Learning for Air Traffic Management: guidelines for future AI systems

Cocchioni, Matteo; Bonelli, Stefano; Westin, C. A. L.; Borst, C.; Hilburn, B

DOI
10.1088/1742-6596/2526/1/012105

Publication date
2023

Document Version
Final published version

Published in
Journal of Physics: Conference Series

Citation (APA)
Cocchioni, M., Bonelli, S., Westin, C. A. L., Borst, C., & Hilburn, B. (2023). Learning for Air Traffic Management: guidelines for future AI systems. *Journal of Physics: Conference Series*, *2526*(1), Article 012105. https://doi.org/10.1088/1742-6596/2526/1/012105

Learning for Air Traffic Management: guidelines for future AI systems

Learning for Air Traffic Management: guidelines for future AI systems

Cocchioni M.[1], Bonelli S.[1], Westin C.[2], Borst C.[3], Bang M.[2] and Hilburn B.[4]

[1]Deep Blue, 00185, Rome, Italy
[2]Linköping University, 581 83, Linköping, Sweden
[3]Delft University of Technology, Delft, The Netherlands
[4]Center for Human Performance Research, Voorburg, The Netherlands

Abstract. The SESAR-funded Modern ATM via Human / Automation Learning Optimisation (MAHALO) project recently completed two years of technical work exploring the human performance impacts of AI and Machine Learning (ML), as applied to enroute ATC conflict detection and resolution (CD&R). It first developed a hybrid ML CD&R capability, along with a realtime simulation platform and experimental User Interface. After a series of development trials, the project culminated in a pair of field studies (i.e., human-in-the-loop trials) across two EU countries, with a total of 35 operational air traffic controllers. In each of these two field studies, controller behaviour was first captured in a pre-test phase, and used to train the ML system. Subsequent main experiment trials then experimentally manipulated within controllers both Conformance (as either a personalised-, group average-, or optimized model) and Transparency (as ether a baseline vector depiction, an enhanced graphical diagram, or a diagram-plus-text presentation). The proposed paper presents guidelines on the design and implementation of ML systems in Air Traffic Control, derived from the results and lesson learned from the Simulations, as well as the qualitative feedback received from the controllers themselves.

1. Introduction

The SESAR-funded Modern ATM via Human / Automation Learning Optimisation (MAHALO) project recently completed two years of technical work exploring the human performance impacts of AI and Machine Learning (ML), as applied to enroute ATC conflict detection and resolution (CD&R). MAHALO started with a simple but profound question: Should we be building ML that matches the style and strategies of the air traffic controller (a construct the team termed strategic conformance) [1], or should we be building ML that is transparent and explainable to the controller? The two constructs represent contrasting approaches in aiding the human in developing an understanding of the automation workings and behavior. In turn, a better understanding can support appropriate trust and the use of automation. Moreover, the two constructs were envisioned to interact depending on the level of personalization and transparency, resulting in different user responses. These are shown in Fig. 1 below.

		TRANSPARENCY	
		Low	High
CONFORMANCE	Low	**Stupid AI:** *"It's doing a strange thing, and I don't understand why…"*	**Peculiar AI:** *"It's doing a strange thing, but I understand why…"*
	High	**Confusing AI:** *"It's doing the right thing, but I don't understand why…"*	**Perfect AI:** *"It's doing the right thing, and I understand why…"*

Figure 1. Strategic conformance and transparency can vary independently.

MAHALO first developed a hybrid ML CD&R capability (using both Supervised- and Reinforcement Learning models), along with a realtime simulation platform and experimental User Interface (UI). After a series of development trials, the project culminated in a pair of field studies (i.e., human-in-the-loop trials) across two EU countries, with a total of 36 operational air traffic controllers. In each of these two field studies, controller behaviour was first captured in a pre-test phase and used to train the strategic conformal ML system. Subsequent main experiment trials then experimentally manipulated within controllers both Strategic conformance of ML models (as either a personalised-, group average-, or optimized model) and conflict resolution advisory transparency (as either a baseline vector depiction, an enhanced graphical diagram, or a diagram-plus-text presentation). Dependent variables included objective performance / behavioural measures (e.g. acceptance of advisories) and self-reported data (rated workload, survey responses). A detailed account of the experimental setup and results has been reported in [2]. Results showed significant effects of strategic conformance on controllers' response to advisories. More specifically, we found that controllers responded more positively to an advisory if the advisory better matched the controller's preferred separation distance. We did not find any main effects of advisory transparency. Advisory transparency was found to interact with strategic conformance. With added transparency, i.e., more information about the system target separation distance in conflict resolution, it became easier for the controller to compare the advisory with his/her own preferences. A better match resulted in a more positive response.

