How to learn from the resilience of Human–Machine Systems?

Kiswendsida Abel Ouedraogo, Simon Enjalbert, Frédéric Vanderhaegen

How to Learn from the Resilience of Human-Machine Systems?

Kiswendsida Abel Ouedraogo [1,2,3], Simon Enjalbert [1,2,3], Frédéric Vanderhaegen [1,2,3]

[1]Univ Lille Nord de France, F-59000 Lille, France
[2]UVHC, LAMIH, F-59313 Valenciennes, France
[3]CNRS, FRE 3304, F-59313 Valenciennes, France

{kiswendsidaabel.ouedraogo, simon.enjalbert, frederic.vanderhaegen}@univ-valenciennes.fr

Abstract: This paper proposes a functional architecture to learn from resilience. First, it defines the concept of resilience applied to Human-Machine System (HMS) in terms of safety management for perturbations and proposes some indicators to assess this resilience. Local and global indicators for evaluating human-machine resilience are used for several criteria. A multi-criteria resilience approach is then developed in order to monitor the evolution of local and global resilience. The resilience indicators are the possible inputs of a learning system that is capable of producing several outputs, such as predictions of the possible evolutions of the system's resilience and possible alternatives for human operators to control resilience. Our system has a feedback-feedforward architecture and is capable of learning from the resilience indicators. A practical example is explained in detail to illustrate the feasibility of such prediction.

Keywords: Human-Machine Systems, Resilience, Learning process, Feedback/Feedforward control

1. INTRODUCTION

Resilience is related to the ability of a material to recover from a shock or disturbance. Resilience is a relatively new field of research, although the concept has been used in physics for Charpy impact tests throughout nearly all of the XXth century. The concept of resilience has also been developed in the field of ecology and is used to characterise natural systems that tend to maintain their integrity when subjected to disturbances (Ludwig *et al.*, 1997). It has generated a lot of interest in different scientific communities and has been applied to psychology, psychiatry (Goussé, 2005), sociology, economy, biology (Orwin and Wardle, 2004; Pérez-España and Sánchez, 2001), computer sciences (Chen *et al.*, 2007; Nakayama *et al.*, 2007; Luo and Yang, 2007), and automation (Tianfield and Unland, 2004; Neema *et al.*, 2004; Numanoglu *et al.*, 2006).

Psychological resilience is linked to the invulnerability theory (*i.e.*, the positive capacity of people to cope with trauma and to bounce back). Biological or ecological resilience is based on the theory of viability (*i.e.*, the ability for an organism to survive after disruption). The resilience of industrial systems is linked to on-line safety management, faced with known or unknown situations. It differs from the traditional off-line safety analysis process, which aims at foreseeing undesirable situations and proposing schemes to avoid their occurrence or protect the system from their consequences. From the organisational and safety management viewpoints, resilience is the capacity of a system to survive, adapt and face unforeseen changes, even catastrophic incidents. There are many formal definitions of resilience, but most of them suppose the existence of functional capacities in order to make a system resilient: the capacity to recognise, adapt to, and absorb changes.

When a Human-Machine System (HMS) doesn't have sufficient resources or competences to control such functions, it cannot be resilient, or its resilience may be managed by another HMS. Another strategy can be applied: learning to face new or unknown situations. HMS decision-makers have to make sense of these kinds of situations and identify alternatives to control them. When the management of these situations is successful, the HMS is resilient.

This paper focuses on the positive control of new, unknown, unexpected or surprising situations and on the possibility of learning from resilience. It proposes a functional architecture for learning from resilience indicators and their evolution. An example applied to a cockpit and its four-person flight crew illustrates the feasibility of such learning.

The rest of this paper is organised as follows. Section 2 focuses on the concept of resilience applied to HMS and the indicators for assessing it. Section 3 presents an original method to learn from other resilience indicators. Section 4 provides an example of learning from resilience. Section 5 offers our conclusions and prospects for future research.

2. PENDING ISSUES ABOUT RESILIENT HMS

One of the first substantive publications on resilience as applied to engineering was "Resilience Engineering: Concepts and Precepts" [Hollnagel *et al.*, 2006]. The basic concepts behind resilience engineering are developed, but at the present stage, resilience engineering has several fundamental problems: 1) there is no appropriate definition of resilience, and 2) the differences between resilience and other similar concepts (*e.g.*, robustness, reliability) are not clarified. These problems need to be addressed in order to advance resilience engineering and transform a theoretical concept into an applied science by defining a quantitative method that can measure system resilience.

2.1. Definition of HMS Resilience

Wreathall (2006) defined resilience as "the ability of an organisation (system) to keep, or recover quickly to, a stable state, allowing it to continue operations during and after a major mishap or in the presence of continuous significant stresses". As a resilience definition, Wreathall's definition lacks a distinction of resilience from robustness (Zhang and Lin, 2010; Wang *et al.*, 2010). Both terms are related to the ability of a system to keep functioning faced with disturbances.

Zhang and Lin (2010) further defined resilience as a system property about how the system can still function to the desired level when it suffers from partial damage. This definition was able to distinguish resilience from robustness: for a robust system, the physical structure of the system is still intact, whereas for resilient system, the physical structure is damaged (Gao, 2010). This definition sees resilience as a system's post-damage property (*i.e.*, the system's ability to recover its functions faced with damage). In essence, a resilient system contains characteristics of a robust system in that it is the magnitude of the disturbance that differentiates resilient system from robust one.

We aim to apply the resilience concept to HMS, with human operators as an unpredictable source of both reliability and errors. We distinguish a robust system from resilient one, based on the nature or typology of the threats/disturbances, as defined by Westrum (2006): robust systems deal with regular and, at a certain level, irregular threats, whereas resilient systems manage unknown situations (*e.g.*, unexpected or unprecedented disturbances).

Assuming that optimal performance level exists (*i.e.*, an initial nominal HMS state or a baseline), after any disturbance, either internal disturbance (*e.g.*, human errors or technical failures) or external disturbances (*e.g.*, environmental events), the HMS performance may be degraded. Several scenarios can be envisioned:

- If the HMS is capable of returning to the initial nominal performance (*i.e.*, known disturbance situations), the system can be defined as resistant;
- If the HMS is capable of recovering from a disturbance and stabilizing at another "acceptable" performance level, which is an unoptimal performance due to controlling unknown situations (*e.g.*, unexpected or unprecedented disturbances), the system can be defined as resilient;
- If the HMS is not capable of recovering from a disturbance (*i.e.*, not an acceptable performance) or stabilizing itself, the system is neither resistant nor resilient.

Human operators and machines in the HMS cooperate to ensure an optimal operation, and they are potentially available resources to make the HMS resilient. These resources need particular capacities, and some methods exist to make a system resilient. One of these methods is related to the learning process.

2.2. Resilience and methods

In order to be resilient, a system or an organisation required the following four qualities (Steen and Aven, 2011):

- The ability to anticipate risk events;
- The ability to monitor what is going on, including its own performance;
- The ability to respond to unplanned events (regular, irregular or unprecedented) in a robust or flexible manner; and
- The ability to learn from experience.

Resilient systems are supposed to adapt to unplanned events with their ability to anticipate failures, to control disturbances, to react and to recover from these events (Figure 1). The system also has the possibility of learning from its reactions to unplanned events (*i.e.*, successes and failures). Thus, the design of resilient systems can be based on some principles, such as the five principles defined by Zhang and Lin (2010), which mainly highlight the need of:

- A certain degree of functional redundancy,
- A controller for redundancy and learning management,
- A sensor for monitoring the whole system's performance,
- A predictor for predicting potential threats or analysing potential vulnerabilities of the system, and
- An "actuator" for implementing changes or training.

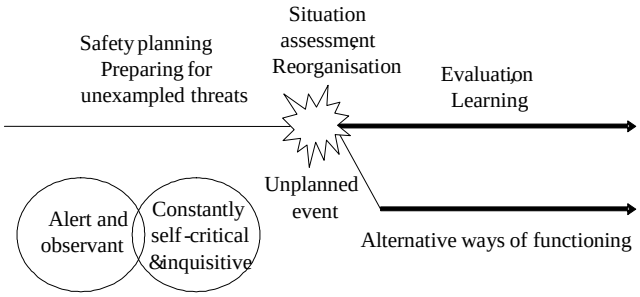

Fig. 1. Resilient organisation (adapted from Hollnagel, 2006).

