## MOMIT
### MULTI-SCALE OBSERVATION AND MONITORING OF RAILWAY INFRASTRUCTURE THREATS

**Grant Agreement No: 777630**
**Project Acronym: MOMIT**

**Project Title:**

**Multi-scale Observation and Monitoring of railway Infrastructure Threats**

# Technical note:  ADAClassifier - Classify ADAs automatically

**This Document is composed by 33 pages, including attachments**

**Last revision date: 2018-07-26**

**Dissemination Level: CO = Confidential**

**MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats**
*Project funded by S2R JU under the European Union's H2020 R&I programme*

**GA No: 777630**

# Deliverable information

## Document Configuration

| | |
|---|---|
| **Programme Name:** | Shift2Rail |
| **Call:** | H2020-S2RJU-2017 |
| **Topic:** | S2R-OC-IP3-03-2017 |
| **Grant Agreement Number:** | 777630 |
| **Project Acronym:** | MOMIT |
| **Project Title:** | Multi-scale Observation and Monitoring of railway Infrastructure Threats |
| **Template Version:** | v1.2 |
| **Deliverable Number:** | |
| **Title of the Deliverable:** | Technical note: ADAClassifier - Classify ADAs automatically |
| **Work-Package:** | WP3 |
| **Due Date:** | 2108-08-31 (M12) |
| **Issue Date:** | 2018-07-26 |
| **Revision of the document:** | v1.0 |
| **Dissemination Level:** | CO = Confidential |
| **Responsible Beneficiary:** | CTTC |
| **Prepared by:** | José Navarro (CTTC) |
| **Reviewed by:** | María Cuevas, Michele Crosetto (CTTC) |
| **Contributors:** | |

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
GA No: 777630
*Project funded by S2R JU under the European Union's H2020 R&I programme*

# Revision History

| Revision | Date | Author | Organisation | Description |
|---|---|---|---|---|
| v1.0 | 2018-07-26 | José Navarro | CTTC | First draft issue |
| | | | | |
| | | | | |
| | | | | |

# Acknowledgement

# Statement of originality

*This document contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.*

# Disclaimer

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

# Table of Contents

**MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats**
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

# List of Tables

**MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats**
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

# List of Figures

*Dissemination Level: CO* = Confidential

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

# Acronyms and Abbreviations

| | |
|---|---|
| **ADA** | Active Deformation Area |
| **DTM** | Digital Terrain Model |
| **GUI** | Graphical User Interface |
| **PS** | Persistent Scatterer |
| **PSI** | Persistent Scatterer Interferometry |

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
GA No: 777630
*Project funded by S2R JU under the European Union's H2020 R&I programme*

# Technical note: ADAClassifier - Classify ADAs automatically

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
GA No: 777630
*Project funded by S2R JU under the European Union's H2020 R&I programme*

# 1 Executive Summary

This technical note presents ADAClassifier, a tool to classify the ADAs (Active Deformation Areas) identified by ADAFinder. Several categories are checked in this classification process, as landslides, sinkholes, land subsidences and constructive or consolidation settlements. The basics of the algorithm used are briefly described, as well as the kind and characteristics of the input, output and configuration files. The Graphical User Interface (GUI), although very concisely, is also explained. The batch version of the tool is also briefly described.

*Dissemination Level: CO* = Confidential

# 2 The ADAClassifier tools

The ADAClassifier (family of) tool(s) is targeted at the classification of Active Deformation Areas (ADA), trying to determine whether these fall in one or more categories, namely landslides, sinkholes, land subsidences and constructive or consolidation settlements.

The application has been developed in C++ for efficiency reasons. Both a command line and a Graphical User Interface (GUI) are offered to the user. The GUI simplifies noticeably the operational procedure; the command line version may be used to integrate this tool in a batch production line, if necessary.

These tools have been designed to accept the output of ADAFinder (see [RD1]), which is a standard Esri shapefile. Extra inputs, essentially different kind of inventories, must be also standard (polygon) Esri shapefiles. The output of the process is also stored using this format. These files are easily managed by a wide number of Geographic Information System (GIS) tools. Among these, ArcGIS and QGIS. For a description of the shapefile format, please see [RD3].

Additionally, Digital Terrain Models (DTM) are used to help the process of classification. In this case, a slight variation in the format in which these files are stored is accepted to make the application a bit more flexible. Nonetheless, the available options in the case of DTMs are limited.