The proposed document focuses on the insights and design guidelines of ML systems for use in Aviation and ATC in particular derived from results of the recent field studies as they addressed the impacts of strategic conformance and transparency on controller behaviour and survey responses. Special attention is dedicated to lessons learnt, in terms of both designing and evaluating ML for ATC applications.

2. Guidelines for future AI systems in ATC

The MAHALO project explored ways how ML/AI systems could be integrated into ATC tasks and what their impacts are on controller acceptance, workload and system understanding. Twenty guidelines were distilled based on empirical insights obtained from the experiments, feedback from controllers and workshop results. The guidelines are divided in five categories:

1. ML/AI design
2. Personalisation
3. Transparency
4. Human-Computer Interaction (HCI)
5. General.

2.1. ML/AI design

ML/AI techniques can offer several benefits in finding solutions to traffic problems for which no analytical solution exists by considering multiple long-term (and sometimes competing) goals. Such ML-based optimisation, however, seems more appropriate for pre-tactical phases (e.g., multi-sector planning and airspace management), featuring a high degree of uncertainty, than for tactical operations in which controllers are faced with solving ad-hoc sector perturbations featuring relatively lower degrees of uncertainties. Additionally, when it is expected that humans need to collaborate with computerised agents capable of making decisions, it is often required that the system behaves consistently and is therefore predictable, as devised by the Human-Centred Automation (HCA) school of thought. ML solutions are generally governed by probabilities and therefore less predictable than conventional deterministic CD&R algorithms.

Since the required amount of training data for ML/AI systems is often underestimated, a guideline is to design the ML/AI system at different levels of complexity, such that a fallback option is available when the highest complexity levels (in terms of a number of states and actions) appear to be unfeasible with the limited amount of training data.

Table 1. Guidelines for future AI systems in ATC: ML/AI Design

N.	Guideline
1.1	Future AI systems for ATC should investigate which ML models are best suited for balancing individual preferences and optimization approaches.
1.2	A large amount of data, collected over a longer time period, must be collected in order to facilitate ML generated personalized outputs.

2.2. Personalisation

MAHALO and its predecessor MUFASA [3] have demonstrated that personalisation helps in making an ML/AI-based advisory system more acceptable and easier to work with (e.g., faster response times). Personalisation can therefore be seen as a way to overcome some of the challenges associated with integrating ML/AI techniques in ATC from the perspective of human-machine collaboration. A prerequisite for personalisation in decision making is that sufficient intra-controller consistency and inter-controller variability exist in terms of actions/clearances. This requires a sufficiently large dataset to determine such existence before personalisation makes sense. MAHALO demonstrated that there is sufficient basis for personalisation in ATC decision making. The disadvantage of such personalisation is, however, that supervised ML/AI techniques aimed at modelling and mimicking an individual human controller could make the ATC system suboptimal. It is therefore recommended that the performance of personalised advisory systems is evaluated against target Key Performance Indicators, irrespective of sufficiently large intra-controller consistency and inter-controller variability.

Table 2. Guidelines for future AI systems in ATC: Personalisation

N.	Guideline
2.1	The development of future personalized AI systems for ATC requires end users' involvement in model development to ensure that the model captures what operators consider important for problem solving in the target task.
2.2	If ML models are to be trained on individual data, the model requires a lot of data from individuals to derive a solid and stable understanding of how that individual works, and what that individual's problem-solving preferences are. Model

development should consider the use of synthetic (i.e. generated) data for training, to augment other data sources.

2.3 A suitable individual preference parameter for personalizing CD&R systems in conflict resolution choices is target separation distance.

2.4 Future ATC systems that are more personalized may lessen the need for them being transparent. A personalized system does not require high transparency, it reduces the need for transparency.

2.5 Individual preferences for parameters considered in conflict resolution decisions can be expected to vary between controllers. E.g. aircraft choice appears important for some but not others.

