Enhancing Safety Assessment of New Tools in Human in the Loop Simulation

Visualising separation performance

EUROCONTROL

INTRODUCTION

How do we ensure a new tool or procedure is safe?

An implementation of a new tool or procedure in the air traffic management field is routinely preceded with an extensive testing and validation process. This is traditionally performed using a human in the loop simulation (HILS) that aims to identify the specific problems emerging from an application of a new tool. The traditional methods used in simulations, although providing practical insight into human performance, do not address safety directly. In addition, simulations provide fairly limited opportunities to investigate safety issues occurring in the wider operational environment. Therefore, the validation process of a new safety tool when based on real time simulation results should use additional safety indicators. Traditional safety indicators collected in real time simulations are focused only on the failures, ignoring the rich data sample of successful operations without any specific occurrences.

The simulation environment resembles operational conditions but does not reflect all operational problems. Adverse events are hard to simulate due to the limitations of the platform. Moreover, in the case of simulating adverse events or 'non-nominal conditions' which might trigger safety-related occurrences, the impact and associated stress would not be perceived by the controller in the same way as in the real life. The usually applicable measurements such as short term conflict alert or separation losses show only a limited picture of the safety level in the simulation. Therefore there is a need for additional measurements, sensitive to the variation of safety levels in a real time simulation environment.

If we want to 'add safety' to the system, as in SESAR, how do we demonstrate that safety has really increased?

Considering the improvement of En Route ATM safety there are two areas of focus. One is collision avoidance, ensuring that when separation is lost, it can be recovered, whether by controller or pilot intervention, in the last minutes before collision. The other focus is avoiding losses of separation in the first place, ensuring that conflicts are resolved early enough that they do not result in a loss of standard separation (LoS).

If substantive improvements in safety are to be gained, such as a factor of ten as outlined in SESAR, then safety must focus on this second, 'upstream' area. This means focusing on separation behaviour, on the positive aspects of controller actions, making this more robust, more resilient. A new tool or procedure which 'adds' safety will most likely be working in this 'upstream' area or layer of protection. This indicates that the validation process for "upstream" tools (such as Eurocontrol's Tactical Controller Tool (TCT) or NATS' IFACTS tool) should be based on a methodology capable of capturing the positive safety influence of the tool in question. The investigation should have a closer look at controller behaviour in the same time scale as the new tool operates, for example in a tactical environment 4 - 10 min in advance of Short Term Conflict Alert (STCA) or separation loss.

An advantage of this approach is that the data from this stage are profuse when compared to counting losses of separation (which are rare). Focusing upstream can be used as a primary source for evaluating whether a new tool increases or decreases the 'safety margin' (Figure 1).

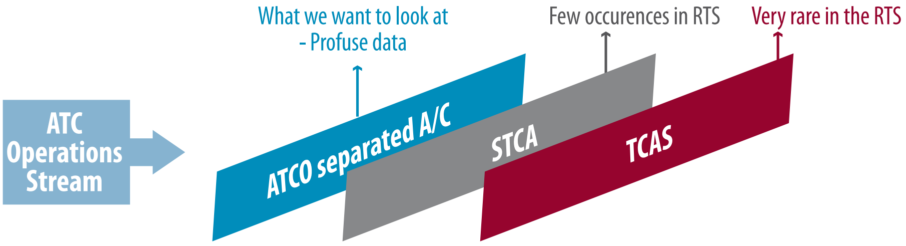

Figure 1: Focusing on 'upstream' safe separation performance.

HOW CAN EVIDENCE FOR SAFETY IMPROVEMENT BE DEMONSTRATED?

In the en-route, radar environment assessment of safety level is based on the losses of separation i.e failures in separation provision as indicated by STCA or TCAS/ACAS alarms. These rare events however do not give an appropriate representation of the safety level of the system. Arguing safety levels based on a fragment of data (Figure 2) does not tell us how well separation is performed, nor give ideas on how to improve it. This is a 'Resilience' concept: since 'failure' data is so rare in ATM, it is better to try and learn from what controllers do right. This is also called 'Positive Safety'.

