HUMAN IN THE LOOP TO ASSESS 4D TRAJECTORY MANAGEMENT WITH CONTRACT-OF-OBJECTIVES

Sandrine Guibert *, Laurent Guichard *, Jean-Yves Grau**

*EUROCONTROL Experimental Centre, France, ** INEOVA, Varna, Bulgaria,

Key words: ATM, punctuality, Contract of Objectives, global efficiency, accountability

Abstract

The Contract-based Air Transportation System (CATS) Project introduces a new way of managing trajectories, using mutually agreed objectives, called Contract of Objectives (CoO). The CoO is a formal and collaborative commitment of ATM actors (i.e. airspace users, Air navigation services providers (ANSP), airports), to the conduction of each flight. It establishes a sequence of spatial and temporal windows which constituted milestones to be met during the flight execution, reconciling then planning and execution of flights. These 4D intervals are called the Target Windows (TWs) and are the fundamental elements to organize the future ATM. CATS proposes, through applying the CoO, one of the possible implementations of the SESAR business trajectory.

The present paper provides an overview of the validation foreseen and presents the results of the second experiment on the evaluation of the CoO concept between ATCOs and aircrew, carried out end of 2009, in SkyGuide premises, through a Human-in-the-Loop (HIL) simulation. The assessment, following E-OCVM, focuses on system performances, safety, and human acceptability.

1 Introduction

The NextGen and SESAR programs [1][4] plan fundamental changes in the air traffic operations in the US and in Europe to reach ambitious objectives. Emphasis is given to performance and cost efficiency, and both initiatives advocate a paradigm shift towards trajectory based operations.

Air transport business stimulates national economies, global trade and tourism [[2]]. Business imperatives will always push for cutting costs, and stronger competition and liberalization will continue to present a challenge for businesses, with an opportunity for new cost-models (e.g., low-cost airlines). The air transport supply-chain as a whole, therefore, needs to become more cost-efficient. Since the Air Transport System (ATS) supply-chain is a complex one involving many partners (such as airports, airlines and ANSPs), these business imperatives will have to be supported and shared by everyone, even if their interests or cost-models are different. The future ATM system should integrate ground and airborne segments more closely, respect schedule integrity, and enhance interoperability.

One possible mechanism to formalize the Business Trajectory advocates by SESAR[4] and to promote a high collaborative and system-wide approach, allowing global optimisation and local constraints integration, has been proposed by the Contract-based Air Transportation System (CATS) project through the Contract of Objectives (CoO) [5]. The CATS project is co-funded by the European Commission through the Sixth Framework Program.

The proposed paper will detail results of the assessments carried out, regarding firstly the human performances (i.e., workload, situation awareness and acceptability) and secondly regarding the 3 Key Performance Areas (i.e. Safety, Efficiency and Capacity).

2 Concept overview

The CATS concept has already been presented in 26th ICAS conference in 2008 [3]. The CoO is represented by the commitment of each actor to deliver a particular aircraft inside temporal and spatial intervals, called TWs. These commitments are agreed upon all involved...
actors for specific transfer of responsibility areas (e.g. between 2 ACCs). Then, each actor will be fully accountable for its own achievements. The ultimate objective of the CoO is punctuality at the destination, while improving the system efficiency and predictability by means of enhanced collaboration between air transport actors.

![Figure 1: Contract of objectives](image)

Instead of precise 4D points, the TW is expressed in terms of temporal and spatial intervals. They are defined on the basis of transfer of responsibility areas (Figure 1). Their sizes and locations reflect negotiated objectives resulting from downstream constraints, such as punctuality at the destination, runway capacity, congested en-route areas or aircraft performance. TWs provide room for manoeuvre to ensure resilience in case of disruption and conflict management and, lastly, impose constraints only if necessary.

3 Validation overview

The aim of the CATS Project is to assess the CoO and associated TWs by involving the major actors in the supply chain.

