



HAL
open science

Evaluation de la performance individuelle en combat aérien dans le Serious Gaming

Tara Potin, Julien Donnot

► **To cite this version:**

Tara Potin, Julien Donnot. Evaluation de la performance individuelle en combat aérien dans le Serious Gaming. ERGO'IA 2023, Oct 2023, Bidart, France. hal-04405060

HAL Id: hal-04405060

<https://hal.science/hal-04405060>

Submitted on 19 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Evaluation de la performance individuelle en combat aérien dans le *Serious Gaming*

Tara Potin

Faculté des STAPS
Bordeaux, FRANCE
tara.potin@etu.u-bordeaux.fr

Julien Donnot

Centre d'Expertise Aérienne Militaire
Mont-de-Marsan, FRANCE
julien.donnot@intradef.gouv.fr

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright is held by the owner/author. ERGO'IA 2023, October 10--13, 2023, Bidart, France. ISBN: 979-8-4007-0910-4

Résumé

Contributing to the design of the Future Air Combat System (SCAF) by preparing human/AI collaboration in aerial operations is one of the many challenges of artificial intelligence within the French Air and Space Force. To ensure the optimal use of AI, it is essential to acquire the ability to assess individual and collective performance. This study proposes a methodology for evaluating and understanding individual performance in simulated and sporting aerial combat. The analyses focused on classifying different missile shots, particularly through replays of aircraft and missile trajectories. Results confirm the effectiveness of the method in assessing individual performance both in relation to match outcomes and to match with the concept of aggressiveness. The discussion addresses the study's limitations, possible method developments integrating collaboration with an AI, and its transferability to operational aerial combat.

Mots-clés

Performance ; simulation ; aviation militaire ; évaluation ; combat aérien.

Introduction

Les technologies d'Intelligences Artificielles (IA) sont, de nos jours, au cœur de tous les enjeux aussi bien sociaux, politiques que économiques. L'IA est un

concept relativement englobant et plus ou moins abstrait dans lequel un grand nombre de technologies différentes sont regroupées. Il peut aussi bien s'agir de simples algorithmes que de très performants assistants virtuels. Les premiers n'ont pas vocation à communiquer avec l'humain, contrairement aux seconds cités. Bien que principalement développées dans le domaine civil, ces technologies sont destinées à être intégrées dans le monde militaire [1].

En effet, le développement de l'IA est d'autant plus important au sein de l'Armée de l'Air et de l'Espace que le Système de Combat Aérien du Futur (SCAF) est pensé sur le principe de la collaboration humain/IA [2]. Ce projet a pour ambition de concevoir un système des systèmes dont le principe est de mener le combat collaboratif dans lequel l'IA devrait être totalement intégrée.

L'IA destinée à servir dans un contexte de combat aérien soulève plusieurs questions. Tout d'abord, sa performance ne peut pas être exclusivement considérée sur la base de ses capacités mais plutôt par la manière dont elle sera utilisée par l'humain. Cet état de fait conditionne directement son développement et par conséquent, la nature des données sur lesquelles se fonde son apprentissage. Ensuite, dans un contexte de combat aérien, les décisions à prendre et les actions à mener dépendent des activités de l'adversaire (e.g., nombre, localisation) ainsi que des variations de l'environnement (e.g., météo, activités civiles). Pour assister au mieux l'équipage, l'IA de combat doit être capable d'identifier les informations essentielles, de les traiter rapidement et de proposer des solutions à l'équipage en affichant le niveau de fiabilité si possible le plus élevé possible pour ainsi garantir une parfaite adaptation à l'environnement changeant.

De ce fait, pour pouvoir ajuster le comportement d'une IA à l'activité d'un équipage de combat, il s'avère primordial d'analyser la performance combinée de ce

couple humain/IA. Cette analyse consiste à identifier les déterminants de la performance pour la comprendre. Dans la présente étude, il nous est apparu judicieux de nous concentrer sur l'évaluation de la performance de l'humain seul pour ensuite être en mesure d'établir les comparaisons avec la performance combinée.