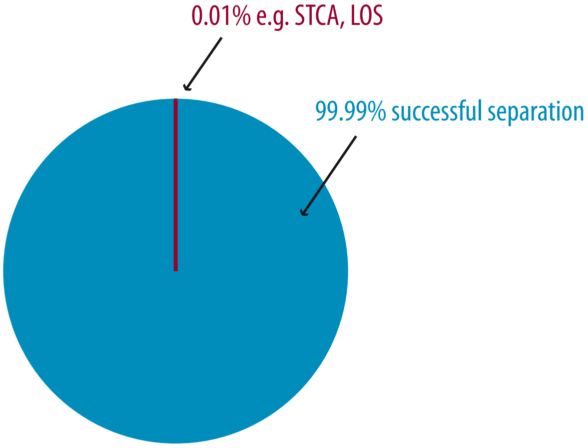

Figure 2: Learning opportunities when focusing on failure vs. successful performance.

More learning about how to achieve safer performance can be achieved by focusing on upstream data - the time before safety events are registered. The data sample of standard operations is much greater and can tell more about controller working methods.

The Air Traffic Controllers (ATCOs), while managing the traffic in European en-route airspace based on ATS surveillance systems, maintain the standard separation minima of five nautical miles and one thousand feet. However in addition to a standard separation required by the regulation, they often apply a personal safety buffer that is not formalised but gives additional time/space to recover in case of critical incidents. Less personal buffer space leaves less time, which starts to load risk, and will have a negative impact on safety. Personal buffers could be affected by a new safety tool, but typical safety measurement methods would not detect decreases in the safety buffer. To have confidence that a new safety tool will improve safety we need to identify how the controllers operate with it. The investigation should focus on the same time scale in which the tool operates. The analysis should take into account the following:

- Standard operations of separation maintenance;
- Adverse or rare events;
- Impact on safety buffers;
- Overall picture of achieved separation.

We propose a method that can complete the current safety measures and provide a holistic picture of controller separation performance.

The Separation Performance Visualiser (SPV) can categorise the controller's interventions with respect to the time before predicted loss of separation (LoS) as well as risk of collision (proximity) of the predicted potential losses of separation.

SPV predicts the aircraft trajectory and dynamically defines the closest approach distance between two aircraft. If the predicted closest approach distance in the horizontal or vertical plane is less than the separation minima, then the situation will be classified as a potential loss of separation. When a potential loss is detected, the tactical controller' actions to resolve that conflict are recorded.

