

**Procédé d'évaluation de la conformité d'un système de pistage à l'aide d'un
nombre restreint de mesures et dispositifs associés**

5 **DOMAINE DE L'INVENTION**

La présente invention concerne un procédé d'évaluation de la conformité d'un système de pistage à un ensemble d'exigences. La présente invention se rapporte également à un module d'évaluation permettant la mise en œuvre du procédé et un système de pistage comportant un tel module d'évaluation.

10

ARRIERE-PLAN TECHNOLOGIQUE DE L'INVENTION

L'invention proposée se situe dans le domaine du contrôle de trafic aérien.

Le contrôle de l'espace aérien est assuré par un suivi de l'ensemble des aéronefs. Il s'agit de superviser le trafic aérien pour prévenir les collisions entre aéronefs et contrôler la circulation, tant en vol de croisière qu'aux abords des aéroports – décollage et atterrissage.

15

Pour mettre en œuvre un tel contrôle, les contrôleurs aériens utilisent un système de surveillance du trafic aérien basé sur un pistage des avions qui s'appelle un système de pistage. Le système de pistage est capable en temps réel d'estimer à partir de données provenant de capteurs la meilleure estimation possible de la position, du cap et de la vitesse de chaque aéronef dans une région d'information de vol donnée.

20

La région d'information de vol est souvent désignée par l'acronyme FIR renvoyant à la dénomination anglaise correspondante de « *Flight Information Region* ».

L'ensemble des données estimées forment une piste.

25

La qualité de l'estimation de la piste (pistage) varie en fonction du paramétrage du système de pistage, de la qualité des données acquises par les capteurs ou encore de données environnementales comme la topologie du terrain ou la météo.

A ce titre, il est souhaitable pour un contrôleur aérien de connaître la qualité de l'estimation des pistes reconstruites par le système de pistage.

30

Pour cela, il est connu d'utiliser les exigences de la norme ESASSP qui est détaillée dans le document intitulé « *EUROCONTROL Specification for ATM Surveillance System Performance* » (Volume 1) dont l'ISBN est 978-2-87497-022-1 et qui a été publié en mars 2012.

35

Plus précisément, il est évalué un ensemble d'exigences définis par cette norme ESASSP par comparaison entre les sorties du système de pistage et une estimation de la trajectoire idéale des aéronefs.

Toutefois, le calcul des exigences de la norme impose d'obtenir une grande quantité de points de mesure, typiquement entre 50 000 et 100 000.

Cela pose un problème en pratique pour obtenir une évaluation à la volée et notamment détecter les éventuelles défaillances en cours de fonctionnement.

5

RESUME DE L'INVENTION

Il existe donc un besoin pour un procédé permettant d'évaluer la qualité du pistage effectuée par un système de pistage avec un nombre réduit de mesures.

10 A cet effet, la description décrit procédé d'évaluation de la conformité d'un système de pistage à un ensemble d'exigences, au moins une exigence étant une exigence obligatoire et au moins une exigence étant une exigence recommandée, chaque exigence imposant qu'une grandeur physique soit supérieure ou égale à un seuil ou qu'une grandeur physique soit inférieure ou égale à un seuil, le procédé d'évaluation étant mis en œuvre par ordinateur, le procédé d'évaluation comportant au moins les étapes suivantes :

- 15 - acquisition d'un ensemble de valeurs, l'ensemble de valeurs comprenant les valeurs accessibles au système de pistage sur un intervalle de temps prédéfini,
- pour chaque exigence, calcul d'une valeur de robustesse, la robustesse étant calculée comme la valeur p d'un test statistique indiquant si l'exigence considérée est respectée ou non,
- 20 - pour chaque exigence, comparaison de chaque valeur de robustesse à un seuil prédéfini selon un critère de comparaison respectif, et

- détermination des exigences auxquelles le système de pistage est conforme, les exigences déterminées comme conformes étant les exigences pour lesquelles, à l'étape de comparaison, il a été déterminé que le critère de comparaison est respecté.

25 Selon des modes de réalisation particuliers, le procédé d'évaluation présente une ou plusieurs des caractéristiques suivantes, prise(s) isolément ou selon toutes les combinaisons techniquement possibles :

- le procédé comporte, en outre, pour chaque valeur de robustesse ne respectant pas le critère de comparaison, une étape de détermination de la cause du non-respect du critère de comparaison, la cause étant soit que l'exigence n'est pas respectée soit que le nombre de valeurs est trop faible pour garantir un seuil de fiabilité prédéfini d'évaluation de la conformité à l'exigence considérée.
- 30 - lorsque la cause déterminée est un nombre de valeurs trop faible, le procédé comporte une étape de déduction d'un nombre de valeurs supplémentaires à acquérir pour obtenir une fiabilité d'évaluation supérieure au seuil de fiabilité prédéfini.
- 35

- le procédé comporte une répétition des étapes d'acquisition, de calcul, de comparaison jusqu'à ce que le nombre de valeurs soit permette d'obtenir une fiabilité d'évaluation supérieure au seuil de fiabilité prédéfini soit atteigne un nombre maximal prédéfini avec une fiabilité d'évaluation inférieure au seuil prédéfini.
- 5 - le procédé comporte, en outre, une étape d'émission d'une alerte pour les exigences non respectées ou pour les exigences dont le nombre de valeurs des étapes d'acquisition, de calcul, de comparaison a atteint le nombre maximal prédéfini avec une fiabilité d'évaluation inférieure au seuil prédéfini.
- le nombre maximal est compris entre 40000 et 60000.
- 10 - lors de l'étape de calcul, il est calculé une intégrale d'une fonction bêta incomplète régularisée.
- lors de l'étape de calcul, une p-valeur est égale au résultat de l'intégrale et une autre p-valeur est une fonction linéaire du résultat de l'intégrale.
- lors de l'étape de calcul, il est calculé une fonction de répartition de la distribution normale centrée réduite.
- 15 - le procédé comporte, en outre :
 - une étape d'application d'une fonction d'évaluation propre à chaque exigence aux valeurs p pour obtenir des valeurs d'évaluation, et
 - une étape d'utilisation des valeurs d'évaluation pour obtenir un score global évaluant la conformité du système de pistage à l'ensemble des exigences.
- 20 - le nombre d'exigences est supérieur ou égal à 20.
- le système de pistage fournit des estimations de mesures à partir de données antérieures issues de plusieurs capteurs, l'ensemble de valeurs comprenant les estimations de mesure.
- 25 La description décrit aussi un module d'évaluation de la conformité d'un système de pistage à un ensemble d'exigences, au moins une exigence étant une exigence obligatoire et au moins une exigence étant une exigence recommandée, chaque exigence imposant qu'une grandeur physique soit supérieure ou égale à un seuil ou qu'une grandeur physique soit inférieure ou égale à un seuil, le module d'évaluation étant propre à :
 - 30 - acquérir un ensemble de valeurs, l'ensemble de valeurs comprenant les valeurs accessibles au système de pistage sur un intervalle de temps prédéfini,
 - pour chaque exigence, calculer une valeur de robustesse, la robustesse étant calculée comme la valeur p d'un test statistique indiquant si l'exigence considérée est respectée ou non,
 - 35 - pour chaque exigence, comparer chaque valeur de robustesse à un seuil prédéfini selon un critère de comparaison respectif, et

- déterminer des exigences auxquelles le système de pistage est conforme, les exigences déterminées comme conformes étant les exigences pour lesquelles, à l'étape de comparaison, il a été déterminé que le critère de comparaison est respecté.

La description propose également un système de pistage comprenant un module d'évaluation.

Dans la présente description, l'expression « propre à » signifie indifféremment « adapté pour », « adapté à » ou « configuré pour ».

BREVE DESCRIPTION DES DESSINS

Des caractéristiques et avantages de l'invention apparaîtront à la lecture de la description qui va suivre, donnée uniquement à titre d'exemple non limitatif, et faite en référence aux dessins annexés, sur lesquels :

- la figure 1 est une représentation schématique d'un exemple de système de pistage en interaction avec un ensemble de capteurs,
- la figure 2 est un ordinogramme d'un exemple de mise en œuvre d'un procédé d'évaluation de la conformité du système de pistage de la figure 1 à un ensemble d'exigences, et
- la figure 3 illustre graphiquement la variation de la p-valeur en fonction de la valeur observée pour une exigence donnée à titre d'exemple.

DESCRIPTION DETAILLEE DE MODES DE REALISATION PREFERES

CAS GENERAL

La figure 1 illustre schématiquement un système de pistage 10 en interaction avec un ensemble 12 de capteurs.

Le système de pistage 10 est propre à observer un espace aérien pour déterminer les trajectoires des aéronefs passant dans un espace observé.

Cela rend possible d'effectuer un contrôle du trafic aérien dans l'espace considéré.

Pour cela, le système de pistage 10 est propre à collecter des données provenant de l'ensemble des capteurs et à analyser les données collectées.

Le système de pistage 10 émet en sortie des données calculées.

Les données calculées sont, par exemple, des trajectoires d'aéronefs dans l'espace observé, c'est-à-dire des données de position, de vitesse ou de cap pour des aéronefs passant dans l'espace observé.

Le système de pistage 10 forme avec l'ensemble 12 de capteurs un système de surveillance du trafic aérien.

Un tel système est souvent désigné sous l'acronyme SSTA.

L'ensemble 12 de capteurs est propre à obtenir des données dans l'espace observé.

Selon l'exemple décrit, l'ensemble 12 de capteurs comporte une unité ADS 14, une unité WAM 16 et un radar 18.

5 L'unité ADS 14 est un système de surveillance coopératif pour le contrôle du trafic aérien et d'autres applications connexes. Un avion équipé d'une unité ADS 14 détermine sa position par un système de positionnement par satellite (GPS) et envoie périodiquement cette position et d'autres informations aux stations au sol.

10 L'abréviation ADS renvoie à la dénomination anglaise correspondante de « *Automatic Dependent Surveillance* » signifiant littéralement « *surveillance de dépendance automatique* ».

Une telle unité ADS 14 est parfois aussi appelée unité-B, l'abréviation ADS-B renvoyant à la dénomination anglaise correspondante de « *Automatic Dependent Surveillance-Broadcast* » signifiant littéralement « *surveillance de dépendance automatique-multidiffusion* ».

15 Une unité WAM 16 utilise les données de plusieurs capteurs pour obtenir la localisation d'un avion.

L'abréviation WAM renvoie à la dénomination anglaise correspondante de « *Wide Area Multilateration* » signifiant littéralement « *multilatéralisation à zone étendue* » et désigne une technologie de surveillance d'avions basée sur le principe de la différence de temps d'arrivée qui est utilisée dans un aéroport.

