|  |  |
| --- | --- |
|  | AKaDo 2.0 Guide utilisateur |
|  |  |
| 26/04/2017 | ANALYSE DE LA QUALITE DE LA DONNEE POUR AVDTH |
|  |  |

AKaDo 2.0 Guide utilisateur

ANALYSE DE LA QUALITE DE LA DONNEE POUR AVDTH

Sommaire

[Introduction 2](#_Toc474301212)

[Installation 3](#_Toc474301213)

[UTILISATION 8](#_Toc474301214)

[JEU de contrÔles 11](#_Toc474301215)

[Anapo 29](#_Toc474301216)

[Remerciements 35](#_Toc474301217)

# Introduction

## Contexte

L'utilitaire AKADO réalise automatiquement un ensemble de tests sur les données et produit des tableaux synthétiques qui dressent un bilan plus ou moins détaillé des anomalies détectées. Un résumé final présente les pourcentages d'occurrences des erreurs qui restent à corriger.

Historiquement, l'utilitaire AKADO traitent des bases de données compatibles avec la dernière évolution du modèle AVDTH (version 3.4) et de la version précédente 3.3. La suite des prédicats évalués n'a pas évolué depuis la version 3.3 de l'application. L'évolution majeure de la dernière version est une traduction en anglais et l'ajout de commentaires dans le rapport d'analyse. Bien que de nombreux contrôles soient effectués en temps réel à la saisie, et qu'à la demande, des outils complémentaires réalisent des tests croisés, il apparaît à l'usage, que le volume des données dans les bases AVDTH a un impact sur le nombre et la qualité des corrections effectuées. Les bases de données « à destination de T3+» ne doivent pas comporter d'erreurs majeures; c'est une garantie à laquelle s'engage la technologie AVDTH.

Dans le cadre du projet de « Journal de pêche électronique (ERS)», nous avons décidé de redévelopper AKaDo pour en faire un logiciel de contrôles générique et modulaire afin qu'il puisse, selon les besoins, faire des tests sur plusieurs bases.

Cette version produit un document de type « feuille de calcul » facilitant la lecture et permettant d’améliorer le traitement des erreurs.

## Prérequis

Être un utilisateur de la base de données AVDTH, « Acquisition et Validation des Données Thonières », créé par l’Observatoire Thonier (OT) de l’Institut de recherche pour le développement (IRD).

Les utilisateurs de la base AVDTH sont : France, Ghana, Maurice, Sénégal, Seychelles et Espagne (1er octobre 2013).

## Divers

Cette version d’AKaDo est développée par Julien Lebranchu. Contact : [l'équipe support de l'OT](mailto:support.obs.thonier@listes.ird.fr?subject=[AKADO]%20).

Le guide a été créé le 15 mars 2016.

La date de dernière modification est le 12 mars 2019 par Julien Lebranchu.

Révision numéro: 269.

# Installation

## Préambule

Pour utiliser le logiciel AKaDo, JAVA doit être préalablement installé. Vous pouvez télécharger la dernière version de JAVA à [https://www.java.com/en/](http://www.microsofttranslator.com/bv.aspx?from=en&to=fr&a=https%3A%2F%2Fwww.java.com%2Fen%2F). Vous devez ensuite suivre les instructions pour l'installer.

## Procédure

Exécutez le programme d'installation d'**AKaDo** en double-cliquant sur le fichier « akado-avdth-installer-XX-standard.jar », où XX est le numéro de version; et suivez les instructions pour installer ou mettre à jour **AKaDo**.

1. Presser « Suivant » pour continuer.
2. Lire le *README*, et presser « Suivant » pour continuer.
3. Accepter la licence d'**AKaDo**, puis appuyez sur « Suivant » pour continuer.
4. Sélectionner les packages à installer (seulement la documentation, les fichiers *Licence* et *README* ne sont pas obligatoires).
5. Sélectionnez le chemin d'installation.
6. Le panneau présente des informations sur l'installation. Appuyez sur « Suivant » pour continuer et commencer l'installation.
7. L’installation est en cours.
8. L’installation est terminée.

## Procédure de désinstallation

Pour désinstaller **AKaDo**, vous pouvez suivre la procédure :

1. Vous devez cliquer sur le raccourci du programme de désinstallation s‘il existe dans le menu du système d'exploitation, ou vous pouvez également exécuter le programme de désinstallation qui se trouve dans le répertoire de l'application. Si vous n'avez pas modifié ce répertoire au moment de l'installation :
   * Sous Windows, le fichier de désinstallation devrait être dans C:\Obs\_Thonier\AKaDo2\Uninstaller ;
   * Sous Linux, le fichier de désinstallation devrait être dans /opt/AKaDo2/Uninstaller.
2. Uniquement si nécessaire, vous pouvez supprimer les fichiers de configuration en utilisant le script adapté à votre système d’exploitation :
   * Sous Windows, par défaut, le script est C:\Obs\_Thonier\AKaDo2\delete-config.bat.
   * Sous Linux, par défaut, le script est /opt/AKaDo2/delete-config.sh