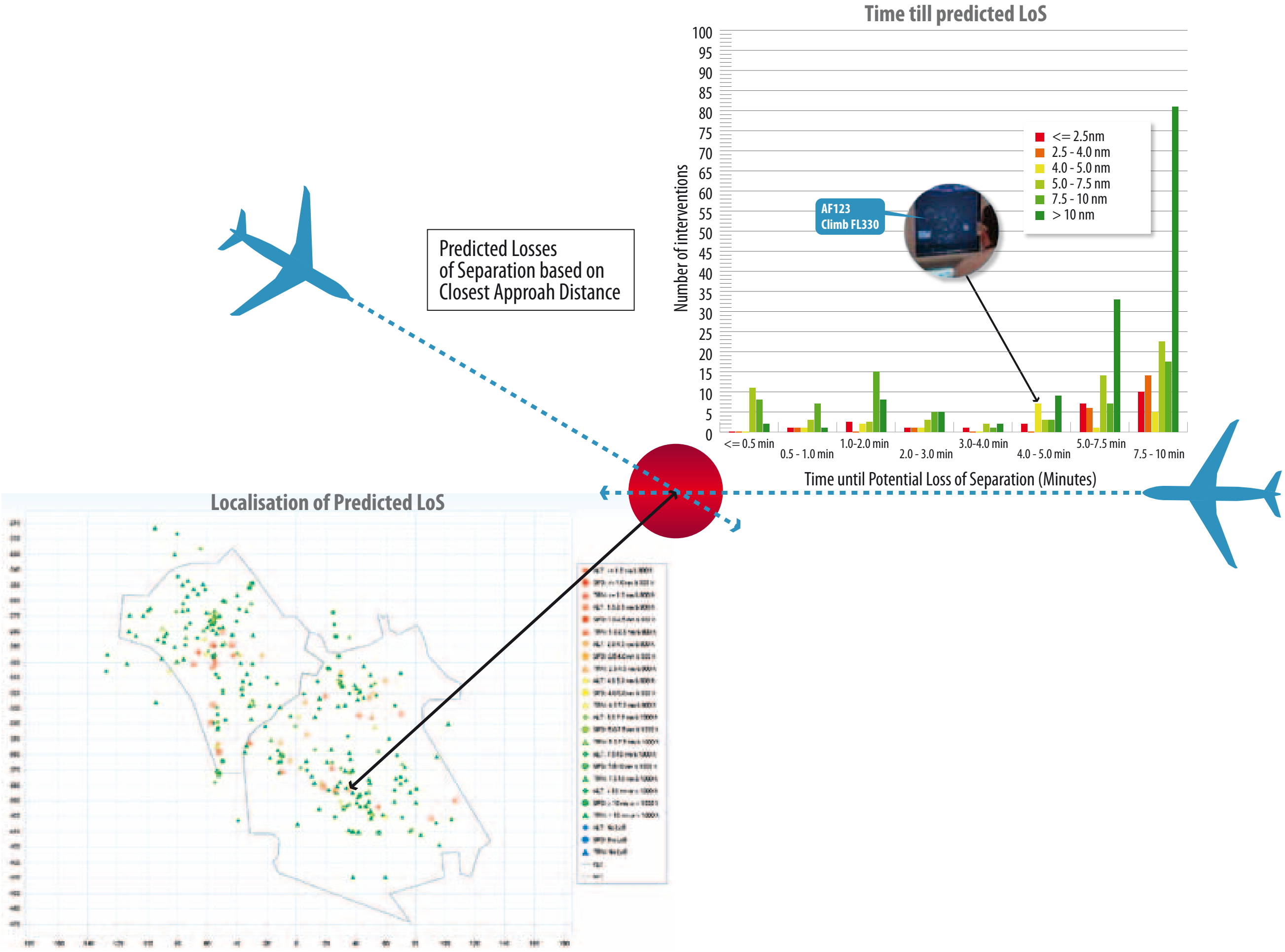

Figure 3: Separation Performance Visualiser – what it measures.

SPV measures the following indicators of separation performance (Figure 3):

- Intervention type to resolve predicted conflict (Heading instruction, Speed change, Altitude change) [What the ATCO does, and how he/she resolves the conflict];
- Time of intervention before potential loss of separation [when the ATCO intervenes];
- Risk of collision of predicted losses of separation [How close the aircraft would fly in absence of controller intervention];
- The actual separation between the aircraft before and after the controller's intervention;
- Localisation of controllers' interventions [where the ATCO intervenes];
- Localisation of predicted losses of separation [where the potential 'hotspots' are];
- Conflicts geometry [analysis of the conflicts both in horizontal and vertical space].

HOW CAN EVIDENCE FOR SAFETY IMPROVEMENT BE DEMONSTRATED?

The Separation Performance Visualiser (SPV) calculates instantaneously the possible losses of separation and their risk of collision. The SPV also registers the controller interventions such as turns, altitude and speed changes together with the time and distance to predicted losses of separation. In Figure 3 loss of separation is based on 5 Nm (standard minimum separation for European en–route airspace under ATS surveillance systems). The interventions presented in the bar graphs are calculated with reference to the 5 Nm limit. The colours of the bar represent the risk of collision of potential losses of separation, starting from red as a highest risk to green as lowest risk occurrences.

