Reliability Engineering and System Safety 139 (2015) 17–32

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

A Simulator for Human Error Probability Analysis (SHERPA)

Valentina Di Pasquale *, Salvatore Miranda, Raffaele Iannone, Stefano Riemma

Department of Industrial Engineering, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano, SA, Italy

ARTICLE INFO

Article history:
Received 8 September 2014
Received in revised form
30 January 2015
Accepted 3 February 2015
Available online 24 February 2015

Keywords:
Modelling
Simulation
Human Reliability Analysis
Human error probability
Performance shaping factor

ABSTRACT

A new Human Reliability Analysis (HRA) method is presented in this paper. The Simulator for Human Error Probability Analysis (SHERPA) model provides a theoretical framework that exploits the advantages of the simulation tools and the traditional HRA methods in order to model human behaviour and to predict the error probability for a given scenario in every kind of industrial system. Human reliability is estimated as function of the performed task, the Performance Shaping Factors (PSF) and the time worked, with the purpose of considering how reliability depends not only on the task and working context, but also on the time that the operator has already spent on the work. The model is able to estimate human reliability; to assess the effects due to different human reliability levels through evaluation of tasks performed more or less correctly; and to assess the impact of context via PSFs. SHERPA also provides the possibility of determining the optimal configuration of breaks. Through a methodology that uses assessments of an economic nature, it allows identification of the conditions required for the suspension of work in the shift for the operator's psychophysical recovery and then for the restoration of acceptable values of reliability.

1. Introduction

Human error is here to stay [1]. This perhaps obvious statement has a more profound implication if we consider how common human errors are in everyday life and in the working environment. The vast majority of current catastrophes arises by a combination of many small events, system faults and human errors that would be irrelevant individually, but – when combined in a special time sequence of circumstances and actions – can lead to unrecoverable situations [2]. For this reason, wrong and inappropriate human actions are source of great concern and create efficiency and safety issues for every kind of working context.

Valid values are difficult to obtain, but estimates indicate that errors committed by man are responsible for 60–90% of the accidents; the remainder of accidents are attributable to technical deficiencies [3–6]. The percentage of incidents connected with human error in several industries is listed in Table 1. These complex systems tend to show a low probability of a negative incident occurring, but the consequences are high if an event occurs. The accidents are, of course, the most obvious human errors in the industrial systems, but minor faults can seriously reduce the operation performance in terms of productivity and efficiency. For example, human error has a direct impact on productivity because errors affect the rates of rejection of a product, thereby increasing the production costs and reducing subsequent sales [7,8]. Several researchers have focused on the concept of human error in order to understand, evaluate and identify possible actions to limit it.

The evidence that human actions are a source of vulnerability for industrial systems gave birth to the Human Reliability Analysis (HRA), which aims at further examination of the human factor through the prediction of when an operator is more likely to fail. The standard definition of human reliability is the probability that a person will perform according to the requirements of the task for a specified period of time and not perform any extraneous activity that can degrade the system [9].

The starting point of this work was to study the state of the art of current HRA methods, beginning with the quantitative methods of the first generation and the qualitative methods of the second one and extending to the third generation HRA approaches and new dynamic HRA methods. All these methods have the purpose of assessing the likelihood of human error – in working systems, for a given operation, in a certain interval of time and in a particular context – on the basis of models that describe, in a more or less simplistic way, the complex mechanism that underlies the single human action that is potentially subject to error [7,8].

Special attention has been paid to dynamic HRA methods that use cognitive modelling and simulation to produce a data framework that may be used in quantifying human error probability (HEP). Human performance simulation reveals important new data sources and possibilities for exploring human reliability [10]. Many efforts have been recently directed towards simulation, in order to assess human behaviour and calculate the reliability for the performed activity. Trucco

* Corresponding author.
E-mail address: valentina.dipasquale1@gmail.com (V. Di Pasquale).

http://dx.doi.org/10.1016/j.ress.2015.02.003
0951-8320/

Table 1
Estimates of human error in various sectors as percentages of all failures.

Sectors	Human error (%)
Automobile	65
Heavy truck	80
Aviation	70–80
Jet transport	65–85
Air traffic control	90
Maritime vessels	80–85
Chemical industry	60–90
Nuclear power plants (US)	50–70
Road transportation	85

and Leva [11,12] developed a new probabilistic cognitive simulator (PROCOS) for approaching human errors in complex operational frameworks, while Mosleh and Chang [13–18] have presented an Information, Decision, and Action in Crew (IDAC) context model for Human Reliability Analysis. Despite the efforts of HRA experts to develop an advanced method, many of the limitations and problems of these approaches have not yet been resolved due to the complexity of human nature and the difficulty in predicting and simulating human behaviour.

This paper proposes a new use of HRA methodologies for HEP calculation through implementation of a Simulator for Human Error Probability Analysis (SHERPA), which aims to predict the likelihood of operator error not only about the performed activity but also as a function of time working. In this way, the simulator can dynamically analyse a whole shift to identify the moments of highest operator unreliability in order to organise the scheduling of breaks. The most important objective of this work is to provide a simulation module for the evaluation of human reliability that can be used in a dual manner [8], as follows:

- In the preventive phase, as an analysis of the possible situations that may occur and as evaluation of the percentage of pieces discarded by the effect of human error.
- In post-production, to understand the nature of the factors that influence human performance in order to reduce errors.

SHERPA assesses human reliability and uses this information to estimate the results obtained from different configurations and distributions of work breaks, thereby offering the possibility of determining the optimal configuration of breaks, in terms of both duration and distribution in shifts.

2. State-of-the-art HRA methodologies

The study of HRA is approximately 50 years old and has always been a hybrid discipline, involving reliability personnel, engineers and human factors specialists or psychologists [1] The goals defined by Swain and Guttmann in discussing the Technique for Human Error Rate Prediction (THERP) approach, one of the first HRA methods developed, are still valid: the objective of a Human Reliability Analysis is to evaluate the operator's contribution to system reliability. More precisely, the aim is to predict human error rates and to evaluate the degradation to human–machine systems likely to be caused by human errors in association with equipment functioning, operational procedures and practices, and other system and human characteristics which influence the system behaviour [2].

HRA has three basic functions: the identification of human errors, the prediction of their likelihood, and, if required, the reduction of their likelihood. To achieve this, HRA techniques estimate the probability of human error of the overall risk of the system, which can be expressed in the following simple formula [19]:

$$HEP = (\text{Number of errors occurred})(\text{Number of opportunities for error}) \quad (1)$$

The central tenet of HRA is that the HEP estimation process must be reasonably accurate, or at least conservative (i.e. tending more towards pessimistic estimates of failure probability rather than optimistic ones). If it is not accurate, or if it is too optimistic, then the risk may be underestimated.

The development of human reliability methods occurred over time in three stages. The first one lasted twenty years (1970–1990) and was the first human reliability method generation that focused on human error probabilities and operational human error [20]. First generation methods include 35–40 methods for human reliability, many of which are variations on a single method [6]. Many of these methods – such as Technique for Human Error Rate Prediction (THERP) [4–6,19–22], Accident Sequence Evaluation Programme (ASEP) [1,22] and Human Cognition Reliability (HCR) [5,6,9,22] – have the basic assumption that the natural deficiencies of humans cause them logically to fail to perform tasks, just as is seen with mechanical or electrical components. Each approach of this generation focuses on quantification in terms of success/failure of actions, with less attention paid to in-depth causes and reasons of observable human behaviour, which for these techniques is borrowed from psychological studies in behavioural sciences [8,23]. These traditional approaches determine the human error probability (HEP) by using established tables, human reliability models or expert judgment. The characterisation of human failure modes is usually very simple, such as in terms of *error of omission* and *errors of commission*.

The second phase (1990–2005), known as the second human reliability method generation, focused on human performance factors and cognitive processes. Human performance factors are internal or external and in general are everything that influences human performance, like workload, stress, sociological issues, psychological issues, illness, etc. [20]. The focus of the second generation shifted to cognitive aspects of humans, causes of errors rather than their frequency, study of factor interactions that increase the probability of error, and interdependencies of PSFs [8]. Advanced cognitive models have been developed, which represent the process of logic operator and synthesise the dependence on personal factors. One of the more widely used second generation techniques, Cognitive Reliability and Error Analysis Method (CREAM) [5,6,9,22,24–28] has an operator model that is more significant and less simplistic than in the first generation methods; HEP is derived from four Contextual Control Modes (CoCoMs): Scrambled, Opportunistic, Tactical and Strategic. CoCoM is based on the assumption that human behaviour is guided by two basic principles: the cyclical nature of human cognition and the dependence of cognitive processes from context and working environment. The Standardized Plant Analysis Risk-Human Reliability Analysis method (SPAR-H) [4,20,22,29–35] is instead built on an explicit information processing model of human performance derived from the behavioural sciences literature.

Additionally, second generation considers the context in which humans make errors and derive relative PSFs. A major difference between two generations can be simply stated as consideration of the PSF impact on operators. PSFs in the first generation were mainly derived by focusing on the environmental impacts on operators, whereas the PSFs in the second generation were derived by focusing on the cognitive impacts on operators. The context is carefully incorporated into the behavioural patterns, considering all the factors that may affect human performance. This is evident in SPAR-H, where its eight operational factors can be directly associated with the human performance model and show the human information processing model with which they are associated (see Fig. 1).

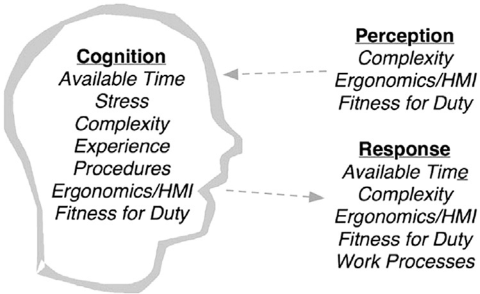

Fig. 1. SPAR-H Performance Shaping Factors in the information processing context [29].

Finally, the third phase, started in 2005 and still in progress, is represented by methods that focus on human performance factor relations and dependencies. While some experts have focused on the development of third generation methods, others have carried out studies on the so-called dynamic HRA, analysed in the next section. The Nuclear Action Reliability Assessment (NARA) method and the Bayesian networks are defined as the only current HRA tools of the third generation.

On one hand, NARA [20,22] recaptures and improves the method of the first generation HEART, trying to overcome some of the limitations of the same, while on the other hand, Bayesian networks [36–38] attempt to overcome some of the limitations of previous methods through qualitative analysis, which emphasises the importance of representing interactions between human actions and the dynamics between them.