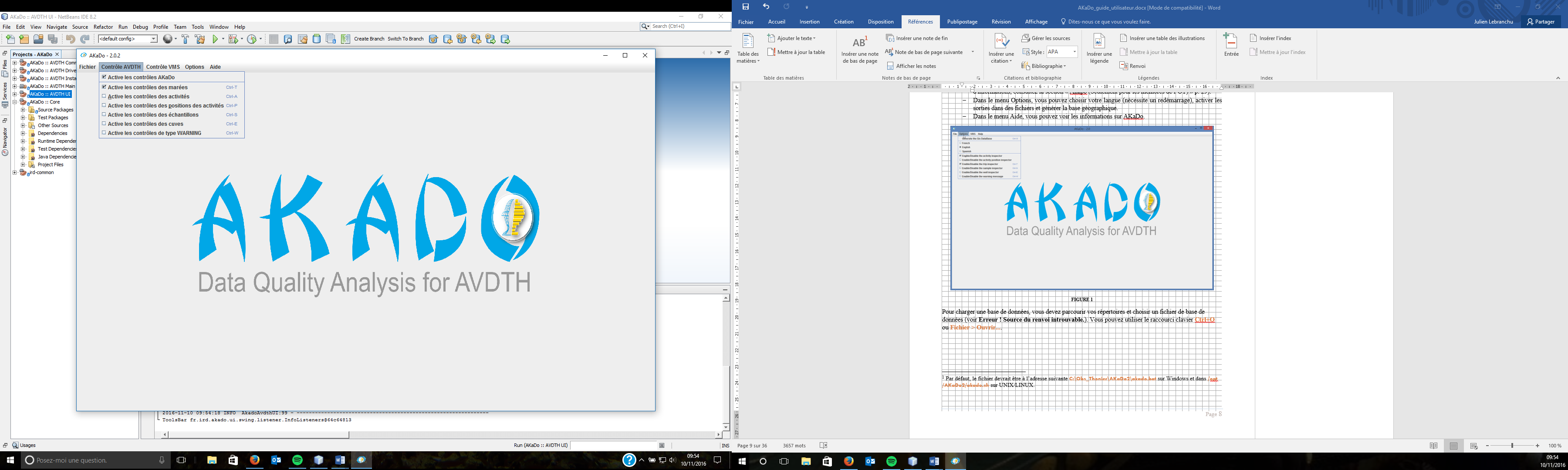
UTILISATION

**Lancer AKaDo**

**Pour lancer AKaDo, vous pouvez exécuter le script qui se trouve dans le répertoire de l'application**[[1]](#footnote-2)**.**

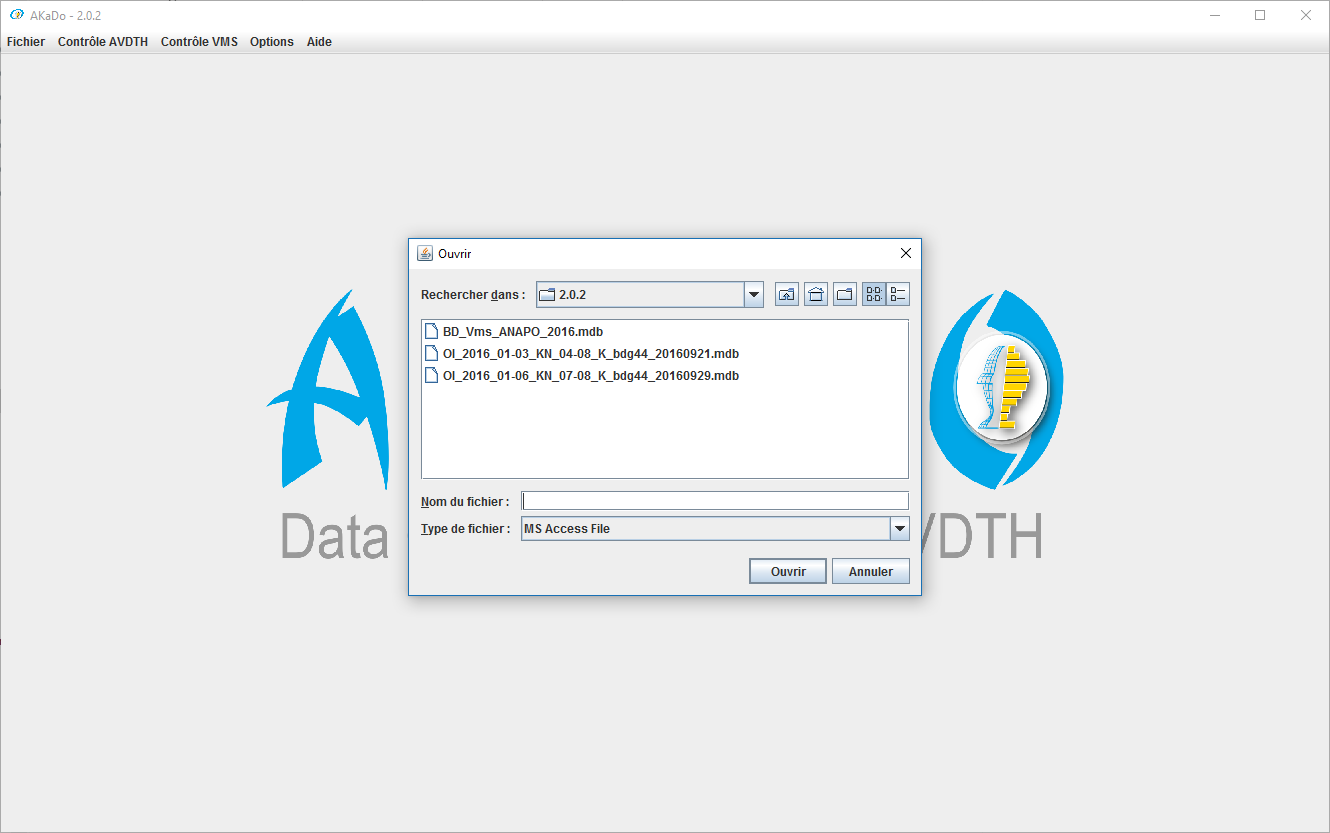
Lorsque vous exécutez AKaDo, l'interface est affichée (voir ci-dessous). Il y a cinq menus : Fichier, Contrôle AVDTH, Contrôle VMS (Anapo), Options et Aide.

* Dans le menu fichier, vous pouvez charger une base de données AVDTH ou quitter l'application.
* Dans le menu Contrôle AVDTH, vous pouvez activer/désactiver tous les inspecteurs.
* Dans le menu Contrôle VMS, vous pouvez gérer les fonctionnalités d’ANAPO (pour plus d'informations, consultez la section « Anapo (Seulement pour les membres de l’OT) » p. 29).
* Dans le menu Options, vous pouvez choisir votre langue (nécessite un redémarrage), activer les sorties dans des fichiers et générer la base géographique.
* Dans le menu Aide, vous pouvez voir les informations sur AKaDo.