On output, ADAClassifier creates a single shapefile containing the results of the classification.

The following sections describe ADAClassifier in more detail. This encompasses the GUI, the formats of the configuration, input and output files as well as guidelines about how to use the two versions of this tool.

## 2.1 The process

ADAClassifier attempts to classify the ADAs identified by ADAFinder, checking if these may be tagged as belonging to one or more categories. These categories are: landslides, sinkholes, land subsidences and constructive or consolidation processes.

All ADAs are checked against all the former categories. Therefore, the same ADA may be tagged as belonging to one or more of these.

Table 1 below summarizes how these checks are performed. Note the red codes (as for instance **Th1** or **Th10(1)**). These are a concise way to refer to several threshold values defined in Table 3 (page 17), where their precise meaning and allowed range of values are defined. Refer to such table for details about the thresholds in the procedure.

| ID type 1 (Landslide) | | | | | |
|---|---|---|---|---|---|
| Is the percentage of intersection of the ADA with the landslide inventory > **Th1**? | Yes | **Landslide** | | | |
| | No | Slope > **Th2** | Horizontal displacement > **Th3**? | Yes | **Potential new landslide** |
| | | | | No | **Not a landslide** |
| | | Slope < **Th2**? | **Not a landslide** | | |

| ID type 2 (Sinkhole) | | | | | |
|---|---|---|---|---|---|
| Is the percentage of intersection of the ADA with the sinkhole inventory > **Th4**? | Yes | **Sinkhole** | | | |
| | No | Saline or carbonate rock/soil > **Th10(1).** | Horizontal displacement > **Th5**? | Yes | **Not a sinkhole** |
| | | | | No | **Potential new sinkhole** |
| | | Other lithologies | **Not a sinkhole** | | |

| ID type 3 (Land subsidence) | | | | | |
|---|---|---|---|---|---|
| Is the percentage of intersection of the ADA with the subsidence inventory > **Th6**? | Yes | **Land subsidence** | | | |
| | No | Quaternary unconsolidated sediments > **Th10(2).** | Yes | Slope > **Th7** | **Not land subsidence** |
| | | | | Slope < **Th7** | **Possible land subsidence** |
| | | | No | **Not land subsidence** | |

| ID type 4 (Constructive settlements – consolidation settlements) | | | | | |
|---|---|---|---|---|---|
| Is the percentage of intersection of the ADA with the infrastructures (building, embankment, bridge, etc.) inventory > **Th8?** | No | **Not constructive settlements** | | | |
| | Yes | The Pearson coefficient measuring the fit of the time series and an inverse exponential function is > **Th9**. | Yes | **Consolidation process.** | |
| | | | No | **Not consolidation process** | |

*Table 1: Procedure to classify the ADAs.*

In Table 1 above, slopes and horizontal displacements are mentioned. These are obtained, respectively, from a DTM and the output of the LOS2HV tool (see [RD2]) which computes the horizontal and vertical components of the ground movement along the Line Of Sight (LOS) of the satellite. These components are stored also in files, which must be read by the application.

There is a noticeable amount of files involved in the classification of the ADAs. Besides the input ADAs themselves (polygons and points) several inventories, a DTM and a file with horizontal components of the ground movement are needed. The normal situation will be that not all these files will be available when processing a specific area. This is why *some of the input files are optional*. See Table 4 for details on mandatory / optional / conditional files.

Consequently, when processing incomplete data sets, not all the classification subprocesses may take place. For instance, the landslide check may not be run in the absence of a landslides inventory file.

ADAClassifier is able to discard some of the classification processes according to data availability, so having only a subset of the required files is not an obstacle for this tool to run. Table 2 below depicts the dependencies between the several classification subprocesses and the data needed to perform these. Or, if the table is read by rows, it describes what are the classification subprocesses

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
GA No: 777630
*Project funded by S2R JU under the European Union's H2020 R&I programme*

that may or may not take place when a file is available or not.

| | Landslides | Sinkholes | Subsidences | Settlements |
|---|---|---|---|---|
| DTM | √ | | √ | |
| Horizontal components | √ | √ | | |
| Landslides inventory | √ | | | |
| Sinkholes inventory | | √ | | |
| Subsidences inventory | | | √ | |
| Geologic inventory | | √ | √ | |
| Infrastructures inventory | | | | √ |

*Table 2: Data versus available classification subprocesses.*

## 2.2 The GUI

Figure 1 and Figure 2 depict the graphical interface of the ADAClassifier (GUI version) tool. Note that Figure 1 depicts the "Options" tab showing the different thresholds involved in the classification process, while Figure 2 shows the "Files" tab, where the (full path and) names of the files available are typed by the user. Such separation in two figures helps to understand the interface.

