|  |  |
| --- | --- |
|  | akado-logo-medium.pngAKaDo 2.0 User Guide |
|  |  |
| 04/10/2016 | DATA QUALITY ANALYSIS FOR AVDTH |
|  |  |

AKaDo 2.0 User Guide

Data quality analysis for AVDTH

Summary

[Introduction 2](#_Toc439163577)

[Installation 3](#_Toc439163578)

[Usage 8](#_Toc439163579)

[Set of controls 11](#_Toc439163580)

[Anapo (OT team only) 1](#_Toc439163581)

[Acknowledgements 2](#_Toc439163582)

# Introduction

## Context

The AKADO tool automatically performs a series of tests on the data and produces summary tables that provide more or less detailed balance anomalies. The new version of AKADO, called “AKaDo 2”, is a fusion of different software developed by the *Tropical Tuna Observatory* team, *i.e.* “BaBys”, “AKADO 4.x and many SQL queries. The purpose of this new version is to prepare the transition with the electronic logbook based on the ERS.

AKaDo runs with compatible databases with the latest evolution of AVDTH model data (version 3.5). All predicates evaluated since the 3.3 version of the AKADO are included in this version. We also introduced many new controls. Like the previous version, the importation in T3+ shall not include major errors; it is a guarantee that engages the AVDTH technology.

The major point of the previous version was a translation in English and inclusion of comments in report analysis. We kept these evolutions and improved report analysis. We have codified each error encountered to perform detailed statistics if needed.

This version of AKaDo produces a spreadsheet document easy to read on screen. The spreadsheet must improve the treatment results with more effective than a web browser interface.

## Prerequisites

To be an user of the AVDTH database, “*Acquisition et Validation des Données Thonières”*, created by the *Tropical Tuna Observatory*, *Institute of research for development*.

France, Ghana, Mauritius, Senegal, Seychelles and Spain AVDTH database (October 1, 2013).

## Miscellaneous

This version of AKaDo is developed by Julien Lebranchu. Contact: [OT support team](mailto:support.obs.thonier@listes.ird.fr?subject=[AKADO]%20).

This guide was created the 8 July 2014.

The date of the last modification is 8 February 2017 by Julien LEBRANCHU, Sete IRD, 04 99 57 32 00.

Revision number: 356.

# Installation

## Preamble

To use AKaDo software, you must have already JAVA installed. You can download JAVA at <https://www.java.com/en/>, and follows the instructions to install it.

## Procedure

Run the AKaDo installer by double clicking the “**akado-avdth-installer-XX-standard.jar**” file where XX is the version number; and follow instructions to install or update AKaDo.

1. Press Next to continue.
2. Read the README, and press Next to continue.
3. Accept the AKaDo’s licence, and press Next to continue.
4. Select the packages you want to install (only the documentation, the Licence and Readme files are not mandatory).
5. Select the installation path, you can choose your directory by browsing your computer.
6. This panel presents information about the installation. Press Next to continue and begin the installation.
7. The installation is progressing.
8. The installation is done.

**How to Uninstall**

To uninstall **AKaDo**, you must follow the procedure:

1. You could click on the Uninstaller shortcut if exists in the start menu or you could run the uninstaller which is in the application directory. If you have not changed this directory :
   * On Windows, the uninstaller file should be in C:\Obs\_Thonier\AKaDo2\Uninstaller;
   * On Linux the uninstaller file should be in /opt/AKaDo2/Uninstaller.
2. Only if required, you could delete the configuration files:
   * On Windows, the script file is C:\Obs\_Thonier\AKaDo2\delete-config.bat.
   * On Linux, the script file is/opt/AKaDo2/delete-config.sh

# Usage

**To launch AKaDo, you could run the script which is in the application directory[[1]](#footnote-1).**

When you run AKaDo, the interface is showed (see Figure 1). There are four menus: File, Option, VMS and Help.

* In the File menu, you can load an AVDTH database or quit the application.
* In the Option menu, you can choose your language (need a restart), you can also turn on/off all inspectors.
* In the VMS menu, you can handle the ANAPO functionalities (for more information see the section “Anapo (OT team only)”, p. 1).
* In the Help menu, you can see the information about AKaDo.