A non-resilient system cannot continue to operate correctly after a major mishap or in presence of continuous stress. Many methods or mechanisms can be used to recover from such situations, such as:

- Using non-affected elements to compensate and accomplish the functions of the affected or degraded parts (Chen *et al.*, 2007; Nakayama *et al.*, 2007; Numanoglu *et al.*, 2006), with what works compensating for what does not work;
- Maintaining the system between the minimum and maximum thresholds of acceptability or perturbation management instead of a stable point or value (Martin, 2005);
- Putting critical elements in redundancy (Luo and Yang, 2007), with the affected elements no longer being solicited and being replaced by redundant components;
- Optimizing the mitigation of the perturbations when the redundancy strategy is too expensive (Neema *et al.*, 2004; Tianfield and Unland, 2004), thus developing fault resilient systems;
- Applying the principles of cooperation (Vanderhaegen, 1997, 1999a; Hsieh, 2009; Zieba *et al.*, 2009), with cooperation between humans and/or artificial agents facilitating the problem-solving for new situations;
- Developing systems or databases based on feedback in order to copy successful practices and learn from failed practices (Vanderhaegen, 2010a);
- Activating the required resources by managing performance evolution and decision-maker autonomy (Zieba *et al.*, 2010, 2011), which can lead to increasing the global system capacity using learning;
- The affected elements learn or re-learn how to work correctly or to work better (Cheveau and Wybo, 2007), with what does not work being reset or rebooted and prepared for future operations;
- Learning vital functions for humans (*i.e.*, the ability to develop attitudes or behaviours ensuring the survival of the human organism) (Marcantoni, 2009).

The last four items on the list concerns the learning processes to make a HMS resilient. The 5 principles mentioned above for resilient system design (Zhang and Lin, 2010) also concern learning management and training. Therefore, the system's learning capacities are thus important features and have to be developed. This is a long-term, variable and dynamic process. It emphasizes the need of HMS to continuously improve their learning capacities.

2.3. Resilience and Learning

Human-machine systems regularly try to anticipate and resist disturbances but may be vulnerable to critical or unexpected disturbances. Therefore, HMS have to manage their knowledge dynamically in order to overcome problems and improve their learning capacities. Vanderhaegen (2010a) has proposed a behavioural model to react and learn from the successful or failed control of known and unknown situations. The model is based on three main activities: prognosis, diagnosis and trial-and-error reasoning (Figure 2).

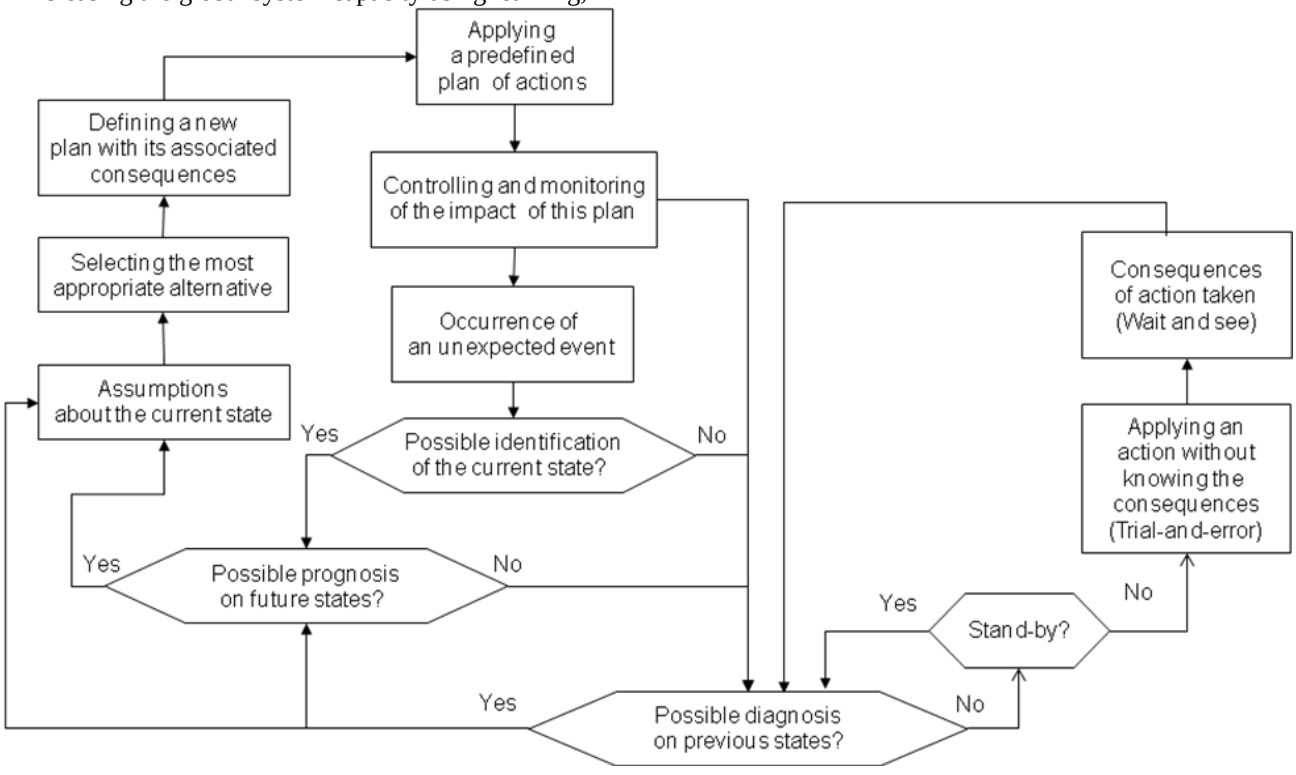

Fig. 2. Behavioural model in response to unexpected events (adapted from Vanderhaegen, 2010b).

The prognosis function leads to the identification of the possible evolutions of the system state, with or without actions. The diagnosis function relates to the explanation of the current system state in terms of the previous states. When this identification is not possible due to unknown system state, trial-and-error reasoning is needed in order to apply an action and to wait and see its consequences. This trial-and-error process tries to understand the current system state and to propose a new adapted action plan.

The management of known or unknown situations or threats requires different "demands" on a resilient system (Westrum, 2006; Hale and Heijer, 2006). For instance, the system can develop a standard response to known situations. However, the unknown situations are more challenging: unprecedented threats that cannot be anticipated may push the HMS outside of its experience. Thus, HMS resilience relates to the system ability to adapt its responses capacities in order to circumvent threats.

Human operators are able to gain experience and to learn by repeating their control tasks, thus improving their behaviour. They can adjust themselves according to the dynamic changes of the HMS they control. This requires adaptive and proactive behaviour (*i.e.*, resilient behaviour) to control the system performances, especially faced with unexpected situations. Resilience is a dynamic process. It is not a static system state, and continuous verifications are necessary in order to qualify the HMS as resilient (Hale and Heijer, 2006). Assessing HMS resilience is required in order to:

- Categorize and compare systems in terms of their resilience characteristics,
- Evaluate the impact of upcoming learning processes in terms of their resilience, and
- Evaluate the possible evolution of HMS resilience.

2.4. Resilience and Measurements

The Charpy impact test gives a first measurement of resilience, related to the ability of a system to recover from a shock. Hollnagel and Woods (2006) argue that resilience itself cannot be measured, but the potential for resilience can be measured. There is still some debate regarding the definition of resilience and its difference with other similar concepts (*e.g.*, robustness, reliability). Since the definition of resilience in the literature is vague or conceptual, its quantification may still be under development.

In order to "engineer" resilience, some objective and quantitative indicators are needed. Objective indicators are based on normative interpretation of the data, and quantitative indicators are measurable metrics that are used to identify when a system's performance changes (Wreathall, 2011). Thus, these indicators can be used to judge whether the system's resilience levels are acceptable.