**Figure 1**

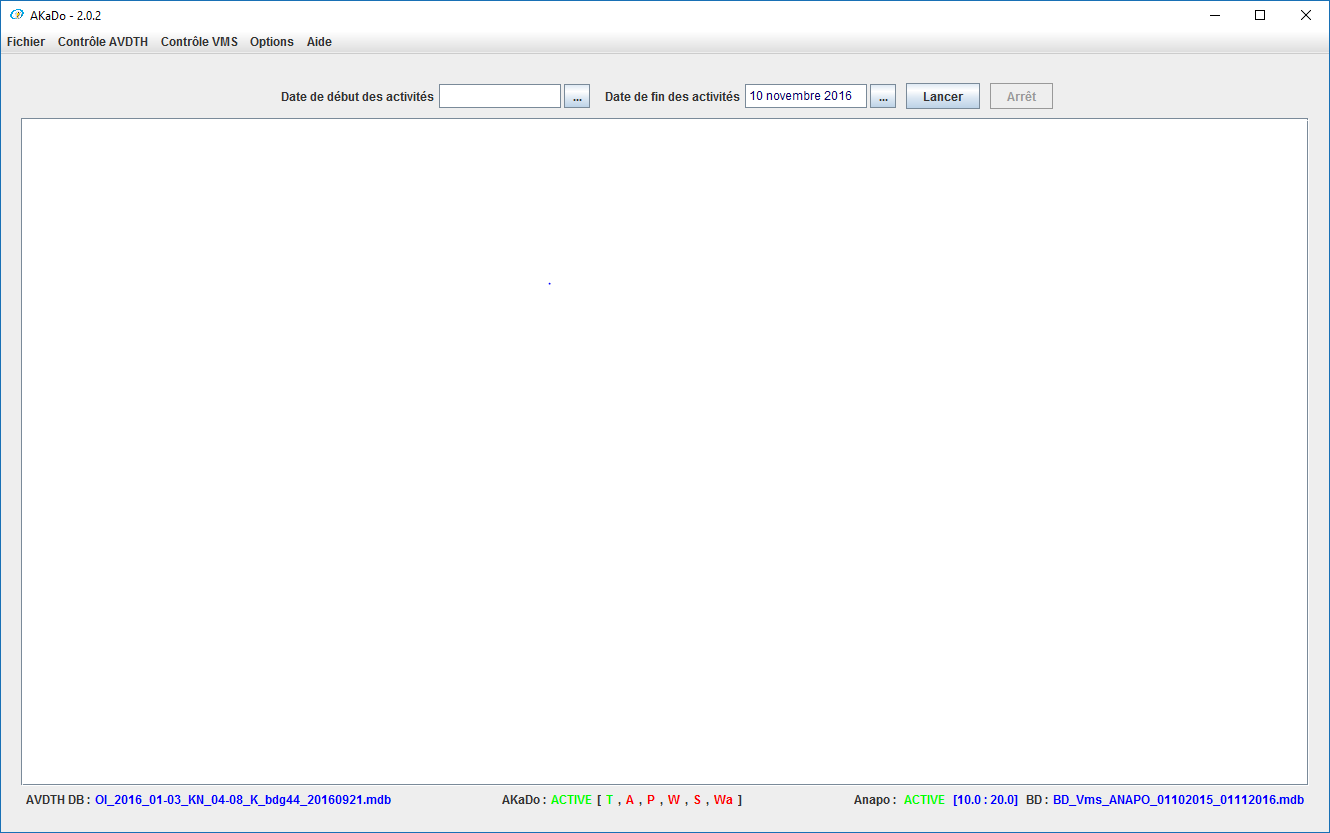
Pour charger une base de données, vous devez parcourir vos répertoires et choisir un fichier de base de données (voir **Figure 2** en page 9). Vous pouvez utiliser le raccourci clavier **Ctrl+O** ou **Fichier > Ouvrir...**.



**Figure 2**

Une fois que vous avez chargé la base de données, il est possible de restreindre la plage des contrôles en sélectionnant une plage temporelle, vous devez appuyer sur le bouton **Lancer** pour exécuter le processus (voir **Figure 3** en page 10). Il est possible d’arrêter le traitement en cliquant sur le bouton **Arrêt.** Au cours du traitement, l'application affiche pour chaque contrôle non valide un message[[2]](#footnote-3). À la fin, un fichier de résultat est généré et le programme affiche le chemin d'accès sur votre système. Ce fichier est une feuille de calcul, et il peut être ouvert avec différents logiciels comme **Microsoft Excel** ou **LibreOffice Calc**[[3]](#footnote-4)**.**

A noter que pour la plage temporelle, il est possible de ne pas remplir les champs, les contrôles se feront sur l’ensemble des données de la base.



Sélection de la plage temporelle des activités

**Figure 3**

**Résultats dans la feuille de calcul**

Pour chaque groupe de contrôles, il y a une feuille dédiée : marée, activité... Par défaut, nous avons réalisé une mise en forme conditionnelle. Par exemple, si le poids de la capture en activité est différent du poids des captures élémentaire la cellule est de couleur rouge (voir **Figure 4**). Il y a deux couleurs attribuées : orange pour les « warnings » et rouge pour les erreurs.



**Figure 4**

# JEU de contrÔles

## Marée

### Activité (Warning)

* Nous vérifions si la marée à au moins une activité et que le flag “F\_ENQ” vaut 1.
* Le flag F\_ENQ est-il cohérent avec la déclaration officielle ? Vérifier la base avec l’application AVDTH -> Débarquements -> Mise à jour des débarquements -> Enquête ?