The same colour coding is applied to a geographical representation ('bird's eye view') of sector activities. Each point shows the localisation of predicted losses of separation and their risk of collision. The analysis of interventions shows where, in terms of time and distance, the aircraft were proceeding.

We might imagine three different controller support tools: tool 1, tool 2 and tool 3. The distribution of interventions for each tool can be different, as showed by Figure 4. The interventions taken while using Tool 1 were mostly taken much in advance of predicted losses of separation. In case of Tool 2 the interventions were taken both closer to predicted loss of separation and in advance. However with Tool 3, the majority of interventions where taken very close to the potential loss of separation. This can mean the controllers are using the tool for "productivity", increasing capacity by running the traffic closer to the safety limits (reducing the safety 'buffer').

Traditional measurements would not show any difference between the tools. From a safety point of view, particularly when trying to demonstrate 'added' or increased safety, the desired operations would be the ones that assure the separation between the aircraft in advance, thus similar to the one achieved with Tool 1. SPV provides the opportunity to look in detail at the working methods applied when using a new tool. It also displays to what extent the safety buffers are affected by new tool implementation. This aspect is especially significant in later stages of validation when adaptation to a new tool becomes more critical.

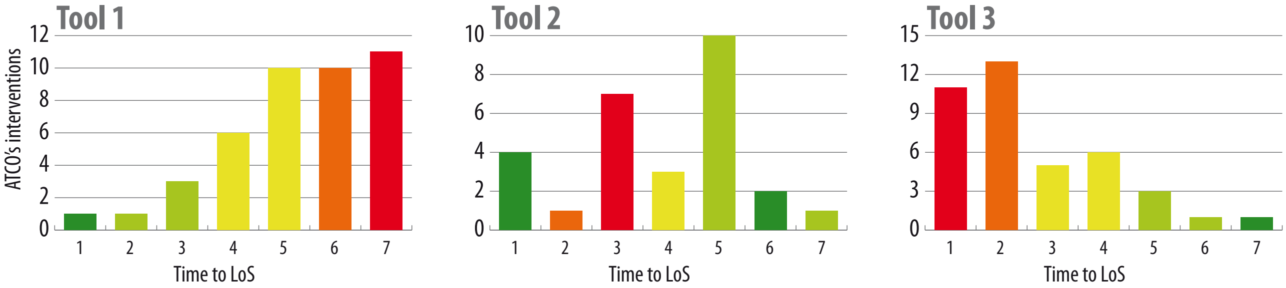

Figure 4: Example of distribution of interventions for three different automated conflict resolution tools (illustrative example).

SEPARATION PERFORMANCE INDICATORS IN THE TOOL LIFECYCLE

The first verification of a new tool or procedure is conducted at an early stage of the development. However the realism of the simulation environment increases with every step of the validation process. Aspects related to ATCO working methods and 'adaptation' can only be captured in later stages of validation. The controllers dealing with a new tool might devise their own working methods or modify those envisaged by the designers.

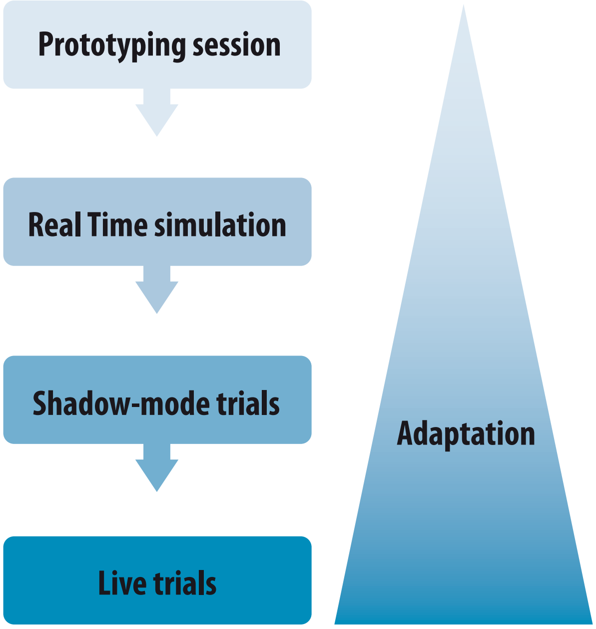

Figure 5: Validation Lifecycle

The performance indicators derived from SPV can evaluate if the tool is used according to its functional intent (how the tool is expected to be used according to the designers).