Développer une méthode d'analyse de la performance sur une activité de combat aérien collaboratif qui n'existe pas encore relève de l'inaccessible. Même si chaque jour des dizaines de pilotes de chasse s'entraînent au combat aérien actuel, cette activité opérationnelle est, d'une part, très difficile d'accès et d'autre part, elle dépend d'un nombre tellement important de variables que cela la rend trop complexe pour être analysée correctement. Par contre, il est de plus en plus fréquent d'observer du combat aérien dans un domaine tel que celui du *serious gaming*. C'est la raison pour laquelle nous avons opté pour le développement d'un premier modèle d'analyse de la performance de l'humain en combat aérien simulé. Notre support sera de suivre et d'évaluer la performance d'une équipe participant à la compétition internationale *SATAL 2023 (Squadron-Air-To-Air-League)* [3]. En effet, notre étude profite de l'opportunité offerte par une entité opérationnelle de l'Armée de l'Air et de l'Espace. La question de savoir comment évaluer la performance d'un pilote avec une IA a été posée au laboratoire Facteurs Humains et Opérations Aériennes du Centre d'Expertise Aérienne Militaire de Mont-de-Marsan.

La compétition *SATAL* est une compétition de combat aérien en *eSport* ouverte à tout public, qui se joue sur le logiciel *DCS World (Digital Combat Simulator World)* [4]. Un match permet l'affrontement de deux équipes, chacune composée de quatre avions de chasse. Chaque patrouille décolle depuis son camp et rejoint ensuite la zone de combat. Cette zone de combat se matérialise par une sorte de bulle dans laquelle, une fois les avions

rentrés ils ne peuvent plus sortir sous peine d'être détruits et donc éliminés. A mesure que le temps s'écoule, la taille de la bulle diminue et rend la survie de plus en plus difficile. Cette spécificité de la bulle est une différence majeure en comparaison d'une activité opérationnelle réelle. Pour gagner, il faut détruire les avions de l'autre équipe. Pour les détruire, plusieurs choix s'offrent aux joueurs : choix de l'avion, choix des missiles, choix de stratégie... La Figure 1 présente la vue disponible à un joueur lors d'un match.



Figure 1: Vue cockpit d'un mirage 2000C dans DCS World.

L'objectif de la présente étude est de proposer un modèle d'évaluation de la performance en combat aérien en nous appuyant sur l'analyse des comportements individuels observés au sein d'une équipe participant à la compétition *SATAL* 2023.

Cadre théorique

Pour orienter l'analyse comportementale, nous avons procédé à l'adaptation d'un modèle générique d'évaluation de la performance développé en interne, le modèle DCP. Selon ce modèle (cf. Figure 2), la performance résulte de la sélection et de la réalisation de gestes techniques en conformité avec une tactique définie, elle-même résultant de la définition d'une stratégie. La stratégie se distingue de la tactique dans le sens où elle fixe un état final à atteindre, là où la

tactique définit les mécanismes par lesquels cet état sera obtenu. Dans le cas de *SATAL*, une stratégie consisterait à chercher à éliminer son adversaire en le repoussant hors de la bulle (et non en lui tirant dessus directement) et ainsi la tactique consisterait à manœuvrer son avion de manière à rester le plus possible au centre de la bulle quitte à être plus vulnérable.

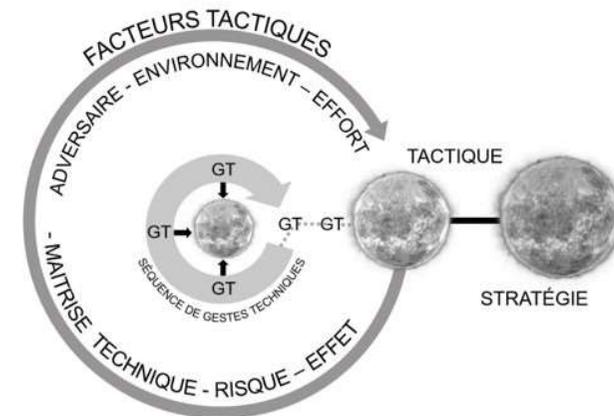


Figure 2: Modèle DCP de l'analyse de la performance sportive.

Pour exploiter le modèle il faut le faire correspondre à l'activité concernée, c'est-à-dire identifier les actions rattachées aux tactiques. Après une première analyse de matchs précédents, nous avons centré nos réflexions sur la procédure de tir. Le nombre de missiles étant limité, les joueurs sont poussés à adopter des tactiques tantôt agressives (ou offensives), tantôt défensives. Cela nous a permis de dresser une liste des différents types de tirs à l'issue desquels un avion est détruit ou repoussé. Trois classes ont été établies en fonction de la destruction effective de l'avion adverse et de l'intention du joueur au moment du tir. La classe A inclut les tirs de performance haute à savoir, ceux qui aboutissent à la destruction de

l'adversaire et conforme à l'intention du joueur au moment du tir. La classe B regroupe les tirs de performance moyenne à savoir, ceux qui aboutissent également à la destruction de l'adversaire mais sans intention initiale du tireur. Enfin, la classe C regroupe les tirs de basse performance, à savoir ceux qui ne font que repousser l'adversaire sans mener à sa destruction. Le tir ami (friendly kill) qui consiste à détruire l'avion d'un co-équipier est considéré à part car il relève d'une mauvaise performance voire handicapante pour son équipe. Le Tableau 1 détaille cette liste.