Par exemple, l'unité WAM 16 collecte les données provenant de plusieurs antennes au sol pour appliquer des calculs mathématiques permettant d'obtenir la position de l'avion.

25 Un radar 18 permet de détecter la présence d'avions dans le ciel et de déterminer leur position. Le radar 18 émet des impulsions électromagnétiques dans le ciel, et la détection et la localisation d'un avion sont obtenues par l'analyse de l'onde réfléchie par l'avion et retransmise dans la direction du radar 18.

Selon l'exemple de la figure 1, le système de pistage 10 comporte un module d'analyse 20, un module d'évaluation 22.

30 Le module d'analyse 20 est propre à analyser les données des capteurs pour estimer des nouvelles données.

Le module d'analyse 20 reçoit ainsi régulièrement des données relatives aux avions de l'environnement observé, notamment leur altitude, leur position horizontale ou leur vitesse.

35 En particulier, le module d'analyse 20 est propre à estimer la trajectoire d'un aéronef.

Un tel module d'analyse 20 est souvent désigné sous la dénomination anglaise de « *tracker* » signifiant littéralement « *suiveur* ».

Le module d'analyse 20 est également propre à estimer des mesures à partir de données antérieures issues de l'ensemble 12 de capteurs.

5 Le module d'évaluation 22 est propre à mettre en œuvre les étapes d'un procédé d'évaluation de la conformité du système de pistage 10 à un ensemble d'exigences.

Selon l'exemple décrit, le module d'évaluation 22 comprend une sous-unité de calcul 24 et une sous-unité d'alerte 26 dont les rôles apparaîtront dans la suite de la description.

10 Un exemple de mise en œuvre du procédé d'évaluation est maintenant décrit en référence à la figure 2.

Le procédé d'évaluation vise à évaluer la conformité du système de pistage 10 à un ensemble d'exigences.

15 Parmi cet ensemble d'exigence, au moins une exigence est une exigence obligatoire et au moins une exigence est une exigence recommandée.

Par définition, une exigence obligatoire est une exigence dont le non-respect entraîne la disqualification du système de pistage 10. Le terme d'exigence nécessaire est parfois utilisé pour désigner ce type d'exigence.

20 Par définition, une exigence recommandée est une exigence dont le non-respect n'entraîne pas la disqualification du système de pistage 10 mais dont le respect est souhaitable. Le terme d'exigence souhaitable est parfois utilisé pour désigner ce type d'exigence.

Selon l'exemple proposé, le nombre d'exigences est supérieur ou égal à 20.

25 Pour la suite, il est supposé que chaque exigence fixe un seuil à respecter. Cela signifie que chaque exigence est formulable sous la forme d'une grandeur physique supérieure ou égale au seuil ou d'une grandeur physique inférieure ou égale au seuil.

30 Selon l'exemple décrit, le procédé d'évaluation comprend une étape d'acquisition E100, une étape de calcul E102, une étape de comparaison E104, une première étape de détermination E106, une deuxième étape de détermination E108, une étape de déduction E110, une étape d'itération E112, une étape d'émission E114, une étape d'application E116 et une étape d'utilisation E118.

Parmi ces étapes, seule l'étape d'émission E114 est mise en œuvre par le sous-module d'alerte 26, les autres étapes étant mises en œuvre par le sous-module de calcul 24.

35 Lors de l'étape d'acquisition E100, le module d'évaluation 22 acquiert un ensemble de valeurs.

L'ensemble des valeurs comprend les valeurs accessibles au système de pistage 10 sur un intervalle de temps prédéfini.

Selon l'exemple décrit, les valeurs accessibles sont les valeurs provenant de l'ensemble 12 de capteurs.

5 En complément, les valeurs comportent aussi les données estimées par le système de pistage 10 et notamment les estimations de mesures obtenus à partir de données antérieures par le module d'analyse.

L'intervalle de temps prédéfini est choisi selon les impératifs liés à l'évaluation de conformité.

10 Par exemple, si on souhaite faire une évaluation régulière, l'intervalle de temps prédéfini correspondra à l'intervalle de temps choisi entre deux évaluations.

On pourra choisir en particulier un intervalle de temps prédéfini entre 10 minutes et 1 heure. L'intervalle de temps prédéfini choisi dépendra en pratique de l'exigence considérée.

15 Comme détaillé ultérieurement, le nombre de valeurs sera usuellement choisi pour permettre une évaluation en quasi-temps réel et est généralement inférieur, voire bien inférieur, au nombre de points de mesures usuellement requis par les normes d'évaluation du respect à des exigences.

Lors de l'étape de calcul E102, le module d'évaluation 22 calcule une valeur de robustesse et ce, pour chaque exigence.

Pour cela, la robustesse étant calculée comme la valeur p d'un test statistique indiquant si l'exigence considérée est respectée ou non.

Un test statistique est un outil mathématique permettant de vérifier si des données permettent de rejeter ou non une hypothèse.

25 Cette hypothèse que l'on cherche à rejeter est dénommée « hypothèse nulle ».

La mise en œuvre du test statistique donne notamment une valeur nommée « valeur p », située entre 0 et 1, qui servira à rejeter ou non l'hypothèse.

La « valeur p » est aussi dénommée « p-valeur » en référence à la dénomination anglaise correspondante de « p-value ».

30 Dans la suite, le terme de « p-valeur » est utilisé.

La p-valeur est définie comme la probabilité, en supposant que l'hypothèse nulle est vraie, d'observer les données, ou des données encore plus « extrêmes ».

Comme cela sera explicité ultérieurement, chaque p-valeur est associée à une métrique dépendant de l'exigence.

35 Cette métrique comporte généralement des moments des distributions de probabilités supposées sur les données acquises.

Dans un tel cas, il peut être favorable de calculer chaque moment dans un premier temps puis les p-valeurs pour éviter d'éventuels calculs redondants.

Lors de l'étape de comparaison E104, le module d'évaluation 22 vient comparer chaque valeur de robustesse à un seuil prédéfini selon un critère de comparaison respectif.

5 Autrement formulé, le module d'évaluation 22 compare la p-valeur à un seuil prédéfini.

Typiquement, le terme de « p-valeur » est comparé à un seuil choisi à 5%.

Cette valeur de 5% est donnée à titre d'exemple, le seuil dépendant des besoins de l'utilisateur du système de pistage 10.

10 Lors de l'étape de détermination, le module d'évaluation 22 détermine les exigences auxquelles le système de pistage 10 est conforme.

Pour cela, le module d'évaluation 22 considère que les exigences déterminées comme conformes sont celles pour lesquelles à l'étape de comparaison E104, le module d'évaluation 22 a déterminé que le critère de comparaison est respecté.

15 Si la p-valeur est inférieure au seuil, il est considéré qu'il est très peu probable d'avoir observé de telles données, sachant que l'hypothèse nulle est vraie. Cela conduit à rejeter l'hypothèse avec une bonne certitude.

Pour mieux expliciter cette idée, il est développé le cas simple d'un tirage de pièce de monnaie.

20 Soit une pièce de monnaie, avec un côté pile, et un côté face. Il est fait l'hypothèse suivante : « la pièce est équilibrée, et peut donc tomber sur pile ou sur face avec une probabilité $\frac{1}{2}$ ».

Selon un premier exemple, la pièce est lancée 10 fois et on obtient 7 piles et 3 faces. Ce sont les données acquises.

25 La question est alors de savoir si l'hypothèse nulle peut être rejetée ou non.

Il est alors calculé la probabilité de ce résultat, ainsi que des résultats plus extrêmes (à savoir, la probabilité qu'on ait eu un nombre de piles entre 7 et 10 inclus). On obtient une probabilité de 0,17, c'est-à-dire une p-valeur de 0,17.

30 Cela signifie que l'étape de calcul E102 conduit à 17% de chances d'obtenir 7 pile ou plus si la pièce était équilibrée.

Cette p-valeur est supérieure à 5%, de sorte qu'il n'est pas possible de rejeter l'hypothèse avec suffisamment de certitude.

Selon un deuxième exemple, la pièce est lancée 10 fois et on obtient 9 piles et 1 face. Ce sont les données acquises.

35 Il est alors calculé la probabilité d'obtenir 9 ou 10 piles, sachant que la pièce est équilibrée. Ceci conduit à une p-valeur de 0,01.

Avec le seuil de 0,05, cela conduit à déterminer que l'hypothèse est à rejeter. Autrement, formulé, il peut être considéré que, sachant la pièce équilibrée, il est très peu probable d'observer autant de pile.

Un troisième exemple peut aussi illustrer le fait que la p-valeur évolue également en fonction du nombre de lancers.

On suppose que 1000 lancers sont réalisés et qu'on obtient 700 piles.

Dans un tel cas, la p-valeur est inférieure à 10^{-5} , ce qui correspond au fait que la probabilité d'obtenir 7 piles ou plus sur 10 lancers (premier exemple) est bien supérieure à celle d'obtenir 700 piles et plus sur 1000 lancers, sachant une pièce équilibrée

Un calcul permet de montrer que l'obtention 527 piles sur 1000 lancers suffit pour conclure avec un seuil de 5% que l'hypothèse nulle est à rejeter, c'est-à-dire que la pièce est biaisée.

Des exemples plus élaborés mais basés sur ce qui vient d'être décrit sont donnés dans la section « Application au cas particulier de la norme ESASSP ». Dans cette section, il est développé des cas particuliers relatifs à des critères d'évaluation du système de pistage 10.

Dans cette section, il sera expliqué comment dériver à partir d'une exigence donnée l'expression du calcul de la p-valeur.

Plus précisément, une exigence donnée correspond généralement au calcul d'une métrique et au respect par cette métrique d'une valeur seuil.

En analysant l'évolution de la p-valeur en fonction de la valeur de l'exigence, il est possible de dériver une expression adaptée pour cette p-valeur.

Une telle construction se fait avant la mise en œuvre du procédé par une analyse dite a priori.

Dans cette analyse, comme exposé ultérieurement, il est également possible de dériver le nombre de valeurs à obtenir pour déterminer chaque exigence pour assurer un niveau de confiance élevé sur les exigences, tout en s'appliquant sur des zones géographiques et des intervalles temporels réduits.

Un tel nombre de valeurs correspond à un seuil de fiabilité prédéfini qui sera utilisé lors de la deuxième étape de détermination E108.

Lors de la deuxième étape de détermination E108, le module d'évaluation 22 détermine cause du non-respect du critère de comparaison pour chaque valeur de robustesse ne respectant pas le critère de comparaison.

Selon l'exemple décrit, il existe deux causes possibles.