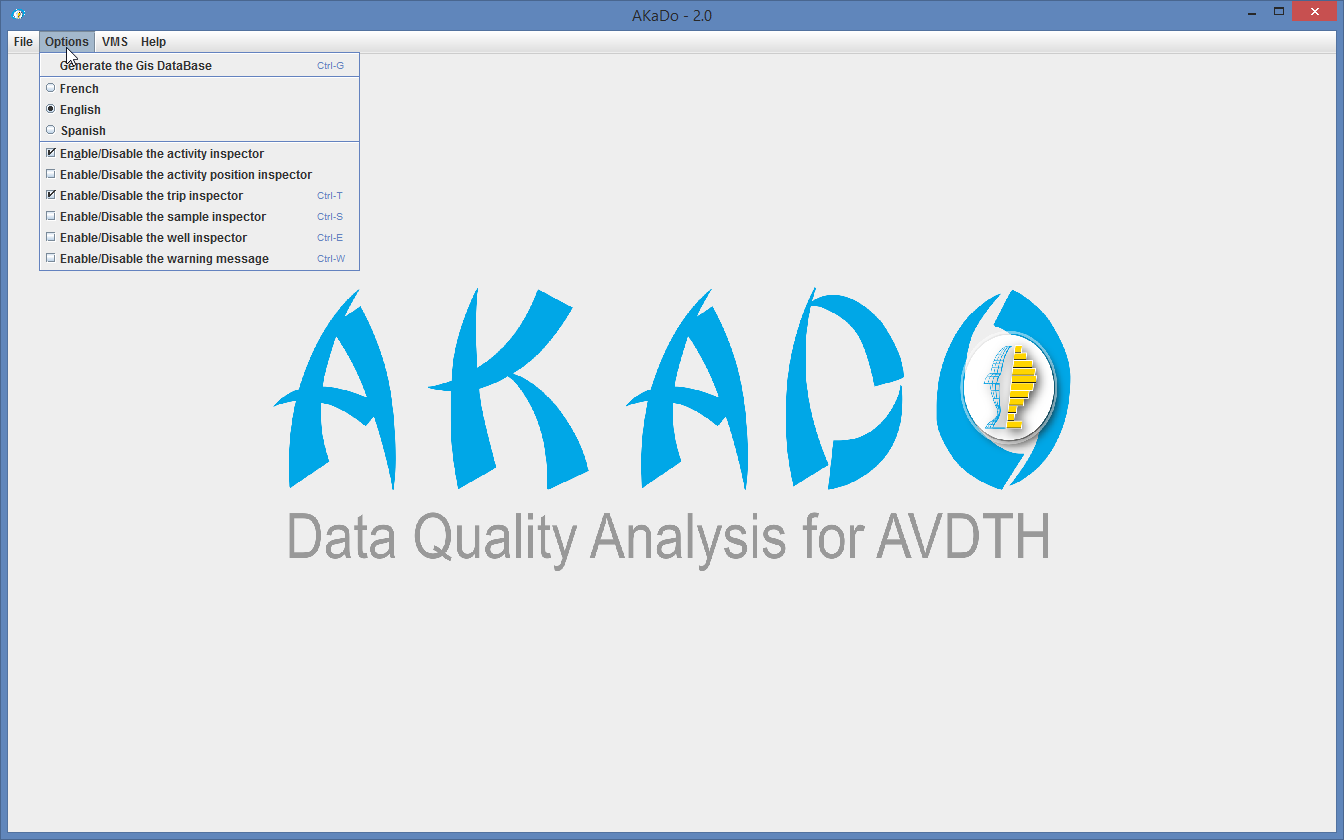


Figure 1

For loading a database, you must browse yours directories and choose a database file (see Figure 2). You can use the shortcut keyboard Ctrl+O or File>Open....