### Temps de pêche

* Nous vérifions si le temps de pêche est égal à la somme des temps de pêche de chaque activité.
* Si les valeurs ne sont pas équivalentes, vous devez reporter la somme dans le champ « Temps de pêche » de la marée.

### Temps de mer

* Si le temps de mer est égal à la somme des temps de mer de chaque activité.
* Si les valeurs ne sont pas équivalentes, vous devez reporter la somme dans le champ « Temps de mer » de la marée.

### Capacité du navire (Warning)

* Nous calculons la capacité totale en tonnes du navire, puis vérifions si cette capacité est supérieure au poids total débarqué.
* Si le poids total débarqué est supérieur à la capacité du navire, vous devez vérifier si le « poids débarqué » est correct.

### Poids total débarqué

* Nous vérifions si le poids total débarqué est cohérent avec les lots commerciaux.
* Si les valeurs ne sont pas égales, vous devez reporter la valeur de la somme des lots commerciaux dans le champ « Poids débarqué » de la marée.

### Distance (Warning)

* Nous vérifions si la distance est comprise entre 0 et la distance maximale obtenue à partir de la vitesse maximale et du nombre d’activités.

### Couverture temporelle

* Nous vérifions si les activités sont temporellement continues durant la marée.
* Vous devez vérifier dans le journal de pêche si un jour n’est pas manquant.

### Limite temporelle

* Nous vérifions si les limites temporelles, le premier et le dernier jour, de la marée sont cohérentes avec les activités.
* Vous devez vérifier si les dates de départ et d’arrivée correspondent dans le journal de pêche et les documents du débarquement.

### Port

* Nous vérifions si le port de départ de la marée est identique au port d’arrivée de la marée précédente.
* Vérifier si tous les journaux de pêches ont été saisis.
* Vérifier auprès du capitaine si des sorties n’ont pas été effectuées entre temps.

### Raising Factor (Info)

* Nous calculons le « raising factor » pour toutes les marées (incluant les marées avec un débarquement partiel).
* Si le ratio n’est pas entre 0.9 < R < 1.1 (Débarquement/Capture), vous devez vérifier la valeur de débarquement partiel.

#### Exploitation des erreurs



La figure ci-dessus met en évidence plusieurs erreurs issues de différents contrôles[[4]](#footnote-5) :

* Dans la colonne « Departure Harbour », les ports mis en évidence correspondent à un port de départ différent de celui d’arrivée de la marée précédente.
* Dans la colonne « Has logbook », il est important de vérifier la cohérence de la valeur 0 (FAUX) et 1 (VRAI) avec les informations des dates d’activités.
  + Si la valeur est à 1 alors la date de la première activité doit être renseignée.
  + Et inversement, si la valeur est à 0, alors la date de la première activité ne doit pas être renseignée.
* Dans la colonne « First Activity Date », les dates mises en évidence indiquent une différence entre la date de départ de la marée et la date de la première activité de la marée (« Departure Date »).
* Dans la colonne « Last Activity Date », les dates mises en évidence indiquent une différence entre la date de débarquement de la marée et la date de la dernière activité de la marée (« Landing Date »).



La figure ci-dessus met en évidence plusieurs erreurs issues de différents contrôles[[5]](#footnote-6):

* Dans la colonne « Time At Sea », la mise en évidence de la valeur « Temps en mer » indique une différence avec la somme des valeurs « Temps en mer » des activités de la marée.
* Dans la colonne « Fishing Time », la mise en évidence de la valeur « Temps de pêche » indique une différence avec la somme des valeurs « Temps de pêche » des activités de la marée.
* Dans la colonne « Landing Weight », la mise en évidence de la valeur « Poids débarqué » indique une différence avec la somme des poids débarqués des lots commerciaux de la marée.

## Activité

### Contexte de pêche

* Nous vérifions si le type de banc et l’association sont cohérents.
* Si le type de banc est « banc objet » (code 1), alors il doit y avoir au moins une association, et l’identifiant doit être une des valeurs suivantes 10, 60, 81 ou entre 20 et 28 ;
* Si le type de banc est « banc libre » (code 2), alors l’identifiant de l’association, si elle existe, ***ne doit pas*** être une des valeurs suivantes 10, 60, 81 ou entre 20 et 28.

### Opération

* Nous vérifions si l’opération associée à l’activité est cohérente avec les autres informations.
* Si l’identifiant de l’opération est 0, alors la valeur du poids total débarqué doit être égale à 0 ;
* Si l’identifiant de l’opération est 1 ou 2, alors la valeur du poids total débarqué doit être différente de 0 ;
* Si l’identifiant de l’opération est 12, 13 ou 14, alors la valeur du poids total débarqué doit être égale à 0.

### Position

* Nous vérifions si la position de l’activité est dans l’océan, au port ou à terre, et si la position et l’océan sont cohérents.
* Si la position est à terre, vous devez vérifier la latitude, la longitude et le quadrant.
* Si la position et l’océan sont différents, vous devez vérifier ces champs avec le journal de pêche.

### Quadrant

* Nous vérifions si le quadrant et la position de l’activité sont cohérents, i.e.
* Si la valeur du quadrant est 3 ou 4, l’océan doit être « Atlantic Ocean ».
* Si la valeur du quadrant est 3 ou 4, la position doit être localisée « Atlantic Ocean ».
* Vous devez vérifier les champs : océan, quadrant et position.
* Nous vérifions les positions limites (longitude = 0 ou latitude = 0) et le quadrant., i.e.

|  |  |  |  |
| --- | --- | --- | --- |
| LATITUDE | LONGITUDE |  | QUADRANT |
| 0 (Par convention NORD) | **OUEST** |  | **4** |
| 0 (Par convention NORD) | **EST** |  | **1** |
| NORD | **0 (Par convention Est)** |  | **1** |
| SUD | **0 (Par convention Est)** |  | **2** |

### Poids total des captures

* Nous vérifions si le poids total de captures est cohérent avec les captures élémentaires.
* Si les valeurs sont différentes, vous devez reporter la somme des captures élémentaires dans l’activité.