The CATS concept assessment, following European Operational Concept Validation Methodology (E-OCVM) [6], is conducted by two main means:

- Systemic validation, which highlights the impacts for the overall ATS on safety and risk management, cost benefits, and legal consequences;
- Operational validation which analyses how the proposed CoO and the associated TWs impact the operators' performance regarding selected Key Performance Areas (KPAs) defined by SESAR [3].

Operational validation is led by three successive Human-In-the-Loop (HIL) experiments which focus on different validation objectives:

- **HIL-1.** Evaluation of the impact of the CoO between Air Traffic Controllers (ATCOs)
- **HIL-2.** Evaluation of the impact of the CoO between ATCOs and aircrew
- **HIL-3.** Evaluation of the renegotiation process involving ATM actors (airlines, airports and ANSPs) incase of CoO is not fulfilled.

This paper presents the results of the operational validation HIL-2 experiment.

4 Human in the Loop 2 experiment

4.1 Objectives

HIL-2 was carried out from 10 days in October 2009 in Skyguide simulation room, and designed from the HIL-1 results [[8]]. The simulation devices encompassed coupled controller working position and cockpit simulators. One of the hypothesis of this experiment was to prove that “shared information can connect the air and ground elements to benefit to the overall system” [[9]].

The HIL-2 aim was to ascertain that:

- CoO implementation allows safe operations;
- TWs integrate flexibility to cope with uncertainty;
- The ATCOs' and aircrews' working methods deriving from CoO execution are acceptable;
- CoO execution does not impact the ATCOs' and aircrews' performance;
- CoO execution does not impact the ATCOs' and aircrews' activity;
- Collaboration between ATCOs and aircrews is high;
- CoO is still manageable with growth of traffic as foreseen in the 2020.
4.2 Experiment Variables
Two independent variables were manipulated during the experiment: Target Windows and traffic loads.

Two conditions, with and without Target Windows, were measured.

Two traffic loads were measured during the experiment: 2008 traffic level and 2020 forecast traffic. The expected level of traffic in 2020 was determined by the EUROCONTROL Statfor services. Traffic is expected to increase by 40% in 2020 in the measured area. Several traffic scenarios were designed and their difficulties have been controlled.

4.3 Measurements
Two kinds of measurements were collected during this experiment: system performance, and human performance.

The aim of the system performance evaluation is to assess whether the CATS benefits are delivered as proposed. From the stakeholders concerns and SESAR performance framework [[9]], four of the SESAR KPAs were identified as potentially improved by CoO and associated TW introduction: capacity, safety, efficiency, and predictability.

The human performance objective is to see whether the contribution of the human to overall system performance is within expected capabilities (workload, situation awareness, working methods, feasibility, acceptability, etc.). Different methods and techniques were used, such as observations, recorded data, questionnaires and self-assessments, as presented in Table 1.

<table>
<thead>
<tr>
<th>Human performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workload</td>
</tr>
<tr>
<td>Instantaneous Self-assessment of Workload (ISA) – ATCO</td>
</tr>
<tr>
<td>NASA-TLX – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Post-run debriefing – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Post experiment questionnaire – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Situation awareness</td>
</tr>
<tr>
<td>Situation Awareness for SHAPE Questionnaire (SASHA-Q) – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Post-run debriefing – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Post experiment questionnaire – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Activity</td>
</tr>
<tr>
<td>Over-The-Shoulder (OTS) observation – ATCO</td>
</tr>
<tr>
<td>Post-run debriefing – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Post experiment questionnaire – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Collaboration</td>
</tr>
<tr>
<td>Communication duration – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Communication content – ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Post-run debriefing –ATCO &amp; Pilot</td>
</tr>
<tr>
<td>Post experiment questionnaire – ATCO &amp; Pilot</td>
</tr>
</tbody>
</table>

Table 1. HIL-2 experiment measurements

4.4 Experimental Environment
The airspace chosen for this experiment was two en-route sectors (Milan MI1 and Geneva KL1) at the border of two ACCs. A total of 4 controllers and 2 pilots participated in the CATS HIL-2. The controllers were from Roma ACC and Brindisi ACC (ENAV) and all had over 10 years of qualified experience and were working as controllers in en-route sectors. The pilots had all retired in the last six months from operational service at Air France. Both were A320 captains with more than 8,000 flying hours on glass cockpit aircraft.