Note that in Figure 1 some codes written in **red** have been used to refer to the available fields (such as **Th1** or **Th9**). These are the threshold codes used in Table 1, codes that correspond to the parameters included in the application's options file, which are described in section 2.3 and Table 3. Refer thus to these section and table for details.

Other codes written in **red** are shown in Figure 2. Those starting with the letters FI (like **FI03**) refer to the several files involved in the computation. Please refer to section 2.3 and Table 4 for a full description of these fields and the corresponding files. Additionally, there are two extra types of codes in this figure, those starting with the letter W and with the letters "BU":

- W01 – This section of the interface window contains a log area, where the different messages

*Dissemination Level: CO =* Confidential

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*
GA No: 777630

issued by the application are shown.

- W02 – Dialog window appearing only when the process is finished. It reports the status of the process (either successful or unsuccessful).
- BU01 – Button to quit the application.
- BU02 – Button to load an existing ADAClassifier options file. Use this button to read an options file (see section 2.3 for a complete description of these) and set its values on the screen instead of typing all the values in the different fields shown by the GUI.
- BU03 – Button to save an ADAClassifier options file. Once that all the fields have been filled by the user (both Options and Files tabs) it is possible to save their values in an options file for later use. Use this button to do it.
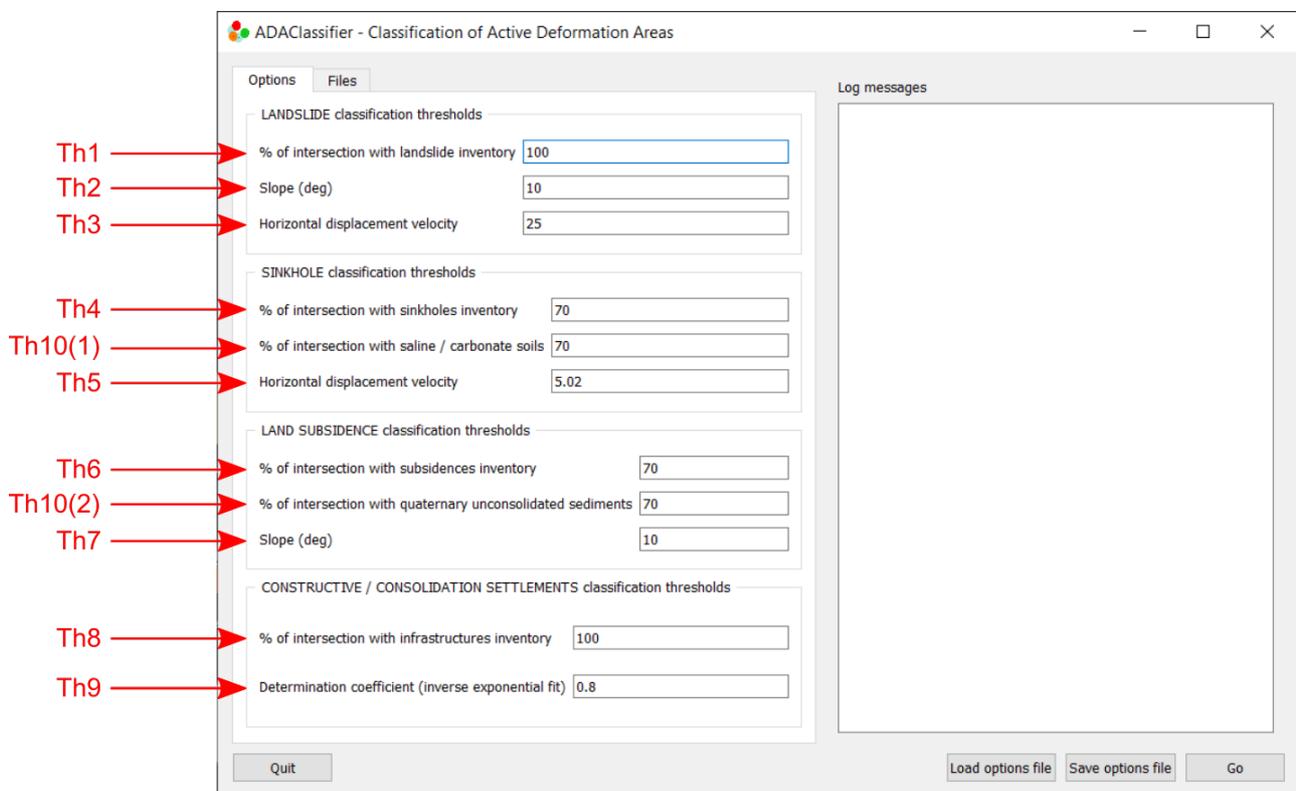- BU04 – Use this button to start the actual classification process.