2.1. Simulation and modelling for dynamic HRA methods

The dynamic HRA is the last generation for HRA approaches. Cacciabue [2] has outlined the importance of simulation and modelling of human performance for the field of human reliability. Specifically, simulation and modelling address the dynamic nature of human performance in a way that has not been possible in most HRA methods. Boring et al. [45] posits that the key to dynamic HRA is not in the development of specific methods but in the using of cognitive modelling and simulation to produce a data framework that may be used in quantifying likelihood of human error [10]. Simulator experiments can produce important basic information for the HRA method development and data for informing the use of existing HRA methods. Simulators allow the study of variations in context and how this impacts human performance [39].

As depicted in Fig. 2, simulation and modelling may be used in three ways to capture and generate data that are meaningful to HRA [10].

- The simulation runs produce logs, which may be analysed by subject matter experts and used to form an estimate of the likelihood of human error. This approach builds heavily on expert estimation techniques that are commonly used in HRA.
- The simulation may be used to produce estimates of performance shaping factors, which can be quantified to produce human error probabilities based on specific HRA methods. For example, Boring [41] postulated a mapping of performance measures produced by the MIDAS simulation system to the eight influencing factors utilised by the SPAR-H method.
- A final approach is to set specific performance criteria by which the virtual performers in the simulation are able to succeed or fail at given tasks. A common performance criterion is time to complete a task, whereby failure to complete the task within a prescribed limit is considered unsatisfactory performance. Iterations of the task that systematically explore the range of human performance can ultimately produce a frequency of failure (or success).

No modelling or simulation tool currently exists that completely or perfectly combines all elements of simulation-based HRA. Significant work is, however, already underway. At present, interest is growing in the fusion of simulation and modelling with HRA (e.g. [11,18,39]). A list of the main simulation projects and some of their main features are reported in Table 2. Some of these, such as CES (Cognitive Environment Simulation) and COSIMO (Cognitive Simulation Model), have been developed in the nuclear field and are computer simulation methods that could potentially be useful, but no use or development is evident since the late 90s [22,40].

Unlike CES and COSIMO, the environment simulation MIDAS (Man Machine Integration Design and Analysis system) was developed in 1986 in field of aerospace and aeronautic and has seen ongoing developments and applications over the years. Among the latest the integration efforts with HRA is the use of SPAR-H performance shaping factors [41] Another system, the Information, Decision, Action in Crew context (IDAC) model [13–18], combines a realistic plant simulator with a cognitive simulation system capable of modelling PSFs. Three generic types of operators are modelled: decision maker (e.g. shift supervisor), action taker (e.g. operators at the control panel), and consultant (e.g. resource experts in the control room). Due to the variety, quantity and detail of the input information, as well as complexity of applying its internal rules, the IDAC model is best implemented through a computer simulation such as the Accident Dynamics Simulator (ADS) environment.

The PROCOS simulator [11,12] instead attempts integration of the quantification of first generation HRA methods in safety assessment (e.g. THERP) with a cognitive evaluation of the operators involved in the context under examination. Its focus is mainly in conveying a quantitative result, comparable to those of a traditional HRA method and taking into account a cognitive analysis of the operator as well. As a further step, the simulator considers the evaluation of error management as part of the overall assessment from the same cognitive point of view. PROCOS does not imply the development of a detailed model for the operator–context interaction; the context is taken into account mainly through the input coming from the PSA framework to which it belongs, and through the use of performance shaping factors, as proposed in traditional HRA methods.

Simulators implemented over time are, above all, cognitive simulators; their aim is to simulate operator or crew behaviour in terms of correct and incorrect actions. These simulations model the operator's thought processes, and offer potentially powerful ways of determining how human operators will respond in emergency scenarios, typically in complex environments such as nuclear power plants [44]. The cognitive simulators developed to date have been mainly used for qualitative analysis and they have not found substantial applications in the quantitative risk assessment framework. The models are sometimes not easy to understand and therefore are not used by HRA specialists that have not been directly involved in their development [11].

2.2. Shortcomings and limitations in HRA methods

Over 70 human reliability tools were developed since 1960 for the same aim: human error quantification. Every method has the same purpose but uses different methodological frameworks, priority, operator models and performance shaping factors. HRA methods and simulation tools, proposed over the years, have not always been particularly useful to the purpose for which they were developed. Four main sources of deficiencies can be identified in current HRA methods [4].

- Lack of empirical data for model development and validation;
- lack of inclusion of human cognition (i.e. need for better human behaviour modelling);
- large variability in implementation (i.e. HRA parameters are different depending on the method used); and
- heavy reliance on expert judgment in selecting PSFs, and use of these PSFs to obtain the HEP in Human Reliability Analysis.

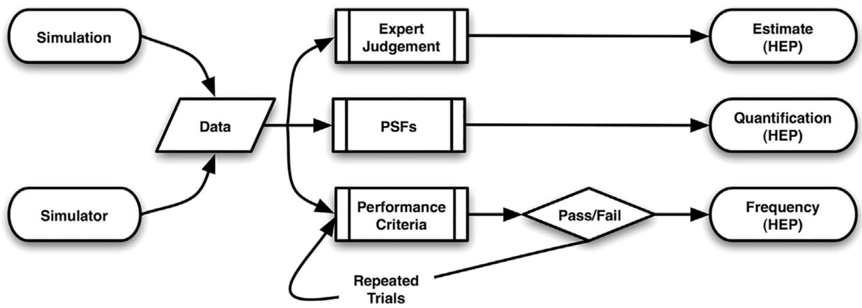

Fig. 2. Simulation and modelling in Human Reliability Analysis [10].

Table 2
Review of main simulators developed for simulating human behaviour in HRA field.

Simulator	Type	Description	Field
CES (Cognitive Environment Simulation) [46]	Quantitative and qualitative	It simulates the behaviour of a control room operator in a nuclear power plant in emergency scenarios. The purpose of CES was to emulate the way in which operators decide to respond, and then use the generated responses as a basis for quantification [40]. Developed using artificial intelligence programming.	Nuclear
COSIMO (Cognitive Simulation Model) [42]	Quantitative and qualitative	It simulates the behaviour of an operator reproduced through the Fallible Machine model by Reason, coupled with a model for the system specific for the system to be considered. Study the operator actions in abnormal plant conditions (accident scenarios) in a nuclear power plant [11,12,42].	Nuclear
ADS-IDAC (Accident Dynamics Simulator – Information, Decision, and Action in Crew context model) [13-17]	Quantitative	It is developed for probabilistic prediction of the responses of the nuclear power plant control room-operating crew during an accident for use in probabilistic risk assessments [13]. The operator response spectrum includes cognitive, emotional and physical activities during the course of the accident. Within the crew context, each individual operator's behaviours are simulated through a cognitive model under the influence of a number of explicitly modelled performance-influencing factors.	Nuclear
MIDAS (Man Machine Integration Design and Analysis system) [43]	Quantitative	It is an integrated suite of software components developed to aid designers and analysts to apply human factor principles and human performance models to the design of complex human–machine systems in aviation. It can simulate the behaviour of a pilot for civil aviation or an air traffic controller. The model of the operator is based on Rasmussen's model [11,12,43].	Aviation
PROCOS (Probabilistic Cognitive Simulator) [11]	Quantitative	It supports Human Reliability Analysis in complex operational contexts. It integrates cognitive human error analysis with standard hazard analysis methods (Hazop and event tree) by means of a *semi static approach* [11,12]. The simulation model comprised two cognitive flow charts reproducing the behaviour of a process industry operator. The simulator allows analysis of both error prevention and error recovery.	General
SYBORG (Simulation System for Behaviour of an Operating group) [47]	Qualitative	It simulates a group of nuclear power plant operators. It needs input coming from a specific plant simulator. It highlights some possible combinations of operator errors and plant condition that can lead to accident sequences; it proposes different strategies to improve the collaboration within the group [11,12].	Nuclear

In addition to having these four limitations, many methods are also deeply qualitative – they analyse all kinds of human error but do not provide useful numerical results. The quantification method is weak, and the quantitative results are unsubstantiated for many second generation methods that pay attention only on the responses of humans in accident scenarios. Other tools, such as THERP, include levels of detail that may be excessive for many assessments. The existing HRA tools allow very thorough evaluations of human behaviour in high-risk environments but can be resource intensive and time consuming.

First and second generation methods have focused on human error quantification and failure mode identification, while providing very detailed and realistic frameworks and approaches for calculation of HEP and human response. Many of these, however, have been developed for a specific context (e.g. aviation or nuclear power plants) and application (e.g. single operator or team simulation). Methods such as THERP or CREAM were born as approaches for nuclear power plants and consider only the typical accident scenarios in this context, so much effort needs to be applied in different fields, such as manual assembly or manufacturing systems [25,26]. In the same way, the major HRA simulation tools, described in the previous section, are adapted to specific fields, such as aviation and the control rooms of nuclear power plants. The model specificity can also be considered a weak point since it means that the models cannot be easily applied to analysis of tasks that differ from the task they have been developed for.

Despite the efforts of HRA experts, many of the limitations and problems of these approaches have not yet been resolved, due to the complexity of human nature and the difficulty of predicting and simulating human behaviour. In particular, Mosleh and Chang [18] have analysed the limitations of the existing HRA methods and outlined the guidelines for future methods, emphasising the importance of having methods that:

- identify human response (errors are the main focus);
- estimate response probabilities;
- identify causes of errors to support development of preventive or mitigating measures;
- have explicit role for ‘context’ both in error identification and probability estimation;

- are applicable by different users for different problems; and
- are traceable, consistent and repeatable.

3. Simulator for Human Error Probability Analysis (SHERPA)

The purpose of each HRA method must be to assess human behaviour and to quantify error probability, in order to reduce and prevent possible conditions of human error in a working context. Existing methods, as previously seen, do not always pursue this aim in an efficient way, but every method or simulator has its own strength. This paper proposes a new HRA model that exploits the advantages of the simulation tools and the traditional method HRAs to predict the likelihood of operator error, for a given scenario, in every kind of industrial system or other type of working environment. Three HRA elements converge into the model:

- task classification in one of the generic tasks proposed by HEART methods (first generation HRA);
- performance shaping factors analysis of SPAR-H methods (second generation HRA); and
- dynamic implementation using computer simulation (dynamic HRA).

The aspiration for the Simulator for Human Error Probability Analysis (SHERPA) model is not that it be a new HRA method in the long list of existing ones, but that it provides a theoretical framework that addresses the problem of human reliability in a different way from most HRA methods. Human reliability is estimated here as function of the performed task, performance shaping factors and also time worked, with the purpose of considering how reliability depends on the task and on the working context, as well as on the time that the operators have already spent at their work. Moreover, contextual factors of the second generation method (SPAR-H) allow careful evaluation of the working environment in order to identify the most negative factors. The model is able to provide for the following functions:

- estimating human reliability, as function of time, of work environment conditions, of physical and mental employee condition and of rest breaks distribution during the shift;
- assessing the effects due to different human reliability levels, through evaluation of processes, activities or tasks performed more or less correctly; and
- assessing the impact of context on human reliability, via performance shaping factors.