La première cause possible correspond au fait que l'exigence n'est pas respectée.

En effet, il peut suffire de peu de points pour qu'on sache déjà que l'exigence ne sera pas respectée même si le nombre de points de mesures serait plus élevé.

La deuxième cause correspond au fait que le nombre de valeurs acquises est trop faible pour garantir un seuil de fiabilité prédéfini d'évaluation de la conformité à l'exigence considérée.

Cela peut être dû à un trafic d'avions inférieur à ce qui est attendu dans la zone géographique considérée (heure creuse ou grève par exemple).

Il est supposé pour la suite que, pour au moins une exigence, la cause déterminée est la deuxième cause.

Lors de l'étape de déduction E110, le module d'évaluation 22 calcule le nombre de valeurs supplémentaires à acquérir pour obtenir une fiabilité d'évaluation supérieure au seuil de fiabilité prédéfini.

Pour cela, le module d'évaluation 22 utilise le nombre de points assurant un respect seuil de fiabilité, ce nombre de points ayant été calculée avant la mise en œuvre du procédé.

Le module d'évaluation 22 vient donc lire la valeur dans une mémoire du système de pistage 10 et en déduit, par exemple, par soustraction le nombre de valeurs supplémentaires à acquérir.

Un tel calcul peut être réalisé de manière plus élaborée en calculant le nombre de points nécessaires pour passer les exigences, tout en prenant en compte dynamiquement les valeurs observées jusque-là. Cela correspond à un recalcul en temps réel.

Cela est maintenant illustré pour des exemples développés dans la section « application au cas particulier de la norme ESASSP ».

Ainsi, pour le cas des exigences de type probabilité de mise à jour, il est supposé que pour le moment observé k^* valeurs parmi n^* valeurs acquises, il est possible de calculer une p-valeur ρ^* sur le test qui vise à vérifier si la valeur minimale p_r requise est dépassée de manière statistiquement significative.

Il est alors possible de calculer le nombre minimal de valeurs supplémentaires à acquérir pour que l'exigence soit passée avec un seuil donné. Cela correspond à supposer tous les prochaines valeurs comme bonnes, et de donner la valeur m minimale telle que $m + k^*$ succès parmi $m + n^*$ rapports permettent de passer l'exigence avec une p-valeur suffisamment basse.

Ce nombre correspond au nombre minimal de nouveaux points (nouvelles valeurs) nécessaires pour rejeter l'hypothèse nulle.

Alternativement ou en complément, le module d'évaluation 22 modifie la discrétisation spatio-temporelle des zones sur lesquelles le calcul de la p-valeur est réalisée.

5 Cette discrétisation spatio-temporelle est généralement fournie avec le calcul de la p-valeur associée à l'exigence considérée, mais peut être un peu modifiée sans altérer la qualité de l'évaluation de l'exigence.

Lors de l'étape d'itération E112, au moins les étapes d'acquisition, de calcul et de comparaison sont réitérées jusqu'à vérifier une condition.

Cela est indiqué par la flèche 120 sur la figure 2.

10 D'autres étapes comme l'étape de déduction E110 peuvent également être itérées.

Dans l'exemple proposé, la condition est de respecter soit une première condition, soit une deuxième condition.

La première condition est d'obtenir une fiabilité d'évaluation supérieure au seuil de fiabilité prédéfini.

15 Cela s'obtient en dépassant le nombre minimal de valeurs supplémentaires à acquérir si celui a été déterminé une seule fois ou lorsque le nombre minimal de valeur supplémentaires à acquérir est déterminé comme nul à une itération de l'étape de déduction E110.

20 La deuxième condition est d'atteindre un nombre maximal prédéfini avec une fiabilité d'évaluation inférieure au seuil prédéfini.

Cela correspond au cas où la p-valeur est trop haute pour que l'exigence puisse être remplie avec suffisamment de certitude, mais où les valeurs observées jusque-là semblent corroborer le fait que l'exigence est remplie (formellement dans l'exemple de mise à jour décrit précédemment: $\frac{k^*}{n^*} > p_r$, mais n^* est trop faible pour qu'on ait assez de certitude).

25 Il est alors possible de calculer le nombre de points nécessaires pour passer l'exigence avec une assez bonne certitude, en supposant que le ratio de succès restera le même. Formellement, en notant $p^* = \frac{k^*}{n^*}$, on cherche le nombre m de rapports (valeurs acquises) tel qu'observer p^*m succès parmi m rapports permette d'obtenir une p-valeur inférieure au seuil requis.

30 En pratique, le nombre maximal prédéfini est souvent compris entre 40000 et 60000.

Lors de l'étape d'émission E114, le module d'évaluation 22 émet une alerte pour les exigences non respectées ou pour les exigences dont le nombre de valeurs des étapes d'acquisition, de calcul, de comparaison a atteint le nombre maximal prédéfini avec une fiabilité d'évaluation inférieure au seuil prédéfini.

Ces cas correspondent au fait que ni la première condition ni la deuxième condition de l'étape d'itération E112 ne sont respectées.

L'alerte est, par exemple, une alerte sonore ou une alerte visuelle émise à destination d'un contrôleur aérien.

5 Lors de l'étape d'application E116, le système applique une fonction d'évaluation propre à chaque exigence aux valeurs p pour obtenir des valeurs d'évaluation.

Chaque fonction d'évaluation propre à chaque exigence est une fonction d'utilité.

Comme explicité plus haut, la p -valeur notée ρ est interprétée comme signifiant que l'exigence est satisfaite si l'hypothèse H_0 est rejetée et donc si la p -valeur ρ est inférieure à un seuil ε .

10 Il est possible de convertir la p -valeur en fonction d'utilité donnant une valeur comprise entre 0 et 1 en fonction de la valeur de la p -valeur.

Par exemple, la fonction d'utilité est une fonction linéaire par morceaux valant 1 dès que $\rho \leq \varepsilon$ (cas correspondant à une conformité à l'exigence) puis qui décroît linéairement jusqu'à 0 pour des valeurs de la p -valeur ρ supérieure au seuil ε .

La fonction linéaire est, par exemple, la suivante :

$$1 - \rho$$

Alternativement, il peut être envisagé de normaliser la fonction d'utilité décroissante en incluant une phase d'apprentissage supervisé permettant d'apprendre cette fonction.

20 Avantageusement, il est utilisé une fonction de normalisation décroissante à seuils.

Par exemple, le module d'évaluation 22 pourra imposer que si la p -valeur est supérieure à une valeur à définir par un expert, alors l'utilité est nulle.

La valeur définie par l'expert est déterminée avant la mise en œuvre du procédé, de sorte qu'elle est accessible au module d'évaluation 22 par lecture de sa valeur dans une mémoire du système de pistage 10.

25 Ensuite, le module d'évaluation 22 utilise les valeurs d'évaluation pour obtenir un score global lors de l'étape d'utilisation E118.

Le score global évalue la conformité du système de pistage 10 à l'ensemble des exigences.

30 Un tel score est souvent dénommé QoS en référence à la dénomination anglaise de « Quality of Service » signifiant littéralement « qualité de service ».

Pour cela, le module d'évaluation 22 peut mettre en œuvre des techniques différentes.

35 Par exemple, avec la fonction d'utilité précédente, l'utilité vaut 1 si la p -valeur est inférieure au seuil et avec très forte confiance que l'exigence soit passée, et a une valeur strictement inférieure à 1 sinon.

Ce faisant, le module d'évaluation 22 peut utiliser une fonction d'agrégation.

Selon un exemple, la fonction d'agrégation est une fonction compensatoire strictement croissante.

En appliquant uniquement cette technique au cas des exigences obligatoires, cela
5 permet d'obtenir un score QoS égal à 1 si toutes les exigences obligatoires sont passées avec une forte confiance et strictement inférieure à 1 si au moins une exigence obligatoire n'est pas complètement respectée.

Selon un autre exemple, l'étape d'utilisation E118 comprend l'apprentissage supervisé de fonction d'évaluation de la conformité à une exigence sous forme d'une
10 fonction linéaire par morceaux, la fonction linéaire comportant au moins trois morceaux, de préférence moins de six morceaux, pour obtenir des fonctions d'évaluation apprises, l'application des fonctions d'évaluation apprises au système de pistage pour obtenir des valeurs d'évaluation, et l'utilisation des valeurs d'évaluation pour obtenir un score global évaluant la conformité du système de pistage 10 à l'ensemble des exigences.

Le module d'évaluation 22 calcule alors le score global pour chaque type d'exigence
15 par agrégation des fonctions d'évaluation de toutes les exigences du type considéré, le score global évaluant la conformité du système de pistage à l'ensemble des exigences étant obtenu par agrégation des scores globaux par type d'exigence.

En variante, au lieu de ce calcul, l'étape d'utilisation E118 comprend l'apprentissage
20 supervisé de fonctions d'agrégation des valeurs d'évaluation, la fonction d'agrégation évaluant la conformité du système de pistage à une partie des exigences de l'ensemble d'exigences correspondant à une grandeur particulière, et- une étape d'utilisation comprenant l'utilisation des valeurs d'évaluation et des fonctions d'agrégation apprises pour
25 obtenir une valeur pour chaque partie des exigences de l'ensemble d'exigences correspondant à une grandeur particulière et l'utilisation des valeurs obtenues lors de l'utilisation ainsi que des fonctions d'agrégation pour obtenir le score global.

Ces techniques sont également applicables au cas des exigences recommandées.

Selon une autre approche, le module d'évaluation 22 agrège différemment les
30 valeurs QoS^O (valeur pour les exigences obligatoires) et QoS^R (valeur pour les exigences recommandées) pour obtenir la QoS globale. Les scores normalisés sont ici interprétés comme le niveau de confiance selon lequel les exigences sont passées.

Compte tenu du caractère obligatoire des exigences sous-jacentes à QoS^O , il faut
que la moindre défaillance d'une exigence obligatoire soit répercutée sur le score global et ne puisse pas être compensée par les exigences recommandées.

35 Pour cela, dans cette approche, le module d'évaluation 22 utilise les formules suivantes :

$$QoS = \begin{cases} \frac{QoS^O}{2} & \text{si } QoS^O < 1 \\ \frac{1 + QoS^R}{2} & \text{si } QoS^O = 1 \end{cases}$$

Dans les techniques précédentes, les fonctions d'agrégation sont compensatoires, ce qui implique qu'un mauvais respect d'une exigence peut être compensée par un bon respect d'une autre exigence.