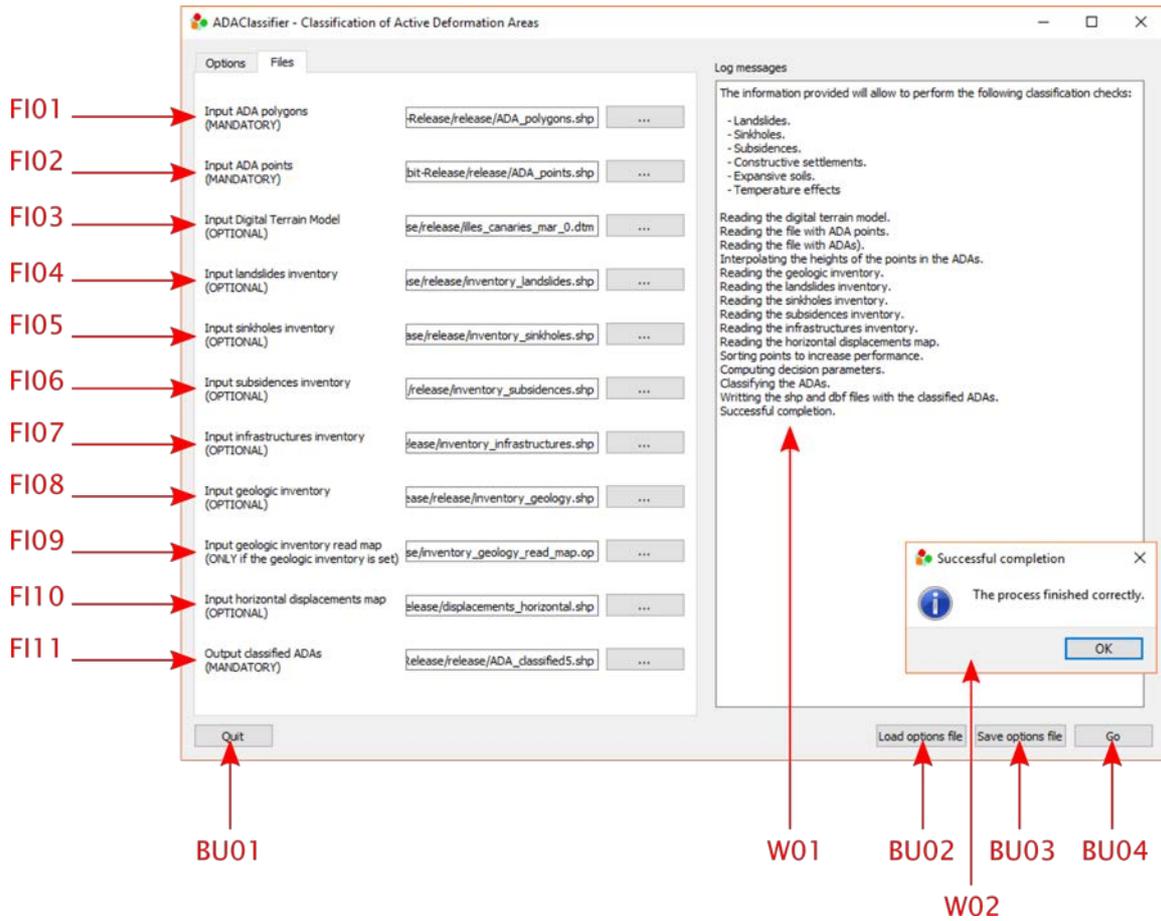*Figure 1: ADAClassifier GUI: the thresholds in the Options tab.*

*Figure 2: ADAClassifier GUI: the files (Files tab) and other controls.*

## 2.3   The defaults (options) file

When using the GUI version of the tool, a defaults file residing in the same directory than the application's executable is loaded to provide with some initial, reasonable values for the thresholds in the Options tab. This is not the case of the command-line tool, which must always be fed with a full path and name to an ADAClassifier options file as an input parameter. This defaults / options file is a plain-text one defining the values of all the options controlling the behaviour of the application.