Classification	Type de tir	Indice Ki
Classe A	Perfect Kill (PK)	K++ i+
	Bubble Out Crash (BOC)	K+ i+
	Crash	K+ i+
Classe B	R Max Kill (RMK)	K++ i-
	FKT Autre	K++ i-
	Failed Perfect Crash (FPC)	K+ i-
	Turn Away Crash (TAC)	K+ i-
Classe C	Turn Away Miss (TAM)	K- i+
	Failed Kill Miss (FKM)	K- i-
Hors classification	Friendly Kill	K++ i-

Tableau 1: Classification des tirs missile. L'indice KI indique une destruction directe par un missile (K++), une destruction indirecte comme une collision ou une sortie de bulle (K+), un évitement du missile (K-), une intention initiale conforme au résultat du tir (I+) ou non conforme (I-).

Les facteurs tactiques du modèle DCP constitueront les éléments à renseigner pour confirmer la correspondance d'une séquence de jeu avec un des tirs de la liste. Par exemple, le crash d'un avion avec le sol pourrait être causé dans certains cas par un manque de maîtrise de la trajectoire et dans d'autres cas serait inévitable suite à la pression imposée par le tir d'un adversaire. Malgré tout, le modèle DCP ne permet pas d'accorder un barème de points aux différents types de tirs.

Le modèle d'Ibañez et al. (2019) [5] affiche une certaine compatibilité avec le modèle DCP, notamment en tenant compte de la prise de décision, de l'exécution technique et de l'efficacité finale. Les auteurs ont mis en place un système de décompte de points dans le domaine du Basket Ball sur la base de ces trois composantes pour chaque action de jeu. Un système de décompte de points inspiré de celui d'Ibañez et al. (2019) permet une quantification numérique de la performance.

Méthode

L'étude a consisté à analyser les données des matchs réalisés par l'équipe « Raptors & Aggressors Alliance ». Il s'agit de personnel militaire issu de différentes spécialités mais dont aucun membre n'est pilote de chasse. Au total, douze matchs seront examinés.

Chaque tir de chaque joueur sera identifié et classé grâce au logiciel Tacview [6]. Ce logiciel permet de rejouer la totalité du combat en visualisant les trajectoires des avions ainsi que celles des missiles tirés (cf. Figure 3) tout en disposant des paramètres de vols (e.g, altitude, hauteur, vitesse, cap...)

L'agressivité sera évaluée par la combinaison du pourcentage d'occupation de l'espace aérien adverse, le déclenchement du premier tir du match et enfin, le fait que le pilote, au moment de tirer, a une position de domination (altitude de vol plus haute) face à son adversaire.



Figure 3: Rejeu d'une mission sur le logiciel Tacview.

Résultats

Au total, 491 tirs ont été analysés au cours de 18 combats aériens. La « Raptors & Aggressors Alliance » comptabilise seize défaites pour deux victoires. Parmi les dix tirs possibles, seuls le PK, le BOC, le FPC, et le FKM ont été réalisés. On observe logiquement que dans les combats gagnés, les équipes effectuent significativement plus de PK et de tirs réussis (PK, BOC et FPC) et moins de FKM que dans les combats perdus (p -value < 0,001). Dans la majorité des combats (82%), la stratégie préférentielle a consisté à détruire les aéronefs par du PK. Mais dans une minorité de cas, une autre stratégie a été appliquée en s'appuyant sur le BOC (6%) ou la stratégie n'a pas pu être correctement appliquée mais l'objectif a été atteint (FPC, 6%).

En ce qui concerne la notion d'agressivité, les résultats montrent qu'un plus grand pourcentage de temps d'occupation de la zone aérienne adverse n'assure pas une plus grande probabilité de victoire. De plus, les tests statistiques ne montrent pas de différence significative entre le nombre de tirs réussis en position de domination et de non domination. Enfin, l'équipe de la « Raptors & Aggressors Alliance » a déclenché le premier tir pour 89% des combats. Sur les deux combats victorieux, ils ont déclenché le premier tir dans 100% des cas et dans 88% des cas pour les seize combats perdus.

Concernant l'environnement de combat, les tests n'ont pas révélé de différence significative entre les vols de jour (12 combats) et ceux de nuit (6 combats).