The tool also provides the possibility of determining the optimal configuration of breaks through use of a methodology that, together with assessments of an economic nature, allows identification of conditions that are required during the shift for the suspension of work for the operator's psychophysical recovery and then for the restoration of acceptable values of reliability [8]. SHERPA focuses on the quantitative aspect of HRA to obtain a significant numerical result, and it does not explore in depth the cognitive aspect of human error, as most HRA methods do. Instead, it includes the error likelihood in process simulation through information processing model of human performance using the SPAR-H approach. The most important objective of the work has been to realise a model for the evaluation of human reliability that can obtain useful information about human reliability for every kind of work task. SHERPA can be used in the preventive phase, as an analysis of the possible situations that may occur and for the evaluation of the percentage of scraps or non-compliant processing due to human error and in post-production to understand the nature of the factors that influence human performance in order to reduce errors.

The proposed model was not created for a particular industry or application and therefore can be easily applied to contexts that vary widely. For example, the module can equally represent manual processing or assembly, by varying the input variables such as performed task, level of contextual factors, or physical and mental employee condition. Simulators and tools similar to the one proposed do not exist today, either from the theoretical point of view or from the point of view of the analysis carried out. The next section describes the SHERPA logic and key steps in detail for determining human error probability.

3.1. Structure and logic of the SHERPA model

The theoretical model for predicting human reliability combines the technique for the quantification of nominal HEP of first

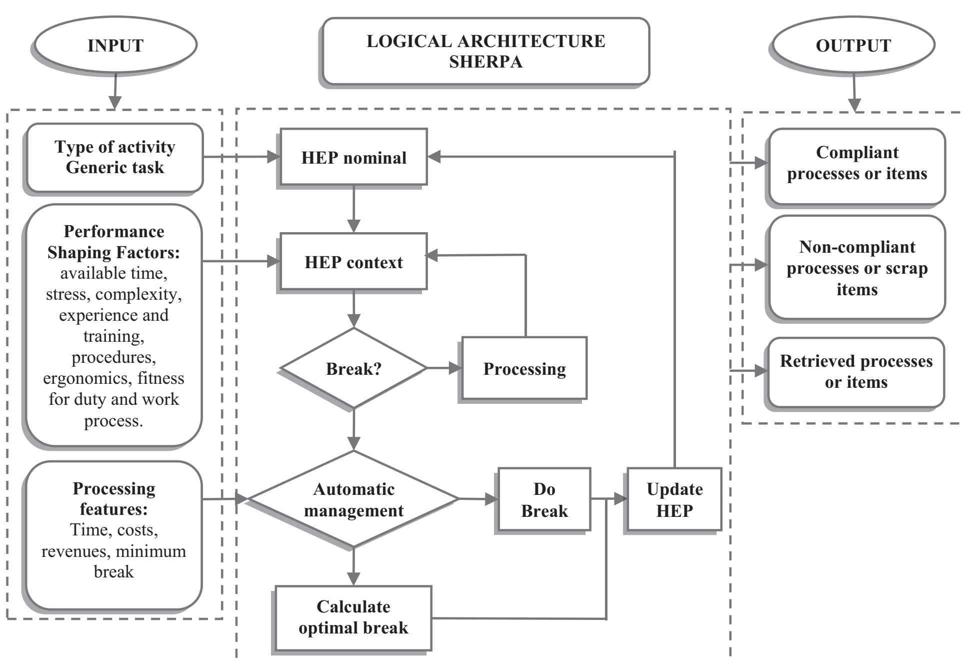

Fig. 3. Logical architecture of SHERPA model.

generation method HEART (Human Error Assessment and Reduction Technique) [19,21,22] with the evaluation of contextual factors of second generation method SPAR-H (Standardized Plant Analysis Risk-Human Reliability Analysis) [20,22,29–35]. This second phase requires a significant effort for the identification and quantification of overall performance shaping factors (PSFs) proposed by the SPAR-H method. Fig. 3 shows the model logic, its phases and the inputs and outputs of the module.

The starting analytical basis for the assessment of human errors in SHERPA is the determination of HEP, followed by quantification of PSF influences on the initial value of HEP. The module receives, as input, the type of activity and generic task, each of which is connected to an appropriate probability distribution that describes HEP as a function of time. Once a generic task has been selected, the human error probability (HEP) is influenced by the performance shaping factors for the current task.

The reliability operator, and the complementary probability of error, are used in a generic work process simulation, in order to indicate the probability that the operator will be able to carry out activities without errors in a specific scenario. The error represents non-compliant work, identified by a scrap piece. The module HRA determines the output number of good pieces, scraps and pieces recovered, on the basis of the working conditions given as input (e.g. working conditions, type of activity, environmental conditions, etc.). The concept of quality defects and scrap is not limited to manufacturing processes, but extends to a wider range of working environments, ranging from services to medical field.

SHERPA is implemented as a simulation template, developed in Arena 14.0©. The Arena template appears as a block flowchart that can be used in various types of simulation models and allows the assessment of human error probability without excessive time consumption. The template is shown in Fig. 4. It can be inserted in the processing flow and it simulates the task performed by operator during a whole shift, considering the time required to complete the task as well as calculating the operator's human reliability (HR) and human error probability (HEP).

Different scheduling of breaks can also be assigned and can be simulated in the shift, considering that a break determines the operator's recovery and the consequent increase in reliability. The main dialogue box, show in Fig. 5, allows access to all the masks for entering data required for simulation; in this way, every input can be defined. The SHERPA dialogue box provides access to several other windows for data entry. Fig. 6 shows, in the form of a tree, the sub-dialogue boxes that are linked to the main dialogue of the module HRA.

3.2. Determination of nominal HEP

The determination of human error probability is needed in any HRA approach and is always the first step to be carried out for a correct Human Reliability Analysis. The preliminary analysis of the model requires advance knowledge of the probability with which an operator can make mistakes, and therefore assumes probability distributions of HEP as functions of time and type of operation to perform, which describe the variations in human performance. The probability distribution that best describes the error distribution is the Weibull [8], whose analytical expression is

$$HEP_{nominal} = 1 - e^{-\alpha t^{\beta}} = 1 - \int_0^t f(\tau)d\tau \tag{2}$$

where the two parameters α and β change the scale and shape of the curve, whose probability density function $f(t)$ is the following:

$$f(t) = \alpha \cdot \beta \cdot t^{(\beta-1)} \cdot e^{-\alpha t^{\beta}} \tag{3}$$

This distribution function of error probability increases with time and minimum probability of error; thus the highest reliability occurs at time zero, in the initial part of the shift. However, the natural process of adaptation for a typical human for a given operation results in a lower reliability in the initial part of the shift – a sort of transient phase in which the error rate must decrease to reach the minimum value in the first hour of processing. This problem requires modification of the Weibull formula, assuming a trend that first increases and then decreases the HR function, as shown in Fig. 7.

The function has also been assumed to have a minimum value of error probability in the first hour of processing and a maximum value at the eighth hour of work during an eight-hour shift. With these assumptions, the probability distribution of error is adapted as follows:

$$\begin{cases} HEP_{nominal}(t) = 1 - k \cdot e^{-\alpha \cdot (1-t)^{\beta}} & \forall\ t \in [0;1] \\ HEP_{nominal}(t) = 1 - k \cdot e^{-\alpha \cdot (t-1)^{\beta}} & \forall t \in]1;\infty[\end{cases} \tag{4}$$

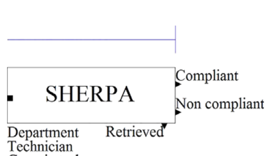

Fig. 4. SHERPA template user interface.

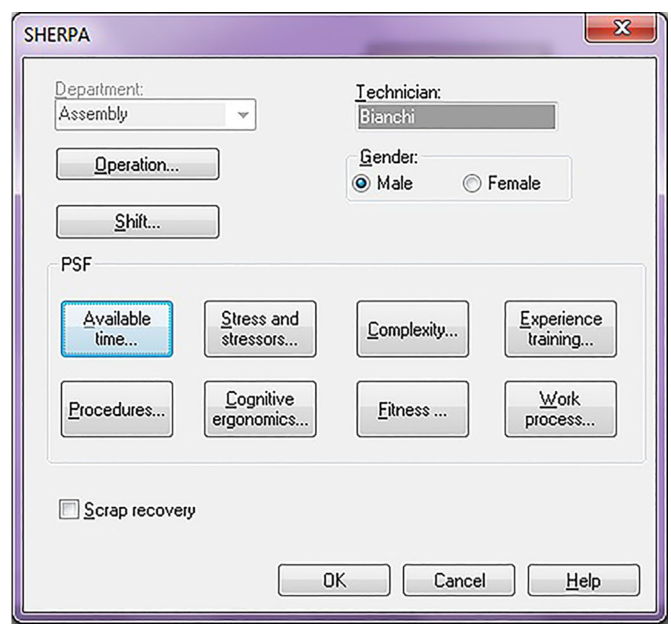

Fig. 5. Main dialogue SHERPA template.

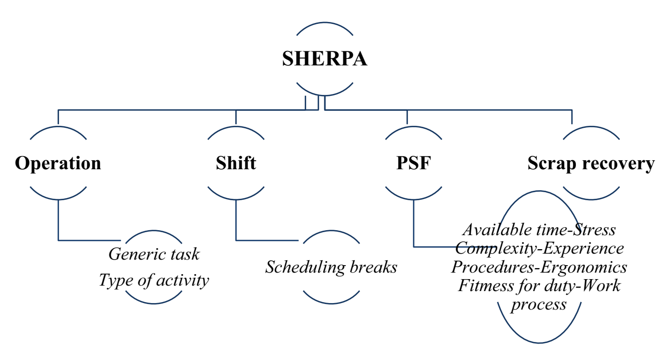

Fig. 6. Connections tree between sub-dialogue boxes and dialogue man.

The value of k will be calculated according to the value that the curve takes for $t=1$, β will be chosen in the function of the form that we want to give to the curve, while α can be calculated by imposing the value of the function for $t=8$, as previously described. The values obtained are shown in Table 3.

As noted above, the first generation HRA methods, such as THERP and HEART, focused on the quantification of nominal HEP. The second generation does not give great importance to the formal quantification of HEP, but often uses standard values, as in the case of SPAR-H, to allow greater focus on the influence of PSFs. For SHERPA, the best choice is the Human Error Assessment and Reduction Technique method, designed to be a quick and simple method applicable to any situation or sector in which human reliability is important.