The simulation environment used was made up of two coupled simulators:

- SkyGuide simulator, with the standard Geneva services and tools. Specific HMI for TWs display and associated tools have been developed by SkySoft ATM (Figure 3).

Figure 3. SkyGuide ATC Simulator and HMI
• A320 cockpit "flight simulator 2004". Specific HMI for TWs display were developed by SkySoft ATM on the Navigational Display (Figure 4).

![Figure 4. Cockpit Simulator](image)

In each run, 4 "flight simulator 2004" aircraft were piloted by the two pilots (2 aircraft by run, and for each pilot). The other aircraft were handled by automatic pseudo pilots that execute the controller instructions. In order to avoid decreasing the ATCOs' workload too much, a data link device has been implemented at the ATCO working position. The data link integrates latency delays between the instructions and their execution by aircraft.

4.5 Experimental plan

Given the independent variables, the CATS experiment followed a 2 (traffic loads) x 2 (TW conditions) repeated measurements design for controllers, resulting in 4 experimental conditions with eight repeated measurements for each condition.

For the pilots, the CATS experiment followed a 2 (TW conditions) repeated measurements design, resulting in 2 experimental conditions with 64 repeated measurement (2 aircraft by run and for each pilot in 16 runs).

The experiment lasted 10 days, and the period timetable encompassed training and familiarisation, six days for performing the 16 experimental runs, and final debriefing.

Simulation runs were conducted on the basis of three runs per day. Each run ran for about 70 minutes, 30 minutes added for filling in questionnaires.

5 Results

HIL-2 results are reported firstly for human performances and secondly for system performance.

5.1 Human Performance

5.1.1 Workload

Controllers

ATCO workload was measured through two subjective methods: Instantaneous Self-Assessment of Workload (ISA) and NASA-Task Load indeX (NASA-TLX) [[10]]. The purpose was to measure the impact of TW management on the controller workload by comparing two similar traffic management situations: one without TWs and one with TWs. The results were subjected to a Wilcoxon test. Examples of the results obtained are described Figures 5 & 6.

![Figure 5. ISA sector KL1 planner controllers' results](image)

![Figure 6. NASA_TLX sector MI1 executive controller's results](image)
ISA and NASA-TLX results are mutually consistent and provide similar results. There is no significant difference (p<0.05) between the "without TW" and "with TW" conditions whatever the traffic load, although the controllers perceived during the post-run debriefings and post-experiment questionnaire the TWs as an additional task, slightly impacting traffic management. This result is observed whatever the control position (executive or planner) and whatever the controlled sector (KL or MI).

There is a significant difference (p<0.05) between the two traffic load conditions whatever the control position and the controlled sector. The analysis of median values shows the workload of 2008 traffic conditions was always lower than the 2020 traffic load conditions.

**Pilots**

Pilot workload was measured through NASA-TLX. The purpose was to measure the impact of TW management on the pilots' workload by comparing the cruise phases into the measured sectors with and without TWs. The results were subjected to a Wilcoxon test. Example of results obtained is given Figure 7.

![Figure 7. NASA pilots' results](image)

**Situation awareness**

**Controllers**

Controllers’ situation awareness was evaluated through SASHA questionnaires [[11]] and also tackled during post–run questionnaires. The results were subjected to a Wilcoxon test and an example is shown Figure 8.

![Figure 8. SASHA-Q sector KL1 executive controller results](image)

Situation awareness is not significantly impaired by the implementation of TWs for KL1 (Geneva) sector planner and executive controllers and the MI (Milan) executive controller, whatever the traffic loads. For the MI (Milan) planner controller, there is a significant difference between the "without TW" and "with TW" conditions whatever the traffic load. However, the levels of situation awareness are always high whatever the experimental conditions. The controllers' feeling on situation awareness is that TW information increases the traffic picture, although there was no statistically significant difference compared with the SASHA questionnaire results.