5 Dans certains cas, il peut être souhaité que cette compensation ne soit pas possible notamment en ce qui concerne les exigences obligatoires. Dès qu'au moins une exigence obligatoire n'est pas respectée, il ne serait alors plus possible de compenser ce non-respect même si toutes les autres exigences sont respectées.

10 Pour cela, il peut être utilisé un autre type de fonction d'agrégation appelées t-normes.

Ce sont des opérateurs binaires $T(x, y) \in [0, 1]$ où x, y sont également à valeur dans $[0, 1]$.

Un opérateur binaire $T: [0, 1] \times [0, 1] \rightarrow [0, 1]$ est une t-norme s'il satisfait aux propriétés suivantes :

- 15 • (élément neutre) $T(1, x) = x$ pour tout $x \in [0, 1]$,
- (commutativité) $T(x, y) = T(y, x)$ pour tout $x, y \in [0, 1]$,
- (monotonie) $T(x, y) \leq T(u, v)$ pour $0 \leq x \leq u \leq 1$ et $0 \leq y \leq v \leq 1$, et
- (associativité) $T(x, T(y, z)) = T(T(x, y), z)$ pour tout $x, y, z \in [0, 1]$.

20 La fonction T satisfait à certaines propriétés. Cette fonction fournit une évaluation pessimiste par rapport à une fonction compensatoire car $T(x, y) \leq \min(x, y)$, c'est-à-dire qu'il n'est jamais possible d'avoir une note meilleure que la pire note. Plusieurs mauvaises notes viennent empirer la note globale.

25 Il peut être montré que l'opérateur binaire $T: [0, 1] \times [0, 1] \rightarrow [0, 1]$ est une t-norme stricte et archimédienne si et seulement si il existe une fonction $f: [0, 1] \rightarrow [0, \infty)$ continue, strictement décroissante avec $f(1) = 0$ tel que $T(x, y) = f^{-1}(\min(f(x) + f(y), f(0)))$.

Par définition, une t-norme T est archimédienne si les deux conditions suivantes sont satisfaites :

- T est continue, et
- $T(x, x) < x$ pour tout $x \in (0, 1)$.

30 Cet opérateur binaire T est facilement transformé en opérateur n-aire :

$$T(x_1, \dots, x_m) = f^{-1}\left(\min\left(\sum_{i=1}^m f(x_i), f(0)\right)\right)$$

où :

- $x_1, \dots, x_m$ sont les scores normalisés des p-valeurs des exigences obligatoires,
- $T(x_1, \dots, x_m)$ est le score QoS^0 , et
- la fonction f est à construire par apprentissage supervisé.

5 En variante, la somme $\sum_{i=1}^m f(x_i)$ est remplacée par une moyenne pondérée ordonnée, ce qui s'écrit mathématiquement comme :

$$T(x_1, \dots, x_m) = f^{-1} \left(\min \left(\sum_{i=1}^m w_i f(x_{\pi(i)}), f(0) \right) \right)$$

où π est une permutation sur $\{1, \dots, m\}$ ordonnant les scores comme suit :

$$x_{\pi(1)} \geq x_{\pi(2)} \geq \dots \geq x_{\pi(m)}$$

10 Le procédé qui vient d'être décrit permet donc d'obtenir une évaluation robuste de la conformité du système de pistage 10 à un ensemble d'exigences.

De fait, plutôt que d'avoir une simple valeur binaire (critère rempli ou non), le procédé donne accès à une estimation d'une marge avec laquelle le critère est rempli, c'est-à-dire une estimation de la confiance qu'on peut avoir en l'estimation.

15 En outre, en présence de trop peu de valeurs pour avoir une estimation fiable, le critère sera toujours considéré comme non-rempli, ce qui est souhaitable dans un contexte critique comme le trafic aérien, où il est mieux d'être alarmiste à tort que trop confiant.

De plus, le procédé donne la certitude que le résultat ne peut pas varier instantanément de « non-rempli » à « rempli », la p-valeur évoluant de manière lisse avec les nouvelles données. On peut cependant avoir un passage au-dessus et en-dessous du seuil, auquel cas l'alarme est lancée, mais il reste toujours une certaine marge à parcourir avant que l'on atteigne effectivement la violation du critère.

De plus, le procédé utilise un nombre de points plus faibles que ceux recommandés dans la norme.

25 Cela permet de réduire la quantité de mémoire occupée ainsi que le stress sur les composants matériels du système informatique 10.

Le procédé présente également l'avantage que l'évaluation est adaptative.

En particulier, le procédé réduit grandement la durée d'acquisition des données pour un calcul fiable de score, et donc de la conformité ou non des critères. Ainsi, les données 30 utilisées sont moins vieilles, et l'estimation de la qualité de service est donc plus réactive.

Cela est bénéfique lors de l'apparition d'un problème puisqu'en cas de dégradation rapide, la métrique se dégrade plus vite, et donc l'opérateur est informé plus rapidement.

La réciproque est également vraie, lorsqu'un problème est solutionné, le retour à la normale se fait plus rapidement.

Cela signifie que le procédé permet de prendre en compte les dérives extrêmes.

Lorsque la dérive est mineure, avec le nombre de données fixé a priori, le procédé
5 n'est pas capable de valider l'exigence avec une forte confiance. On peut alors augmenter temporairement le nombre de points pour absorber cette dérive mineure passagère. Cela évite de remonter une alerte qui n'a pas lieu d'être. Pour augmenter le nombre de points de mesure, il est soit augmenté la durée de recueil des données (si cela est autorisé), soit regroupées des zones géographiques. Ces deux mesures sont temporaires.

10 Lorsque la dérive est majeure, dès le début de recueil des données sur un intervalle de temps, le procédé permet de se rendre compte qu'il n'y a aucune chance de satisfaire à l'exigence compte tenu du nombre trop important de mauvaises données. Au lieu d'attendre la fin de l'intervalle de recueil d'information, le procédé remonte une alerte au plus tôt à l'opérateur.

15 Le procédé définit aussi un mécanisme de contrôle du nombre de points de mesure, ce mécanisme de contrôle pouvant être statique (déterminé a priori) ou/et dynamique (adapté à partir des valeurs mesurées).

Ainsi, il est possible de mettre en œuvre le procédé en temps réel sur un intervalle de temps pouvant être d'une longueur fixe configurée à l'avance ou dynamique en fonction
20 du débit d'acquisition des valeurs. Cela permet de détecter la présence d'une anomalie du système de pistage en temps réel.

Dans chacun des exemples décrits, chaque module ou sous-module est réalisé chacun sous forme d'un logiciel, ou d'une brique logicielle.

L'ensemble des modules ou sous-modules est alors réalisé c'est-à-dire sous forme
25 d'un programme d'ordinateur, également appelé produit programme d'ordinateur, il est en outre apte à être enregistré sur un support, non représenté, lisible par ordinateur. Le support lisible par ordinateur est par exemple un medium apte à mémoriser des instructions électroniques et à être couplé à un bus d'un système informatique. A titre d'exemple, le support lisible est un disque optique, un disque magnéto-optique, une mémoire ROM, une
30 mémoire RAM, tout type de mémoire non-volatile (par exemple FLASH ou NVRAM) ou une carte magnétique. Sur le support lisible est alors mémorisé un programme d'ordinateur comprenant des instructions logicielles, mémorisé dans une mémoire exécutable par un processeur.

En variante non représentée, chaque module ou sous-module est réalisé sous forme
35 d'un composant logique programmable, tel qu'un FPGA (de l'anglais *Field Programmable*

Gate Array), ou encore d'un circuit intégré, tel qu'un ASIC (de l'anglais *Application Specific Integrated Circuit*).

L'invention concerne toute combinaison techniquement possible des modes de réalisation précédemment décrits.

5

APPLICATION AU CAS PARTICULIER DE LA NORME ESASSP

Présentation des critères

Le procédé qui vient d'être décrit peut s'appliquer pour toute type de norme mais sera illustré dans ce qui suit à travers l'exemple de la norme ESASSP qui est détaillée dans le document intitulé « EUROCONTROL Specification for ATM Surveillance System Performance » (Volume 1) dont l'ISBN est 978-2-87497-022-1 et qui a été publié en mars 2012.

L'objectif de cette norme est d'assurer la qualité requise pour éviter des collisions. Cette qualité vise à une meilleure séparation entre les avions dans le trafic aérien (civil). Afin de pouvoir augmenter la capacité aérienne, il est primordial de disposer d'outils de pistage permettant d'assurer une séparation de plus en plus fine entre les avions. Deux normes ont été définies pour assurer deux séparations en distance : 5 NM et 3 NM.

La norme ESASSP définit 22 contraintes nommées de R1 à R22, c'est-à-dire 22 métriques spécifiques. Les 22 contraintes sont expliquées sommairement dans ce qui suit.

La première contrainte R1 concerne l'intervalle de mesure de la probabilité de mise à jour (dénommée en anglais « Measurement Interval for Probability of Update »). La première contrainte R1 se décompose en deux exigences : une exigence obligatoire et une exigence recommandée.

La deuxième contrainte R2 concerne la probabilité de mise à jour de la position horizontale (dénommée en anglais « Probability of update of horizontal position »). La deuxième contrainte R2 se décompose en deux exigences : une exigence obligatoire et une exigence recommandée.

La troisième contrainte R3 concerne le ratio de temps où l'avion n'a pas été pisté (dénommée en anglais « ratio of missed 3D position involved in long gaps »). La troisième contrainte R3 correspond à une exigence obligatoire.

La quatrième contrainte R4 concerne l'erreur quadratique en position horizontale (dénommée en anglais « Horizontal position RMS error »). L'erreur quadratique est souvent considérée comme une erreur principale représentative de la précision des mesures. La quatrième contrainte R4 se décompose en deux exigences : une exigence obligatoire et une exigence recommandée.

La cinquième contrainte R5 concerne un ratio en erreurs corrélées consécutives (dénommée en anglais « Consecutive correlated error ratio »). La cinquième contrainte R5 correspond à une exigence recommandée.

5 La sixième contrainte R6 concerne l'intervalle maximum de temps avec une proximité proche (dénommée en anglais « Max delta time in close proximity »). Une telle contrainte correspond au fait que, sur un radar, des pistes très proches qui ne sont pas rafraichies au même instant sont présentes. Dans la présentation au contrôleur aérien, il ne faut pas qu'il y ait deux pistes côte-à-côte qui aient des dates de mise à jour éloignées. Cela pourrait amener à des risques de collision, c'est-à-dire des problèmes de sécurité graves.
10 La sixième contrainte R6 correspond à une exigence recommandée.

La septième contrainte R7 concerne la probabilité de mise à jour de la pression en altitude (dénommée en anglais « Probability of update of pressure altitude »). La septième contrainte R7 correspond à une exigence obligatoire.