In addition, local implementation of a new tool requires specific adjustments which might lead to unexpected 'drift' from the design intent of the initial concept. SPV indicators will also be sensitive to such changes. When the tool is implemented in an operational environment, SPV allows detection of local adaptation of ways of working. The SPV indicators can also detect the changes in separation provision related to skills degradation over time.

IS THE NEW TOOL OR PROCEDURE GOING TO MAKE ATM SAFER OR MORE DANGEROUS?

The air traffic system is known to be resilient, and safety is assured through compensation of different factors. In many cases, the controller's skills prove to be sufficient to provide safe separation even during equipment failures. Even if the new tool decreased safety, the controllers might be able to cope with it under simulation conditions. Traditional safety indicators such as STCA or losses of separation might not change at all, however the controller behaviour would adapt to maintain safety in the new conditions i.e with the tool. This 'compensation' behaviour might be sustainable for short periods, for example in a simulation, but lead to hazards in live operations.

A further consideration might be a conflict detection tool that provides automated conflict resolutions. The controllers who gained confidence with support of the tool are able to manage a higher load of traffic, and in the same way improve the capacity of the sector. However, we would also need to know what would happen in case of failure of the tool. Would the controllers be able to revert to their previous performance levels, or would this failure condition prove very unsafe? In this case, we need to know how the tool is affecting the controller's underlying strategies and skills for conflict detection and resolution. If it is changing the ATCO's skills, then 'reversion' to previous ways of working will be harder.

Photo: © Stefan Meyer

SEPARATION PERFORMANCE VISUALISER INDICATORS

Type of interventions

The type of interventions graph shows the controllers' actions: altitude, headings and speed changes in relation to time to potential loss of separation. Depending on the sector and traffic complexity, the controllers might prefer to re-route the aircraft or to change altitude. Controllers' actions are linked to time demonstrating how long in advance the traffic was processed. In Figure 6 the majority of the interventions are turns (headings), followed by altitude changes that occurred more than 10 min prior to potential loss of separation.

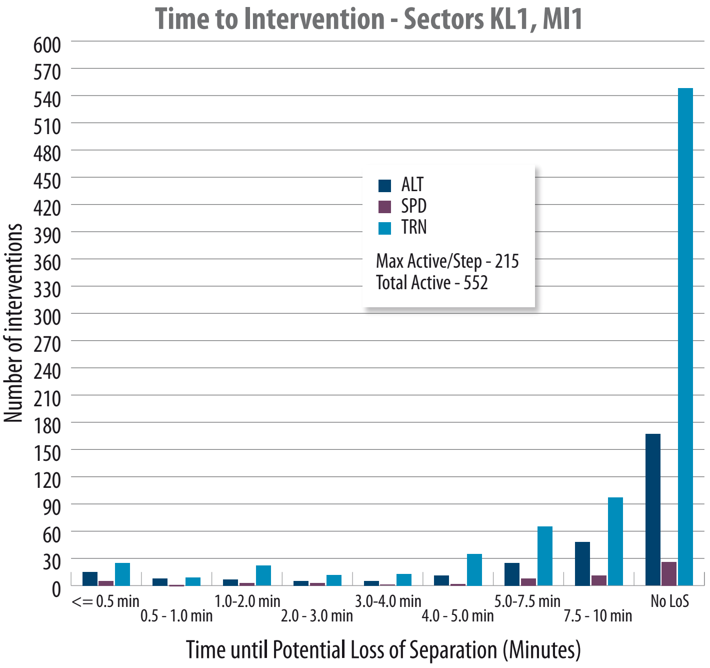

Figure 6: Type of Interventions

Time of intervention

Time of interventions groups all types of controllers' actions in relation to time of potential loss of separation (LoS) and risk of potential LoS. In Figure 7, the bars indicate the number of actions whereas colour represents risk of potential LoS that the action was related to. Red indicates the predicted loss of separation with higher risk while green with lower risk. In the presented example, the actions to solve most critical occurrences took place 5 minutes in advance of predicted loss of separation.