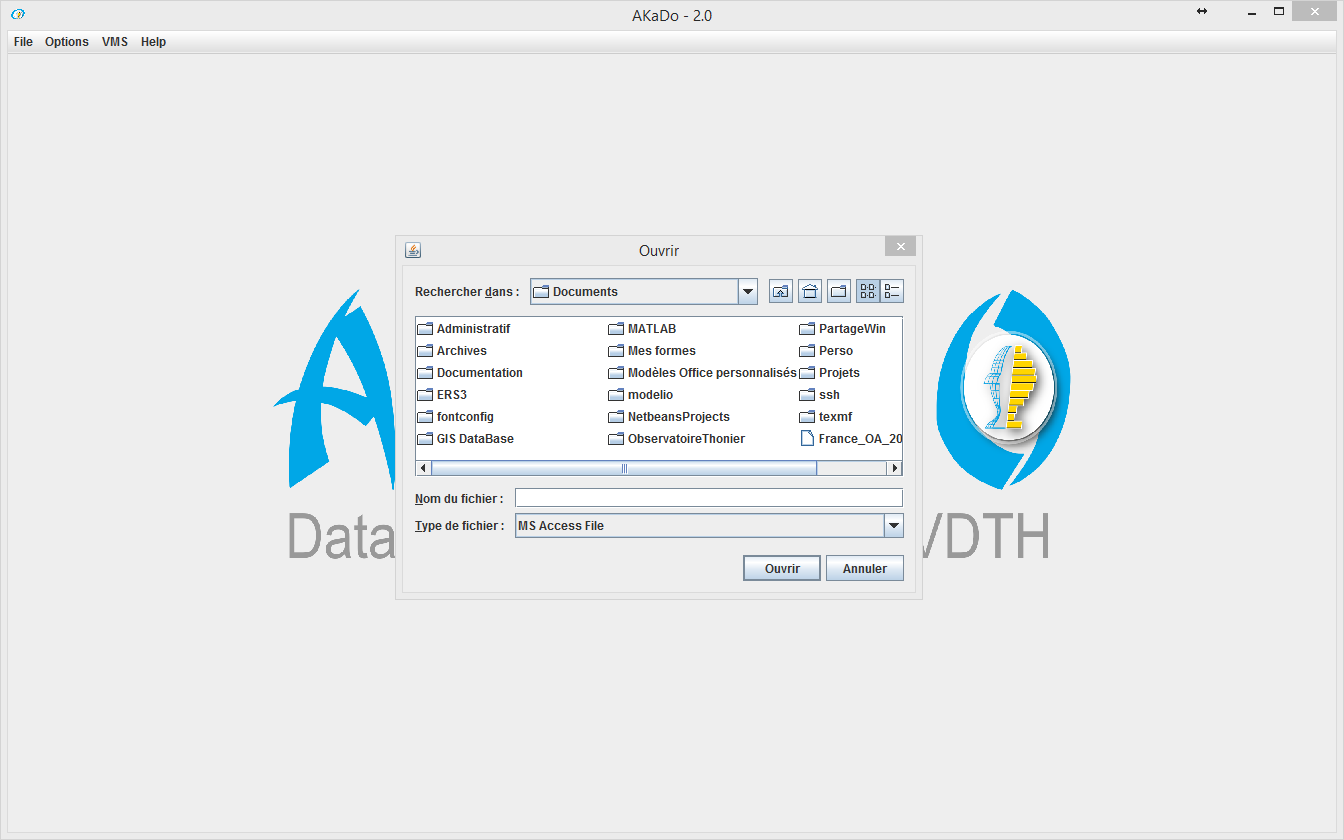
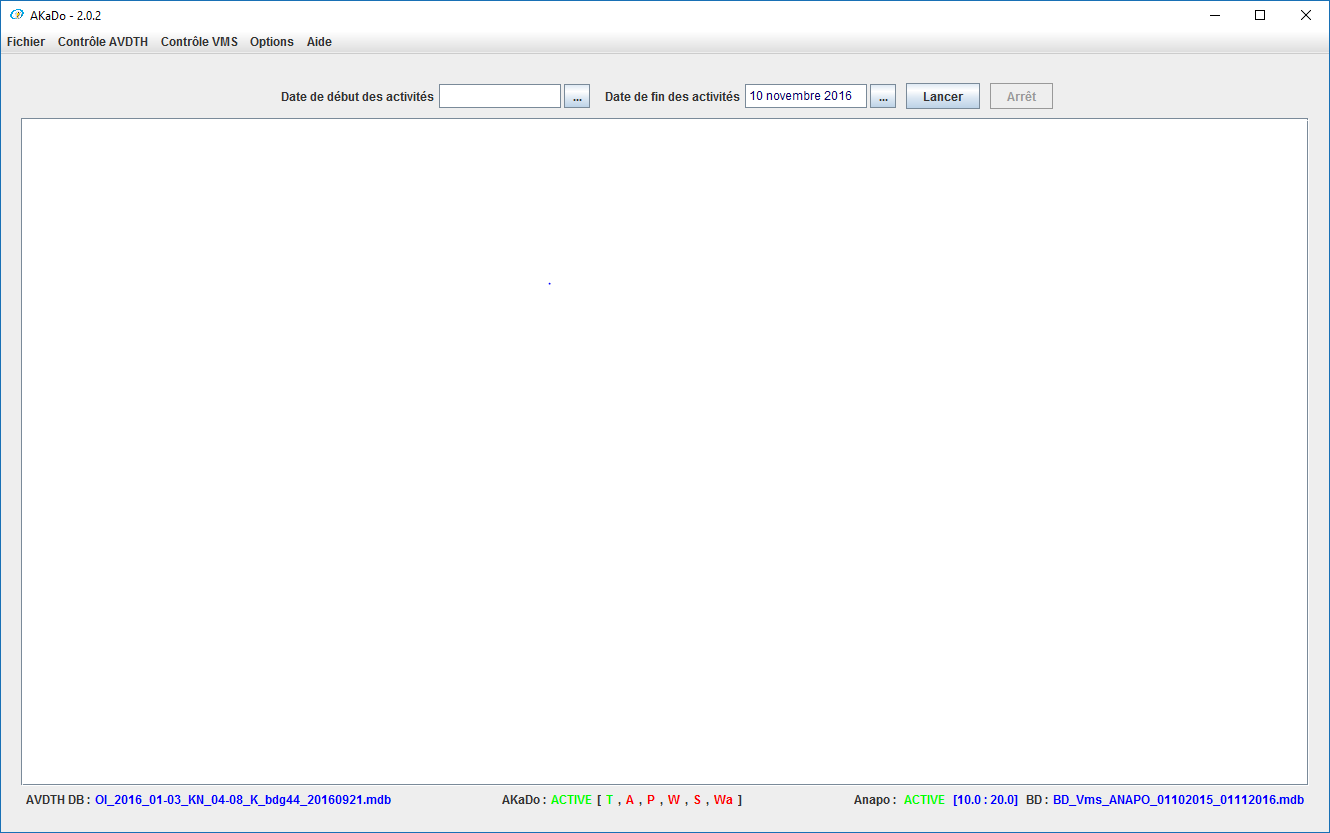


Figure 2

Once you have loaded the database, you can select the temporal range in purpose to limit the activities number that are treated, and then you must press the start button to run the process (see Figure 3). You can stop the process with the stop button (see Figure 3). During the processing, the application prints for each invalid control a message[[2]](#footnote-2). At the end, a result file is generate and the program prints the path on your system. The spreadsheet can be opened with different software like Microsoft Excel or LibreOffice Calc[[3]](#footnote-3).