15 La huitième contrainte R8 se rapporte à l'âge moyen des données transmises de la pression en altitude (dénommée en anglais « Average data age of forwarded pressure altitude »). La huitième contrainte R8 correspond à une exigence obligatoire.

La neuvième contrainte R9 se rapporte à l'âge maximal des données transmises de la pression en altitude (dénommée en anglais « Max data age of forwarded pressure altitude »). La neuvième contrainte R9 correspond à une exigence obligatoire.

20 La dixième contrainte R10 se rapporte au rapport de données transmises de la pression en altitude qui sont incorrectes (dénommée en anglais « Ratio of incorrect forwarded pressure altitude »). Il est à noter que la dixième contrainte R10 est une estimation de l'erreur correspondant à une erreur parasite. La dixième contrainte R10 correspond à une exigence obligatoire.

25 La onzième contrainte R11 concerne les erreurs non signées de la pression en altitude (dénommée en anglais « Pressure altitude unsigned error »). Il est à noter que la onzième contrainte R11 est une estimation de l'erreur de type erreur principale. La onzième contrainte R11 correspond à une exigence obligatoire.

30 La douzième contrainte R12 concerne le délai d'apparition des indicateurs d'urgence (dénommée en anglais « Delay of apparition of the emergency indicator / SPI report »). La douzième contrainte R12 correspond à une exigence obligatoire.

La treizième contrainte R13 se rapporte au délai de changement dans l'identification de l'avion (dénommée en anglais « Delay of change in Aircraft Id »). La treizième contrainte R13 correspond à une exigence obligatoire.

35 La quatorzième contrainte R14 concerne la probabilité de mise à jour avec une identification correcte (dénommée en anglais « Probability of update of aircraft identity with

correct value »). La quatorzième contrainte R14 se décompose en deux exigences : une exigence obligatoire et une exigence recommandée.

La quinzième contrainte R15 se rapporte au rapport d'identification incorrecte de l'avion (dénommée en anglais « Ratio of Incorrect Aircraft identity »). Il est à noter que la
5 quinzième contrainte R15 est une estimation de l'erreur de type erreur parasite La
quinzième contrainte R15 correspond à une exigence obligatoire.

La seizième contrainte R16 concerne l'erreur quadratique du taux de descente ou de montée (dénommée en anglais « Rate of climb/descent RMS error »). Dans ce contexte, l'erreur estimée est du type erreur principale. La seizième contrainte R16 correspond à une
10 exigence recommandée.

La dix-septième contrainte R17 concerne l'erreur quadratique de vitesse de poursuite (dénommée en anglais « Track velocity RMS error »). Dans ce contexte, l'erreur estimée est du type erreur principale. La dix-septième contrainte R17 correspond à une exigence recommandée.

La dix-huitième contrainte R18 concerne l'erreur quadratique de l'angle de vitesse de poursuite (dénommée en anglais « Track velocity angle RMS error »). Dans ce contexte, l'erreur estimée est du type erreur principale. La dix-huitième contrainte R18 correspond à une exigence recommandée.

La dix-neuvième contrainte R19 concerne la densité de fausses cibles non corrélées (dénommée en anglais « Density of uncorrelated false target reports »). Dans ce contexte, l'erreur estimée est du type erreur parasite. La dix-neuvième contrainte R19 correspond à une exigence recommandée.

La vingtième contrainte R20 concerne le ratio horaire de fausses pistes proches de vraies pistes (dénommée en anglais « Hourly rate of false tracks close to true tracks »). Dans ce contexte, l'erreur estimée est du type erreur corrélée. La vingtième contrainte R20 correspond à une exigence recommandée.

La vingt-et-unième contrainte R21 concerne la continuité (dénommée en anglais « Continuity »). La vingt-et-unième contrainte R21 correspond à une exigence recommandée.

La vingt-deuxième contrainte R22 concerne des investigations manuelles (analyses manuelles des résultats) qui doivent être menées lorsque les contraintes R2, R4, R12 ou R13 ne sont pas atteintes. La vingt-deuxième contrainte R22 se décline selon deux exigences : une exigence obligatoire et une exigence recommandée.

De manière synthétique, le tableau 1 suivant peut être établi, la présence d'une croix
35 indiquant que la présence d'une exigence associée :

Contraintes	Nom anglais utilisé dans la norme	Unité	Exigence obligatoire	Exigence recommandée
R1	Measurement Interval for Probability of Update	s	X	X
R2	Probability of update of horizontal position	%	X	X
R3	ratio of missed 3D position involved in long gaps	%	X	
R4	Horizontal position RMS error	m	X	X
R5	Consecutive correlated error ratio	%		X
R6	Max delta time in close proximity	s		X
R7	Probability of update of pressure altitude with correct value	%	X	
R8	Average data age of forwarded pressure altitude	s	X	
R9	Max data age of forwarded pressure altitude	s	X	
R10	Ratio of incorrect forwarded pressure altitude	%	X	
R11	Pressure altitude unsigned error	%	X	
R12	Delay of apparition of the emergency indicator / SPI report	s	X	
R13	Delay of change in Aircraft Id	s	X	
R14	Probability of update of aircraft identity with correct value	%	X	X
R15	Ratio of Incorrect Aircraft Identity	%	X	
R16	Rate of climb/descent RMS error	m/s		X
R17	Track velocity RMS error	m/s		X
R18	Track velocity angle RMS error	degrés		X
R19	Density of uncorrelated false target reports	nombre		X
R20	Number per hour of false tracks close to true tracks	nombre		X
R21	Continuity	%		X
R22	Investigations		X	

Tableau 1 : Association pour chaque contrainte de sa dénomination en anglais, de l'unité associée ainsi que la présence d'une exigence obligatoire et d'une exigence recommandée

5 En résumé, la norme ESASSP est un ensemble de 14 exigences obligatoires et de 12 exigences recommandées, ce qui représente 26 exigences au total, ces 26 exigences étant estimées selon 22 métriques distinctes.

Calcul des p-valeurs associées aux critères

10 Pour le calcul des p-valeurs, les critères définis dans l'ESASSP peuvent être regroupés en plusieurs familles. Ces familles sont au nombre de cinq et sont exposées dans ce qui suit.

 La première famille regroupe les critères liés à des probabilités de mise à jour, ces critères étant les critères R2, R7 et R14.

15 Ces critères R2, R7 et R14 ont en commun de porter sur une proportion entre le nombre de rapports reçus contenant un certain élément (position horizontale, altitude-pression, et identité de l'aéronef respectivement) parmi l'ensemble des rapports reçus (qui peut contenir ou non l'élément en question). Ces critères imposent que la valeur de la proportion soit supérieure à une valeur seuil.

20 La deuxième famille regroupe les critères liés à la racine de l'erreur quadratique moyenne, ces critères étant les critères R4, R6, R16, R17 et R18.

 En effet, les critères de la deuxième famille portent sur un calcul de distance euclidienne moyenne entre une valeur et une mesure de référence, souvent calculée a posteriori par le SSTA. Ces critères imposent que la valeur de la proportion soit inférieure
25 à une valeur seuil.

 La troisième famille regroupe les critères liés à des ratios, ces critères étant les critères R3, R5, R10, R11 et R15 et exigent qu'un rapport entre deux grandeurs présente une valeur inférieure à une valeur prédéfinie.

30 La quatrième famille regroupe les critères liés à des moyennes, ces critères étant les critères R8, R12 et R13.

 La cinquième famille regroupe les critères ne relevant pas des quatre familles déjà présentées. Il s'agit des critères R19, R20 et R9.

 Il peut être remarqué que ces familles ne regroupent pas tous les critères. De fait, le critère R1 est un paramètre du système de pistage 10 qui est donc fixé a priori, le critère
35 R9 n'est pas un critère, mais un filtre sur la validité des données, le critère implique de ne

pas avoir d'incident sur des périodes de plusieurs années (ce qui ne peut être évalué sur une courte période) et le critère R22 n'est pas un critère quantitatif.

Il est maintenant décrit comment calculer les probabilités d'obtenir une nouvelle valeur ne respectant pas l'exigence sachant les valeurs acquises, c'est-à-dire les p-valeurs pour chacune des familles précitées.

Première famille

Pour la première famille, il est d'abord exposé comment déterminer les p-valeurs pour le critère R2 impliquant une probabilité de mise à jour simple.

Le critère R2 implique une métrique calculée comme le rapport entre un nombre de portions de trajectoires observées (noté n) et le nombre de ces trajectoires qui contiennent une position horizontale (noté k). La probabilité est donc :

$$P = k/n$$

Où :

- n est le nombre de points de mesures (par exemple le nombre de tronçons de taille identique dans toutes les trajectoires, pour le critère R2) et
- k est le nombre de points de mesure ayant une certaine propriété (par exemple le nombre de tronçons de trajectoires dans lesquels le système de pistage 10 a reçu au moins un update de position horizontale de l'avion, pour le critère R2).

Il est supposé que la valeur de k est une réalisation d'une variable aléatoire N_R qui suit une loi binomiale de probabilité p , avec n essais. Cela suppose l'indépendance des observations successives. La règle exige que cette probabilité soit au-dessus d'un certain seuil (noté p_r).

Selon un mode de réalisation simple, l'hypothèse nulle peut être choisie comme :

$$H_0 : p \leq p_r$$

Le calcul de la p-valeur notée ρ peut alors être mis en œuvre en utilisant la formule :

$$\rho = P(N_R \geq k | H_0)$$

Où N_R est une variable aléatoire suivant une loi binomiale de probabilité p , avec n essais.

Cela correspond à la formulation suivante : « sachant que la probabilité réelle ne remplit pas le critère, quelle est la probabilité que le système de pistage 10 observe un nombre égal ou supérieur de rapports contenant une position horizontale ? ». Si cette valeur est suffisamment faible, il est considéré que l'observation de k n'est probablement pas issue de la chance, et que le seuil est donc passé avec une forte probabilité.

Selon un mode de réalisation plus élaboré et notamment dans le contexte d'un système critique, il est préférable d'imposer des marges plus grandes.

Pour cela, l'hypothèse nulle devient $H_0 : p = p_r$ et le calcul de la p-valeur comme pour le mode de réalisation simple.

5 Cela correspond à la formulation suivante : « sachant que la probabilité réelle remplit tout juste le critère, quelle est la probabilité que l'on observe un nombre supérieur ou égal de rapports contenant une position horizontale ? ».

Ainsi, il est plus difficile pour le système de remplir le critère, et on s'assure donc une plus grande robustesse du système.