When the GUI version of ADAClassifier is started, the defaults file is read and only some of its values – the thresholds in the Options tab, not the file names -  are shown on the screen[1].

The rationale behind this decision is that while working on a project, the set of parameters controlling the application will be more or less stable – so these are not changed frequently – while the files to process may change more often – so their names must always be provided by the user.

At any rate, all the parameters, regardless whether they are file names or values, may be modified by the user using the GUI.

The command-line version of ADAClassifier, on the contrary, loads the whole set of values in the

---

[1] Changing the contents of the default options file, located in the same directory than the application's executable, the default values shown in the GUI will change accordingly.

options file, ignoring none at all.

---

**Remember that. when using the GUI version of the tool, the default options file must reside in the same directory than the tool's executable. The name of such defaults file must be exactly `ADAClassifier_default_configuration.op`.**

---

The details concerning the syntax of the ADAClassifier defaults file is described in [RD4]. The available options that must be included in such file may be found in section 2.3.1.

### 2.3.1    The available options

The options that must be present in an ADAClassifier options file are summarized in Table 3 and Table 4, containing, respectively, the description of regular parameters (the thresholds) and the files intervening in the process. This separation into two tables just pursues simplicity.

Note that ADAClassifier expects to find **all** the options listed in the aforementioned tables. Missing parameters or incorrect values of these will make the application complain about the error. Even in the case of **optional** files (Table 4) a value must be provided**. In these cases, the value of the file name must be a single dash (-).**

The parameters in Table 3 have been grouped according to the classification subprocess these belong to (see section 2.1). Additionally, the **threshold codes** used in Table 1 and Figure 1 are included in Table 3 too written in red, so it is possible to match the thresholds in the GUI, the names of these parameters in the options file and their actual meaning.

| Label | Description & comments |
|---|---|
| `THRESHOLD_INTERSECTION_LANDSLIDE` | Minimum percentage of intersection to consider that an ADA overlaps with the **landslides** inventory. The percentage is compared against the percentage of intersection between the polygon in the inventory that overlaps most with the points in the ADA. |
| | Floating point value. Range: [0..100]. |
| | Classification subprocess: landslides. Table 1 threshold code: **Th1**. |
| `THRESHOLD_SLOPE_LANDSLIDE` | Minimum slope (in degrees) of the plane adjusted to the points in the ADA to keep considering such ADA as a landslide candidate. |
| | Floating point value. Range: (0..90). |
| | Classification subprocess: landslides. Table 1 threshold code: **Th2**. |
| `THRESHOLD_DISP_HORIZONTAL_LANDSLIDE` | Minimum value of the horizontal component of the displacement of the ADA to keep considering such ADA as a landslide candidate. |
| | Floating point value. Range: [0..∞]. |