Select the temporal range of the activities

Figure 3

## Results in the spread sheet

For each group of controls, there is a dedicated file: Trip, Activity... By default, we have realised some conditional formatting. For example, if the catch weight in activity is different to the elementary catches weight then the cell is coloured red (see Figure 4). There is two colours: warnings are in orange and errors are in red.



Figure 4

# Set of controls

**Trip**

### Activity (Warning)

* In this case, we check if the trip has at least one activity and the flag “F\_ENQ” equals 1.
* Is flag F\_ENQ consistent with the official declaration? We check database with application AVDTH -> Landings -> Trip file update -> Logbook?

### Fishing Time

* In this case, we check if the fishing time in trip equals the sum of fishing time in all activities.
* If the values are not equal, you must report the sum value in the “fishing time” in the trip.

### Time at Sea

* In this case, we check if the time at sea in equals the sum of time at sea all activities.
* If the values are not equal, you must report the sum value in the “time at sea” in the trip.

### Landing (Warning)

* We calculate the capacity in tons from the vessel, we check if the capacity is superior of the landing total weight.
* If the landing total weight is superior of the capacity, you must check if the “landing weight” in the trip is correct.

### Landing Total Weight

* We check if the landing total weight is consistent with the elementary landing weight.
* If the values are not equal, you must report the sum value in the “landing weight” in the trip.

### Loch (Warning)

* We check if the value is between 0 and the maximum distance which is calculated from the maximum speed and the number of activities.

### Recovery Time

* We check if the activities are continuous during a trip.
* You must examine if a day is missing in logbook.

### Temporal Limit

* We check if the temporal limits, the first and last day, of the trip are consistent with activities.
* You must check Departure and Arrival date between logbook and landing documents.

### Total Catch Weight

* We check if the total catch weight from activities and elementary catches are equals.
* If the values are not equal, you must check the catch weight in the “Activity” sheet of the result file.

### Harbour

* We check if the departure harbour of the trip and the return harbour of the previous trip are consistent.
* Check if all logbooks have been entered.
* Check with the master if trips have not been in the meantime.

### Raising Factor (Info)

* We calculate the raising factor for all trips (included trip with partial landings).
* If the ratio is not between 0.9 < R < 1.1 (Landing/Catches), you must inspect the flag partial landing.

#### Exploitation OF errOrs



The figure above highlights several from errors of various controls[[4]](#footnote-4):

* In the column 'Departure Harbour', the harbours highlighted correspond to a different harbour between the landing harbour of a previous trip and the departure habour of this trip.
* In the column "Has logbook", it is important to check the consistency of the value 0 (false) and 1 (true) with the information from the dates of activities.
  + If the value is 1 then the date of the first activity must be populated.
  + And in return, if the value is 0, then the date of the first activity should not be populated.
* In the "First Activity Date" column, the dates highlighted indicate a difference between the departure date of the trip ('Departure Date' column) and the date of the first activity.
* In the "Last Activity Date" column, the dates highlighted indicate a difference between the landing date of the trip ('Landing Date' column) and the date of the last activity.



The figure above shows several errors of various controls[[5]](#footnote-5):

* In the column "Time at Sea", the highlighted value indicates a difference with the sum of the values 'Time at sea' for all activities of the trip.
* In the column 'Fishing Time', the highlighted value indicates a difference with the sum of the values 'Fishing time' for all activities of the trip.
* In the column "Landing Weight", the highlighted value indicates a difference with the sum of the landed weight of commercial of the trip.

**Activity**

### Fishing Context

* We check if the school type and the fishing context are consistent.
* If the school type is an artificial school (code 1), then it must at least one fishing context;
* If the school type is an artificial school (code 1), then the fishing context code must be equal to 10, 60, 81 or between 20 and 28;
* If the school type is a free school (code 2), then the fishing context code must ***not*** be equal to 10, 60, 81 or between 20 and 28.

### Operation

* We check if the operation associated with activity is consistent with other information.
* If the operation code is 0, then the value of the total catch weight must be equal to 0;
* If the operation code is 1 or 2, then the value of the total catch weight must not be equal to 0;
* If the value of the total catch weight isn’t equal to 0, then the operation code is 1 or 2;
* If the operation code is 12, 13 or 14, then the value of the total catch weight must be equal to 0.

### Position

* We check if the position activity is in ocean or inland, and if the position activity and ocean are consistent.
* If the position is in land, you must inspect the latitude, longitude and quadrant fields.
* If the position activity and the ocean field are not equal, you must examine these fields and correct it.