There is currently no measurement for assessing a HMS's resilience. Therefore, this paper proposes a quantitative resilience indicator. In other domains, several resilience indicators based on the evolution of performance have been proposed in the literature. In order to present these indicators, Figure 3 presents an example of changes in system safety when faced with a disturbance. The baseline in this figure represents totally safe conditions, and the minimum acceptable threshold indicates an acceptable safety level for designers. E_{max} is the maximum amplitude of the disturbance's effect on safety, and E_j is the amplitude of the disturbance's effect on safety at time T_j.

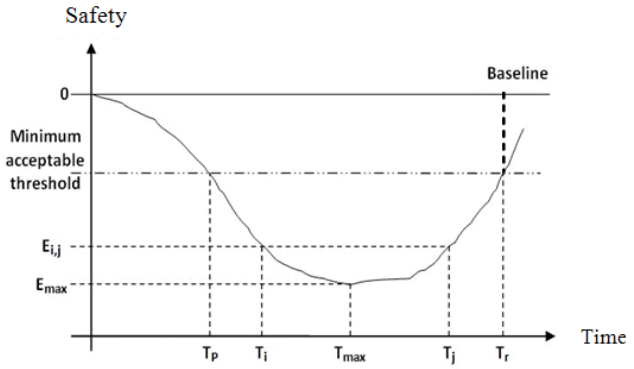

Fig. 3. Measured resilience in the literature.

Resilience has been evaluated by estimating the maximum intensity of an absorbable force (E_{max}) without perturbing the system's functions. In the ecological domain, Orwin and Wardle (2004) have linked resilience with the measurement of the instantaneous and maximal disturbance:

$$resilience_{T_j} = \left(\frac{2 \times |E_{\max}|}{|E_{\max}| + |E_j|} \right) - 1 \qquad (1)$$

The instantaneous resilience varied between 0 and +1, where the value +1 corresponds to the maximal resilience when the disturbance's effects were recovered ($E_j=0$). This indicator does not consider the recovering time: two systems with different recovering time Δt_1 and Δt_2 cannot have the same resilience. For Luo and Yang (2002), the measurement of the resilience is linked with the speed of recovery from a disturbance:

$$resilience = T_r - T_p \qquad (2)$$

This computer communications indicator does not consider the disturbance effect on the system. However, if a system can handle a large number of disturbances, this system may be more resilient.

Pérez-España and Arreguín-Sánchez (2001) have calculated the resilience of ecological system by the opposite of the tangent of the ratio between the resistance and the recovery time of a disturbance:

$$resilience = \tan^{-1}\left(\frac{1}{\frac{|E_{\max}|}{T_r - T_p}}\right) \quad (3)$$

The instantaneous resilience varied between 0° and 90°, where the value 90° corresponds to the maximal resilience level. Their resilience indicator is performed within an interval (*i.e.*, during the disturbance effect on the system) and does not take the general evolution of the system into account. For an application to the Human-Machine Systems, there is a lack of such instantaneous resilience indicators. All the proposed indicators mentioned above do not consider simultaneously disturbance period, effect, and recovery speed. All propose an instantaneous value without considering their possible evolutions.

Wang *et al.* (2010) proposed an indicator for a company information system's resilience based on the maximum recovery ability of the system. For a partially damaged company information system, the resilience is defined as:

$$resilience = \max \sum_{i=1}^{m} z_i \frac{d_i}{c_i} \quad (4)$$

where (d_i) represents the demand time for the recovery of function i ($i=1,\ldots,m$), m is the number of function of the system, and (c_i) is the completion time. (z_i) is the weight of the function i, which represents the importance of this function based on the particular features of the system among all functions and is subject to the following constraint:

$$\sum_{i=1}^{m} z_i = 1 \quad (5)$$

When all functions can be recovered within the demand time, the resilience value will be larger than 1. The larger the resilience value, the more resilient a given system.

The indicator of Wang *et al.* (2010) is limited in that it considers the resource reallocation for different recovery solutions, which supposes the number of functions and the number of recovery solutions are known. However, HMS are supposed be dealing with unknown situations, so their indicator is not appropriate. Other indicators are thus required for studying HMS resilience.

3. HOW TO LEARN FROM RESILIENCE

For studying HMS, several other authors have proposed resilience indicators. They assess resilience not only for the safety impact but also for other relevant system criteria. They are used as inputs or outputs for a functional architecture of a system able to learn from such indicators.

3.1. Resilience indicators for the learning process

Assessing HMS resilience requires evaluating two classes of indicators:

- The performance stability indicator on a given time interval (*i.e.*, the time period during which the performance improves or stays the same), and
- The HMS performance indicators related to the consequences of human actions in order to compare performance levels between two dates (*i.e.*, the current one and a past one from, for example, a sampling period).

In order to assess the system's resilience, the system safety factor *S(t)* is determined by the cumulative effects of the possible factors that can affect it (*e.g.*, speed, braking distance, driver awareness) (Gu *et al.*, 2009). For instance, resilience can be assessed between times t_b and t_e (Figure 4). The time t_b is the beginning of disturbance effect (*i.e.*, the safety indicator is below a minimum acceptable threshold), and the time t_e is the result of the recovering process (*i.e.*, the end of the unacceptable performance).

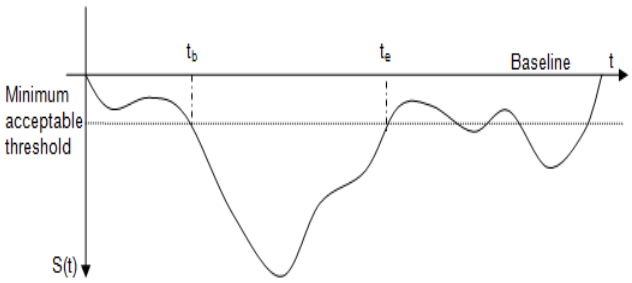

Fig. 4. Evolution of safety indicator for assessing resilience.

Based on this safety indicator, Enjalbert *et al.* (2009) have proposed a local resilience evaluation:

$$local\ resilience = \frac{dS(t)}{dt} \quad (6)$$

The local resilience is an instantaneous measurement of resilience. Its value depends on the effect of disturbance on the system: it can be negative if the performance decreases or positive if the system recovers from the disturbance. The global resilience is the integral of local resilience over a period of time (Enjalbert *et al.*, 2009):

$$global\quad resilience = \int_{tb}^{te} local\ resilience = \int_{tb}^{te} \frac{dS(t)}{dt} \quad (7)$$

Table 1 presents the possible evaluations that can be obtained for the times t_i (during safety performance decrease), t_{max} (maximum effect of disturbance) and t_j (during safety performance recovery). The time t_p is the beginning of disturbance effect (*i.e.*, the safety indicator is below a minimum acceptable threshold), and the time t_r is the result of the recovering process (*i.e.*, the end of the unacceptable performance), as shown in Figure 3.

Table 1: Resilience evaluation classification

times	t_p	t_i	t_{max}	t_j	t_r
local resilience	$S'(t_p)$	$S'(t_i)$	0	$S'(t_j)$	$S'(t_r)$
global resilience	0	$\int_{tp}^{ti} S'(t)$	$\int_{tp}^{t\max} S'(t)$	$\int_{tp}^{tj} S'(t)$	$\int_{tp}^{tr} S'(t)$

This evaluation (equations 6 and 7) is from a mono-criterion (*i.e.*, safety criterion) viewpoint. It is used to produce a multi-criteria viewpoint. These new criteria of system evaluation concern particular human or machine behaviours or their effects, or the occurrence or the consequences of external perturbations (*e.g.*, the instantaneous or anticipated human workload (Vanderhaegen, 1999b, 1999c), the number of human errors (Vanderhaegen *et al.*, 2004), and the quality or production of services (Polet *et al.*, 2009)). Thus, several performance criteria, related to the system safety, the human workload and the team mission are defined and described in detail in section 4.2.

Based on these criteria, by applying the equations 6 and 7, we can assess a multi-criteria resilience (*i.e.*, local and global) of a HMS:

- Local_resilience_on_safety is the local resilience evaluation based on the safety indicator,
- Global_resilience_on_safety is the global resilience evaluation based on the safety indicator,
- and so on for all considered criteria.