HEART uses eight general categories to classify operator tasks, but only six have been chosen for the proposed model. The categories shown in Table 3 can represent a wide range of work activities from simple to more complex ones, from ones with a very high error rates to those more reliable, thanks to the presence of automatic systems of supervision. This range of activities allows the module to apply the model to very different working environments without any kind of restrictions.

For each category, Fig. 8 shows the performance of human reliability function (HR) and the probability of human error (HEP), based on the changes in the Weibull distribution.

The information about the task to be performed by the operator can be inserted into the dialogue box Operation. Action or diagnosis can be chosen in this window (Fig. 9), according to the SPAR-H method, where the tasks are divided into action (implementations of actions/processes simple or complex) and diagnosis (interpretation of system status and decision making in case of need). The window allows specification of the type of activity and the generic task performed by the operator. Such information is necessary to identify the nominal reliability curve of the operator in order to quantify the total HR in the simulation.

3.3. Influence of the performance shaping factors

The performance shaping factor is determined by the individual characteristics of the human being, the environment, the organisation or the activity that enhances or decreases human performance and increases or decreases the likelihood of human error. The PSFs allow all environmental and behavioural factors that affect human performance to be taken into account.

The SPAR-H method is used to quantify the performance shaping factors in SHERPA. While many HRA methods have often proposed a large number of PSFs, even as many as fifty, SPAR-H attempts to provide a reasonable coverage of the influence spectra of human performance in a reasonable minimum number of PSFs. The eight PSFs are: available time, stress, complexity, experience and training, procedures, ergonomics, fitness for duty and work process. The decision to use only eight PSFs is based on a review of the available HRA methods and the behavioural sciences [34]. Unlike most of the HRA methods, SPAR-H recognises that a number of PSFs can have both positive and negative effects on performance.

As shown in Fig. 10, the probability of error increases with the growth of the negative influence of the PSFs, while, on the contrary,

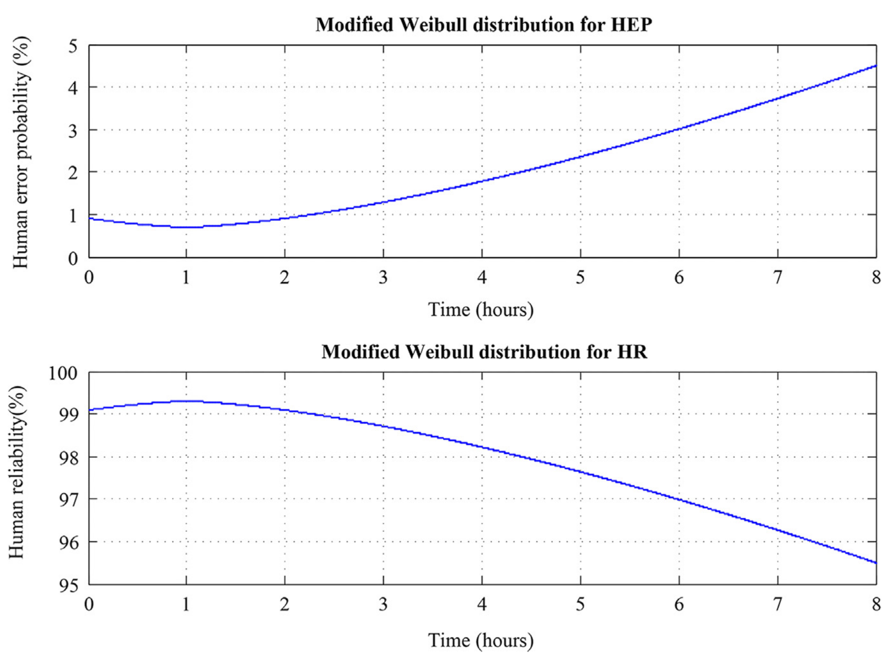

Fig. 7. Trend in modified human error probability and human reliability.

Table 3
Coefficient values for every generic task [19].

	Generic task	Limitations of unreliability for operation (%)	k	α	β
1	Totally unfamiliar	35/97	0.65	0.1660762	1.5
2	Complex task requiring high level of comprehension and skill	12/28	0.88	0.0108352	1.5
3	Fairly simple task performed rapidly or given scant attention	6/13	0.94	0.0041785	1.5
4	Routine, highly-practised	0.7/4.5	0.993	0.0021068	1.5
5	Completely familiar, well-designed, highly practised, routine task	0.008/0.9	0.9992	0.0004838	1.5
6	Respond correctly to system command even when there is an augmented or automated supervisory system	0.0001/0.09	0.99991	$4.813^{*}10^{-5}$	1.5

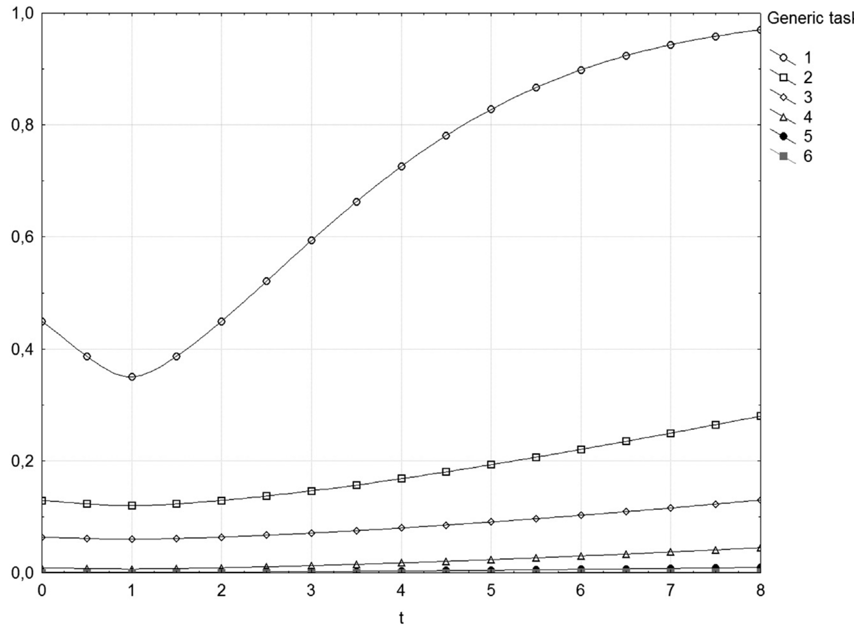

Fig. 8. Trend in HEP for every generic task.

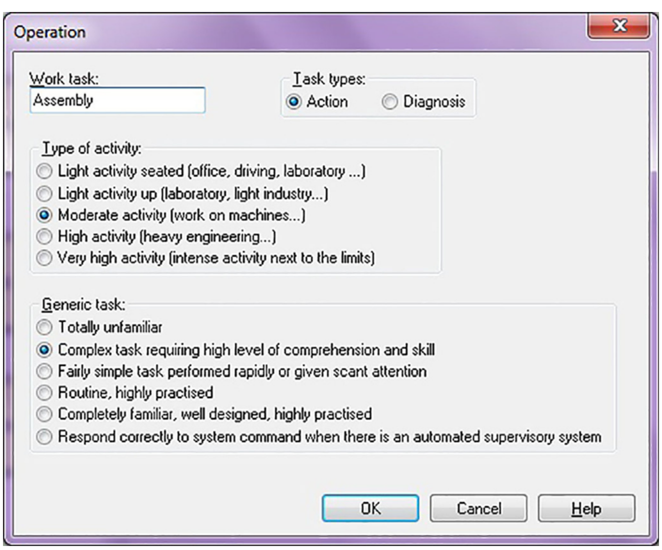

Fig. 9. Dialogue box for operation data entry.

the error probability decreases as the positive influence of the PSFs grows. When the influencing factor represents a positive effect, it corresponds to a value less than one; therefore, the multiplication of a nominal HEP with this value is used to decrease the overall HEP. When the PSF instead represents a negative effect, it corresponds to a value greater than one and the multiplication of a nominal HEP with this positive integer serves to increase the HEP. The impact of the context is assessed using the following adjustment factor [35].

$$HEP_{contextual} = HEP_{nominal} \cdot PSF_{composite} / HEP_{nominal} \cdot (PSF_{composite} - 1) + 1 \quad (5)$$

where $PSF_{composite}$ is calculated as the product of the evaluations of all the PSFs [35].

$$PSF_{composite} = PSF_1 \times \ldots \times PSF_x \times \ldots \times PSF_8 \quad (6)$$

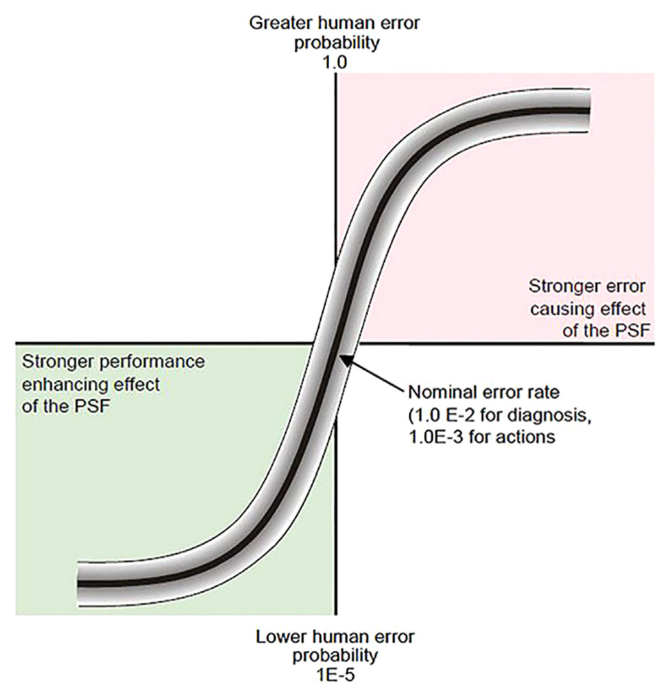

Fig. 10. Ideal mean HEP as a function of the influence of performance shaping factors [34].

The strength of SPAR-H is in providing a guide for assigning numerical weights for the PSFs; the multiplier values for every PSF are reported in Table 4.

The multiplier values were attributed by analysts of the method, on the basis of several studies carried out on nuclear power plants. In order to align the evaluation of PSFs in our model, we standardised the multipliers shown in Table 5. This standardisation changes the values of the multipliers for each generic task compared to the average value of the class, maintaining them equal to a nominal level.

The multiplier value is attributed to some PSFs as a direct input of the level (e.g. experience is directly established as a low, nominal or

Table 4
Multipliers of the performance shaping factors separated between action and diagnosis [29–31,35].