### Quadrant

* We check if the quadrant and the position activity are consistency, i.e.
* If the quadrant value is 3 or 4, the ocean value must be “Atlantic Ocean”.
* If the quadrant value is 3 or 4, the position must be located in “Atlantic Ocean”.
* You must verify the ocean field, the quadrant field and the position field.
* We check if the position and/or the latitude with the zero value and the quadrant are correct.

|  |  |  |  |
| --- | --- | --- | --- |
| LATITUDE | LONGITUDE |  | QUADRANT |
| 0 (NORTH by convention) | **WEST** |  | **4** |
| 0 (NORTH by convention) | **EAST** |  | **1** |
| NORTH | **0 (East** by convention**)** |  | **1** |
| SOUTH | **0 (East** by convention**)** |  | **2** |

### Weight

* We check if the total catch weight is consistent with elementary catches.
* If the values are not equal, you must report the sum value of elementary catches weight in the activity.

#### Exploitation OF errOrs

****The figure above shows several errors of various controls[[6]](#footnote-6):

* In the column “Operation”, the highlighted value indicates that you must controlled the operation and the catch weight.
* In the column "Catch weight", the highlighted value indicates a difference with the sum of the weights of the elementary catch of the activity.
* In the columns "School type" and "Fishing context", the highlighted value indicates an inconsistency between the columns, for example the association.
* In the ' Ocean AO/IO ' column, the value is calculated from the positions activity, and there is a highlighting of the associated value in the column «Ocean» when they differ.
* In the column "Latitude", the highlighted value indicates that the position and the quadrant are incorrect.
* In the column "Longitude", the highlighted value indicates that the position and the quadrant are incorrect.
* In the column "In Land", the highlighted value indicates that the position is an activity in land.
* In the column 'Temperature', the highlighted value indicates that the temperature is not between 15 and 32.

**Sample**

### Activity

* We check if the activity information for each sample well is consistent.
* You must examine the following information: the date, the activity number, the quadrant, the latitude, the longitude and the school type

### Length class

* We check if the length class is consistent with each length class of species (L=80cm for YFT and BET and L=42cm for ALB).

### Species

* We check if the species sampled is authorised.

### Measure

* We check if the sample species number is consistent with the measure number.

### Position

* We check if the position activity for each sample well is consistent.

### Sample without Measure

* We check if the sample has at least one measure.

### Sample without Species

* We check if the sample has at least one species.

### Sample without Trip

* We check if the sample is linked at one trip.

### Super Sample

* We check if the sub-sample number is consistent.

### Well

* We check if the sample is linked at one trip.

### Ratio of little and big fish

* We check if the percentage of little and big fish sampled is consistent.

### Weighting

* We examine the weighting information for each sample well is consistent.
* If the vessel engine is a purse seine, the sample weight must be inferior to 100kg.
* If the vessel engine is a purse seine, the weighted weight must be inferior to the weight and the ration between weighted weight and the weight must be superior or equal to 0,85.
* If the vessel engine is a bait boat,

#### Exploitation OF errOrs

****

The figure above shows several errors of various controls[[7]](#footnote-7):

* In the column "Sampling overall weight", the highlighted value means that the column should be 0 if the values for the columns "Sampling weight M10" and "Sampling weight P10" are different from 0, or vice versa.
* In the column 'Trip exist', the highlighted value means that the sample is not bound to a trip.
* In the "Activity" column, the highlighted value means that the sample is not associated with a well.
* In the "Activity" and "Activity position", the highlighted columns mean that the associated activity does not exist. Moreover, we specify if possible the field posing problem using the marker "?" with a blue color.



The figure above shows several errors of various controls[[8]](#footnote-8):

* In the column 'Species', the highlighted values can mean the following errors:
* If the case code is preceded by a "?", this means that the species should not be sampled;
* If the case code is followed by '? LD1', this means that there is an error on the classes of sizes.

 The maximum size for YFT class (1) and BET (3) is 55cm.

 The maximum size for AB class (4) is 42cm.

* In the "LDLF" column, the highlighted values can mean the following errors:
  + If the LDLF code is framed of "!", it means that the species code is incompatible with this value of LDLF.
  + If the LDLF code is preceded of '? ', this means that the values filled in columns "Sampling weight M10", "Sampling weight P10" and "Sampling overall weight" are inconsistent with this value of LDLF.

 If LDLF equal 1 or 3 while PM10 or overall weight must be greater than 0 or "warning."

 If equal LDLF 2 then M10 or overall weight must be greater than 0 or "warning."

* In the column "Measured count", the highlighted value means the number of measure is different from the number of measured individuals.
* In the column "Has well", the highlighted value means that it has no tank associated with the sample.
* In the column ' Distribution + 10 /-10 ', the highlighted value means that the distribution values - 10 / + 10 are inconsistent with that of the postponed well.

**Well**

### Activity

* We check if the activity information for each well plan is consistent. We examine the date and the number.
* You must examine in the well plan the date and the number of the activity.
* You must check if the logbook is in AVDTH.

### Well without Trip

* We check if the well is associated at one trip existing.
* You must enter the logbook in AVDTH.

### Well without Well Plan

* We check if the well is associated, at least one well plans existing.