The evolution of the resilience can then be assessed recursively. If a perturbation is not recovered, the system is not resilient. Therefore, two recovery levels can be identified:

- Prevention recovery from the perturbation's occurrence, and
- Protection recovery from the consequences of this perturbation.

All these multi-criteria resilience indicators are the inputs or the outputs of our learning system. This system is then able to anticipate the evolution of system resilience in terms of several performance criteria or to propose alternatives to recover from perturbations: prevention and protection recovery processes.

3.2. The system architecture to learn from resilience

Our learning system, shown in Figure 5, uses a reinforced iterative formalism for sequential learning for HMS resilience:

- Iterative learning – learning built on prior knowledge to predict the future evolution of indicators of resilience;
- Reinforced learning – for a given iteration *i*, the correct prediction assessment consists of comparing the real resilience values ($O_{[i]}$) with the predicted one ($O^*_{[i]}$); the knowledge gain is integrated and synthesized in a reduced number of cases, allowing the correction or reinforcement of the global knowledge database;
- Sequential learning – learning built on a chronological sequence of events.

This system has a feedforward-feedback architecture and is able to learn from resilience indicators. These indicators are the possible inputs ($I_{[1,\ldots,i]}$) of our learning system, which is able to produce several outputs ($O^*_{[i+1,\ldots,n]}$) (*e.g.*, the prediction of the possible evolutions of the system resilience, the possible alternatives for human operators to control resilience).

The controllers or decision-makers (either automated or human) need feedback information about the actual state of the controlled process to satisfy their safety management objectives. Thus, Leveson (2004) has proposed an accident model STAMP (Systems-Theoretic Accident Model and Processes), where the system is a dynamic process that is continually adapting, based on feedback loops of information and control, to accomplish its goals and to react to changes in the system and its environment. In fact, an inadequate or missing feedback can lead the system into hazards and accidents.

In our structure, the feedback process recovers possible erroneous knowledge, refines knowledge or creates new knowledge (Vanderhaegen, 2010). The feedforward process assesses the possible future decisions in terms of the current system states and the management of the previous states. Thus, the feedforward-feedback mechanism uses current knowledge related to previous activities in order to calculate the future ones (Ouedraogo *et al.*, 2010a, 2010b). Therefore, at iteration *i*, the input vectors ($I_{[1,\ldots,i]}$) can contain a chronological sequence of resilience values (*e.g.*,

$I_{[1,\ldots,i]}$={*local*_resilience_on_*safety*(t_1,...t_i),
g*lobal*_resilience_on_*safety*(t_1,...t_i),
*local*_resilience_on_*workload*(t_1,...t_i),
g*lobal*_resilience_on_*workload*(t_1,...t_i),
*local*_resilience_on_*mission*(t_1,...t_i),
*global*_resilience_on_*mission*(t_1,...t_i)}).

This input vectors ($I_{[1,\ldots,i]}$) dimension is related to *i*, the number of iterations; so in every iteration, the input dimension is increased.

Then, the system can complete the sequence by predicting the other resilience values (*e.g.*, for $n>i$

$O^*_{[i+1,\ldots,n]}$={*local*_resilience_on_*safety*(t_{i+1},...t_n),
g*lobal*_resilience_on_*safety* (t_{i+1},...t_n),
*local*_resilience_on_*workload* (t_{i+1},...t_n),
g*lobal*_resilience_on_*workload*(t_{i+1},...t_n),
*local*_resilience_on_*mission*(t_{i+1},...t_n),
*global*_resilience_on_*mission*(t_{i+1},...t_n)}).

The output vectors $O^*_{[i+1,\ldots,n]}$ dimension is also related to *i*, so in every iteration, the output dimension is decreased.

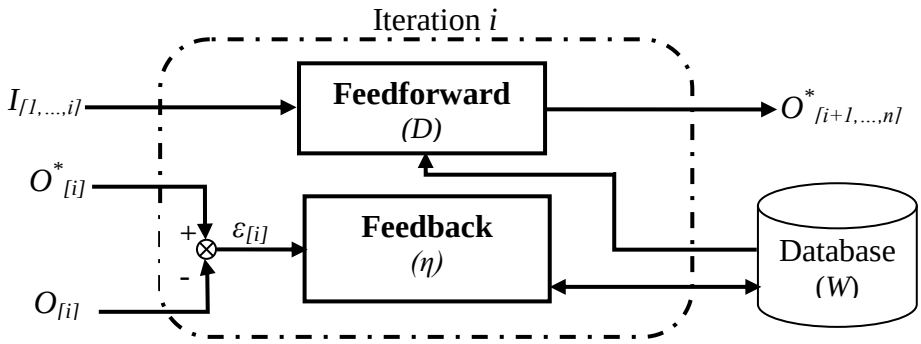

Fig. 5. Iterative and sequential reinforced learning structure.

As the criteria are related to the interactions between humans and the system, and to the system structure and it operation, our future research has the objective of predicting human operator actions in order to determine the appropriate alternatives of human operations, for example, in the decision-making system. So, at iteration *i*, the input vectors ($I_{[1,\ldots,i]}$) can contain a chronological sequence of resilience values and the system may predict the human operators' possible alternative actions plan as outputs ($O^*_{[i+1,\ldots,n]}$).

The prediction application is developed in C++ based on the Kohonen self-organization map. The interested reader can consult Zhang *et al.* (2004) and Vanderhaegen *et al.* (2009) for a more detailed discussion of the implementation. This system requires an euclidean distance (*D*) in equation 8 that identifies in the knowledge database (*W*), the vector ($w_{[1,\ldots,i]}$) which lies closest to the input vector ($I_{[1,\ldots,i]}$).

$$\forall\left(w_{[1,\ldots i]}\right)\in W$$

$$D(I_{[1,\ldots,i]},w_{[1,\ldots,i]}) = \left[\sum_{1}^{i}\left(I_{[1,\ldots,i]}{}^{T} - w_{[1,\ldots,i]}{}^{T}\right)^2\right]^{1/2} \quad (8)$$

The Feedforward part of the architecture may predict $O^*_{[i+1,\ldots,nj}$, which is based on the minimum of the euclidean distances: $min(D(I_{[1,\ldots,i]}, w_{[1,\ldots,i]}))$.

Through the Feedback part of the architecture, the database (*W*) is then incremented by values of $w_{[i]}$ at the *i*th iteration. Based on the current error ($\varepsilon_{[i]}$), which is the difference between the observed resilience values $O_{[i]}$ at the *i*th iteration and the corresponding predicted output values ($O^*_{[i]}$) obtained at the iteration *i-1*, database (*W*) content is reinforced thanks to Equation 9:

$$w_{[i]} = w^*_{[i]} + \eta\varepsilon_{[i]} \quad (9)$$

where $w^*_{[i]}$ is the predicted value of $O^*_{[i]}$ at the (*i-1*)th iteration and parameter η includes the learning rate and a neighbourhood threshold function.

4. APPLICATION OF OUR ARCHITECTURE TO A MILITARY AIR TRANSPORTATION SYSTEM

Our system architecture was validated with a military air transportation system, involving a simulated cockpit with a four-person flight crew.

4.1. Experimental Protocol

The experiments were performed with the In-flight Refuelling Group of the Istres air base (France). Military teams, working in small 4-person groups, are trained together and are brought to make many decisions in uncertain situations. Their activities are reproducible through a flight simulator: a BC-135 Boeing during an in-flight refuelling.

The experimental scenario was inspired by a real incident. Initially, smoke accompanied by a burning smell appears in the cabin. Then, a series of failures without apparent links occurs (*e.g.*, frost on the windows, loss of fuel indications, overheating transformer, smoke). The aircraft is over the ocean and cannot land. The problem is an electrical failure and is located on the specific area of a generator. Its fuse, which is less visible, has blown. In fact, all the failed components have the same origin, but expert opinion remains divided between two possible causes. Thus, the team has to face an ambiguous or uncertain situation.

Facing these repetitive failures, the team has to make sense of the situation in order to apply the correct procedures. They are not supposed to know the recovery rules, but they have all the manuals to identify the recovery rules. Several criteria are identified to assess such behaviours in terms of resilience.