#### Exploitation des erreurs

****

La figure ci-dessus met en évidence plusieurs erreurs issues de différents contrôles[[6]](#footnote-7):

* Dans la colonne « Operation », la mise en évidence indique qu’il faut vérifier le type d’opération et la capture.
* Dans la colonne « Catch weight », la mise en évidence de la valeur indique une différence avec la somme des poids des captures élémentaires de l’activité.
* Dans les colonnes « School type » et « Fishing context », la mise en évidence indique une incohérence entre les colonnes, par exemple l’association
* Dans la colonne « Ocean AO/IO », la valeur est calculée à partir des positions de l’activité, et il y a une mise en évidence de la valeur associée dans la colonne « Ocean » quand elles diffèrent.
* Dans la colonne « In Land », la mise en évidence indique que la position de l’activité est à terre.
* Dans la colonne « Temperature », la mise en évidence indique que la température n’est pas comprise entre 15 et 32.
* Dans la colonne « Latitude », la mise en évidence indique la position et le quadrant ne sont pas cohérents.
* Dans la colonne « Longitude », la mise en évidence indique la position et le quadrant ne sont pas cohérents.

## Echantillon

### Activité

* Nous vérifions si les informations pour chaque échantillon sont cohérentes avec l’activité associée.
* Vous devez examiner les informations suivantes : la date, le numéro d’activité, le quadrant, la latitude, la longitude et le type de banc.

### Classe de taille

* Nous vérifions si la classe de taille est cohérente avec la classe de taille de l’espèce (L=55cm pour YFT et BET, et L=42cm pour ALB).

### Espèces

* Nous vérifions si l’espèce échantillonnée est autorisée.

### Mesure

* Nous vérifions si le nombre d’individus mesurés est cohérent avec le nombre de mesures.

### Position

* Nous vérifions si la position de l’activité pour chaque échantillon est cohérente.

### Echantillon sans mesure

* Nous vérifions si l’échantillon a au moins une mesure.

### Echantillon sans Espèces

* Nous vérifions si l’échantillon a au moins une espèce.

### Echantillon sans Marée

* Nous vérifions si l’échantillon est lié à une marée.

### Super Echantillon

* Nous vérifions si les numéros de sous-échantillons sont cohérents.

### Cuve

* Nous vérifions si le numéro de cuve de l’échantillon existe.

### Ratio de petit et de gros poissons

* Nous vérifions si les pourcentages de petits et de gros poissons sont cohérents.

### Pondération

* Nous examinons que la pondération pour chaque échantillon est cohérente.

#### Exploitation des erreurs

****

La figure ci-dessus met en évidence plusieurs erreurs issues de différents contrôles[[7]](#footnote-8):

* Dans la colonne « Sampling global weight », la mise en évidence de la valeur signifie que la colonne devrait être à 0 si les ou les valeurs associées pour les colonnes « Sampling weight M10» et « Sampling weight P10 » sont différentes de 0, ou inversement.
* Dans la colonne « Trip exist », la mise en évidence signifie que l’échantillon n'est pas lié à une marée.
* Dans la colonne « Activity », la mise en évidence signifie que l’échantillon n'est pas associé à une cuve.
* Dans les colonnes « Activity » et « Activity position », la mise en évidence signifie l’'activité associée n'existe pas. Par ailleurs, nous précisons si possible le champ qui pose problème à l’aide du marqueur « ? » et avec une mise en évidence de couleur bleue.



La figure ci-dessus met en évidence plusieurs erreurs issues de différents contrôles[[8]](#footnote-9):

* Dans la colonne « Species », la mise en évidence peut signifier les erreurs suivantes :
  + Si le code espèce est précédé d’un « ? », cela signifie que l’espèce ne devrait pas être échantillonnée ;
  + Si le code espèce est suivi de « ?LD1 », cela signifie qu’il y a une erreur sur les classes de tailles renseignées.
    - La classe de taille maximum pour YFT (1) et BET (3) est 55cm.
    - La classe de taille maximum pour ALB (4) est 42cm.
* Dans la colonne « LDLF », la mise en évidence peut signifier les erreurs suivantes :
  + Si le code LDLF est encadré de « !! », cela signifie que le code espèce est incompatible avec cette valeur de LDLF.
  + Si le code LDLF est précédé de « ? », cela signifie que les valeurs renseignées dans les colonnes « Sampling weight M10», « Sampling weight P10 » et « Sampling global weight » sont incohérentes avec cette valeur de LDLF.
    - Si LDLF égal 1 ou 3 alors P10 ou Poids global doit être supérieur à 0 sinon « warning ».
    - Si LDLF égal 2 alors M10 ou Poids global doit être supérieur à 0 sinon « warning ».
* Dans la colonne « Measured count », la mise en évidence signifie que le nombre de mesure est différent du nombre d’individus mesurés.
* Dans la colonne « Has well », la mise en évidence signifie qu’il n’y a pas de cuve associée à l’échantillon.
* Dans la colonne « Distribution +10/-10 », la mise en évidence signifie que les valeurs de répartition -10/+10 sont incohérentes avec celle de la cuve reportée.

## Cuve

### Activité

* Nous vérifions pour chaque plan de cuve s’il est cohérent avec la date et numéro d’activité.
* Vous devez examiner dans le plan de cuve la date et le numéro d’activité.
* Vous devez vérifier si le journal de pêche est dans AVDTH.

### Cuve sans Marée

* Nous vérifions si la cuve est associée à une marée existante.
* Vous devez saisir le journal de pêche dans AVDTH.

### Cuve sans plan de cuve

* Nous vérifions si la cuve est associée avec au moins un plan de cuve.