#### Exploitation OF errOrs

#### 

The figure above shows several errors of various controls[[9]](#footnote-9):

* In the column 'Trip exist', the highlighted element indicates that the well is not linked to a trip.
* In the column 'Well Plan exist', the highlighted element indicates if the well is missing.
* In the column "Activity exist", the highlighted element indicates that the activity associated with the well (regarding the date and the number of activity) dont exist.

# 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).

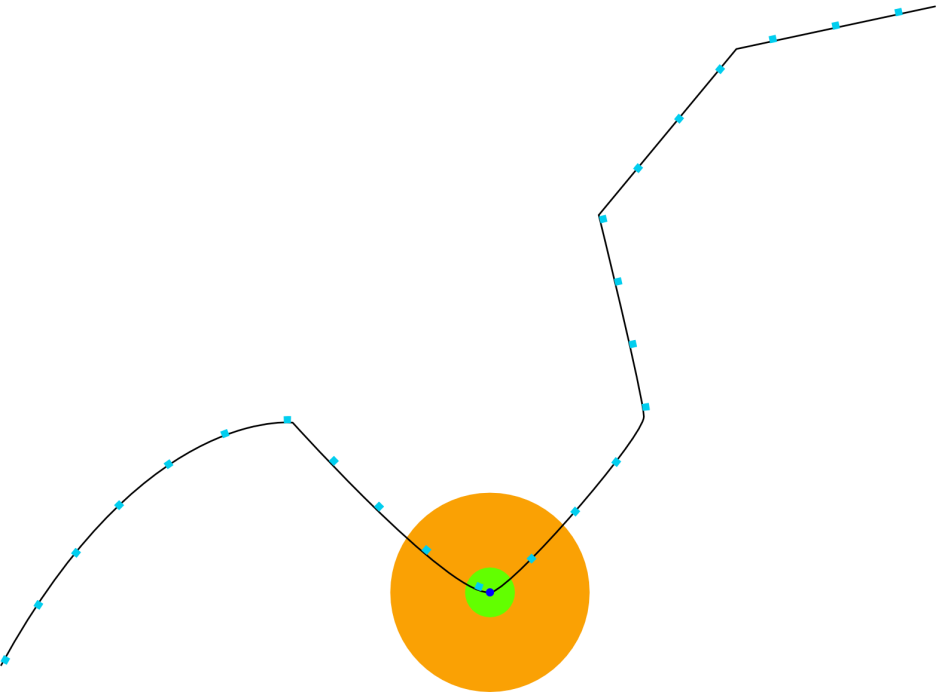


Figure 5 CASE where the POSITION is valid

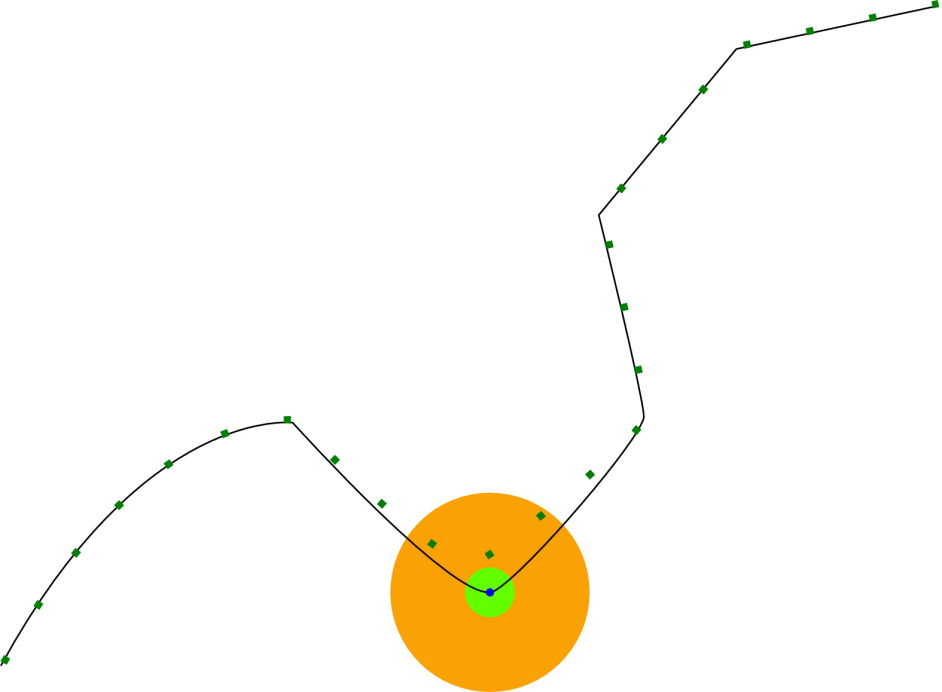


Figure 6 Case wehre the position is in the limit zone

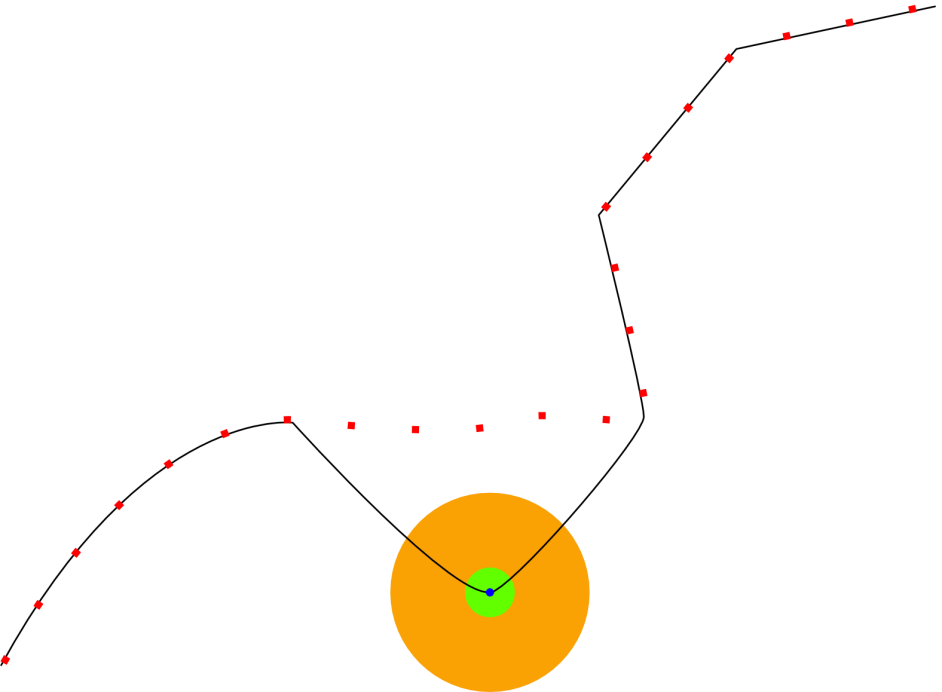


Figure 7 case where the position is invalid

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)