4.2. Performance criteria for resilience assessment

Several criteria were defined in order to evaluate the general HMS evolution: criteria related to system safety, human workload, and the team mission. These criteria are the main factors that concern system performance under major mishap. The safety criterion relates to system performance level faced with failures. Initially, the system is supposed to be almost 100% safe. Multiple sequential disturbances (*e.g.*, frost on the windows, overheated transformer with smoke, loss of fuel indications) occur at time (T_j) may decrease this system safety level ($S_s(t)$),

depending on the effect of these disturbances. $E_{i,j}$=*α, β* and so on, where $E_{i,j}$ equals the effect E of a disturbance event *i* at time T_j, and *α & β* are the disturbance effect values (given in percentages).

Initially, the maximum safety level is given by:

$$S_s(t) = S_{s\_init}(t) = 100\% \quad (10)$$

For the occurrence of a single event, the maximum safety level is expressed as:

$$S_s(t) = S_s(t) - \begin{cases} \alpha \text{ if frosted on the windows} \\ \beta \text{ if overheating transformer} \\ etc \end{cases} \quad (11)$$

For the occurrence of multiple events, the maximum safety level is assessed as follows:

$$S_s(t) = S_s(t) - \alpha - \beta - \cdots \quad (12)$$

If the system safety level decreases below the minimum acceptable threshold (*e.g.*, $S_s(t)$ = 90%), the HMS will try to recover from the disturbance. As a result, the safety level increases, but a performance loss still remains. This recovery process from an event occurrence is assessed as follows:

$$S_s(t) = S_s(t) + \begin{cases} \alpha - \Delta\alpha \\ \beta - \Delta\beta \\ etc \end{cases} \quad (13)$$

where $\Delta\alpha$ or $\Delta\beta$ ... (*e.g.*, Δ=0,1) are the performance losses after a recovery from disturbance effect E= *α, β*, and so on.

The human workload criterion is linked to the number of interactions between the staff members (*i.e.*, communication frequency) and between the staff and the technical system (*e.g.*, standard procedures, applied actions). Initially, the team is free to do or not do anything. Their available workload capacity is maximal and given by:

$$S_w(t) = S_{w\_init}(t) = 100\% \quad (14)$$

In order to accomplish their mission, the team may have more, or less, work to do because of the disturbance occurrences and the team members may increase their workload. In fact, they are supposed to apply the standard procedures and actions, to increase their communication frequency in order to find the appropriate actions and to make and validate checklists, for example. These behaviours aim at overcoming the initial or current problems.

The available workload capacity level decreases according to the frequency of team communications, the standard procedures performed and the frequency of the applied actions. This workload criterion is then based on the number and demand of the communications, procedures and actions:

$$S_w(t) = S_w(t) - \xi \times ((\text{number of communications}) + (\text{number of procedures}) + (\text{number of actions})) \quad (15)$$

where ξ is a scaled factor; with ξ=*0,1*% based on the maximal available workload capacity (equation 14) and the number of interactions: when the number of interactions (communications, procedures, actions) reach 1000, the available workload criterion downs to 0%; *i.e.*, $S_w = S_{w\_init} - \xi$*(1000) = 100% – 0,1%*1000 = 0%.*

The mission criterion is the respect of the landing time and the aircraft's landing location. Without any disturbance, the team is supposed to respect the mission. The mission success can be evaluated in terms of the landing time and the landing location. Due to the occurrence of unexpected events, the team can make some changes in landing time or location. Some penalties may be applied depending on disturbance effects, in terms of landing delay and/or landing location:

$$S_m(t) = S_m(t) - (\text{disturbance\_effect\_landing\_time}) - (\text{disturbance\_effect\_landing\_location}) \quad (16)$$

Therefore, the mission criterion can be seen as the correlation between system safety (*e.g.*, change of the landing location to improve system safety) and its workload demand (*i.e.*, procedures and actions applied to overcome a problem). The evolution of these criteria and the associated local and global resilience can then be evaluated during the entire mission.

4.3. Evolution of the performance criteria

Initially, the system is supposed to be almost 100% safe. The team is free to do or not do anything. Their available workload capacity is maximal (*i.e.*, 100%), and the team is supposed to respect the given mission (*i.e.*, 100%). Faced with repetitive failures, the team has to make sense of the situation in order to overcome the problems. To evaluate the evolution of the given criteria, we consider a set of actions performed by the flight crew in response to each abnormal situation, thus contributing, at least theoretically, to the system's resilience.

The safety criteria depend on the occurrence of disturbances (equations 10-13): its amplitude, its occurrence time, the crew's recognition and the crew's recovery time. Table 2 presents the disturbance amplitudes defined based on the disturbance's criticity: from "*overheating transformer with smoke"* (30%) to *"frost on the windows" (10*%).

Table 2: Events effects on safety criteria for Team 1

Time (T_i)	**Disturbance Events** (*i*)	**Effect (%)** (*α, β, …*)
00:06:02	*Frosted on the windows*	10
00:11:50	*Overheating transformer with smoke*	30
00:24:24	*Loss of fuel indications*	10
00:28:52	*Problem with breakers (T1&T2)*	15
00:51:00	*Problem with "boost pomp"*	15
01:01:40	*Problem with breaker in phase C*	20

The available workload (equations 14 & 15) depends on the standard procedures performed, and in cases of a disturbance occurrences, the number of interactions (*e.g.*, communications, procedures, actions) between operators and the system. The mission criteria (equation 16) depend on the actions and decisions affecting the landing time or location, such as securing the flight by re-routing to a location X, help from air traffic controller (ATC), queuing analysis, and/or changing flight plan. The disturbances listed in Table 2 also affected the mission criteria at certain level. All these "disturbances" are classified according to their effects on mission criteria, from *"re-routing to a location X"* (30%) to *"rapid test"* (5%).

4.4. Evolution of the local and global resilience

If the situation is not understood and not recovered during the landing, the situation can turn into a disaster (*i.e.*, a crash). In cases when this disaster occurs, all the values of the criteria (safety, mission and workload) would be equal to 0. Based on the definitions of the criteria in the previous section, Figure 6 gives an example of the evolution of the safety, mission and workload criteria for a team. Initially, the system is supposed to be almost 100% safe, the team are supposed to respect the given mission (100%), and there is no workload demand.

Due to the perturbations, the system safety level decreases so the team has to communicate in order to perform actions and/or procedures to keep the criteria higher than the minimum acceptable threshold. Therefore, the available workload level decreases too, and the team may not be able to respect the initial mission in terms of landing time. With more and more perturbations, the main criteria, and thus the safety level, decrease below the minimum acceptable threshold, defined by team in terms of the safety level: 90%. The team changes the landing location to improve the system's safety level, thus increasing the team members' workload, to overcome the problem. By the end of the mission, the team manages to have a safe landing, thus increasing the safety level and available workload to 100%, but they changed the landing location so the mission cannot attain 100%.

