers' fatigue levels and mental symptoms. Marit Policy Manag 2016;43:329–43.
[20] Fan S, Blanco-Davis E, Zhang J, Bury A, Warren J, Yang Z, Yan X, Wang J, Fairclough S. The role of the prefrontal cortex and functional connectivity during maritime operations: an fNIRS study. Brain Behav 2021;11:e01910.
[21] Hou X, Liu Y, Sourina O, Muellerwittig W. CogniMeter: eEG-based emotion, mental workload and stress visual monitoring. Int Conf Cyberworlds 2015:153–60.
[22] Lafont A, Rogé J, Ndiaye D, Boucheix JM. Driver's emotional state and detection of vulnerable road users: towards a better understanding of how emotions affect drivers' perception using cardiac and ocular metrics. Transp Res Part F Traffic Psychol Behav 2018;55:141–52.
[23] Yan L, Zheng K, Wu C, Wen J, Zhu D, Shi J. A special laboratory method for inducing driving anger: based on the virtual scene. In: Proceedings of the transportation research board 94th annual meeting; 2015.
[24] Fan S, Yan X, Zhang J, Wang J. A review on human factors in maritime transportation using seafarers' physiological data. In: Proceedings of the 4th international conference on transportation information and safety (ICTIS. IEEE; 2017. p. 104–10.
[25] Fishburn FA, Norr ME, Medvedev AV, Vaidya CJ. Sensitivity of fNIRS to cognitive state and load. Front Hum Neurosci 2014;8:76.
[26] Causse M, Chua Z, Peysakhovich V, Del Campo N, Matton N. Mental workload and neural efficiency quantified in the prefrontal cortex using fNIRS. Sci Rep 2017;7:1–15.
[27] Foy HJ, Runham P, Chapman P. Prefrontal cortex activation and young driver behaviour: a fNIRS study. PLoS One 2016;11:18.
[28] Bergmann B, Wiedemann J, Zehrt P. Description and training of competence for failure diagnosis. Zeitschrift Fur Arbeits-Und Organisationspsychologie 1995;39:146–56.
[29] Ohtsuka T, Yoshimura S, Kawano R, Fujiie M, Ujita H, Kubota R. Nuclear-power-plant operator performance analysis using training simulators - operator performance under abnormal-plant conditions. J Nucl Sci Technol 1994;31:1184–93.
[30] Ki JS. Interactive training simulator for aerial working platform in a virtual environment. Comput Appl Eng Educ 2011;19:733–8.
[31] Borghetti BJ, Giametta JJ, Rusnock CF. Assessing continuous operator workload with a hybrid scaffolded neuroergonomic modeling approach. Hum Factors 2017;59:134–46.

[32] Iqbal MU, Shahab MA, Choudhary M, Srinivasan B, Srinivasan R. Electroencephalography (EEG) based cognitive measures for evaluating the effectiveness of operator training. Process Saf Env Prot 2021;150:51–67.
[33] Shahab MA, Iqbal MU, Srinivasan B, Srinivasan R. Metrics for objectively assessing operator training using eye gaze patterns. Process Saf Environ Prot 2021;156:508–20.
[34] Song YS, Baek JH, Hwang DS, Lee JW, Lee YJ, Park HJ, Choi JH, Yang HK. Physiological status assessment of locomotive engineer during train operation. J Electr Eng Technol 2014;9:324–33.
[35] Hausdorf PA, Risavy SD. Predicting training and job performance for transit operators. Int J Sel Assess 2015;23:191–5.
[36] Alfes CM, Manacci CF. Taking simulation to new heights: designing a flight simulation center. Clin Simul Nurs 2014;10:442–5.
[37] Jo S, D'agostini E, Kang J. From seafarers to e-farers: maritime cadets' perceptions towards seafaring jobs in the industry 4.0. Sustainability 2020;12(19):8077.
[38] Dos Santos I, Franca JEM, Santos LFM, Haddad AN. Allocation of performance shaping factors in the risk assessment of an offshore installation. J Loss Prev Process Ind 2020;64:104085.
[39] Lisanti G, Karaman S, Pezzatini D, Del Bimbo A. A multi-camera image processing and visualization system for train safety assessment. Multimed Tools Appl 2018;77:1583–604.
[40] Dhalmahapatra K, Maiti J, Krishna OB. Assessment of virtual reality based safety training simulator for electric overhead crane operations. Saf Sci 2021;139:105241.
[41] Brandenburger N, Hormann HJ, Stelling D, Naumann A. Tasks, skills, and competencies of future high-speed train drivers. Proc Inst Mech Eng Part F J Rail Rapid Transit 2017;231:1115–22.