This feeling was justified by the specific information displayed for the TWs. This information allowed the controller to manage the flight plan more effectively than with the information currently displayed. TW information gave the controller more details, particularly regarding the exit conditions, to deal
efficiently with the traffic, taking into account the overall constraints of the flight (not only the local sector constraints).

**Pilots**

Pilot situation awareness was evaluated through pilot-suitable SASHA-Q version, and also tackled during post-run questionnaires. The results were subjected to a Wilcoxon test and the Figure 9 described one of them.

![Figure 9. SA pilots results](image)

There is no significant difference (p<0.05) between the ‘without TW’ and ‘with TW’ conditions. This result is strengthened by the pilots’ feeling gathered from the debriefings and the post-experiment questionnaire. Pilots stress that TW data is easy to perceive and understand.

**5.3 Usability and acceptability**

**Controllers**

The number of control instructions given increased when traffic load increased, but independently of TW use. When safety may be impaired, the executive systematically applies a separation solution without considering the TW constraints. This reinforces the fact that safety concerns remain the first priority in the controller's mind.

Unanimously, controllers strongly agreed that TWs are easy to use whatever the control position (executive or planner controller). This feeling was widely expressed during the post-run debriefings. Controllers quickly became familiarized with the concept, and were autonomous at their control working positions. This feeling was reinforced by the fact that the controllers found TW management easy to learn.

**Pilots**

Although the flight simulator was a one-seat cockpit, pilots feel that TW management requires more communications between the captain and the first officer. Pilots quickly became familiarized with the concept, and were autonomous in managing the TWs in the cockpit. This impression was reinforced by the fact that pilots found TW management easy to learn.

The high level of usability assessed during the simulation is a strong point of the concept for its future development. Nevertheless, they did express one criticism on the TW display: the navigational display does not allow for an easy display of TWs in the long-term time horizon. The Multiple Control and Display Unit (MCDU) will meet this need and provide aircrew with a good degree of anticipation.

**5.4 Collaboration between controllers and pilots**

On the basis of the debriefings and questionnaires, ATCOs and pilots feel that TW management does not require more communications between them. ATCOs and pilots felt that this feeling derives from the fact that the same TW data were shared by the cockpit and the control working position. Consequently, all requests or instructions are understood in the context of the TW data, without specifically saying it. When a request or instruction is not understood properly by the ATCO and/or the pilot, he clarifies the reference to the TWs.

ATCOs and pilots have no difficulty in finding vocabulary in order to communicate about TWs. Nevertheless, all expressed the need for a new specific phraseology. Use of terms "due to Target Windows" seems accurate. By saying "due to Target Windows", the controller and the pilots immediately understand the context of a request or an instruction.

By adding information on the flight plan data, the TWs improve the aircraft intent representation that ATCOs and pilots share. This
heightens the feeling that collaboration is improved.

TWs are deemed positive for the collaboration process between ATCOs and pilots. The cost of this cooperation in terms of workload is rated as being without additional workload by controllers, and as increasing a little bit the workload by pilots.

Finally, pilots and ATCOs agree that in normal operative conditions, communications about TWs may be supported by data-link. Voice communication has to be kept for emergencies or situations where there is a lack of understanding.

5.5 System performance

5.5.1 Safety

The current level of ATM safety is high. Incidents and occurrences are rare, and most of the time, losses of separation do not occur during an exercise. The HIL-1 experiment [7] showed the lack of validity and sensitivity of the short-term conflict alert count, and recommended a more sensitive approach. For the HIL-2 experiment, safety was evaluated through aircraft separation performance.

Separation performance was assessed using the Separation Performance Tool (SPT) designed by EUROCONTROL [12] and dedicated questions in post-run questionnaires for controllers and pilots.