10 Pour ce critère, et sous ces hypothèses, la p-valeur se calcule :

$$\rho = 1 - I_{1-p_r}(n - k + 1, k)$$

où I est la fonction bêta incomplète régularisée.

La courbe de la p-valeur ainsi obtenue est représentée dans la figure 3 pour l'exigence R2. L'axe des x représente le ratio $\hat{p} = \frac{k}{n}$, l'axe des y représente la p-valeur. Pour
15 rappel, l'exigence R2 demande à ce que ratio soit supérieur ou égal à $p_r = 0.97$, matérialisé par le trait vertical pointillé. Le trait horizontal pointillé correspond au seuil 0,05 couramment utilisé sur les p-valeurs. Autrement dit, si la p-valeur est en dessous de cette barre, alors on a une très forte confiance que l'exigence R2 est passée.

Chaque courbe correspond à un nombre de points. Chaque courbe prend la forme
20 d'une sigmoïde décroissante dont le point d'inflexion est atteint en $p_r = 0.97$. On remarque que la courbe est d'autant plus pentue autour de ce point d'inflexion que le nombre de point est élevé.

Cela signifie qu'avec beaucoup de points, il y a une très forte confiance que l'exigence R2 est passée dès lors que l'on est à peine supérieur au seuil $p_r = 0.97$. En
25 revanche, pour un nombre faible de points, il faut être significativement au-dessus du seuil pour être confiant que l'exigence est passée.

Une telle courbe peut être réalisée pour chacune des exigences et conduit aux résultats qui seront détaillés dans la suite. Dans un souci de simplicité, ces courbes ne sont pas introduites dans ce qui suit pour simplifier le propos mais les résultats obtenus utilisent
30 ces courbes qui ont été réalisées par la Demanderesse.

Dans certains cas, il est requis que la valeur calculée pour un critère soit atteinte pour toutes les trajectoires pistées. C'est le cas notamment pour le critère « recommandé » du critère R2.

Dans ce cas, il est utilisé les valeurs, qui sont calculées trajectoire par trajectoire. Cela signifie que, pour une trajectoire donnée, les valeurs sont calculées uniquement sur les points appartenant à cette trajectoire.

5 Soit ρ_t la p-valeur associée à une trajectoire t donnée et T l'ensemble des trajectoires.

Dans le cas du critère R2, il s'agit de vérifier que, pour tout $t \in T$, la probabilité de mise à jour sous-jacente p_t vérifie $p_t > 0.97$.

Il est ainsi calculé pour une trajectoire t la p-valeur notée ρ_t du test défini dans le paragraphe précédent, pour lequel il est utilisé uniquement les informations des rapports
10 qui concernent la trajectoire t .

Il vient ainsi :

$$\rho = P(\exists t \in T, N_R^t \geq \hat{k}^t | H_0)$$

Cela peut se réécrire comme :

$$\rho = 1 - P(\forall t \in T, N_R^t < \hat{k}^t | H_0) = 1 - \prod_{t \in T} P(N_R^t < \hat{k}^t | H_0)$$

15 En supposant que les probabilités sont indépendantes suivant les différentes trajectoires, il vient :

$$\rho = 1 - \prod_{t \in T} (1 - P(N_R^t \geq \hat{k}^t | H_0))$$

Ainsi, la p-valeur du test visant à vérifier que toutes trajectoires vérifient $p_t > 0.97$ se calcule comme:

$$20 \quad \rho = 1 - \prod_{t \in T} (1 - \rho_t)$$

Il est à noter que cette démarche est applicable à toutes les p-valeurs subséquentes pour lesquelles on veut vérifier que l'ensemble des trajectoires remplissent un critère donné.

25 Il est ensuite exposé comment déterminer les p-valeurs pour les critères R7 et R14 impliquant une probabilité de mise à jour double.

Des métriques impliquant une probabilité de mise à jour double se calculent de la même manière que pour les métriques impliquant probabilité de mise à jour simple à une différence près. Plus précisément, la probabilité s'écrit :

$$P = z/k$$

30 Où :

- k est le nombre de réalisations de référence (par exemple, pour le critère R7, le nombre d'intervalles de temps où une mise à jour de l'altitude est reçue), et

- z est le nombre de réalisations ayant une certaine propriété (par exemple, pour le critère R7, le nombre d'intervalles de temps où une mise à jour non erronée de l'altitude)

Dans ce cas, les nombres k et z sont considérés comme issus de lois binomiales. Cela suppose l'indépendance des observations successives, et le fait que ces observations soient identiquement distribuées.

Selon un mode de réalisation particulier, le nombre k correspond toujours au nombre de rapports contenant une altitude-pression. Quant au nombre z , il s'agit, parmi ces k rapports, de ceux dont la valeur de l'altitude-pression est correcte (dans le cas du critère R7) ou un identifiant d'aéronef (critère R14).

Ainsi, le nombre k suit la loi binomiale N_R du paragraphe précédent, avec une probabilité inconnue p_R , et le nombre z suit une loi binomiale de paramètres k et $\hat{p}$.

Dans ce contexte, la valeur $\hat{p}$ est inconnue et il s'agit de vérifier que sa valeur est supérieure à un seuil donné p_r .

En prenant toujours l'hypothèse nulle $H_0 : \hat{p} = p_r$, la p-valeur se calcule alors comme suit:

$$\rho = 1 - \int_{l=0}^1 I_{1-lp_r}(n-z+1, z) \frac{\binom{n}{k} l^k (1-l)^{n-k}}{\int_0^1 \binom{n}{k} t^k (1-t)^{n-k} dt} dl$$

En pratique, l'intégrale pourra être calculée par des méthodes numériques comme un tirage de Monte-carlo ou une méthode d'approximation par des rectangles.

20

Deuxième famille

Pour la deuxième famille, il s'agit de calculer la p-valeur pour les erreurs quadratiques moyennes racines.

Un critère de la deuxième famille est donc une métrique qui est une erreur quadratique moyenne entre une valeur observée et une valeur de référence. Le critère est rempli lorsque cette métrique est inférieure à une valeur donnée.

Soit EQMR l'erreur quadratique moyenne racine. Cette erreur quadratique s'écrit :

$$EQMR = \sqrt{\frac{1}{n} \sum_{t=1}^n E_t}$$

Où :

- E_t est l'erreur quadratique sur une grandeur donnée pour la t -ième mesure. Par exemple pour le critère R4, la grandeur est l'erreur de la position horizontale de l'avion.

30

En supposant que tous les erreurs quadratiques E_t sont identiquement distribuées et indépendantes, et que la loi correspondante a une moyenne μ^2 et une variance σ^2 (inconnues), le critère requiert que la moyenne μ^2 soit inférieure à une certaine valeur μ_r^2 , qui est le seuil de l'exigence.

5 Soit $\hat{\mu}^2$ la moyenne observée des erreurs quadratiques E_t .

En faisant l'hypothèse nulle : $H_0 : \mu^2 = \mu_r^2$, le théorème central limite permet d'obtenir la relation suivante :

$$\rho = P(EQMR < \mu_r \mid H_0) = \Phi\left(\frac{\hat{\mu}^2 - \mu_r^2}{\frac{\hat{\sigma}}{\sqrt{n}}}\right)$$

Où :

10 • Φ est la fonction de répartition de la distribution normale centrée réduite.

Pour certains critères, les métriques sont calculées comme le minimum sur différentes trajectoires de l'erreur quadratique moyenne calculée sur chaque trajectoire. C'est par exemple du critère R4.

L'approche est alors la même que précédemment.

15

Troisième famille

Pour la troisième famille, comme pour la première famille, il existe deux types de critères dans cette famille, un premier type dit ratio simple rassemblant les critères basés sur une métrique impliquant une variable divisée par une constante et un deuxième type dit ratio double rassemblant les critères basés sur une métrique impliquant une variable divisée par une autre variable. Pour les deux types, il s'agit que la métrique soit inférieure à un seuil prédéfini.

Il est d'abord exposé comment calculer la p-valeur pour des critères basés sur un ratio simple, comme c'est le cas pour les critères R3, R5 et R11.

25 Ce calcul est assez similaire au cas de la probabilité de mise à jour simple puisque les ratios sont de la forme :

$$R = \frac{k}{n}$$

Avec :

- 30 • n un nombre donné (par exemple, le nombre de mises à jour de pistes reçus pour le critère R5), et
- k une variable aléatoire qui suit une loi binomiale de paramètres n et p (par exemple, en ce qui concerne le critère R, le nombre de mises à jour de

pistes pour lesquelles au moins 3 mises à jour consécutives de positions horizontales présentent une erreur du même côté de la trajectoire).

Il s'agit de tester si p est sensiblement inférieur à p_r , la valeur requise, au vu des observations.

5 Ainsi, en prenant comme hypothèse nulle, $H_0 : p = p_r$ et, pour cette métrique, et sous ces hypothèses, la p-valeur se calcule selon la formule suivante :

$$\rho = I_{1-p_r}(n - k + 1, k)$$

Il peut être noté ici la symétrie avec la probabilité de mise à jour simple. Ceci s'explique par le fait que, dans un cas, il s'agit de maximiser la probabilité de mise à jour simple alors qu'ici, il s'agit de minimiser les ratios simples.

Pour ce qui concerne les ratios doubles, la démarche suivante peut être suivie. Cette démarche est notamment adaptée pour les critères R10 et R15.

Dans le cas des ratios doubles, la métrique s'écrit sous la forme suivante :

$$R = \frac{N_I}{N_V}$$

15 Où :

- N_I est le nombre de réalisations de référence (par exemple, le nombre de mises à jour de pistes contenant une altitude barométrique, pour le critère R10), et
- N_V est le nombre de réalisations ayant une certaine propriété (par exemple, le nombre de mises à jour de pistes contenant une altitude barométrique incorrecte, pour le critère R10).

Pour ces métriques, il s'agit de vérifier que le ratio de deux variables binomiales est inférieur à un seuil. Par analogies aux ratios simples, la p-valeur se calcule comme le symétrique de celles calculées pour les probabilités de mise à jour double :

$$25 \quad \rho = \int_{l=0}^1 I_{1-p_r l}(N_T - N_I, 1 + N_I) \frac{\binom{n}{k} l^{N_V} (1-l)^{N_T-N_V}}{\int_0^1 \binom{n}{k} t^{N_V} (1-t)^{N_T-N_V} dt} dl$$

Où :

- l joue le rôle de la probabilité observée de la valeur intermédiaire (à savoir la probabilité sous-jacente de la loi qui détermine la valeur N_V), comme p_R l'était dans le cas des probabilités de mise à jour double.