This graph can also be useful to identify the impact on personal safety buffers applied by the controllers. More interventions taken to resolve severe LoS (marked in red and orange) close to border of separation minima, indicate a small buffer. This means that if anything goes wrong, a LoS is likely (the controller may not have a 'Plan B' which will be effective in time.)

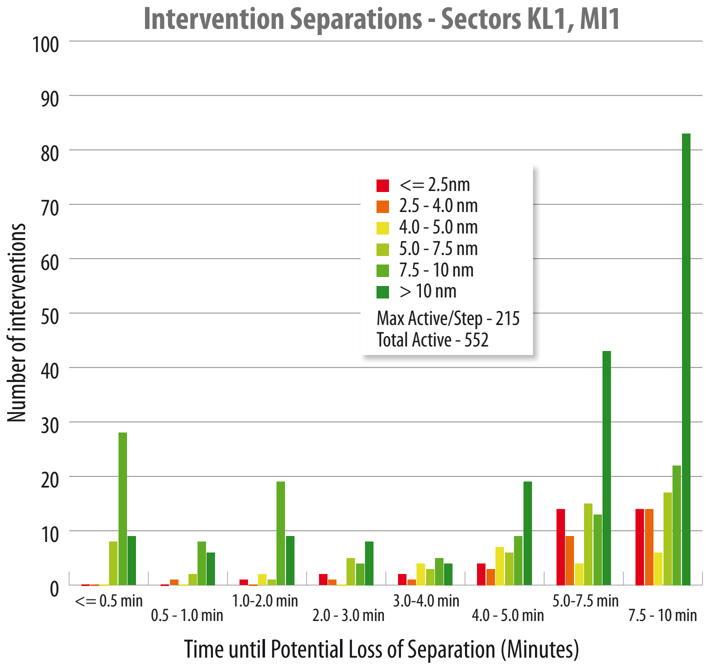

Figure 7: Time of interventions

Global view of actual separation

Two further questions relating to separation performance are how much time aircraft spend 'at risk' in a sector, and how efficient are controllers at reducing aircraft 'exposure' to risk? In Figure 8, the actual separation represents the total flight time of all aircraft in the sector (in flight minutes). The actual separation (in blue) is contrasted with the predicted separation (in purple) calculated based on aircraft prediction track. The purple bars show how the controllers initially planned the traffic, in a few cases fairly close. The blue bars however show that the controllers then performed better, resolving conflicts earlier. Thus, in Figure 8 the majority of the traffic maintained separation of more than 10NM and 1000Ft. The global picture of separation shows the percentage of traffic that was actually separated without any risk of potential losses of separation. This shows good ATM practice by controllers.