Discussion

L'obtention d'un score et d'un classement à l'issue d'un match fait que la performance globale est déjà évaluée. Le travail présenté a surtout consisté à caractériser la performance grâce à la classification des tirs. Sur la totalité des tirs, seulement 11% n'ont pas pu être classés. Cette absence de classification provient de la difficulté de capturer dans certains cas l'intention du pilote par les communications. Sachant que l'absence de communication résulte d'un problème technique, la méthode n'est pas directement en cause. Pour autant, il est difficile de juger de l'efficacité globale de la méthode concernant la population et l'activité visée en raison du manque de variabilité des données. Suivre un plus grand nombre d'équipes participant à SATAL 21 permettrait d'obtenir plus de variabilité. En effet, des équipes mieux classées exploitent probablement des stratégies se traduisant par la réalisation des autres tirs identifiés dans la classification.

Pour envisager le transfert de cette méthode à une activité réelle de combat rapproché, il est nécessaire d'identifier ses différences avec celle de la compétition SATAL. La principale différence réside dans l'absence de bulle dans l'activité réelle, dans le sens où le choix des tactiques et des stratégies des pilotes dépend directement de cette contrainte. Dans SATAL, la stratégie d'amener l'adversaire à être détruit en le poussant à l'extérieur de la bulle n'est pas possible en activité réelle. D'ailleurs, il est important de distinguer l'objectif, qui détermine un état final recherché objectivable, de la stratégie et de la tactique. La stratégie correspond à une logique d'enchaînement d'événements dont le dernier événement amène à l'état final recherché, donc l'atteinte de l'objectif. La tactique regroupe les actions à mener pour provoquer les événements attendus. Le fait que les événements sont

prévisibles et les actions sont observables, permet de vérifier si la tactique sert correctement la stratégie.

L'analyse des actions mène à la caractérisation de la performance c'est-à-dire, à identifier les indicateurs pertinents pour permettre au pilote de percevoir les bonnes informations au bon moment et ainsi, prendre les bonnes décisions. Certaines des informations, pourtant disponibles, sont difficilement exploitables en temps réel par l'humain (e.g, l'altitude, la vitesse et la distance de l'équipier privilégié). Néanmoins, elles revêtent un caractère crucial du fait de leur impact sur la survivabilité du pilote. Le Centre d'Expertise Aérienne Militaire (CEAM) développe actuellement la brique IA *Alphawingman* qui consiste à synthétiser l'ensemble des informations inexploitablement par l'humain en un indice de survie en temps réel beaucoup plus pertinent pour le cerveau. De plus, l'expérimentation présentée ici s'inscrit dans une étude plus large visant à comparer les performances du pilote avec et sans l'aide de l'IA. Les résultats de cette étude détermineront l'usage spécifique de cette assistance. En effet, soit l'IA endosse temporairement un rôle d'instructeur en aidant le pilote à maîtriser l'interprétation de données complexes, soit le pilote se repose en permanence sur cette assistance au risque de devenir dépendant. Cette problématique pose la question de l'évolution de l'expertise pour laquelle le pilote acceptera à l'avenir de déléguer au système le traitement d'informations déterminantes pour sa performance.

Références

- [1] Elysée. 2023. Transformer nos armées : le Président de la République présente le nouveau projet de loi de programmation militaire. [<https://urlz.fr/IHk3>].
- [2] Défense Nationale. 2021. Europe de la défense : l'émergence d'une autonomie stratégique. [<https://www.defnat.com/e-RDN/vue-article-cahier.php?carticle=106>].
- [3] DCS World Events. 2023. SATAL23 - Survivor Series. [<https://forum.dcs.world/topic/310311-satal23-survivor-series-by-dcsworldevents/#comment-5208849>].
- [4] Digital Combat Simulator. [<https://www.digitalcombatsimulator.com/fr/>].
- [5] Ibáñez, S. J., Martínez-Fernández, S., Gonzalez-Espinosa, S., García-Rubio, J., and Feu, S. 2019. Designing and Validating a Basketball Learning and Performance Assessment Instrument (BALPAI). *Frontiers in Psychology* 10, 1595 (2019), <https://doi.org/10.3389/fpsyg.2019.01595>.
- [6] Tacview. [<https://www.tacview.net/>].

Remerciements

Les auteurs tiennent à remercier le commandant Frédéric Coubard pour son investissement à tous les niveaux de l'étude, aussi bien dans la préparation du sujet que pour l'acculturation des auteurs, la mise à disposition des données et bien évidemment ses conseils avisés.