30 L'expression précédente de la p-valeur correspond à intégrer sur toutes les valeurs possibles de l , pondérées par la vraisemblance de l au vu de la valeur de N_V .

Quatrième famille

Pour la quatrième famille, les critères impliquent des métriques de la forme :

$$M = \frac{1}{n} \sum_{i=1}^n X_i$$

Où :

- 5 • X est une variable d'intérêt (par exemple, l'âge de l'information de la pression barométrique lorsque celle-ci est envoyée, pour le critère R8), donnée par le rapport i , et
- n est le nombre de rapports considérés.

Le critère est rempli lorsque ces moyennes soient inférieures à un seuil, noté μ_r .

10 En supposant les variables X identiquement distribuées et indépendantes sur les rapports, et un nombre de rapports suffisant, il peut être appliqué le théorème centrale limite.

Soit μ la moyenne et σ l'écart type (inconnus) de la variable X , $\hat{\mu}$ la moyenne et $\hat{\sigma}$ l'écart type observés. Comme indiqué précédemment, il s'agit que μ est plus petit que μ_r .

15 En faisant l'hypothèse nulle $H_0 : \mu = \mu_r$, par le théorème centrale limite, il vient :

$$\rho = P(M \leq \mu_r \mid H_0) = \Phi\left(\frac{\hat{\mu} - \mu_r}{\hat{\sigma}/\sqrt{n}}\right)$$

Où :

- Φ est la fonction cumulative de la distribution normale centrée réduite.

20 **Cinquième famille**

Pour la cinquième famille, chacun des critères restants sont examinés.

En ce qui concerne le critère R19, il est considéré, une métrique correspondant au nombre de faux rapports obtenus pour chaque intervalle d'une heure et pour chaque zone de 900 NM².

25 Cela donne une distribution empirique (représentable sous la forme d'un histogramme), pour chaque intervalle de temps.

Soit D cette distribution, M sa moyenne et S son écart-type.

Cette métrique est calculée sur une heure, cela donne au minimum 450 intervalles de temps du fait que la longueur d'intervalle de temps est d'au plus 8 secondes.

30 De plus, la métrique est cumulative.

Ceci implique que :

$$K = \sum_{i \in I} d_i$$

Où :

- d_i désigne une réalisation de D sur un intervalle i ,
- I désigne l'ensemble des intervalles i qui composent l'heure d'observation,
- et
- K désigne le nombre de fausses pistes observées durant l'heure d'observation, sur une zone prédéfinie de 900 NM² données.

En considérant une seule zone de 900 NM², pour N intervalles dans une heure, le critère correspond à un maximum de n_r fausses pistes.

Il peut alors être défini une distribution empirique renormalisée D^* . Par définition, D^* est équivalente à D , mais les valeurs sont multipliées par $\frac{n_r}{M \times N}$.

Il vient ainsi :

$$K^* = \sum_{i \in I} d_i^*$$

Où :

- d_i^* désigne une réalisation de D^* sur un intervalle i , et
- K^* est une variable aléatoire de la variable d'intérêt (nombre de fausses pistes) sachant que la distribution du nombre de fausses pistes par intervalle est distribuée selon la loi D^* .

L'espérance de K^* est n_r , à savoir le seuil requis.

L'hypothèse nulle est que le nombre de fausses pistes observées dans la zone pendant un intervalle suit la loi D^* .

La p-valeur s'écrit ainsi :

$$\rho = P\left(K^* \leq \sum_{i \in I} n_i \mid d_i^* \sim D^*\right)$$

Où :

- ρ désigne la p-valeur,
- P désigne la probabilité,
- n_i le nombre de fausses pistes observées dans la zone durant l'intervalle i ,
- et
- $\sim$ signifie "suit la loi", et donc ici que d_i^* suit la loi D^* .

Comme le nombre d'intervalles est de plusieurs centaines d'intervalles i , il est possible d'appliquer le théorème central limite, donnant :

$$\rho = \Phi\left(\frac{\sum_{i \in I} n_i - n_r}{n_r S} M \sqrt{N}\right)$$

Où :

- Φ est la fonction cumulative de la distribution normale centrée réduite.

Ainsi, pour chacune des zones z de 900NM^2 , il est possible de calculer la p-valeur sur cette zone ρ_z .

5 Comme précédemment, il peut être vérifié que c'est bien le maximum de ces valeurs sur chacune des zones qui est inférieure au seuil.

Ainsi, la p-valeur globale (sur toutes les zones $z \in Z$ où Z est l'ensemble des zones) se calcule comme suit :

$$\rho_z = 1 - \prod_{z \in Z} (1 - \rho_z)$$

10 En ce qui concerne le critère R20, il est considéré une métrique correspondant au nombre de fausses pistes proches de pistes réelles pour chaque intervalle d'une heure.

Le même raisonnement que pour le critère R19 conduit à l'expression suivante pour la p-valeur :

$$\rho = P\left(K^* \leq \sum_{i \in I} n_i \mid d_i^* \sim D^*\right)$$

15 Où D^* est la distribution empirique renormalisée du nombre de fausses pistes proches de pistes réelles sur chacun des intervalles.

En ce qui concerne le critère R9, il peut être appliqué une stratégie similaire.

Soit :

- D la distribution observée des âges des rapports contenant une altitude-pression,
- D^* la distribution D remise à l'échelle pour avoir une moyenne de 16 secondes (le seuil requis), et
- C^* la fonction cumulative de D^* .

25 Il peut alors être calculé la p-valeur du test d'hypothèse dont l'hypothèse nulle est que la variable de l'âge des rapports contenant une altitude pression suit la distribution D^* .

Il vient ainsi l'expression suivante pour la p-valeur :

$$\rho = 1 - \prod_{a \in R} (1 - C^*(t_a))$$

Où :

- R désigne un ensemble de rapports, chaque rapport comportant une pression altitude,
- a désigne un rapport, et

- t_a désigne l'âge du rapport a et est suppose indépendant des autres âges.

5 **Calcul de la valeur optimale du nombre de points pour la détermination d'un critère**

Les paramètres influençant le nombre optimal de points pour chaque critère en fonction des familles précédentes sont donnés dans le tableau qui suit.

<i>Exigence</i>	<i>Famille</i>	<i>Paramètres influençant le nombre optimal de points</i>
R2	1	- nombre d'erreurs commises - nombre de trajectoires - nombre d'erreurs commises par trajectoire
R3	3	- nombre d'erreurs commises
R4	2	- moyenne - écart-type - nombre de trajectoires - moyenne par trajectoire - écart-type par trajectoire
R5	3	- nombre d'erreurs commises
R6	2	- moyenne - écart-type
R7	1	- nombre d'erreurs commises - nombre de cas où la situation de référence ne se produit pas
R8	4	- moyenne - écart-type
R9	5	- moments sur les lois empiriques
R10	3	- nombre d'erreurs commises - nombre de cas où la situation de référence ne se produit pas
R11	3	- nombre d'erreurs commises
R12	4	- moyenne

		- écart-type
R13	4	- moyenne - écart-type
R14	1	- nombre d'erreurs commises - nombre de cas où la situation de référence ne se produit pas - nombre de trajectoires - nombre d'erreurs commises par trajectoire - nombre de cas où la situation de référence ne se produit pas par trajectoire
R15	3	- nombre d'erreurs commises - nombre de cas où la situation de référence ne se produit pas
R16	2	- moyenne - écart-type
R17	2	- moyenne - écart-type
R18	2	- moyenne - écart-type
R19	5	- moyenne empirique de données passées - écart-type empirique de données passées - nombre d'erreurs commises
R20	5	- moyenne empirique de données passées - écart-type empirique de données passées - nombre d'erreurs commises

Ainsi, pour déterminer le nombre de points pour chaque exigence ainsi que la discrétisation spatio-temporelle, il convient de renseigner les valeurs des paramètres d'influence.

- 5 Pour cela, il est supposé qu'on dispose d'un échantillon D de valeurs acquises normales (devant passer toutes les exigences) représentatives de la situation que l'on étudie (données du même aéroport ou de la même zone En-Route).

Il est alors calculé a priori, pour chacune des exigences, le nombre moyen de points nécessaires pour que la p-valeur puisse atteindre le seuil désiré (si les données le corroborent). En effet, si le nombre de point est trop faible, la p-valeur sera trop haute pour rejeter l'hypothèse nulle avec suffisamment de certitude, et ce même si les données sont

5

« parfaites » (c'est-à-dire qu'elles tendent à valider l'exigence).
Par conséquent, en supposant les données parfaites, il est possible d'identifier le nombre minimal de points nécessaires pour que la p-valeur passe sous le seuil requis.

10

Cela permet ensuite de décider de faire le calcul en temps réel, en se basant sur un nombre de données supérieur. Pour cela, avec l'aide d'un expert, il est déterminé un compromis entre un petit nombre de données, qui permet un système beaucoup plus réactif puisqu'on se contente de moins d'intervalles de temps observés, et un grand nombre de données, qui réduit le dynamisme du système, mais offre plus de précision (et donc une meilleure vision de la situation).

15

Ainsi, l'identification du nombre de points minimum peut être mise en œuvre en mettant en œuvre une série d'opérations.

Lors d'une première opération, il est obtenu la durée de surveillance maximale souhaitée afin d'assurer une réactivité suffisante. Cette durée maximale est fournie par l'utilisateur du système de pistage 10.

20

De cette durée, il est déduit un nombre de points M_{max} maximal que l'on pourra obtenir sur cette durée.

Lors d'une deuxième opération, il est obtenu un échantillon D de valeurs normales, c'est-à-dire un échantillon respectant l'ensemble des exigences. Ces valeurs normales sont appelées données normales dans la suite.

25

Lors d'une troisième opération, il est calculé, pour chaque donnée normale $d \in D$, les valeurs des paramètres influençant le nombre minimal de points pour chaque exigence. Il est déduit le nombre minimal m_i^d de points qu'il faudrait pour passer l'exigence i sur cette donnée d .

En pratique, trois cas vont se présenter.

30

Dans un premier cas, si $m_i^d \leq M_{max}$ pour une exigence i et toute donnée d , alors il s'agit du cas idéal où toutes les données dont on dispose passent les exigences avec une très forte confiance, en utilisant M_{max} données.

On peut également utiliser un nombre de données pour l'exigence i dépendant de $\max_{d \in D} m_i^d$ (auquel une marge est rajoutée).

35

Selon un deuxième cas, $m_i^d \leq M_{max}$ pour la large majorité des exigences i et la majorité des données d , mais $m_i^d > M_{max}$ dans quelques cas isolé.