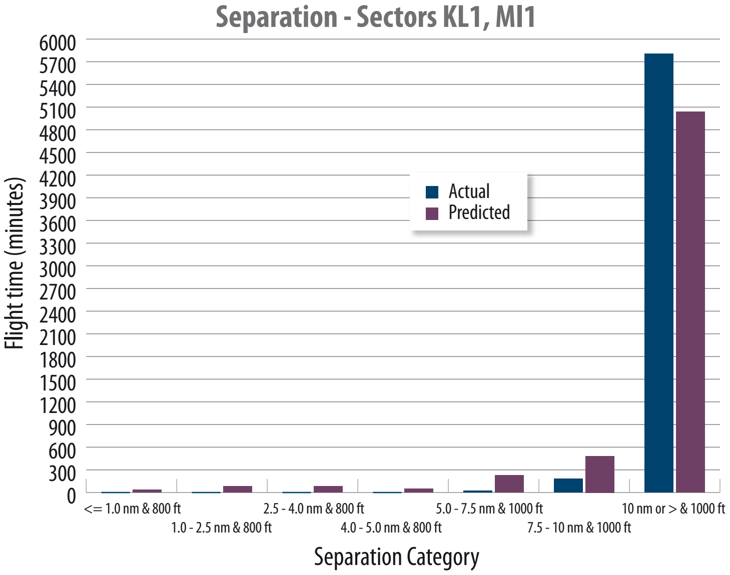

Figure 8: Global separation performance

Location of interventions according to risk categories

Locations of interventions (Figure 9) represents the geographical distribution of actions in the sector. The interventions are categorised according to the type of instruction, distance and time to predicted losses of separation. The graph demonstrates the dense areas, and the main traffic streams in the sectors. This type of analysis is useful to identify the dependencies between adjacent sectors. The risk categories can be adjusted according to the needs of the investigation.

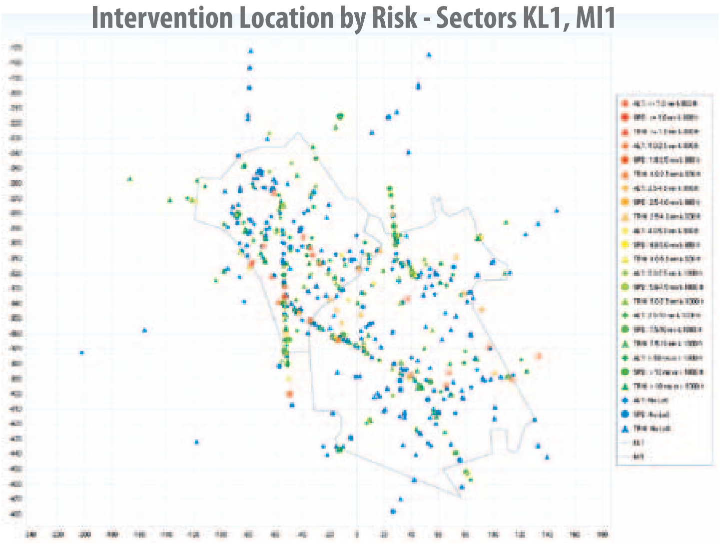

Figure 9: Location of Interventions

Location of predicted losses of separation according to risk categories

Location of losses of separation (see Figure 10) shows the geographical distribution of speed, altitude and heading changes categorised according to risk categories. The graph demonstrates the main 'hotspots' of the sectors. In addition, the complexity of the traffic can be evaluated by analysis of the different horizontal and vertical geometry of the conflicts. The analysis can also help to verify the traffic patterns before the simulation, to ensure equivalent complexity between different patterns used in the same simulation.

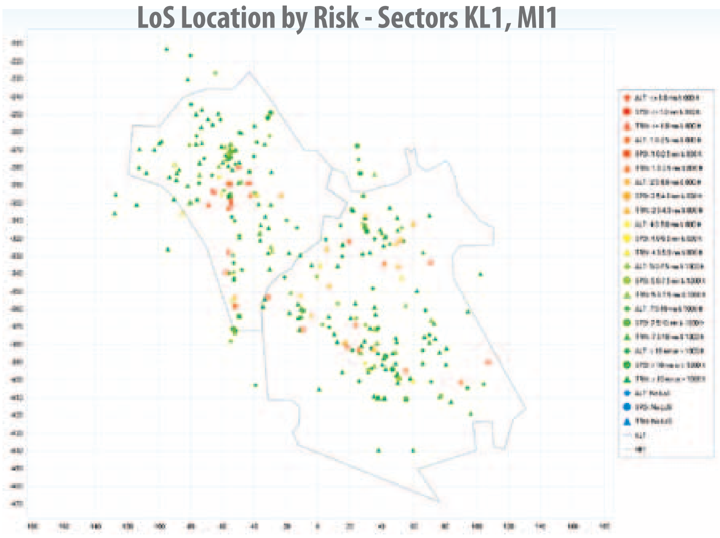

Figure 10: Location of predicted losses of separation

APPLICATIONS OF SPV – DOES IT DELIVER?