The SPT provides the flown and predicted flight times (in minutes) for different separation bands for all the flights. The actual separation represents the total flight time that the aircraft have flown during the exercises, grouped by the closest separation distance between aircraft. Predicted flight time is based on aircraft trajectories without controllers' interventions, and the closest separation distances according to the intended trajectories. Separation above 10 Nm and 1000 ft are grouped together with the traffic that did not risk loss of separation. Loss of separation is when aircraft are < 5 NM and 800 ft.

Separation performance was assessed for the four experimental conditions (with and without TWs for 2008 and 2020 traffic loads). Figure 10 shows the results for the TW-2020 condition for the two measured sectors.

The distribution of the flight time results in the four experimental conditions shows that there is no loss of separation. A high safety level is maintained whatever the traffic load conditions and/or TW conditions.

The majority of the traffic is maintained at more than 10Nm and 1000ft. The controllers successfully separated the aircraft whatever the experimental conditions (traffic load and TWs) because the flight time of flown traffic is greater in the band >10 NM than the flight time of the predicted traffic.

For the two bands between 5 NM and 10 NM, the flight time of the predicted traffic is greater than the flight time of the flown traffic. An interpretation of this result is that the majority of potential conflicts are avoided before these separation limits. Such results are consistent with the hypothesis that controllers have good anticipation in conflict detection and solving. This result is found in the four experimental conditions.

The ATCOs’ and pilots’ feeling from debriefings and post-run questionnaires were consistent with the quantitative data, establishing that safety was not impacted (either positively or negatively) by TW use, even when capacity matched the forecast 2020 traffic load.

5.5.2 Efficiency

Traffic efficiency was assessed through three indicators: flight duration and number of fulfilled TWs.

Flight duration is calculated by comparing the time flown by each aircraft into the sector during the experimental exercises, with a reference time which is the time to fly through
the sector for the same aircraft without any ATCO actions (simulator flying the aircraft, following the flight plan). Flight duration was calculated for the 4 experimental conditions and the Figure 11 shown the results for sector KL.

![Figure 11. Flight Duration in KL (Geneva) Sector](image)

Results show that the traffic load condition does not have an impact on flight duration. Whatever the measured sectors, flight duration is shorter with TWs. The difference is significant for 2008 traffic load and not for 2020 traffic load. The median values of flight duration with TW are closer to the 100 value than without TW, indicating that with the TWs, the aircraft flew closer to the flight plan. This suggests a better adherence to the trajectory. This outcome means that TW use increased traffic efficiency.

The number of aircraft fulfilling their TWs is a strong indicator of the concept validity and traffic efficiency. If controllers are not able to fulfill TWs and if too many aircraft are out of their TWs, the concept will lose its relevance and traffic efficiency will decrease. The percentage of ‘out TW’ aircraft was calculated for the results of each sector in the 2008 and 2020 TW conditions.

The results showed there is no significant difference (p<0.05) between the two traffic load conditions whatever the controlled sector. The median value is always equal to 0 whatever the sectors and the traffic load. The percentages of TW Out are very low and fully compatible with operational use. The renegotiation process has to be initiated very rarely. The number of ‘out TWs’ aircraft was found to be acceptable by all ATCOs.

The data collected during the HIL-2 experiment lead to the conclusion that TWs do not impact traffic efficiency. On the contrary, in an appropriate sector where the sector shape, size and airspace may positively impact the flight duration, an increase in traffic efficiency could be observed.

5.5.3 Capacity

The choice made in the HIL2 experiment for evaluating KPA capacity was to assess two levels of capacity, and not to progressively assess capacity growth and identify the breakpoint. The two levels of capacity were 2008 capacity and 2020 forecast capacity. The results obtained during the experiment, mainly regarding system performance, indicated that the 2020 forecast capacity was properly and safely managed.