Dans un tel cas, il est quantifié le nombre ou pourcentage de données de D qui ne passent pas les exigences avec une forte confiance. L'utilisateur fixe un seuil de tolérance, en dessous duquel on tolère que des données correctes lèvent des alertes. Dans ce cas, on utilise M_{max} données, ou un nombre de données pour le KPI i dépendant de $\max_{d \in D} m_i^d$.

5 Dans le troisième cas, un nombre significatif de données de D ne passent pas les exigences.

Un exemple de nombre significatif dans ce contexte est par exemple la moitié des données totales. En effet, si une mesure sur deux est mauvaise, il peut être considéré qu'il y a un problème de fonctionnement du système de pistage 10 ou du moins qu'il convient de vérifier le fonctionnement de celui-ci.

Dans un tel cas, cela signifie que ces exigences ne sont pas compatibles avec une évaluation en temps réel. C'est par exemple le cas de la détection d'événement rares.

Il peut être noté ici qu'une autre possible de gain en terme de nombre de points de données sans rallonger la durée d'acquisition est d'élargir la zone géographique d'observation. Pour cela, il peut être regroupé deux zones adjacentes par lesquelles 15 passent peu de trafic, afin d'avoir, sur cette zone plus large, suffisamment de données pour estimer précisément la satisfaction ou non des exigences.

Les zones initiales (avant regroupement) sont prédéterminées, notamment choisies par l'utilisateur du système de pistage 10.

20 Dans ce qui suit, il est étudié pour certains critères le nombre de points nécessaires pour obtenir une p-valeur inférieure ou égale à 0,05 en fonction de données dégradant la conformité au critère.

Critère	Seuil à atteindre	Nombre d'erreurs			
		0	1	2	3
R2 plus de 100 points	> 0,97	99	157	208	257
	> 0,99	299	473	628	773
R2 moins de 100 points	> 0,97	99	>100	>100	>100
R3	< 0,005	598	947	1258	1549
R5	< 0,0003	9985	15812	20984	25844
R11 (cas d'un vol à altitude constante)	< 0,001	2995	4742	6294	7752
R11 (cas d'un atterrissage ou d'un décollage)	< 0,015	199	315	418	515

Tableau 3 : Nombre de points nécessaires pour atteindre une p-valeur de 0,05 ou moins pour les critères R2, R5 et R11

5

Dans le tableau 3, une erreur est un rapport problématique tendant à dégrader le critère.

e_R	e_A							
	0	1	2	3	4	5	10	20
0	57	100	138	173	208	241	401	702
1	81	119	154	188	222	254	411	710
2	100	135	169	202	235	267	422	718
3	116	150	183	215	247	278	431	726
4	130	163	195	227	258	289	440	733
5	143	175	206	238	268	299	449	741
10	195	225	255	285	315	344	490	776
20	273	302	331	360	388	417	558	838

Tableau 4 : Nombre de points nécessaires pour atteindre une p-valeur de 0.05 ou moins, pour le critère R7

10

Dans le tableau 4, e_R est le nombre de rapports ne contenant pas de position horizontale et e_A est le nombre de rapports contenant une position horizontale mais ne contenant pas d'altitude-pressure.

e_R	e_A							
	0	1	2	3	4	5	10	20
0	115	200	276	348	418	485	806	1410
1	163	238	300	375	444	510	825	1422
2	200	271	339	405	470	534	844	1438
3	232	299	365	430	493	556	863	1453
4	260	325	389	453	515	577	881	1468
5	285	349	412	474	536	597	898	1462
10	388	448	508	567	626	685	977	1549
20	539	596	654	711	768	825	1107	1666

5

Tableau 5 : Nombre de points nécessaires pour atteindre une p-valeur de 0.05 ou moins, pour le critère R14

Dans le tableau 5, e_R est le nombre de rapports ne contenant pas de position horizontale et e_A est le nombre de rapports.

10

P_R	ϵ_A								
	0	1	2	3	4	5	10	20	100
1.0	3000	4800	6300	7800	9200	10600	17000	29100	ND
0.98	3100	4900	6500	8000	9400	10800	17400	29700	ND
0.96	3200	5000	6600	8100	9600	11000	17700	30300	ND
0.9	3400	5300	7000	8700	10200	11700	18900	32300	ND
0.75	4000	6400	8400	10400	12300	14100	22700	38800	ND
0.5	6000	9500	12600	15600	18400	21100	34000	ND	ND
0.25	12000	19000	25200	31100	36700	42100	ND	ND	ND

Tableau 6 : Nombre de points nécessaires pour atteindre une p-valeur de 0.05 ou moins, pour le critère R10

15

Dans le tableau 6, le nombre de points nécessaire est arrondi à la centaine. En outre, p_R désigne la proportion de rapports contenant une altitude-pression (resp identifiant d'aeronef), e_A est le nombre de rapports contenant altitude-pression invalide et ND signifie que ce nombre est supérieur à 50 000.

5 L'analyse des tableaux 3 à 6 montrent que, dans la plupart des cas, le nombre optimal de points est inférieur à 50 000.

Les seuls cas où le nombre optimal excède 50 000 sont des cas où le nombre d'erreurs est très important.

10 Un grand nombre d'erreurs étant un événement relativement rare, ces tableaux montrent que le procédé permet, dans des fonctionnements normaux, d'obtenir des évaluations de la conformité des critères avec un nombre réduit de mesures.

REVENDEICATIONS

1. Procédé d'évaluation de la conformité d'un système de pistage (10) à un ensemble d'exigences, au moins une exigence étant une exigence obligatoire et au moins
5 une exigence étant une exigence recommandée, chaque exigence imposant qu'une grandeur physique soit supérieure ou égale à un seuil ou qu'une grandeur physique soit inférieure ou égale à un seuil, le procédé d'évaluation étant mis en œuvre par ordinateur, le procédé d'évaluation comportant au moins les étapes suivantes :

- acquisition d'un ensemble de valeurs, l'ensemble de valeurs comprenant les
10 valeurs accessibles au système de pistage (10) sur un intervalle de temps prédéfini,
- pour chaque exigence, calcul d'une valeur de robustesse, la robustesse étant calculée comme la valeur p d'un test statistique indiquant si l'exigence considérée est respectée ou non,
- pour chaque exigence, comparaison de chaque valeur de robustesse à un seuil
15 prédéfini selon un critère de comparaison respectif, et
- détermination des exigences auxquelles le système de pistage (10) est conforme, les exigences déterminées comme conformes étant les exigences pour lesquelles, à l'étape de comparaison, il a été déterminé que le critère de comparaison est respecté.

20

2. Procédé selon la revendication 1, dans lequel, le procédé comporte, en outre, pour chaque valeur de robustesse ne respectant pas le critère de comparaison, une étape de détermination de la cause du non-respect du critère de comparaison, la cause étant soit que l'exigence n'est pas respectée soit que le nombre de valeurs est trop faible pour garantir
25 un seuil de fiabilité prédéfini d'évaluation de la conformité à l'exigence considérée.

3. Procédé selon la revendication 2, dans lequel, lorsque la cause déterminée est un nombre de valeurs trop faible, le procédé comporte une étape de déduction d'un nombre de valeurs supplémentaires à acquérir pour obtenir une fiabilité d'évaluation supérieure au
30 seuil de fiabilité prédéfini.

4. Procédé selon la revendication 3, dans lequel, le procédé comporte une répétition des étapes d'acquisition, de calcul, de comparaison jusqu'à ce que le nombre de valeurs soit permette d'obtenir une fiabilité d'évaluation supérieure au seuil de fiabilité
35 prédéfini soit atteigne un nombre maximal prédéfini avec une fiabilité d'évaluation inférieure au seuil prédéfini.

5 **5.** Procédé selon la revendication 4, dans lequel le procédé comporte, en outre, une étape d'émission d'une alerte pour les exigences non respectées ou pour les exigences dont le nombre de valeurs des étapes d'acquisition, de calcul, de comparaison a atteint le nombre maximal prédéfini avec une fiabilité d'évaluation inférieure au seuil prédéfini.

6. Procédé selon la revendication 4 ou 5, dans lequel le nombre maximal est compris entre 40000 et 60000.

10 **7.** Procédé selon l'une quelconque des revendications 1 à 6, dans lequel, lors de l'étape de calcul, il est calculé une intégrale d'une fonction bêta incomplète régularisée.

8. Procédé selon la revendication 7, dans lequel, lors de l'étape de calcul, une p-valeur est égale au résultat de l'intégrale et une autre p-valeur est une fonction linéaire du
15 résultat de l'intégrale.

9. Procédé selon l'une quelconque des revendications 1 à 8, dans lequel lors de l'étape de calcul, il est calculé une fonction de répartition de la distribution normale centrée réduite.
20

10. Procédé selon l'une quelconque des revendications 1 à 9, dans lequel le procédé comporte, en outre :

- une étape d'application d'une fonction d'évaluation propre à chaque exigence aux valeurs p pour obtenir des valeurs d'évaluation, et
- 25 - une étape d'utilisation des valeurs d'évaluation pour obtenir un score global évaluant la conformité du système de pistage (10) à l'ensemble des exigences.

11. Procédé selon l'une quelconque des revendications 1 à 10, dans lequel le nombre d'exigences est supérieur ou égal à 20.
30

12. Procédé selon l'une quelconque des revendications 1 à 11, dans lequel le système de pistage (10) fournit des estimations de mesures à partir de données antérieures issues de plusieurs capteurs (12), l'ensemble de valeurs comprenant les estimations de mesure.
35

13. Module d'évaluation (22) de la conformité d'un système de pistage (10) à un ensemble d'exigences, au moins une exigence étant une exigence obligatoire et au moins

une exigence étant une exigence recommandée, chaque exigence imposant qu'une grandeur physique soit supérieure ou égale à un seuil ou qu'une grandeur physique soit inférieure ou égale à un seuil, le module d'évaluation (22) étant propre à :

- 5 - acquérir un ensemble de valeurs, l'ensemble de valeurs comprenant les valeurs accessibles au système de pistage (10) sur un intervalle de temps prédéfini,
- pour chaque exigence, calculer une valeur de robustesse, la robustesse étant calculée comme la valeur p d'un test statistique indiquant si l'exigence considérée est respectée ou non,
- 10 - pour chaque exigence, comparer chaque valeur de robustesse à un seuil prédéfini selon un critère de comparaison respectif, et
- déterminer des exigences auxquelles le système de pistage (10) est conforme, les exigences déterminées comme conformes étant les exigences pour lesquelles, à l'étape de comparaison, il a été déterminé que le critère de comparaison est respecté.

15

14. Système de pistage (10) comprenant un module d'évaluation (22) selon la revendication 13.