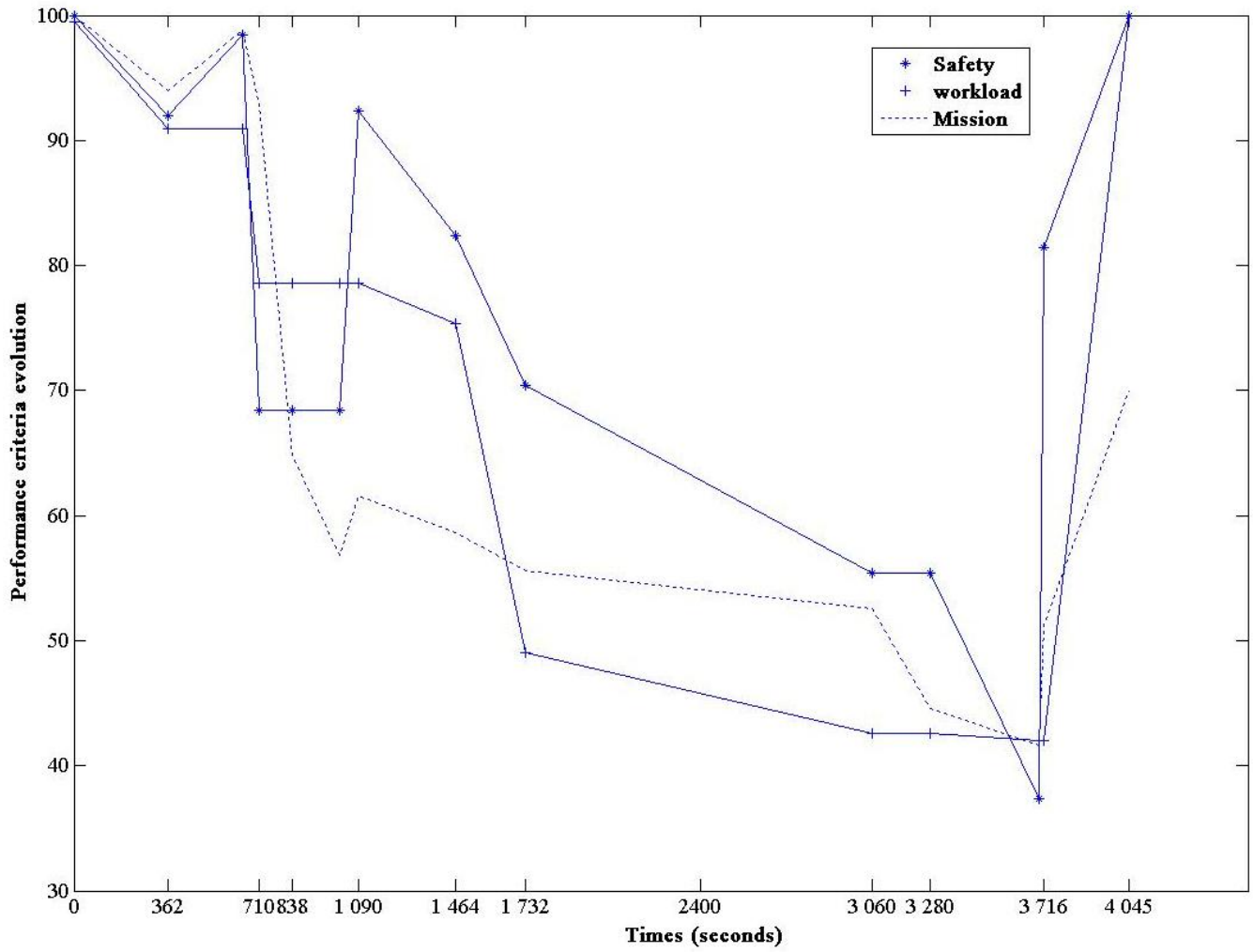

Fig. 6. Evolution of performance criteria (mission, safety and workload) for a team.

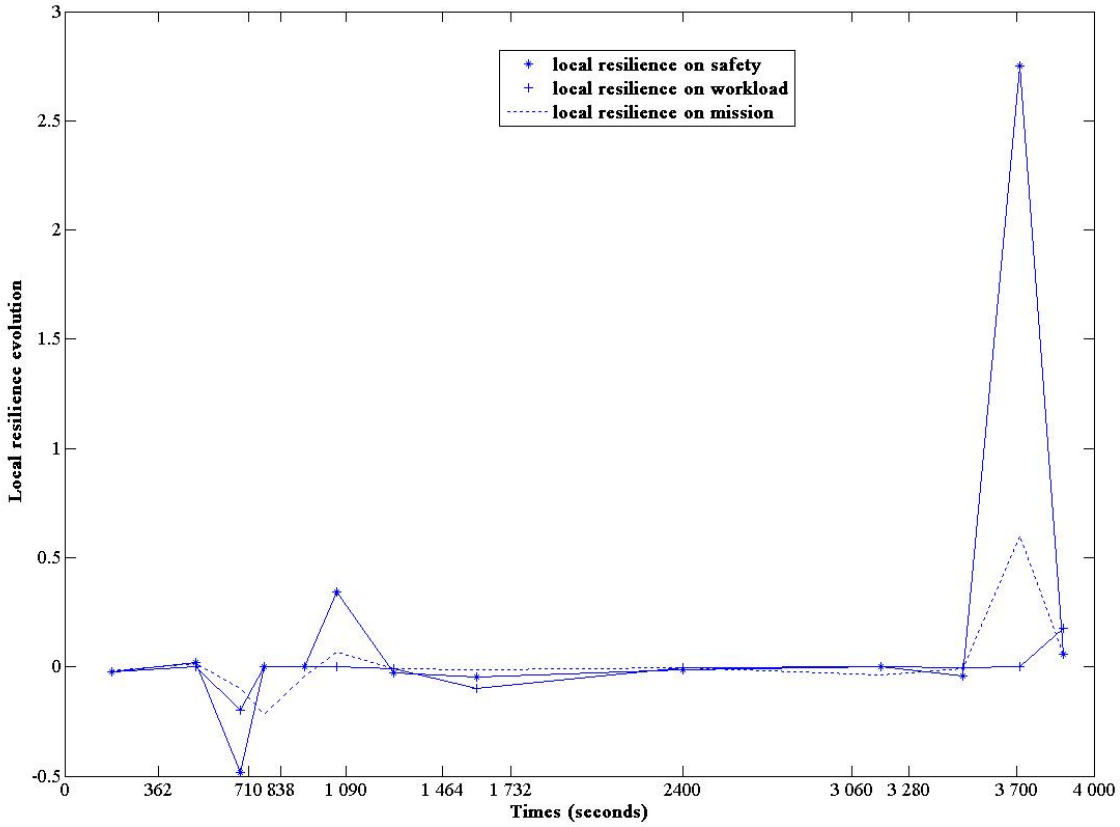

Fig. 7. Evolution of the local resilience for a team.

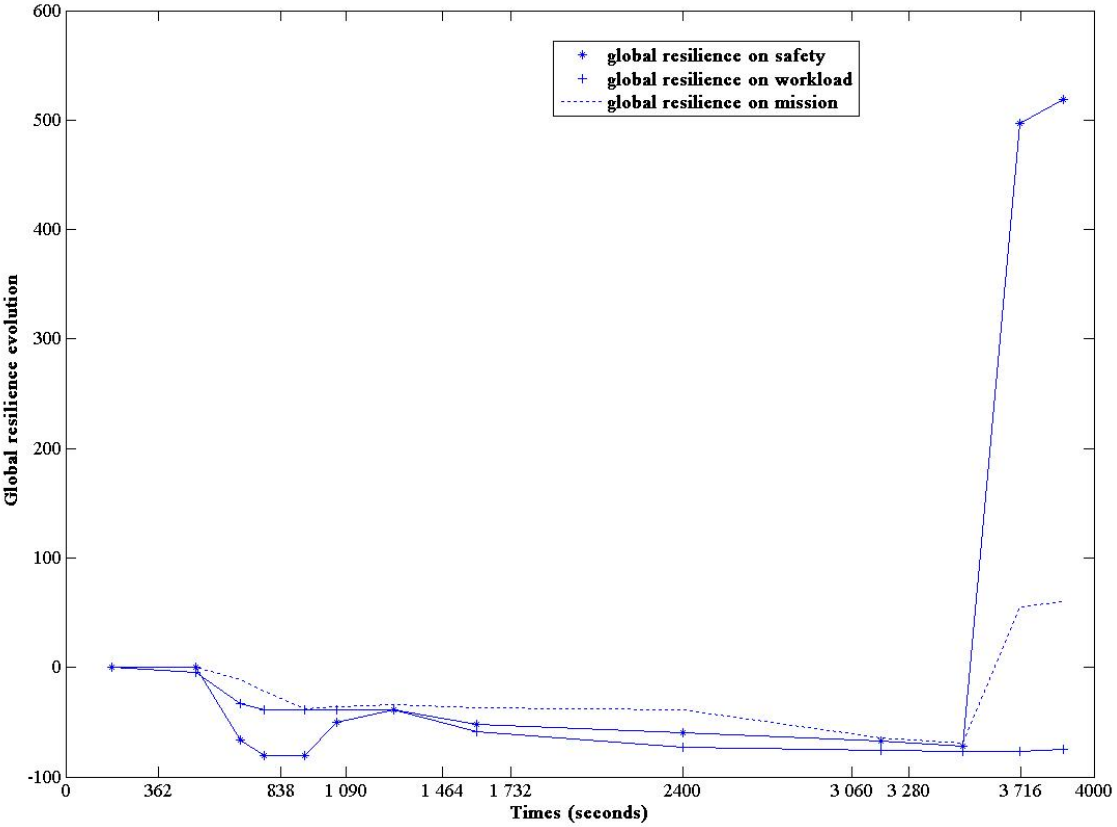

Fig. 8. Evolution of the global resilience for a team.