SPAR-H PSFs	PSF Levels	Multipliers action	Multipliers diagnosis
Available time	Inadequate time	P (failure)=1	P (failure)=1
	Time available=time required Barely adequate time	10	10
	Nominal time	1	1
	Time available > 5 × time required (extra time)	0.1	0.1
	Time available > 50 × time required (expansive time)	0.01	0.01
	Insufficient information	Nominal time	Nominal time
Stress/stressors	Extreme	5	5
	High	2	2
	Nominal	1	1
	Insufficient information	Nominal	Nominal
Complexity	Highly complex	5	5
	Moderately complex	2	2
	Nominal	1	1
	Obvious diagnosis	–	0.1
	Insufficient information	Nominal	Nominal
Experience/training	Low	3	10
	Nominal	1	1
	High	0.5	0.5
	Insufficient information	Nominal	Nominal
Procedures	Not available	50	50
	Incomplete	20	20
	Available, but poor	5	5
	Nominal	1	1
	Diagnostic/symptom oriented	–	0.5
	Insufficient information	Nominal	Nominal
Ergonomics	Missing/misleading	50	50
	Poor	10	10
	Nominal	1	1
	Good	0.5	0.5
	Insufficient information	Nominal	Nominal
Fitness for duty	Unfit	P=1	P=1
	Degraded Fitness	5	5
	Nominal	1	1
	Insufficient information	Nominal	Nominal
	Poor	5	2
Work processes	Nominal	1	1
	Good	0.5	0.8
	Insufficient information	Nominal	Nominal

high level). For other PSFs, the final value of the multiplier is obtained from the weighted average of the multipliers assigned to the single sub-factors, where the weight is assigned by the rating analysts. Thereafter, identified levels and influencing factors are considered for each PSF (see [29–31,35]).

3.3.1. *Available time*

Available time, as a PSF term, can be misleading. In the assessment of the Available Time, SPAR-H does not look solely at the amount of time that is available for a task. Rather, it looks at the amount of time available relative to the time required to complete the task. Available time refers to the amount of time that an operator or a crew has to diagnose and act upon an abnormal event ([29–31,35]). The time available can take on six levels, both positive and negative. The dialogue box allows user to choose between different alternatives, which differ in case of action or diagnosis, as shown in Fig. 11. For detailed information on this and on subsequent PSFs, see [29,30,31,35].

Insufficient information is always present as alternative for all eight performance shaping factors and represents the situation where information is insufficient for assigning a level to a PSF. *Insufficient information* is quantified with the same value as *Nominal*.

In the Available time dialogue box there is a dialogue button for the window Multipliers Available time; this dialogue box allows the values of the PSF multipliers to be changed (Fig. 12). This type of window is present for each PSF; it allows the use of the default values of the multipliers as well as modification of them to suit different operational situations.

3.3.2. *Stress and stressors*

Stress, as used in SPAR-H, specifically refers to the level of undesirable conditions and circumstances that impede the operator in completing a task [29–31,35] Note that the effect of stress on performance is curvilinear – that is, some small amount of stress can enhance performance, and in the context of SPAR-H should be considered nominal, while high and extreme levels of stress will negatively affect human performance. The degradation of performance is the key point when assigning high or extreme stress levels.

For stress, as well as for the complexity and work processes, the value of the multiplier is determined by the presence of more sub-factors. In these cases, the overall PSF value is given by the weighted average of the sub-factor multipliers with respect to the weights that can be reset or assigned from time to time during the insertion of the input. Several environmental and behavioural factors contribute to identify the multiplier:

1. Circadian rhythm;
2. mental stress;
3. pressure time;

Table 5
Modified multipliers due to standardisation.

SPAR-H multipliers	Generic task 1	Generic task 2	Generic task 3	Generic task 4	Generic task 5	Generic task 6
50	21.00	26.00	28.00	34.00	56.00	82.00
20	8.40	10.40	11.20	13.60	22.40	32.80
10	4.20	5.20	5.60	6.80	11.20	16.40
5	2.10	2.60	2.80	3.40	5.60	8.20
3	1.26	1.56	1.68	2.04	3.36	4.92
2	1.01	1.04	1.12	1.36	2.24	3.28
1	1	1	1	1	1	1
0.8	0.34	0.42	0.45	0.54	0.90	0.99
0.5	0.21	0.26	0.28	0.34	0.56	0.82
0.1	0.04	0.05	0.06	0.07	0.11	0.16
0.01	0.004	0.005	0.006	0.007	0.011	0.016

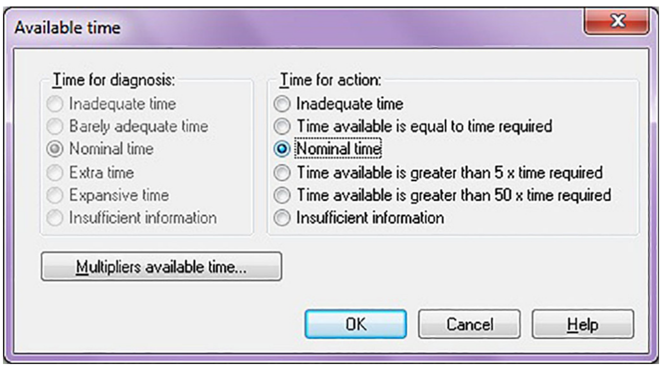

Fig. 11. Dialogue box available time data entry for action.

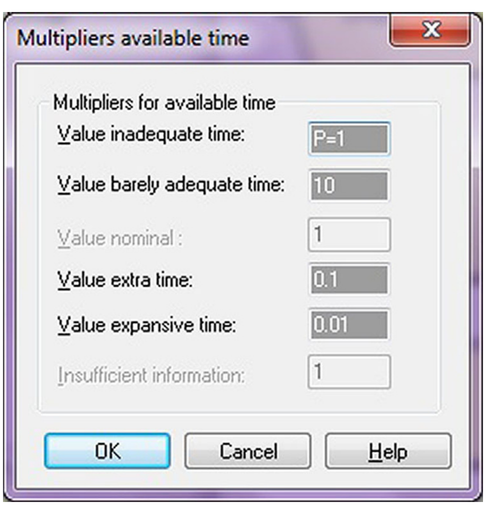

Fig. 12. Dialogue box to change PSF multipliers.

4. workplace;
5. microclimate;
6. lighting;
7. noise;
8. vibration;
9. ionising and non-ionising radiation.

Each of these contributes to the calculation of the total PSF stress through the formula $PSF_{stress} = F_1 \cdot W_1 + \ldots + F_x \cdot W_x + \ldots + F_9 \cdot W_9$, where F_1 is the level assigned to one of the nine factors listed above and W_1 is the weight of each factor between 0 and 1. The weights must respect the condition $\sum_{i=1}^{9} W_i = 1$. The presence of several factors makes the window stress richer and more complex, as evident from Fig. 13.

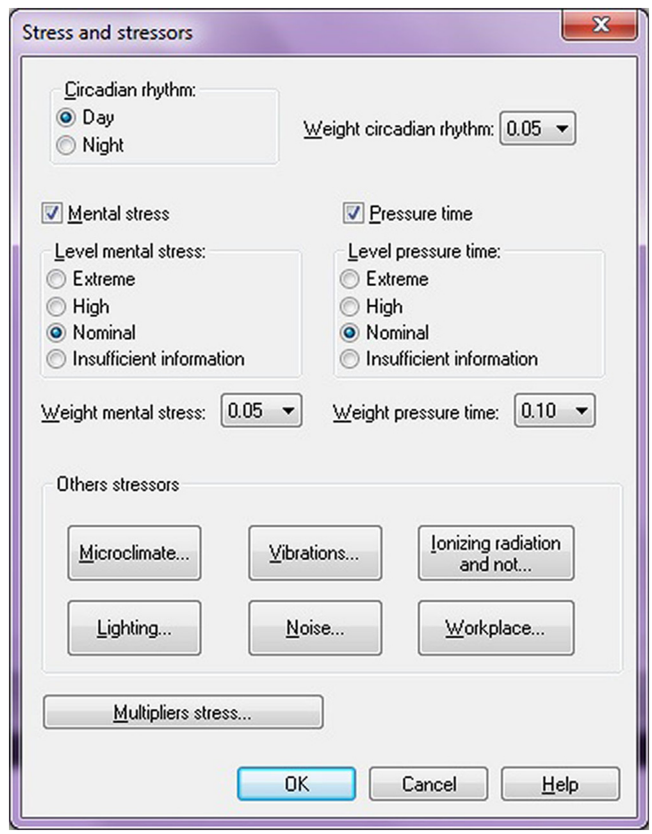

Fig. 13. Dialogue box stress data entry.

3.3.3. Complexity

Complexity refers to how difficult the task is to perform in the given context; it considers both the task and the environment [29–31,35]. A more difficult task has a greater chance for human error. Similarly, a more ambiguous task has a greater chance for human error. The complexity is one of those PSFs that Boring [45] defines as indirect, as it cannot be measured directly. For this reason, the value of the complexity cannot be assigned directly but relies on input from several elements (Fig. 14)

1. General complexity;
2. mental efforts required;
3. physical effort required from type of activity;
4. precision level of the activity; and
5. parallel tasks.

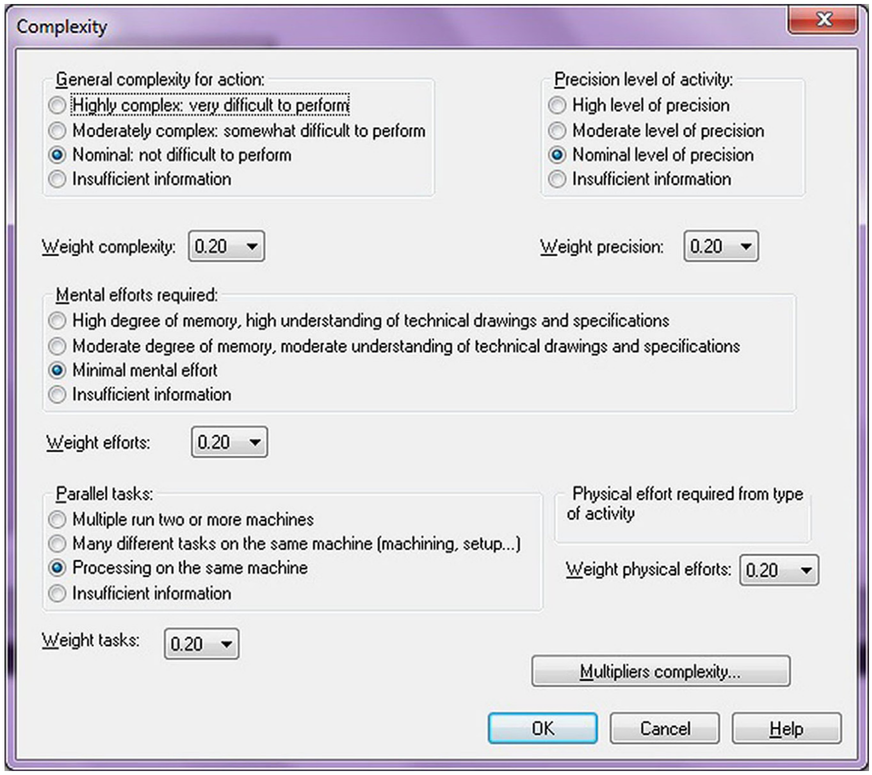

Fig. 14. Dialogue box complexity data entry.