[42] Dijksterhuis C, Brookhuis KA, De Waard D. Effects of steering demand on lane keeping behaviour, self-reports, and physiology. A simulator study. Accid Anal Prev 2011;43:1074–81.
[43] Ngodang T, Murai K, Hayashi Y, Mitomo N, Yoshimura K, Hikida K. A study on navigator's performance in ship bridge simulator using heart rate variability. In: Proceedings of the IEEE international conference on systems, man, and cybernetics. IEEE; 2012.
[44] Liu YS, Lan ZR, Cui J, Krishnan G, Sourina O, Konovessis D, et al. Psychophysiological evaluation of seafarers to improve training in maritime virtual simulator. Adv Eng Inform 2020;44:101048.
[45] Martinez-Marquez D, Pingali S, Panuwatwanich K, Stewart RA, Mohamed S. Application of eye tracking technology in aviation, maritime, and construction industries: a systematic review. Sensors 2021;21(13):4289.
[46] Wu YB, Miwa T, Uchida M. Using physiological signals to measure operator's mental workload in shipping - an engine room simulator study. J Mar Eng Technol 2017;16:61–9.
[47] Du N, Zhou F, Pulver EM, Tilbury DM, Robert LP, Pradhan AK, Yang XJ. Examining the effects of emotional valence and arousal on takeover performance in conditionally automated driving. Transp Res Part C Emerg Technol 2020;112:78–87.
[48] Rakauskas ME, Ward NJ, Boer ER, Bernat EM, Cadwallader M, Patrick CJ. Combined effects of alcohol and distraction on driving performance. Accid Anal Prev 2008;40:1742–9.
[49] Gateau T, Durantin G, Lancelot F, Scannella S, Dehais F. Real-time state estimation in a flight simulator using fNIRS. PLoS One 2015;10:e0121279.
[50] Li Q, Ng KK, Fan Z, Yuan X, Liu H, Bu L. A human-centred approach based on functional near-infrared spectroscopy for adaptive decision-making in the air traffic control environment: a case study. Adv Eng Inform 2021;49:101325.
[51] Lim WL, Liu YS, Subramaniam SCH, Liew SHP, Krishnan G, Sourina O, Konovessis D, Ang HE, Wang LP, Gavrilova ML, Tan CJK, Sourin A. EEG-based mental workload and stress monitoring of crew members in maritime virtual simulator. Transactions on computational science xxxii: special issue on cybersecurity and biometrics. Cham: Springer International Publishing Ag; 2018.
[52] Chang CH, Kontovas C, Yu Q, Yang ZL. Risk assessment of the operations of maritime autonomous surface ships. Reliab Eng Syst Saf 2021;207:107324.
[53] Martinez-Marquez D, Pingali S, Panuwatwanich K, Stewart RA, Mohamed S. Application of eye tracking technology in aviation, maritime, and construction industries: a systematic review. Sensors 2021;21:4289.
[54] Christian H, Dominic H, Ole F, Johannes H, Felix P, Tanja S. Mental workload during n-back task—Quantified in the prefrontal cortex using fNIRS. Front Hum Neurosci 2013;7:935.
[55] Cezar-Vaz MR, Bonow CA, De Almeida MCV, Sant'anna CF, Cardoso LS. Workload and associated factors: a study in maritime port in Brazil. Rev Lat Am Enfermagem 2016;24.
[56] Yoshida M, Shimizu E, Sugomori M, Umeda A. Identification of the relationship between maritime autonomous surface ships and the operator's mental workload. Appl Sci Basel 2021;11(5):233 1.
[57] Yan SY, Wei YY, Tran CC. Evaluation and prediction mental workload in user interface of maritime operations using eye response. Int J Ind Ergon 2019;71:117–27.
[58] Kim H, Yang CS, Lee BW, Yang YH, Hong S. Alcohol effects on navigational ability using ship handling simulator. Int J Ind Ergon 2007;37:733–43.
[59] Cui X, Zhang YJ, Zhou YW, Huang TC, Li ZZ. Measurements of team workload: a time pressure and scenario complexity study for maritime operation tasks. Int J Ind Ergon 2021;83:103110.
[60] Leung AWS, Chan CCH, Ng JJM, Wong PCC. Factors contributing to officers' fatigue in high-speed maritime craft operations. Appl Ergon 2006;37:565–76.
[61] Monteiro TG, Skourup C, Zhang HX. A task agnostic mental fatigue assessment approach based on eeg frequency bands for demanding maritime operation. IEEE Instrum Meas Mag 2021;24:82–8.
[62] Andrei DM, Griffin MA, Grech M, Neal A. How demands and resources impact chronic fatigue in the maritime industry. The mediating effect of acute fatigue, sleep quality and recovery. Saf Sci 2020;121:362–72.