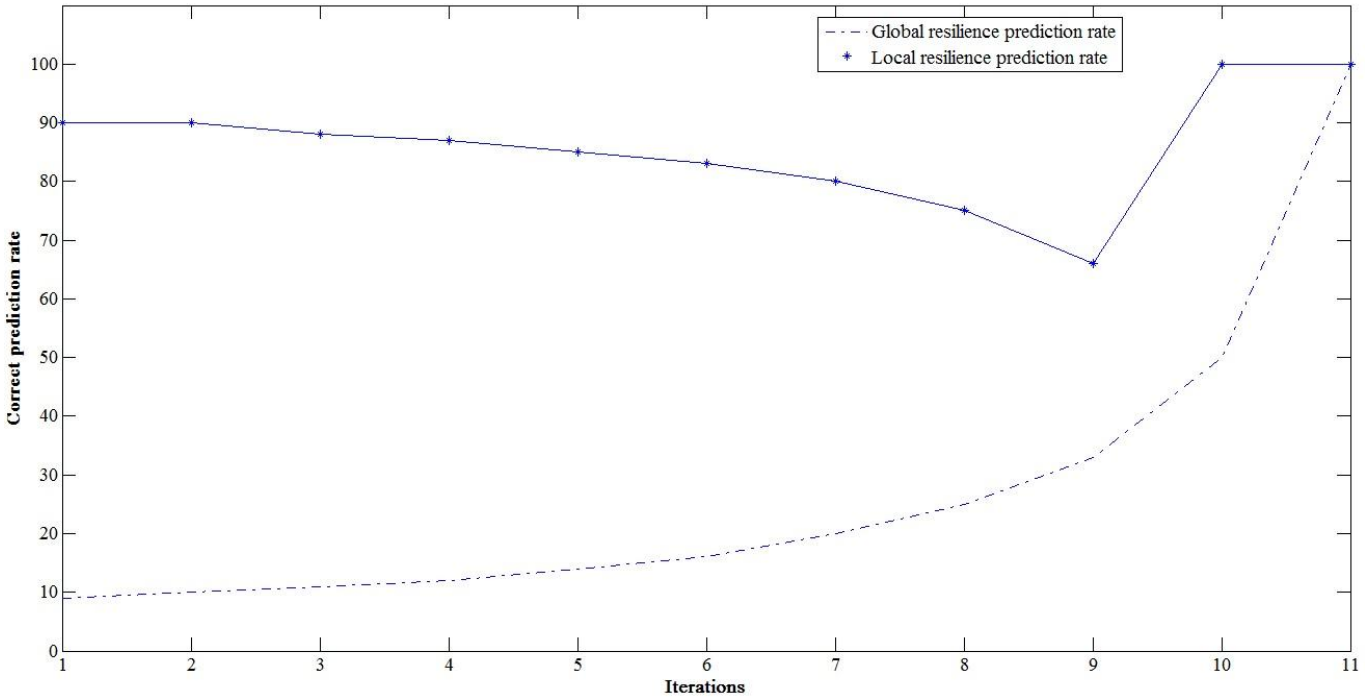

Fig. 9. Results for the local and the global resilience prediction rate.

14 specific times, corresponding to the measurable consequences of perturbation on the Human-Machine System, have been selected. Then, based on the local and the global resilience equations (equations 6 & 7), it can be observed that 13 iterations are needed to complete the evolution of the resilience values.

In Figure 7, the local resilience ($S'(t_i)$) is the derivative value of a criterion (safety, workload or mission) at specific time (t_i), and the global resilience is the integral of local resilience over a period of time or interval, as defined in section 3.1. Thus, we can have a look at the effect of the perturbation when it occurs: the local resilience is negative if the system performance decreases or positive if the system recovers from the disturbance. The capacity of the system to handle perturbation over the time can also be considered: the lower the system's global resilience, the more this system is resilient (Figure 8).

4.5. Results of the learning from resilience indicators

In this simulation, the input vectors ($I_{[1,\ldots,i]}$) contain a partial chronological sequence of resilience values:

$I_{[1,\ldots,i]}$={*local*_resilience_on_*safety*(t_1,...t_i), *global*_resilience_on_*safety*(t_1,...t_i), l*ocal*_resilience_on_*workload*(t_1,...t_i), *global*_resilience_on_*workload*(t_1,...t_i), *local*_resilience_on_*mission*(t_1,...t_i), *global*_resilience_on_*mission*(t_1,...t_i)}.

For initialization, database has been constituted with the first two iterations, and so 11 remaining iterations are needed in order to complete the n=13 iterations of studied experiment. The system complete the sequence by predicting the other resilience values,

- the local resilience prediction, $O^*_{[i+1,\ldots,n]}$={local_resilience_on_safety(t_{i+1},...t_n), local_resilience_on_mission(t_{i+1},...t_n), local_resilience_on_workload(t_{i+1},...t_n)}.
- the global resilience prediction, $O^*_{[i+1,\ldots,n]}$={global_resilience_on_safety(t_{i+1},...t_n) global_resilience_on_mission(t_{i+1},...t_n), global_resilience_on_workload(t_{i+1},...t_n)}.

Figure 9 gives the results related to the prediction quality of our architecture. The prediction rate is a comparison between the real resilience indicator values and the predicted ones.

Both local and global resilience prediction rates converge toward almost 100% after 11 iterations because database stores information and increases in every iteration. Indeed, correct prediction is easier to determine as the number of iteration to predict decreases. However, through all 11 iterations, the local resilience prediction rate stay very good around 90% because values from the calculated indicator do not change very much whereas the global resilience prediction rate may be more challenging because variation of the indicator based on the experiment time are important. Authors should consider a way to integrate periodic measure between iterations in further development of the applied architecture to avoid such a disparity in results.

5. CONCLUSION

The concept of resilience was defined and applied to HMS in terms of managing the system's safety during perturbations. Some series of possible learning behaviours in order to improve the system resilience were proposed. The existing indicators to assess this resilience were also studied. Some valuable indicators to assess the resilience of HMS were defined, and a sequential, iterative and reinforced system able to learn from the temporal evolution of such indicators was proposed. Our system architecture includes a feedforward process to predict the evolution of the resilience indicators and a feedback process to refine the system knowledge, taking into account the inputs and the outputs of the previous iterations. Its implementation is based on the Kohonen model.

A practical example for military air transportation is detailed to illustrate the feasibility of such a system. For resilience assessment, a multi-criteria (*i.e.*, safety, mission and workload) resilience approach was developed in order to monitor the local and the global resilience evolution of the HMS. For predicting the resilience evolution, our feedforward-feedback architecture is evaluated in term of prediction quality.

Future research will refine the definition of the criteria, for example, the workload criterion could take into account the duration of the procedures, actions or communications. Another possible development is to use the possible evolution of the resilience indicators to predict or define the most appropriate alternatives for human operator actions. This can provide an online tool for decision-making or monitoring.

6. ACKNOWLEDGMENTS

The present research was supported by the International Campus on Safety and Intermodality in Transportation, the Nord-Pas-de-Calais Region, the European Community, the Regional Delegation for Research and Technology, the Ministry of Higher Education and Research, and the National Center for Scientific Research. The authors would like to thank the European project, Information Technology for Error Remediation And Trapping Emergencies (ITERATE), in the seventh Framework Program, for the analysis of transport driver behaviours, and the REACT project, financed by DGA (French army), for helping heterogeneous military units to learn and react faced with unexpected events. The authors gratefully acknowledge the support of these institutions.

REFERENCES

Chen C.-M., Lin C.-W., Chen Y.-C. (2007) Adaptive error-resilience transcoding using prioritized intra-refresh for video multicast over wireless networks. *Signal Processing: Image and Communication*, volume 22, 277-297.