Each of these contributes to the calculation of the overall complexity PSF through the formula $PSF_{complexity} = F_1 \cdot W_1 + ... + F_x \cdot W_x + ... + F_5 \cdot W_5$, where F_1 is the level assigned to one of the five factors listed above and W_1 is the weight of each factor between 0 and 1. The weights must respect the condition $\sum_{i=1}^{5} W_i = 1$.

3.3.4. *Experience and training*

This PSF refers to the experience and training of the operator involved in the task [29–31,35]. Included in this consideration are years of experience of the individual or crew, and whether or not the operator/crew has been trained on the type of accident, the amount of time passed since training, the frequency of training, and the systems involved in the task and scenario. In SHERPA, the data on the training and experience of the operator are inserted directly according to following levels:

- Low = less than 6 months of relevant experience and/or training.
- Nominal = more than 6 months of relevant experience and/or training.
- High = extensive experience; a demonstrated master.

3.3.5. *Procedures*

This PSF refers to the existence and use of formal operating procedures for the tasks under consideration [29–31,35] Common problems seen in event investigations for procedures include situations where procedures give wrong or inadequate information regarding a particular control sequence. Another common problem is the ambiguity of steps. Levels used for this PSF in SPAR-H are depicted in Fig. 15.

3.3.6. *Cognitive ergonomics*

Ergonomics refers to the equipment, displays and controls, layout, quality, and quantity of information available from instrumentation, and the interaction of the operator/crew with the equipment to carry out tasks [29–31,35]. Aspects of the human–machine interface are included in this category, as well as the adequacy or inadequacy of computer software. SPAR-H was born in the nuclear field, so ergonomics are mainly oriented to the interaction of a human or group to the instrumentation typical of a control room, such as the display and control buttons. In other kinds of industries, this PSF focuses instead on the ergonomics of the workplace and the equipment used (Fig. 16).

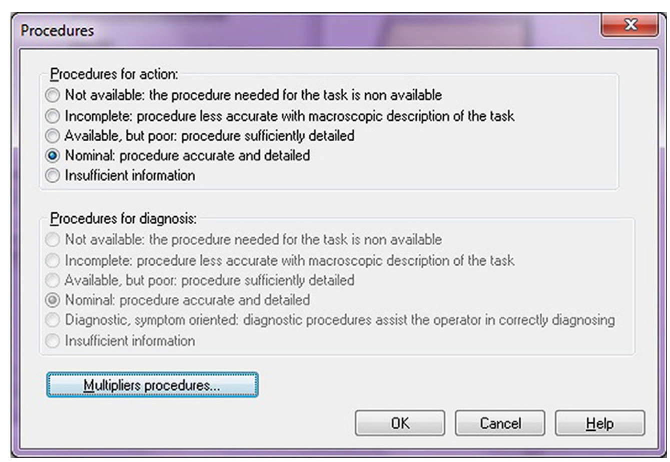

Fig. 15. Dialogue box procedures data entry.

3.3.7. *Fitness for duty*

Fitness for duty refers to whether or not the individual is physically and mentally suited to the task at hand [29–31,35]. This PSF includes fatigue, sickness, drug use (legal or illegal), overconfidence, personal problems and distractions and also includes factors associated with individuals, but not related to training, experience or stress (which are covered by other PSFs). Levels used in SHERPA are:

- Unfit = the individual is unable to carry out the required tasks, due to illness or other physical or mental incapacitation (e.g. having an incapacitating stroke).
- Degraded fitness = the individual is able to carry out the tasks, although performance is negatively affected. Mental and physical performance can be affected if an individual is ill, such as having a fever. Individuals can also exhibit degraded performance if they are inappropriately overconfident in their abilities to perform.
- Nominal = the individual is able to carry out tasks; no known performance degradation is observed. Nominal should also be used when the analyst judges the PSF as not a performance driver.

3.3.8. *Work processes*

Work processes refer to aspects of doing work, including inter-organisational factors, safety culture, work planning, communication and management support and policies [29–31,35]. How work is planned, communicated and executed can affect individual and crew performance. If planning and communication are poor, then individuals might not fully understand the work requirements. Work processes also include any management, organisational or supervisory factors that may affect performance. In this case, the value of the PSF also may not be assigned directly but may be based on multiple elements that are input (Fig. 17)

1. Communication and integration in team work and
2. work processes.

Each of these contributes to the calculation of the total PSF work processes through the formula $PSF_{work\ processes} = F_1 \cdot W_1 + F_2 \cdot W_2$,

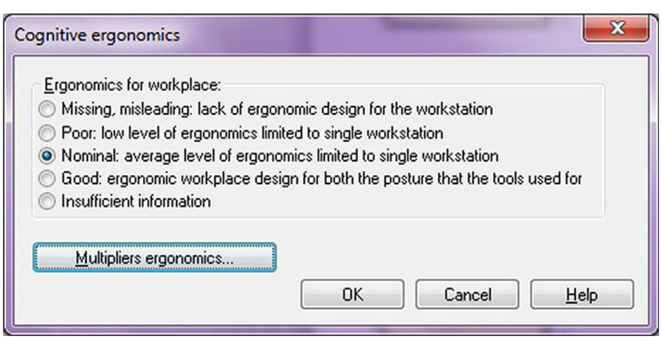

Fig. 16. Dialogue box cognitive ergonomics data entry.

where F_1 is the level assigned to one of two factors listed above and W_1 is the weight of each factor between 0 and 1. The weights must respect the condition $\sum_{i=1}^{2} W_i = 1$.

4. Case study: simulation of manual assembly processing

The proposed SHERPA model has been used to conduct several simulations to evaluate the effect of human reliability as part of manufacturing activity in the prevailing manual content. In order to illustrate the operation of the module described in the previous section, a manual assembly processing was simulated as a case study, in which human reliability is critical due to the high contribution of manual tasks. The experiment was conducted using a simulation model, which reproduces the operator work station involved in manual assembly on an eight-hour shift. The simulation was carried out on an annual basis, taking into account 235 working days, always with the same work shift. The assembly operation was simulated for three different items with random arrival sequences based on a default mix. For each item, processing times, characterised by a triangular distribution, fixed and variable costs and selling prices, as well as overall production mix, were defined. The data described are shown in Table 6.

The SHERPA template is integrated in a specific Arena model in order to allow simulation of established scenarios. Fig. 18 shows the Arena model that provides for the entity creation, the assignment of the attributes required for simulation and implementation of the different production mix chosen. Features of items, production mix and SHERPA inputs can easily be modified for subsequent simulations through the following model. Once the required attributes are assigned, entities are entered into template that allow simulation of the working process and generate the following output: compliant, non-compliant and retrieved items. Of course, the template must be completed with all information concerning the activity and the environmental and behavioural operator conditions, through the dialogue boxes discussed in the previous section.

The first aspect of SHERPA taken into consideration in the case study was the impact of different types of modelled generic tasks. Three scenarios were simulated for each of the six categories, keeping the contextual factors at nominal level, with the composite PSF equal to one. Results for every scenario are shown in Table 7, where the total value of compliant and non-compliant items, their respective percentages and the mean values of HEP nominal and HEP context are reported.

The most interesting aspect of the SHERPA model, however, is its ability to simulate several environmental conditions for the

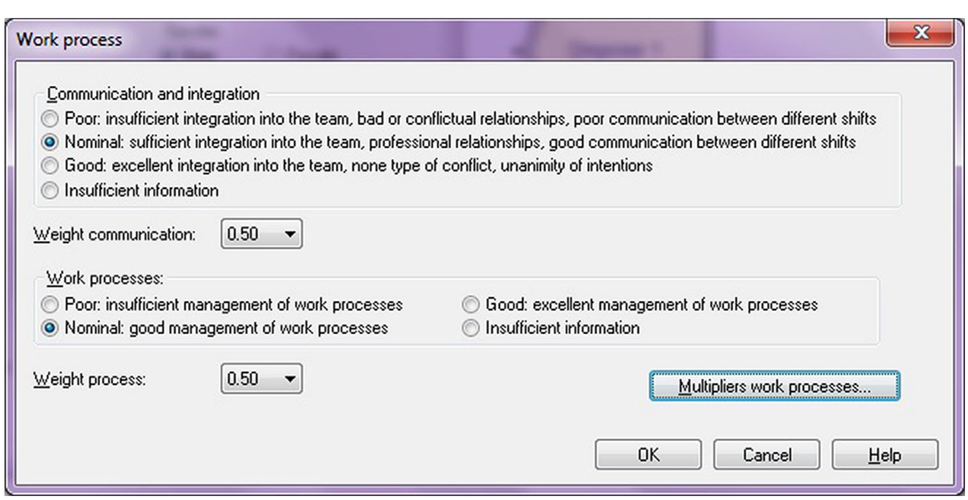

Fig. 17. Dialogue box work process data entry.

same performed activity. The second step of simulation is focused on positive or negative influences of PSFs, keeping constant the type of activity sets equal to generic task three. Table 8 shows the results of simulations carried out by changing, from time to time, only the complexity and procedures level and keeping all other values at the nominal level.

The performance shaping factors do not always have a negative impact on the reliability, but factors such as experience, ergonomics, time available and work processes may lead to the improvement of the reliability and the consequent decrease in the probability of human error. In the case study, two different conditions were tested where the positive effect of experience was tested and then the positive effect of ergonomics was added. Table 9 shows a high human reliability improvement due to the decrease in the value of the composite PSF for scenarios two and three.

In most real cases positive and negative factors coexist and affect the activity carried out by the operator. This condition was simulated considering factors with positive impact such as high experience, and other negative factors, such as moderate stress, poor procedures and poor working processes. In this last simulation, the same conditions are used for every scenario and the results are shown in Table 10.

4.1. Discussion

The SHERPA model allows to simulate numerous scenarios, considering a plurality of conditions and working activities, as evident from the case study, in which the three simulated scenarios, even if fictitious, are representative of actual working environments. The simulator produces in short time results in terms of compliant and non-compliant items and human error probability, as reported in the previous Tables 7–10.