[63] Monteiro TG, Li GY, Skourup C, Zhang HX. Investigating an integrated sensor fusion system for mental fatigue assessment for demanding maritime operations. Sensors 2020;20.
[64] Sandhaland H, Oltedal HA, Hystad SW, Eid J. Effects of leadership style and psychological job demands on situation awareness and the willingness to take a risk: a survey of selected offshore vessels. Saf Sci 2017;93:178–86.
[65] Villafaina S, Fuentes-Garcia JP, Gusi N, Tornero-Aguilera JF, Clemente-Suarez VJ. Psychophysiological response of military pilots in different combat flight maneuvers in a flight simulator. Physiol Behav 2021;238.
[66] Wang HB, Pan T, Si HQ, Li Y, Jiang NQ. Research on influencing factor selection of pilot's intention. Int J Aerosp Eng 2020;2020:113483.
[67] Yan FW, Liu MT, Ding CH, Wang Y, Yan LR. Driving style recognition based on electroencephalography data from a simulated driving experiment. Front Psychol 2019;10:1254.
[68] Yang L, Guan W, Ma R, Li XM. Comparison among driving state prediction models for car-following condition based on EEG and driving features. Accid Anal Prev 2019;133:105296.
[69] Yang L, Ma R, Zhang HM, Guan W, Jiang SX. Driving behavior recognition using EEG data from a simulated car-following experiment. Accid Anal Prev 2018;116:30–40.
[70] Wróbel K, Gil M, Chae CJ. On the influence of human factors on safety of remotely-controlled merchant vessels. Appl Sci 2021;11:1145.
[71] Yoshida M, Shimizu E, Sugomori M, Umeda A. Regulatory requirements on the competence of remote operator in maritime autonomous surface ship: situation awareness, ship sense and goal-based gap analysis. Appl Sci Basel 2020;10(23):8751.
[72] He J, Tian Y, Wu L. A hybrid data-driven method for rapid prediction of lithium-ion battery capacity. Reliab Eng Syst Saf 2022;226:108674.
[73] Yang Y, Li S, Zhang P. Data-driven accident consequence assessment on urban gas pipeline network based on machine learning. Reliab Eng Syst Saf 2022;219:108216.
[74] Vasilikis N, Geertsma R, Visser K. Operational data-driven energy performance assessment of ships: the case study of a naval vessel with hybrid propulsion. J Mar Eng Technol 2022:1–17.
[75] Zhang M, Kujala P, Hirdaris S. A machine learning method for the evaluation of ship grounding risk in real operational conditions. Reliab Eng Syst Saf 2022;226:108697.
[76] Sassaroli A, Fantini S. Comment on the modified Beer–Lambert law for scattering media. Phys Med Biol 2004;49(14):N255–7.
[77] Cui X, Bray S, Reiss AL. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. NeuroImage 2010;49:3039–46.
[78] Cabaneros SM, Calautit JK, Hughes B. A review of artificial neural network models for ambient air pollution prediction. Environ Model Softw 2019;119:285–304.
[79] Saikia P, Baruah RD, Singh SK, Chaudhuri PK. Artificial neural networks in the domain of reservoir characterization: a review from shallow to deep models. Comput Geosci 2020;135:104357.
[80] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15:1929–58.
[81] Anifowose FA, Labadin J, Abdulraheem A. Hybrid intelligent systems in petroleum reservoir characterization and modeling: the journey so far and the challenges ahead. J Pet Explor Prod Technol 2017;7:251–63.
[82] Ju E, Kim KH, Chang KH, Shim JB, Kim CY, Lee NK, Lee S, Park CG. Artificial neural network-based variable importance analysis of prognostic factors related to radiation pneumonitis in patients with lung cancer: preliminary study. J Korean Phys Soc 2019;75:277–82.
[83] Woods K, Bowyer KW. Generating ROC curves for artificial neural networks. IEEE Trans Med Imaging 1997;16:329–37.
[84] Bao JZ, Li Y, Duan ZL, Li TT, Zhang PF. Key factors affecting the quality of maritime education and training: empirical evidence from China. J Navig 2021;74:396–408.
[85] Chen HC, Lu HA, Lee HH. An assessment of job performance of vessel traffic service operators using an analytic hierarchy process and a grey interval measure. J Mar Sci Technol Taiwan 2013;21:522–31.
[86] Zhang MY, Zhang D, Yao HJ, Zhang K. A probabilistic model of human error assessment for autonomous cargo ships focusing on human-autonomy collaboration. Saf Sci 2020;130:104838.
[87] Azevedo-Sa H, Zhao HJ, Esterwood C, Yang XJ, Tilbury DM, Robert LP. How internal and external risks affect the relationships between trust and driver behavior in automated driving systems. Transp Res Part C Emerg Technol 2021;123:102973.