2.6 Controllers are more likely to accept and agree with a personalized system that adapts its recommendations to the individual's preferences.

2.7 There is no added benefit to acceptance or agreement of conflict resolution advisories that are shaped after the group of controllers' preferences in terms of aircraft type, resolution direction, intervention time, and separation distance.

2.8 Future ATC systems should explore personalization mechanisms to benefit system acceptance and agreement.

2.9 Future ATC systems should acknowledge and embrace in the design that controllers differ in their conflict resolution preferences.

2.10 Future ATC systems should consider personalized applications when possible (i.e., taking into account a safety risk assessment).

2.11 Decision support systems capable of providing advisories/recommendations on actions should do so before the operator has made a decision on how to act (note that this can be before the action is implemented).

2.12 What aspects of a system that should be personalized should be driven by the operator's individual preferences in working and problem solving, and in what regards the operator is consistent over time.

2.13 Conformal (fully personalized) advisories should not be the main objective of future AI systems in ATC. That is, the system should not aim to only mimic human behaviour or decision making. Future systems should aim to optimize solutions but consider the individual operator's preferences and adjust the solution when feasible and appropriate. If the system goes against the individual's preferences, the system should be able to provide an explanation for why the system believes its solution to be better than the individual's.

2.3. Transparency

Similar to personalisation, transparency offers a way to increase the acceptance and understanding of (ML-based) advisory systems in ATC. MAHALO undertook an ecological approach in operationalising ML transparency by putting the emphasis on interpretable (visual) representations (here, Solution Space Diagram) rather than explainable ML models found in XAI fields. Controllers seemed to appreciate this approach as it puts ML solutions in the context of the problem that needs to be solved (i.e., traffic conflict). The same representation also served as a decision support tool, allowing controllers to formulate their own solutions and/or nudge the advised ML solution. This raises the discussion on what an operational controller might want and need to understand about the automated system and to what extent. For example, an ATCo might not need a deep insight into ML neural networks at the level of a ML system developer. Additionally, transparency needs are also affected by workload demands – in time-critical situations, an ATCo generally prefers to receive any workable solution and may not want to waste valuable cognitive resources in trying to understand that solution by digging through layers of information. In such cases, information on the resulting aircraft separation targeted by the advisory would be sufficient. Thus, given that transparency needs are likely context dependent and sensitive to operator preferences, we recommend an adaptive approach that allows ATCos to put machine decisions into context and lets them decide upon what they wish to see and when. Note that such an adaptive approach can be regarded as another form of personalisation, namely one that focuses on the preferred information that one wishes to see.

Table 3. Guidelines for future AI systems in ATC: Transparency

N.	Guideline
3.1	Future AI systems for ATC should focus on applying increased transparency for situations where the human and system work differently, and/or where the human has difficulties understanding the system. The need for transparency, and expected benefits, is higher for situations when system behaviour and advisories (e.g. its output) are different from how the human operator prefers to work and solve problems. The need for transparency, and expected benefits, is higher for situations where the user does not understand system behaviour and advisories (e.g. its output).
3.2	Ecological interface design approaches can be used to increase the transparency of presented CD&R advisories by providing information on the constraints and solution possibilities affecting the control problem.
3.3	Future AI systems for ATC should investigate how transparency approaches can be used to improve system design. Increased transparency can benefit understanding of e.g. the system and/or situation, but does not necessarily benefit acceptance of a system and agreement with its advisories. The effect might be the opposite, where increased transparency decreases acceptance and agreement.
3.4	Increased transparency supports a better understanding or e.g., the system, its output, and/or the situation. With improved understanding, the operator can better determine if system behaviour or advisories are appropriate for the problem at hand. The use of the system partly depends on how the system's behaviour or advisory matches the preferences of the individual operator. As such, increased transparency can reveal to what extent the AI and operator work to solve problems similarly or differently.
3.5	Transparency should ideally be individually tailored to facilitate a dialogue between humans and AI. Transparency should be provided in relation to what the user needs to understand, which requires the AI to be able to develop an understanding of the human it is interacting with.