Unlike many HRA methods, SHERPA has been implemented for covering a wide range of working task, for this reason the six modelled categories may represent activities that are more or less reliable. As evident from Fig. 19, the percentage of non-compliant items decreases going from generic task one to task six, due to the increase in the reliability level of each category and the complementary decrease in the human error probability. The generic task one represents the worst activity in terms of reliability; in fact, the average human reliability is approximately equal to 30% in every scenario without taking into account the influence of PSFs. The other categories are higher nominal values of human reliability, reflecting the HEART limitations of unreliability for operation reported in previous Table 3.

The simulated scenarios have been used to assess the behaviour of the template when the PSFs levels vary, i.e. with different contextual conditions, for the same performed activity.

In the Scenario 1, shown in Table 8, the simulations highlight the relationship between the composite PSF and the contextual human error probability; in fact, the value of contextual HEP grows with increases in the composite PSF. Starting from the same nominal HEP value, always kept constant, the performance shaping factors increase variably the HEP contextual according to their multiplier (Fig. 20).

In some cases the variation is limited; for example, when considering a moderate level of complexity there is the increase of approximately 10% compared to the nominal level. The increase grows up to 95%, with a 18.85% of non-compliant items, when the complexity level is extreme. In other cases, the particular environmental or personal conditions can lead to high increases in the probability of error, as in the case of not available procedures where the variations in HEP are larger due to the multiplier theoretically assigned from the SPAR-H method.

A further assessment done in the case study is relative to the positive performance shaping factors. Positive factors lead to a decrease in the final value of the composite PSF and an improvement in the operator reliability compared to the nominal HEP (see Table 9). As evident in Fig. 21, a multiplier of high experience, amounting to 0.28, improves the human error probability, lowering it to the 2.61%. Where two or more positive PSFs are merged, the improvement is even more evident; for example, in the simulation, high experience level and good ergonomics level allow the values of human error to approach nearly zero, equal to 0.75% in both scenarios.

Table 6
Features simulated items.

Features	Item 1	Item 2	Item 3
Processing time (min.)	25	36	45
Setup time (min.)	5	5	5
Price (€)	115	155	200
Fixed cost (€)	52	65	78
Variable cost (€)	18	24	32
SCENARIO 1 (%)	20	5	75
SCENARIO 2 (%)	15	65	20
SCENARIO 3 (%)	50	30	20

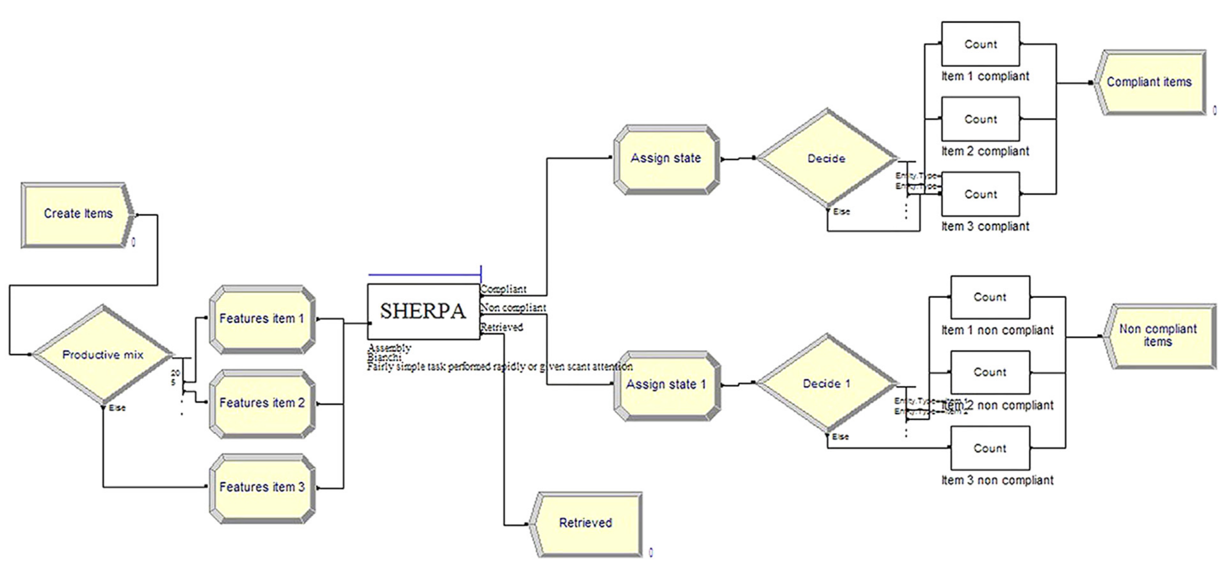

Fig. 18. Assembly model with SHERPA template.

Table 7
Results of the first step of the simulation.

	Generic task 1	Generic task 2	Generic task 3	Generic task 4	Generic task 5	Generic task 6
	SCENARIO 1					
Compliant items	838	2162	2400	2589	2645	2652
Non-compliant items	1814	490	252	63	7	0
Total items	2652	2652	2652	2652	2652	2652
Compliant (%)	**31.60**	**81.52**	**90.50**	**97.62**	**99.74**	**100**
Non-compliant (%)	**68.40**	**18.48**	**9.50**	**2.38**	**0.26**	**0**
Average HEP nominal (%)	68.79	17.90	8.50	2.06	0.39	0.012
Average HEP Context	68.79	17.90	8.50	2.06	0.39	0.012
	SCENARIO 2					
Compliant items	917	2380	2643	2845	2896	2911
Non-compliant items	1995	532	269	67	16	1
Total items	2912	2912	2912	2912	2912	2912
Compliant (%)	**31.49**	**81.73**	**90.76**	**97.70**	**99.45**	**99.96**
Non-compliant (%)	**68.51**	**18.27**	**9.24**	**2.30**	**0.55**	**0.04**
Average HEP nominal (%)	68.73	17.89	8.50	2.06	0.39	0.012
Average HEP context (%)	68.73	17.89	8.50	2.06	0.39	0.012
	SCENARIO 3					
Compliant items	959	2621	2931	3124	3180	3196
Non-compliant items	2238	576	266	73	17	1
Total items	3197	3197	3197	3197	3197	3197
Compliant (%)	**30**	**81.98**	**91.68**	**97.72**	**99.47**	**99.97**
Non-compliant (%)	**70**	**18.02**	**8.32**	**2.28**	**0.53**	**0.03**
Average HEP nominal (%)	68.57	17.84	8.50	2.06	0.39	0.012
Average HEP context (%)	68.57	17.84	8.50	2.06	0.39	0.012

Table 8
Results of simulation for scenario one, while changing complexity and procedures levels.

SCENARIO 1	**Nominal complexity**	**Moderate complexity**	**Extreme complexity**	**Available but poor procedures**	**Incomplete procedures**	**Not available procedures**
Composite PSF	1.024	1.12	2.464	2.867	11.4688	28.672
Compliant items	2396	2372	2150	2106	1311	761
Non-compliant items	256	280	500	546	1341	1891
Total items	2652	2652	2652	2652	2652	2652
Compliant (%)	**90.35**	**89.44**	**81.15**	**79.41**	**49.43**	**28.70**
Non-compliant (%)	**9.65**	**10.56**	**18.85**	**20.59**	**50.57**	**71.30**
Average HEP nominal (%)	8.52	8.52	8.52	8.52	8.52	8.52
Average HEP context (%)	8.70	9.44	18.55	20.92	50.85	71.80

Table 9
Results of simulation for scenarios two and three, changing experience and ergonomics levels.

	SCENARIO 2			**SCENARIO 3**		
	Nominal	**Experience**	**Experience + ergonomics**	**Nominal**	**Experience**	**Experience + ergonomics**
PSF	–	0.28	0.28-0.28	–	0.28	0.28-0.28
Composite PSF	1	0.28	0.078	1	0.28	0078
Compliant items	2643	2828	2886	2931	3107	3171
Non-compliant items	269	84	26	266	90	26
Total items	2912	2912	2912	3197	3197	3197
Compliant (%)	**90.76**	**97.12**	**99.11**	**91.68**	**97.18**	**99.19**
Non-compliant (%)	**9.24**	**2.88**	**0.89**	**8.32**	**2.82**	**0.81**
Average HEP nominal (%)	8.52	8.52	8.52	8.52	8.52	8.52
Average HEP Context (%)	8.52	2.61	0.75	8.52	2.61	0.74

Finally the same kind of task and the same contextual conditions have been used in the three scenarios. The results, in terms of reliability, are very similar to each other, as shown in Table 10, because the difference in production mix translates especially in terms of total units produced, given the different processing times. The error probability, in these cases, is determined more by the type of activity than by performance shaping factors.

The previous analysis underline the major SHERPA features, described in the theoretical model. In particular, its versatility is useful in revealing the environmental and psycho-physical factors which mainly influence the human reliability and may therefore be subject to improvement in order to reduce errors. The template requires, however, additional tests for the validation and the calibration of coefficients and weights.

5. Conclusion

Human error in the workplace can have more or less serious consequences, such as accidents, malfunctions and defects in the

Table 10
Results of last step of simulation.

	SCENARIO 1	*SCENARIO 2*	*SCENARIO 3*
Composite PSF	1.7208	1.7208	1.7208
Compliant items	2269	2489	2754
Non-compliant items	383	423	443
Total items	2652	2912	3197
Compliant (%)	**85.56**	**85.47**	**86.14**
Non-compliant (%)	**14.44**	**14.53**	**13.85**
Average HEP nominal (%)	8.52	8.52	8.50
Average HEP context (%)	13.77	13.77	13.73

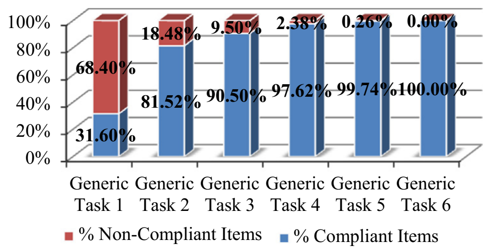

Fig. 19. Percentage of compliant and non-compliant items for generic categories of activities.

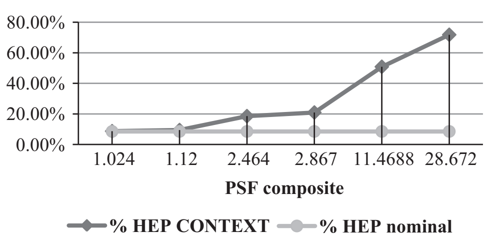

Fig. 20. HEP as function of increasing composite PSF.