5.5.4 Predictability

Predictability indicators are defined by SESAR as the measurement of the trajectory flown against the reference business trajectory. The same applies for the HIL-2 experiment as indicators are used for assessing the efficiency. This means TW did not impair traffic predictability, and may even improve it with an appropriate control sector shape, size and airspace structure.

6 Discussion

6.1 Controllers

The Contract of Objectives concept is manageable with the current 2008 and expected 2020 traffic loads in the two measured sectors, without any impact on traffic safety. Controllers deemed TW management to be feasible and acceptable, although TWs add some constraints when considering conflict resolution. Controllers are more constrained by the heavy traffic load than by TW use. However, TW management involves more information, which increases the perception of workload. But this increase in information is also considered a positive aspect for improving the situation.
awareness. Objectively, quantitative data reveal that TW use has no impact on the workload and situation awareness. Fundamentally, the Contract of Objectives does not modify the way the controllers work. The communication and cooperation processes are not impacted, and the common situation awareness between executive and planner controller is assessed satisfactory. Some slight changes are described for the executive when he (she) has to solve conflicts. He (she) also has to analyse the TW data to find the best solution for conflict resolution. The planner controller is not impacted. Controllers also say that it is easy to keep in mind safety as first priority. Thus, mainly in 2020 conditions, they concentrate on conflict solving, then return to ensuring TW fulfilment. There is no difficulty ensuring that safety is maintained.

6.2 Pilots
Contract of Objectives concept is manageable in the cockpit by the aircrew without any impact on safety. Pilots judged TW management to be feasible and acceptable, although TWs add some constraints as far as their management is concerned. Pilots are slightly constrained by TW use, since they modify the workload and increase collaboration between crew members. The impression of workload is increased and the quantitative data confirm this impression. However, the level of workload remains low. It is not a concern for the flying phases tested during the experiment, but the issue has to be considered for heavy workload phases or emergency situations. However, pilots are aware that safety is always the first goal, and then they are able to give up the TW constraint. Like for ATCOs, TW is also considered as positive in improving situation awareness.

Contract of Objectives does not modify the way the pilots work. The communication and cooperation processes with controllers are not impacted and the common situation awareness with controllers is perceived as being better. This is an important result in terms of the objective of increasing collaboration between crew and ground. Cooperation between the two crew members is deemed to be easy to manage with TWs, although pilots feel this will require a little more work.

6.3 System Performances Results
The results obtained in terms of system performances indicated the capacity expected in 2020 was properly and safely managed. CoO and TW concepts contribute to reaching the expected level of efficiency and predictability.

Safety data, as well as qualitative data obtained through post-run questionnaires and debriefings, are consistent in confirming that safety was not impacted by TW use for the controllers and for the pilots, even when capacity matched the 2020 forecast traffic load.

7 Conclusion
The HIL-2 experiment objectives were to assess the CoO concept and associated TWs for controllers and pilots, to investigate the impact of this concept on their activity and relationships, and to evaluate the operational acceptability from a controller’s and pilot’s point of view.

ATCOs and pilots were very positive about the concept. They all recognize that implementation of such concept will increase the collaboration between crew and ground, as they share not only the same data but also the same objective. The HIL2 experiment results are consistent with HIL1 experiment [7] as far as the ATCO topics are concerned.

This experiment was the second step in the operational assessment planned to validate the CoO concept in the CATS project. This will be followed by the third and last step, dealing with the renegotiation process and its impact on air crews and controllers.

The HIL-2 results [13] show that the CATS concept could be seen as a possible driver for implementing the SESAR Business Trajectory.

8 References
[1] Next Gen Concept of Operations v2.0, 13 June 2007, JPDO
of International Council of the Aeronautical Sciences, Anchorage.


Copyright Statement

The authors confirm that they, and/or their company or institution, hold copyright on all of the original material included in their paper. They also confirm they have obtained permission, from the copyright holder of any third party material included in their paper, to publish it as part of their paper. The authors grant full permission for the publication and distribution of their paper as part of the ICAS2010 proceedings or as individual off-prints from the proceedings.