To date there have been three applications of SPV. The first involved a prototyping simulation for a tactical conflict detection support tool for controllers. The 'normal' measures of safety showed no change. However SPV revealed a reduction in 'safety buffers' with the new tool. The new tool gave controllers more confidence and their turns etc. were left later than they would normally be made. The safety 'margin' decreased.

In a second study carried out with a European ANSP, SPV was applied to investigate an impact of 4-D constraints to be respected during the flight execution, on separation practices in a Real Time Simulation. Analysis of interventions showed that ATCOs adapted their conflict resolution strategy by providing progressive instructions instead of direct ones. The number of instructions as well as the proportions between different instruction types changed – suggesting a change in ATCO working methods with this new procedure. Additionally, the hotspots were more distinctive when 4D constraints were applied, resulting in a higher number of predicted losses of separation, but a better situation awareness by controllers due to hotspot predictability. The STCAs collected during this simulation showed no difference to normal conditions.

In a third study, using live data from a European Centre, it was observed that most interventions to solve predicted losses of separation were taken more than 10 min in advance. The analysis of global performance showed that in fairly dense and challenging airspace 98% of flight times were above 10Nm and 1000Ft, and only 2 % of flight times were in separation categories 5-10Nm and 1000Ft. Such a result demonstrates the resilience of the ATM system.

CONCLUSIONS ON EVALUATION OF SAFETY TOOLS IN REAL TIME SIMULATIONS

The SESAR Safety target for European En Route civil traffic is currently a factor of ten improvement by 2020 over 2005 safety levels. This is not trivial to achieve, nor is it simple to measure. Future tools which can improve safety will have a safety case, but the emphasis can all too easily rest on whether or not such tools themselves fail in some way. While this is necessary and important, there must be equal focus on whether they can actually deliver more safety than is available in the current system. Formal safety case studies cannot hope to model all the potential variations and permutations of separation behaviours seen in European Air Traffic. The best way to view separation behaviour is 'in vivo', i.e. live controller performance in controlled and measurable experiments (real-time simulations).

Such simulations need to be credible/representative of actual (future) operations, including participation by valid operational controllers. Furthermore the simulations must be controlled and follow a rigorous experimental design, testing a new tool or procedure against a baseline. Measures must be sensitive to subtle changes in safe performance, including measuring what is happening to the 'safety buffer'. Last, such studies of the impacts on safety need to be measured at each incremental step of the design and implementation process, from Concept stage (prototyping simulation) to pre-operation (live trial) and full scale operation.

SPV can gather safety evidence to show whether a new tool or procedure is adding safety or making the overall system more dangerous. SPV is therefore a valuable tool in the Safety Toolbox for making a Safety Case, whether for an individual tool, or for a Macro-system upgrade like SESAR.

Further information

Separation Performance Visualiser, developed by EUROCONTROL, has been successfully tested with real time simulation data and data from an operational centre. The SPV is available free of charge under a EUROCONTROL licence. Should you need further information please contact:

ella.pinska-chauvin@eurocontrol.int or
barry.kirwan@eurocontrol.int

Barry Kirwan

Ella Pinska-Chauvin

EUROCONTROL

November 2010 - © European Organisation for the Safety of Air Navigation (EUROCONTROL)

This document is published by EUROCONTROL for information purposes. It may be copied in whole or in part, provided that EUROCONTROL is mentioned as the source and it is not used for commercial purposes (i.e. for financial gain). The information in this document may not be modified without prior written permission from EUROCONTROL.

www.eurocontrol.int