Chen, J., Yea, Y. (2007). Controlled output variance based diagnosis tree for feedforward/feedback control systems. *Chemical Engineering Science 62*. 943-956.

Cheveau F.-R., Wybo J.-L., (2007). Approche pratique de la culture de sécurité : pour une maîtrise des risques industriels plus efficace. *Revue Française de Gestion 174*, pp. 171-198.

Enjalbert S., Vanderhaegen F., Pichon M., Ouedraogo K-A., Millot P. (2010). Assessment of transportation system resilience. *Human Modelling in Assisted Transportation*, Belgirate, Italie.

Gao F. (2010). Resilience analysis and measurement for water supply system. *MS thesis*, University of Saskatchewan, Canada.

Goussé V. (2005). Apport de la génétique dans les études sur la résilience : l'exemple de l'autisme. *Annales Médico-Psychologiques*, Sous presse.

Gu L.., Enjalbert S., Vanderhaegen F. (2009). Human-machine systems resilience – Safety application. *28th European Annual Conference on Human Decision-Making and Manual Control*, Reims, France.

Hale A., Heijer T. (2006). Defining Resilience. In *E. Hollnagel, D.D. Woods, N. Leveson (Eds) Resilience engineering : concepts and Precepts* Ashgate.

Hollnagel E. (2006). Achieving system safety by resilience engineering. *1st Institution of Engineering and Technology International Conference on System Safety*.

Hollnagel E., Woods D.D. (2006). Epilogue : Resilience engineering precepts. In *E. Hollnagel, D.D. Woods, N. Leveson (Eds) Resilience engineering : concepts and Precepts* Ashgate.

Hollnagel E., Woods D.D., Leveson N. (2006). Resilience engineering : concepts and Precepts. Ashgate.

Hsieh F.-S. (2009). Developing cooperation mechanism for multi-agent systems with Petri nets. *Engineering Applications of Artificial Intelligence 22*. 616–627.

Leveson N. (2004). A new accident model for engineering safer systems. Safety Science 42, pp. 237–270

Ludwig D., Walker B., Holling C.S. (1997). Sustainability, stability and resilience. *Conservation Ecology*, Vol. 1, N° 1, Art. 7.

Luo M.-Y., Yang C.-S. (2002). Enabling fault resilience for web services. *Computer Communications*, 25 198-209.

Marcantoni W.S., (2009). Mécanismes cellulaires de l'apprentissage.http://www.unites.uqam.ca/cnc/psy4042/mecanismesneuronaux.pdf [Accessed 17 November 2011].

Martin S. (2005). La résilience dans les modèles de systèmes écologiques et sociaux. Thèse de doctorat, Ecole Normale Supérieure de Cachan, France.

Nakayama H., Ansari N., Jamalipour A., Kato N., (2007) Fault-resilient sensing in wireless sensor networks. *Computer Communication*, 30, 2375-2384.

Neema S., Bapty T., Shetty S., Nordstrom S. (2004). Autonomic fault mitigation in embedded systems. *Engineering Applications of Artificial Intelligence 17*. Pp. 711–725

Numanoglu T., Tavli B., Heinzelman W., (2006) Energy efficiency and error resilience in coordinated and non-coordinated medium access control protocols. *Computer Communications*, 29, 3493-3506.

Orwin K. H., Wardle D. A., (2004). New indices for quantifying the resistance and resilience of soil biota to exogenous disturbances. *Soil Biology & Biochemistry*, 36 1907-1912.

Ouedraogo, K.A., Enjalbert, S., Vanderhaegen, F. (2010a). How to learn from the resilience of Human-Machine Systems? *The 11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems.* Valenciennes, France.

Ouedraogo, K.A., Enjalbert, S., Vanderhaegen, F. (2010b). Principes de résilience et processus d'apprentissage face à l'imprévu. *6ème Conférence Internationale Francophone d'Automatique*. 2-4 Juin., Nancy, France.

Pérez-España H., Sánchez A., (2001). An inverse relationship between stability and maturity in models of aquatic ecosystems. *Ecological Modelling*, 145 189-196.

Polet P, Vanderhaegen F, Millot P (2009) Human behaviour analysis of barrier deviation using the benefit-cost-deficit model. Advances in Human-Computer Interaction. Available on http://www.hindawi.com/journals/ahci/2009/642929

Steen, R., Aven, T. (2011). A risk perspective suitable for resilience engineering. *Safety Science 49*, pp.292–297

Tianfield H., Unland R. (2004). Towards autonomic computing systems. *Engineering Applications of Artificial Intelligence 17*. 689–699

Vanderhaegen F (1997). Multilevel organization design: the case of the air traffic control. *Control Engineering Practice*, **5,** 3, 391-399.

Vanderhaegen F (1999a) Cooperative system organisation and task allocation: illustration of task allocation in air traffic control. *Le Travail Humain*, **63**, 3, 197-222.

Vanderhaegen F (1999b) Multilevel allocation modes - Allocator control policies to share tasks between human and computer. System Analysis Modelling Simulation, 35, 191-213, 1999.

Vanderhaegen F (1999c) Toward a model of unreliability to study error prevention supports. Interacting With Computers, 11, 575-595.

Vanderhaegen F, Jouglet D, Piechowiak S (2004) Human-reliability analysis of diagnosis support cooperative redundancy. IEEE Transactions on Reliability. 53, 4, 458-464.

Vanderhaegen F (2010a). Human-error-based design of barriers and analysis of their uses. Cognition Technology & Work, 12, 133-142.

Vanderhaegen, F. (2010b). Autonomy control of Human-Machine Systems. *The 11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems.* Valenciennes, France.

Vanderhaegen F., Polet P., Zieba S., (2009). A reinforced iterative formalism to learn from human errors and uncertainty. *Engineering Applications and Artificial Intelligence*, 22, 654-659.

Vanderhaegen F, Zieba S., Enjalbert S., Polet P., (2011). A benefit/cost/deficit (BCD) model for learning from human errors. *Reliability Engineering and System Safety*, Volume 96, Issue 7, Pp.757-766.

Wang J.W., Gao F. and Ip W.H. (May 2010). Measurement of resilience and its application to enterprise information systems. *Enterprise Information Systems. Vol. 4, No. 2*, 215-223.

Westrum R. (2006). A Typology of Resilience Situations. In E. Hollnagel, D. Woods and N. Leveson (Eds.): *Resilience engineering: Concepts and precepts.* Aldershot, UK: Ashgate.

Wreathall J. (2006). Properties of resilient organisations: An initial view. In E. Hollnagel, D. Woods and N. Leveson (Eds.): *Resilience engineering: Concepts and precepts.* Aldershot, UK: Ashgate.

Wreathall J. (2011). Monitoring – A Critical Ability in Resilience Engineering. In E. Hollnagel, J. Pariès, D. Woods and J. Wreathall (Eds.): *Resilience Engineering in Practice: A guidebook.*. Aldershot, UK: Ashgate.

Xu, J.-X., Lee, T.H., Zhang, H.-W. (2004a). Analysis and comparison of iterative learning control schemes. *Engineering Applications of Artificial Intelligence 17* Pp. 675–686.

Zhang W.J. and Lin Y. (May 2010). On the principle of design of resilient systems — application to enterprise information systems. *Enterprise Information Systems. Vol. 4. No. 2*, 99-110

Zhang Z., Polet P., Vanderhaegen F., Millot P. (2004). Artificial neural network for violation analysis. Reliability Engineering and System Safety 84, 3–18.

Zieba S, Polet P, Vanderhaegen F, Debernard S (2009) Resilience of a human-robot system using adjustable autonomy and human-robot collaborative control. *International Journal on Adaptive and Innovative Systems, 1, 1*, 13-29.

Zieba S, Polet P, Vanderhaegen F, Debernard S (2010). Principles of adjustable autonomy: a framework for resilient human machine cooperation. *Cognition, Technology and work, 12 (3)*, pp. 193-203.

Zieba S, Polet P, Vanderhaegen F (2011). Using adjustable autonomy and human–machine cooperation to make a human–machine system resilient – Application to a ground robotic system. *Information Sciences, 181, 3*, 379-397.