2.4. Human Computer Interaction

We believe that interaction flexibility is important for ATCo engagement. When humans are expected to play a central role in the system and bear the final responsibility over the safety of operations, human interaction with computerised systems is of paramount importance. That is, bearing responsibility without having authority is not the best position to be in. MAHALO showed one possible way of facilitating interaction by embedding it into existing controller tools. Via a conventional clearance menu, ATCos could not only accept, nudge, or change machine advisories, but also reject them and work with any other aircraft than the one receiving the advisory. Such flexibility was generally appreciated by ATCos as they felt empowered to influence the system in any way they preferred. We believe that offering such flexibility outweighs the (slight) performance decrements that could arise when ATCos change an optimal advisory into a suboptimal one. Note that affording flexibility in interaction can also be regarded as a form of personalisation.

2.5. General

Future ATC systems considering human-machines working together, should acknowledge that what is an optimal solution to a problem depends on the individual human (e.g., preferences, physiological state, stress, fatigue, etc).

3. Discussion

Taken together, the MAHALO project provided several guidelines on the importance and how to incorporate strategic conformance and transparent mechanisms of AI solutions to CD&R in particular, and to problem solving tasks in safety critical systems in general. We concluded twenty design guidelines across five categories. In relation to the hypothesized interaction between strategic conformal and transparent automation, the MAHALO results point to the highest benefits of highly conformal automation (pesonalised) with low transparency. In Fig.1, this is referred to as Confusing automation: "It's doing the right thing, but I don't understand why…" It turns out that as long as the system is doing the right thing, transparency is not required.

The traditional system design approach is to design a "one-size-fits-all" system that all operators have to conform to. This might be appropriate for many standard operating procedures and situations that easily can be optimized. But it makes less sense, however, for time and safety critical situations where the definition of what is an optimal solution is subjective and performance depends more on the operator's capacity to cope with the tasks (and avoid peaks of workload and stress). MAHALO has shown that a person's response to a resolution advisory partly depends on how close that advisory is to the person's solution preferences. Future ATC systems should acknowledge and embrace in the design that controllers differ in their conflict resolution preferences. And future systems should explore strategic conformal approaches for improving human-AI collaboration. However, a system's sole objective should not be to conform to the individual's problem solving preferences. We should seek the best, most efficient, and safest solution when that can be defined. But when appropriate, e.g., when what is optimal is ambiguous, or human acceptance and trust is critical, systems should consider the individual's preferences. Finally, transparency should be individually tailored to facilitate a dialogue between humans and AI. In other words, the need for an explanation of system behavior or reasoning is most valuable when there is a disagreement between the system and human, and the human does not understand the automation output or its behavior. . This dialogue can be adaptive – i.e. initiated by the automation sensing disagreement or that the operator is trying to understand the system, or adaptable – i.e., human initiated – e.g. asking the system "why did you propose that?"

We conclude that future AI systems should support individual controller preferences when generating resolution advisories and when explaining why that solution is advised, but more research is needed on how to implement strategic conformal automation and how it affects user responses. More research is needed to explore the potential utility of personalisation, or tuneable parameters (e.g., target separation distance) that might allow for a hybrid of the optimized and personal model view. Such that controllers could tune certain parameters (and within a certain range) within the confines of an optimal advisory system. Should we strive for personal or optimal solutions? We do not need to choose but could explore both. Personal solution might be suitable for some tasks or situations where the degrees of freedom are larger and an objective optimal does not exist or is very difficult to define. Future research is required to study in what contexts and situations personalized approaches are suitable. Finally, future research is required to explore the potential benefits of advisory transparency on advisory acceptance and system trust What should be made transparent or explainable – what are the needs that operators have? Transparency should be provided in relation to what the user needs to understand, which requires the AI to be able to develop an understanding of the human it is interacting with.

References

[1] Westin, C., Borst, C., & Hilburn, B. (2016). Strategic Conformance: Overcoming acceptance issues of decision aiding automation? IEEE Trans. Human–Mach. Syst., 46(1), 41-52. doi:10.1109/THMS.2015.2482480

[2] C. Westin, B. Hilburn, C. Borst, E-J. van Kampen, T. Nunes, S. Boonsong, M. Cocchioni, and S. Bonelli. Personalizerd and transparent AI support for ATC conflict detection and resolution: an empirical study. In 12th SID, Dec. 5-8, Budapest, Hungary.

[3] Hilburn, B., Westin, C., & Borst, C. (2014). Will controllers accept a machine that thinks like they think? The role of strategic conformance in decision aiding automation. Air Traffic Control Quarterly, 22(2), 115-136.