#### Exploitation des erreurs

#### 

La figure ci-dessus met en évidence plusieurs erreurs issues de différents contrôles[[9]](#footnote-10):

* Dans la colonne « Trip exist », la mise en évidence indique que la cuve n’est pas liée à une marée.
* Dans la colonne « Well Plan exist », la mise en évidence indique si le plan de cuve est manquant.
* Dans la colonne « Activity exist », la mise en évidence indique que l’activité associée à la cuve (vis-à-vis de la date et du numéro d’activité) n’existe pas.

# Anapo

Le principal test que nous réalisons est le calcul de distance entre la position déclarée et les positions VMS de la journée de l’activité avec des valeurs seuils permettant de déterminer s’il y a erreur ou non (cf. les images ci-dessous).

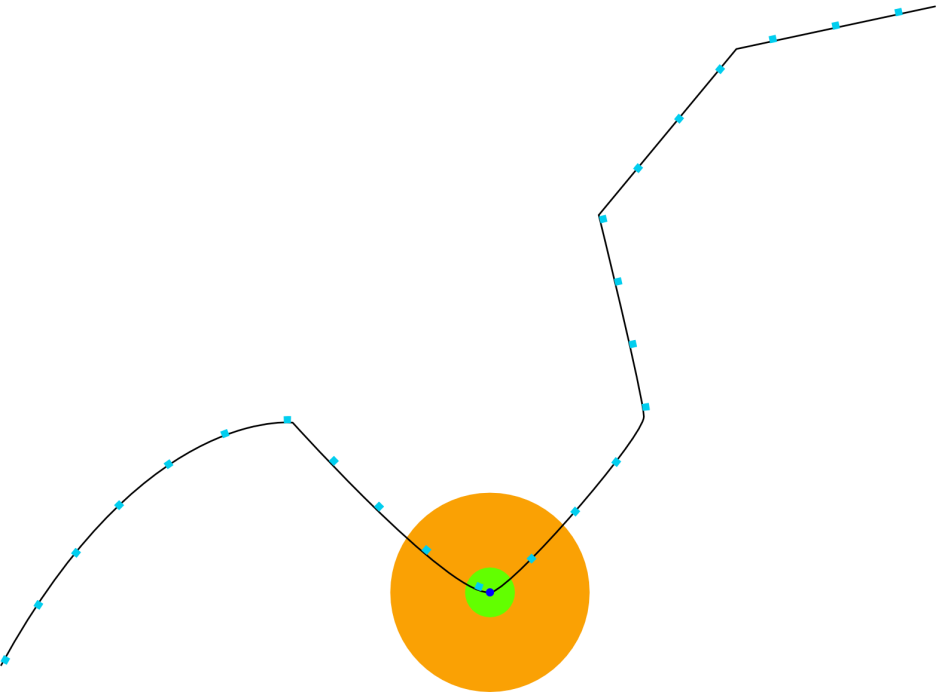


Figure 5 CAS OU LA POSITION EST VALIDE

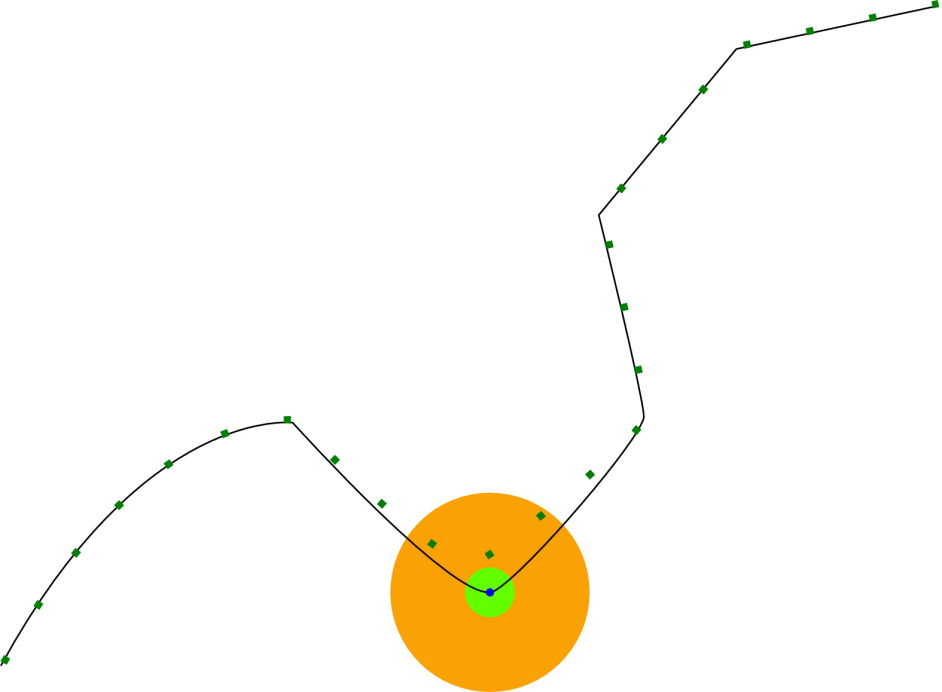


Figure 6 Cas où la position est dans la zone limite

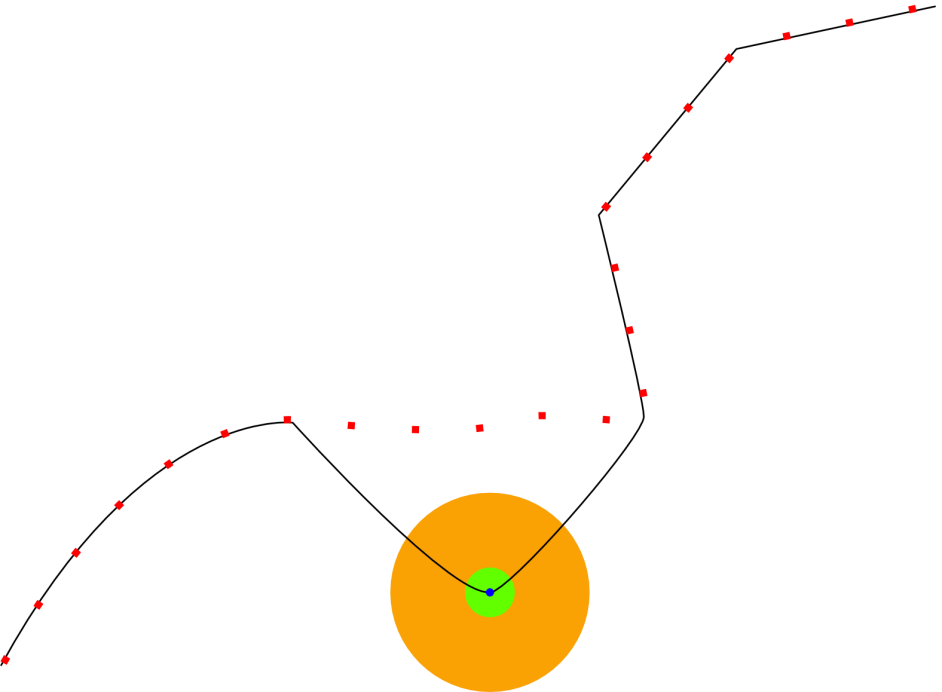


Figure 7 Cas où la position est non valide

Pour activer les contrôles ANAPO, il est nécessaire de réaliser les actions suivantes :

* Activer les contrôles ANAPO le menu Contrôles VMS (cf. la figure ci-dessous)

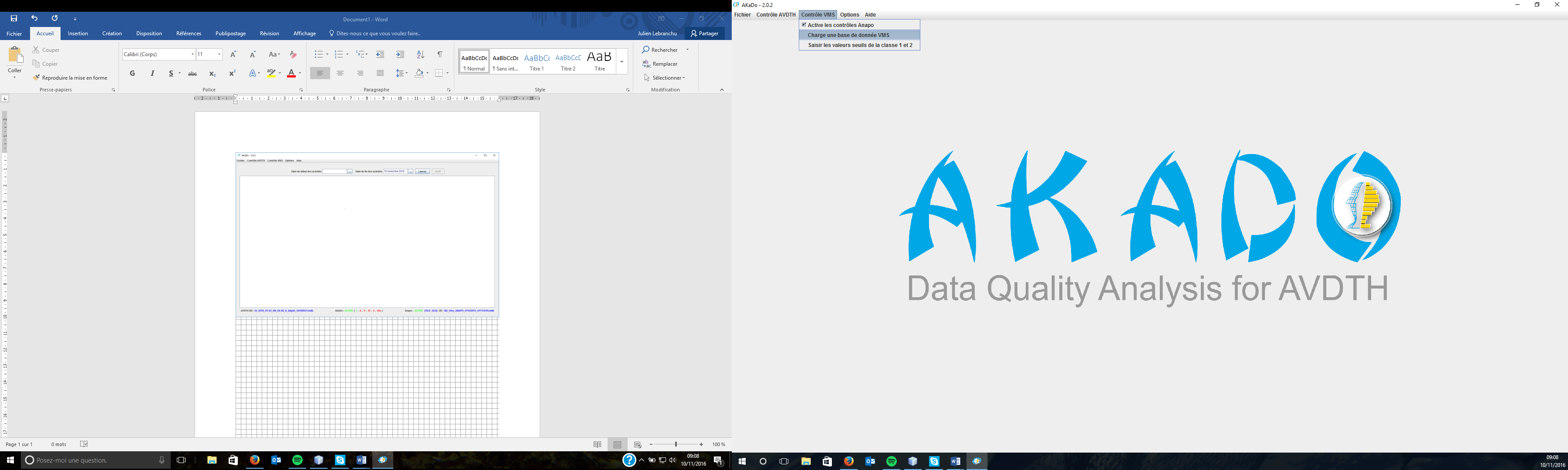


Figure 8

* Configurer les classes de distances dans le menu **Anapo > Saisir les valeurs seuils de la classe 1 et 2** (cf. Figure 9).

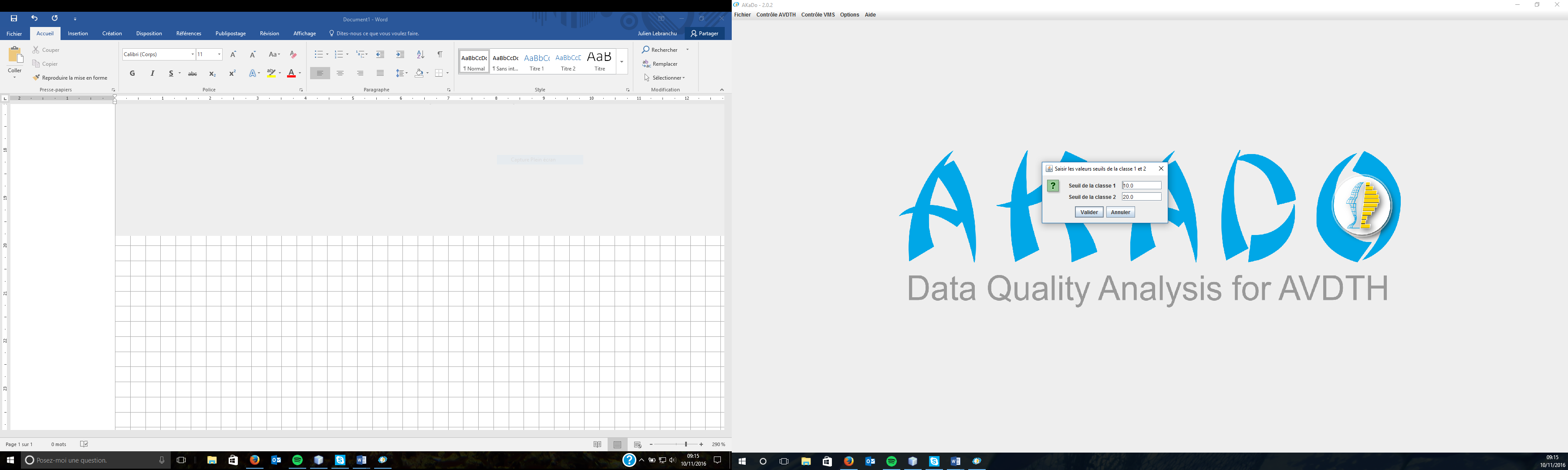


Figure 9

Dans le cas d’une analyse de positions VMS et d’une activité avec une marge d'erreur 10/20 milles, quatre cas se présentent :