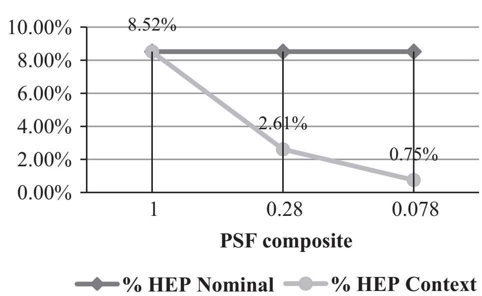

Fig. 21. HEP as function of decreasing composite PSF.

quality of the performed task. The evidence that human actions are a source of vulnerability for industrial systems has led to the birth of many Human Reliability Analysis (HRA) methods, which aim at further examination of the human factor, but they have not always been especially useful for this purpose. The SHERPA model proposed in this paper has as its main objective the provision of a model for quantifying human error probability in any work situation and in every context – quantification that today is hardly possible given the lack of tools similar to that achieved in this work.

Through SHERPA, the concept of human reliability, often dealt with only in theory, is taken up in terms of production capacity (compliant and non-compliant items or retrieved items), and useful information about human reliability can be obtained for every kind of working context. The SHERPA model can be effectively used to evaluate changes in human error probability when changes occur in type of activity, contextual conditions, time spent at work and breaks assigned during the shift. The main advantage of the model lies in its being generic – it is suitable for any environment and working conditions, without limitations related to a particular sector or activity. A large number of scenarios can be simulated without being resource intensive or time consuming. The case study has shown the potential of SHERPA, both in classifying different work activities and in evaluating the impact of performance shaping factors. Furthermore, the ability to change the values of the multipliers makes it easy to modify the weight of each factor PSF, regardless of the values assigned by the SPAR-H method.

The management of breaks provided by the module allows simulation of all distributions of unintended breaks and evaluation of its effect on both the percentage of non-compliance and the economic return. In this way, different scheduling of breaks can be tested and compared, rapidly and with limited costs, in order to choose the best solution for the particular domain of work.

References

[1] Kirwan B. A guide to practical human reliability assessment. London: Taylor & Francis; 1994.
[2] Cacciabue PC. Modelling and simulation of human behaviour for safety analysis and control of complex systems. Saf Sci 1998;28:97–110.
[3] Reason J. Human factors: a personal perspective. In: Human factors seminar, Helsinki, Finland; February 13, 2006.
[4] Griffith CD, Mahadevan S. Inclusion of fatigue effects in Human Reliability Analysis. Reliab Eng Syst Saf 2011;96(11):1437–47.
[5] Madonna M, Martella G, Monica L, Pichini Maini E, Tomassini L. Il fattore umano nella valutazione dei rischi: confronto metodologico fra le tecniche per l'analisi dell'affidabilità umana. Prev Oggi 2009;5:67–83.
[6] De Felice F, Petrillo A, Carlomusto A, Romano U. A review of human reliability analysis: management practices and techniques. Glob J Manag Sci Technol 2012;1(4):131–47.
[7] Di Pasquale V, Iannone R, Miranda S, Riemma S. An overview of human reliability analysis techniques in manufacturing operations. In: Massimiliano Schiraldi, editor. Operations management. InTech-Open Access Publisher; 2013. p. 221–40. http://dx.doi.org/10.5772/55065.
[8] Iannone R., Miranda S., Riemma S. Proposta di un modello simulativo per la determinazione automatica delle pause di lavoro in attività manifatturiere a prevalente contenuto manuale. In: Atti del XXXI Convegno Nazionale Animp, Treviso; October 14–15, 2004, p. 46–60.
[9] Hollnagel E. Cognitive reliability and error analysis method (CREAM). Amsterdam: Elsevier; 1998.
[10] Boring RL. Dynamic Human Reliability Analysis: benefits and challenges of simulating human performance. Risk Reliab Soc Saf 2007;2:1043–9.
[11] Trucco P, Leva MC. A probabilistic cognitive simulator for HRA studies (PROCOS). Reliab Eng Syst Saf 2007;92(8):1117–30.
[12] Leva MC, De Ambroggi M, Grippa D, De Garis R, Trucco P, et al. Quantitative analysis of ATM safety issues using retrospective accident data: The Dynamic Risk Modelling Project. Saf Sci 2009;47:250–64.
[13] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 1: overview of the IDAC model. Reliab Eng Syst Saf 2006;92:997–1013.
[14] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 2: IDAC performance influencing factors model. Reliab Eng Syst Saf 2007;92:1014–40.
[15] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 3: IDAC operator response model. Reliab Eng Syst Saf 2006;92:1041–60.
[16] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 4: IDAC causal model of operator problem-solving response. Reliab Eng Syst Saf 2006;92:1061–75.
[17] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents. Part 5: dynamic probabilistic simulation of IDAC model. Reliab Eng Syst Saf 2006;92:1076–101.
[18] Mosleh A, Chang YH. Model-based Human Reliability Analysis: prospects and requirements. Reliab Eng Syst Saf 2004;83:241–53.
[19] Kirwan B. The validation of three human reliability quantification techniques – THERP, HEART and JHEDI: Part 1: technique descriptions and validation issues. Appl Ergon 1996;27(6):359–73.
[20] Calixto E. Comparing SLIM, SPAR-H and Bayesian network methodologies. Open J Saf Sci Technol 2013;3:31–41.
[21] Kirwan B. The validation of three human reliability quantification techniques – THERP, HEART, and JHEDI: Part 3: practical aspects of the usage of the techniques. Appl Ergon 1997;28(1):27–39.
[22] Bell J, Holroyd J. Review of human reliability assessment methods. London, UK: Health and Safety Executive Research Report RR679; 2009.
[23] Cacciabue PC. Human error risk management for engineering systems: a methodology for design, safety assessment, accident investigation and training. Reliab Eng Syst Saf 2004;83(2):229–40.

[24] Kim MC, Seong PH, Hollnagel E. A probabilistic approach for determining the control mode in CREAM. Reliab Eng Syst Saf 2006;91(2):191–9.
[25] Schemeleva K, Nguyen C, Duriex S, Caux C. Human error probability computation for manufacturing system simulation using CREAM. In: Proceedings of the 9th international conference of modeling, optimization and simulation, Bordeaux, France; June 6–8, 2012.
[26] Le Y, Qiang S, Liangfa S. A novel method of analyzing quality defects due to human errors in engine assembly line. In: Proceedings of the international conference on information management, innovation management and industrial engineering; October 2012. p. 154–7.
[27] Yang Z, Wang J, Rochdi M, Belkacem O. Bayesian modelling for human error probability analysis in CREAM. In: Proceedings of the international conference on quality, reliability, risk, maintenance, and safety engineering, Xi'an, China; June 17–19, 2011.
[28] Konstandinidou M, Nivolianitou Z, Kiranoudis C, Markatos N. A fuzzy modeling application of CREAM methodology for Human Reliability Analysis. Reliab Eng Syst Saf 2006;91(6):706–16.
[29] Galyean WJ, Whaley AM, Kelly DL, Boring RL. SPAR-H step-by-step guidance. Idaho National Laboratory (United States). Funding organisation: US Department of Energy (United States); 2011.
[30] Blackman HS, Gertman DI, Boring RL. Human error quantification using performance shaping factors in the SPAR-H method. In: Proceedings of the human factors and ergonomics society annual meeting. SAGE Publications; 2008: 52(21).
[31] Boring RL, Blackman HS. The origins of the SPAR-H method's performance shaping factor multipliers. In: 13th annual meeting on human factors and power plants and HPRCT. , Sage Publications; August, 2007. p. 177–84.
[32] Gould KS, Ringstad AJ,Van de Merwe K. Human Reliability Analysis in major accident risk analyses in the Norwegian petroleum industry. In: Proceedings of the human factors and ergonomics society annual meeting; 2012.
[33] Van de Merwe K, Øie S, Gould KS. The application of the SPAR-H method in managed-pressure drilling operations. In: Proceedings of the human factors and ergonomics society annual meeting; 2012.
[34] Boring RL. How many performance shaping factors are necessary for Human Reliability Analysis. In: Proceedings of the 10th international probabilistic safety assessment and management conference (PSAM10), Seattle, US; 2010.
[35] Gertman D, Blackman H, Marble J, Byers J, Smith C. The SPAR-H Human Reliability Analysis method. Washington, DC: US: Nuclear Regulatory Commission; 2005 [NUREG/CR-6883].
[36] Droguett EL, Menezes R. Análise da confiabilidade humana via redes Bayesianas: uma aplicação à manutenção de linhas de transmissão. Produção 2007;17(1):162–85.
[37] Droguett EL, Menezes R. Uma metodologia para analise de confiabilidade de sistemas complexos. In: Proceedings of the XXXVIII Simposio brasileiro de Pesquisa operacional; 2006.
[38] Droguett EL, Menezes R, Firmino P. Analise de confiabilidade humana via redes bayesianas. In: Proceedings of the ENEGEP 2005 XXV Encontro Nacional de Engenharia de Produção; 2005.
[39] Bye A, Laumann K, Braarud PØ, Massaiu S. Methodology for improving HRA by simulator studies (PSAM-0391). In: Proceedings of the eighth international conference on probabilistic safety assessment and management (PSAM). ASME Press; 2006.
[41] Boring RL. Modeling Human Reliability Analysis using MIDAS. In: In: Proceedings of the fifth international topical meeting on nuclear plant instrumentation, controls, and human machine interface technology; 2006. p. 1270–4.
[40] Hollnagel E. Reliability analysis and operator modelling. Reliab Eng Syst Saf 1996;52(3):327–37.
[42] Cacciabue PC, Decortis F, Drozdowicz B, Masson M, Nordvik JP. COSIMO: a cognitive simulation model of human decision making and behaviour in accident management of complex plants. IEEE Trans Syst Man Cybern 1992;22(5):1058–74.
[43] Gore BF, Jarvis PA. New integrated modeling capabilities: MIDAS'recent Behavioral Enhancements. SAE Technical Paper; 2005.
[44] Kirwan B. Human error identification techniques for risk assessment of high risk system – part 1: review and evaluation of techniques. Appl Ergon 1998;29(3):157–77.
[45] Boring R, Griffith C, Joe JC. The measure of human error: direct and indirect performance shaping factors. In: Proceedings of the joint 8th annual IEEE conference on human factors and power plants and 13th annual workshop on human performance/root cause/trending/operating experience/self assessment, Monterey, CA; August 26–31, 2007.
[46] Woods DD, Roth EM, People HE. Cognitive environment simulation: an artificial intelligence system for human performance assessment. Washington DC: US: US Regulatory Commission; 1987 [NUREG-CR-4862].
[47] Sasou K, Takano K, Yoshimura S, Haroko K, Kitamura M. Modelling and simulation of operator team behaviour in nuclear power plants. In: Proceedings of the HCI international, Tokyo; 1995.