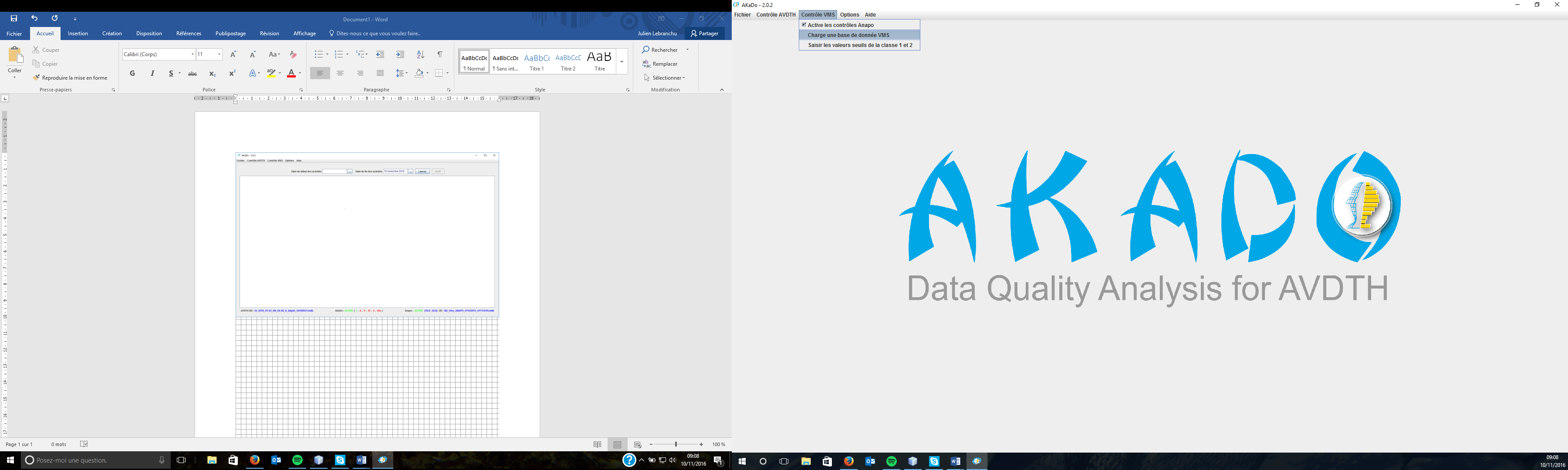


Figure 8

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

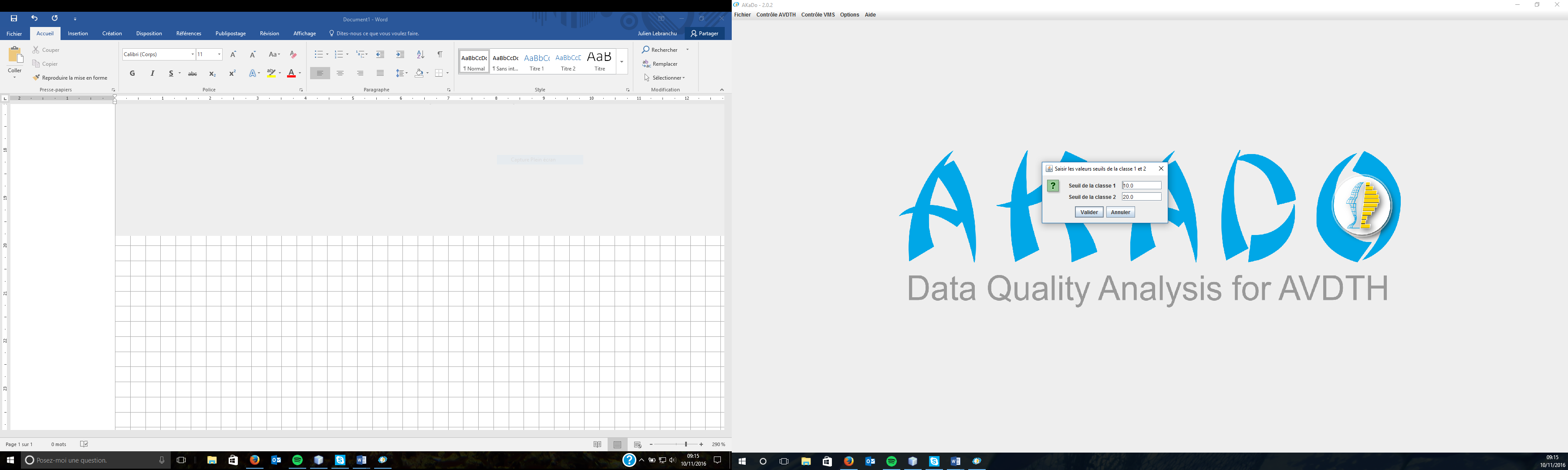


Figure 9

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.

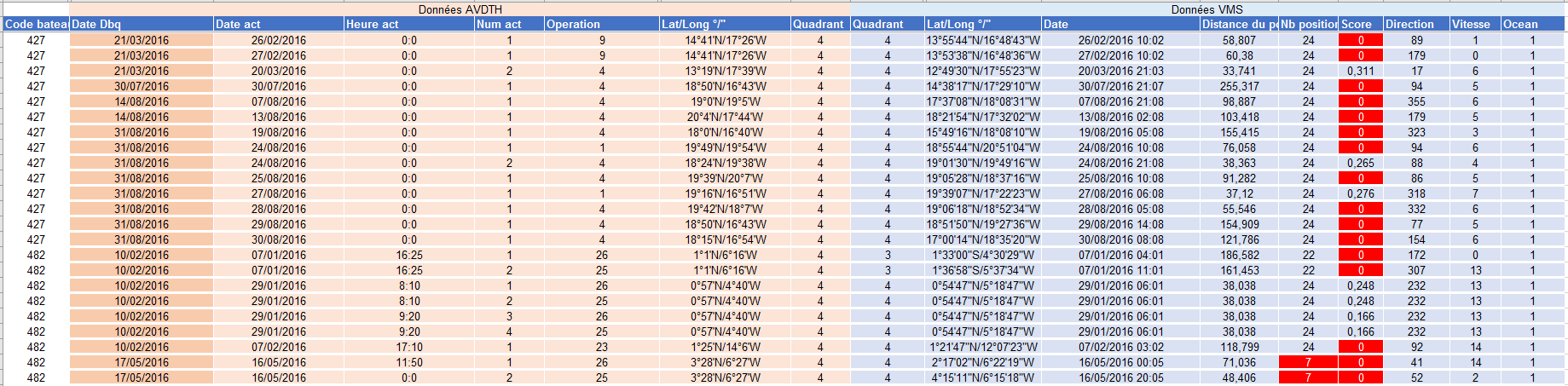


Figure 10

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.

# Acknowledgements

Jean-Jacques Lechauve and Laurent Floc’h for the development of the previous version.

Pascal Cauquil, Jean-Jacques Lechauve for the development of the AVDTH database.

Pierre Chavance, Emmanuel Chassot, Patrice Dewals and Alain Damiano for their scientific and fisheries information.

Pierre Lopez for the design of the logo on the front page.

1. If you have not changed this directory, the file should be in **C:\Obs\_Thonier\AKaDo2\akado.bat** on Windows and in user home **/opt /AKaDo2//akado.sh** on UNIX/LINUX. [↑](#footnote-ref-1)
2. The message can be of three types: INFO, WARNING and ERROR. [↑](#footnote-ref-2)
3. **LibreOffice** is free and open source software; you can download it at <https://www.libreoffice.org/> . [↑](#footnote-ref-3)
4. Note that we have hidden some columns of the output file for readability. [↑](#footnote-ref-4)
5. Note that we have hidden some columns of the output file for readability. [↑](#footnote-ref-5)
6. Note that we have hidden some columns of the output file for readability. [↑](#footnote-ref-6)
7. Note that we have hidden some columns of the output file for readability. [↑](#footnote-ref-7)
8. Note that we have hidden some columns of the output file for readability. [↑](#footnote-ref-8)
9. Note that we have hidden some columns of the output file for readability. [↑](#footnote-ref-9)