1. Si au minimum une des distances entre l'activité et la VMS est comprise entre 0 et 10, alors c'est ok, et donc rien ne sera affiché dans le fichier de sortie.
2. Si au minimum une des distances entre l'activité et la VMS est comprise entre 10 et 20, alors il n’affiche rien dans le fichier de sortie.
3. Si au minimum une des distances entre l'activité et la VMS est comprise entre 20 et 2 fois la borne supérieure soit 40, alors il affiche dans le fichier de sortie chaque position entre 20 et 40 ainsi que le score associé. Ce score permettant de choisir les positions valides.
4. Si aucune position est en dessous de 40, alors il affiche toutes les positions VMS associé à l'activité.

Les sorties d’erreurs sont présentées à l’aide d’une feuille Excel (cf. Figure 10). Il est nécessaire dans la mesure du possible de consulter le capitaine et de consulter le journal de passerelle.

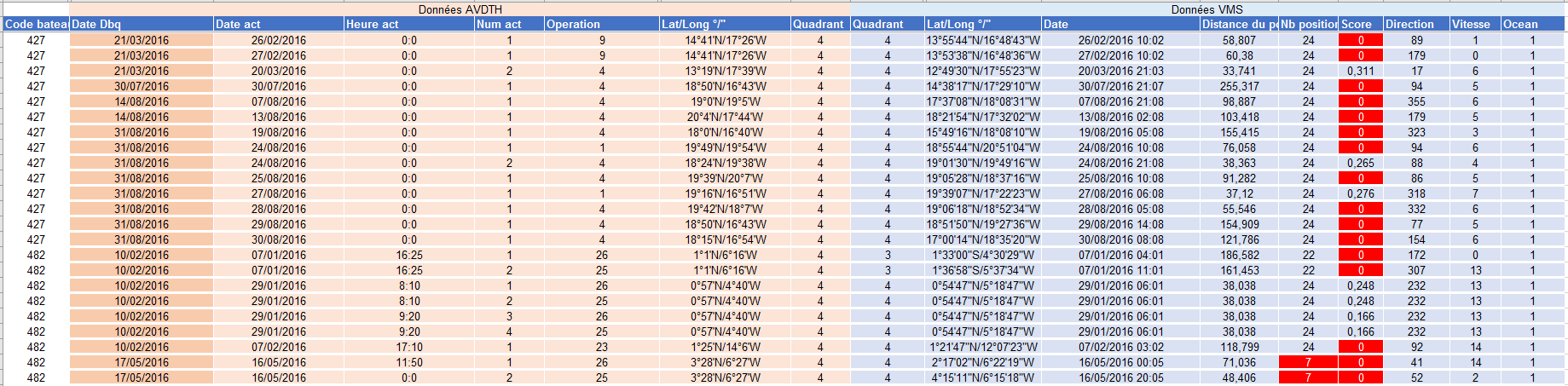


Figure 10

La figure ci-dessus présente les informations importantes suivantes :

* Dans la zone orange, les données des activités liées à la marée d’un navire.
* Dans la zone bleue, les données des VMS liées aux activités.
* Dans la colonne « Distance de la plus proche position », la valeur de distance, en mille nautique, entre la position de l’activité et celle de la VMS.
* Dans la colonne « Score », l’indice de confiance de la position VMS par rapport à celle de l’activité (distance géographique et temporelle). Si le score est nul, alors il est nécessaire de regarder les données ANAPO pour trouver une position adéquate.
* Dans la colonne « Nb position VMS », la mise en évidence indique qu’il manque significativement de positions.

# Remerciements

Jean-Jacques Lechauve et Laurent Floc’h pour le développement des versions précédentes.

Pascal Cauquil, Jean-Jacques Lechauve pour le développement de la base AVDTH.

Pierre Chavance, Emmanuel Chassot, Patrice Dewals et Alain Damiano pour les informations métiers et scientifiques.

Pierre Lopez pour la réalisation du logo de la page de garde.

1. Par défaut, le fichier devrait être à l’adresse suivante **C:\Obs\_Thonier\AKaDo2\akado.bat** sur Windows et dans **/opt /AKaDo2/akado.sh** sur UNIX/LINUX. [↑](#footnote-ref-2)
2. Le message peut être de trois types : INFO, WARNING et ERROR. [↑](#footnote-ref-3)
3. **LibreOffice** est un logiciel libre et open source; il est téléchargeable ici : <https://www.libreoffice.org/> . [↑](#footnote-ref-4)
4. Noter que nous avons masqué certaines colonnes du fichier de sortie pour une meilleure lisibilité. [↑](#footnote-ref-5)
5. Noter que nous avons masqué certaines colonnes du fichier de sortie pour une meilleure lisibilité. [↑](#footnote-ref-6)
6. Noter que nous avons masqué certaines colonnes du fichier de sortie pour une meilleure lisibilité. [↑](#footnote-ref-7)
7. Noter que nous avons masqué certaines colonnes du fichier de sortie pour une meilleure lisibilité. [↑](#footnote-ref-8)
8. Noter que nous avons masqué certaines colonnes du fichier de sortie pour une meilleure lisibilité. [↑](#footnote-ref-9)
9. Noter que nous avons masqué certaines lignes du fichier de sortie pour une meilleure lisibilité. [↑](#footnote-ref-10)