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

| Label | Description & comments |
|---|---|
| | Classification subprocess: landslides. Table 1 threshold code: **Th3**. |
| THRESHOLD_INTERSECTION_SINKHOLE | Minimum percentage of intersection to consider that an ADA overlaps with the **sinkholes** inventory. The percentage is compared against the percentage of intersection between the polygon in the inventory that overlaps most with the points in the ADA. |
| | Floating point value. Range: [0..100]. |
| | Classification subprocess: sinkholes. Table 1 threshold code: **Th4**. |
| THRESHOLD_DISP_HORIZONTAL_SINKHOLE | Minimum value of the horizontal component of the displacement of the ADA to keep considering such ADA as a sinkhole candidate. |
| | Floating point value. Range: [0..∞]. |
| | Classification subprocess: sinkholes. Table 1 threshold code: **Th5**. |
| THRESHOLD_INTERSECTION_SUBSIDENCE | Minimum percentage of intersection to consider that an ADA overlaps with the **land subsidences** inventory. The percentage is compared against the percentage of intersection between the polygon in the inventory that overlaps most with the points in the ADA. |
| | Floating point value. Range: [0..100]. |
| | Classification subprocess: subsidences. Table 1 threshold code: **Th6**. |
| THRESHOLD_SLOPE_SUBSIDENCE | Minimum slope (in degrees) of the plane adjusted to the points in the ADA to keep considering such ADA as a subsidence candidate. |
| | Floating point value. Range: (0..90). |
| | Classification subprocess: subsidences. Table 1 threshold code: **Th7**. |
| THRESHOLD_INTERSECTION_IFS_SETTLEMENT | Minimum percentage of intersection to consider that an ADA overlaps with the **infrasttructures** inventory. The percentage is compared against the percentage of intersection between the polygon in the inventory that overlaps most with the points in the ADA. |
| | Floating point value. Range: [0..100]. |
| | Classification subprocess: constructive / consolidation settlements. Table 1 threshold code: **Th8**. |
| THRESHOLD_EXPONENTIAL_FIT_SETTLEMENT | Minimum Pearson coefficient measuring the quality of the match between the deformation time series for the points in the ADA and an inverse exponential function to accept such kind of fit. |
| | Floating point value. Range: [-1 .. 1]. |
| | Classification subprocess: constructive / consolidation settlements. Table 1 threshold code: **Th9**. |

| Label | Description & comments |
|---|---|
| THRESHOLD_INTERSECTION_GEOLOGY | Minimum percentage of intersection to consider that an ADA overlaps with some kind of elements in the **geologic** inventory. The percentage is compared against the percentage of intersection between the polygon in the inventory that overlaps most with the points in the ADA. |
| | This parameter is, in fact, a **list** of two values. The first one measures the intersection with saline or carbonate soils: the second is used to check the intersection with quaternary unconsolidated sediments. |
| | Floating point values. Range: [0..100]. |
| | First value: Classification subprocess: sinkholes. Table 1 threshold code: **Th10(1)**.<br><br>Second value: Classification subprocess: subsidences. Table 1 threshold code: **Th10(2)**. |

*Table 3: The options controlling the behavior of the application: thresholds.*

Table 4 below describe the options in the options file that refer to file names. Note that, once more, a reference tag, written in **red**, has been included to help to identify the elements in Figure 2 and therefore complete their description.

| Label | Description & comments |
|---|---|
| FILE_ADA_POLYGONS | Type 5 (polygons) Esri shapefile containing the polygons defining the boundaries of the ADAs identified by ADAFinder. |
| | MANDATORY INPUT. |
| | File tag: **FI01**. |
| FILE_ADA_POINTS | Type 1 (points) Esri shapefile containing the points included in the ADAs identified by ADAFinder. |
| | MANDATORY INPUT. |
| | File tag: **FI02**. |
| FILE_DTM | Digital terrain model (header). See section 2.4.2 for details on the format supported. |
| | OPTIONAL INPUT. See Table 2 for information about what classification processes need this file. |
| | File tag: **FI03**. |
| FILE_INVENTORY_LANDSLIDES | Type 5 (polygons) Esri shapefile containing the inventory of landslides. |
| | OPTIONAL INPUT. See Table 2 for information about what classification processes need this file. |
| | File tag: **FI04**. |
| FILE_INVENTORY_SINKHOLES | Type 5 (polygons) Esri shapefile containing the inventory of sinkholes. |
| | OPTIONAL INPUT. See Table 2 for information about what classification processes need this file. |
| | File tag: **FI05**. |

| Label | Description & comments |
|---|---|
| `FILE_INVENTORY_SUBSIDENCES` | Type 5 (polygons) Esri shapefile containing the inventory of subsidences. |
|  | OPTIONAL INPUT. See Table 2 for information about what classification processes need this file. |
|  | File tag: **FI06**. |
| `FILE_INVENTORY_INFRASTRUCTURES` | Type 5 (polygons) Esri shapefile containing the inventory of infrastructures. |
|  | OPTIONAL INPUT. See Table 2 for information about what classification processes need this file. |
|  | File tag: **FI07**. |
| `FILE_INVENTORY_GEOLOGY` | Type 5 (polygons) Esri shapefile containing the geologic map or inventory. |
|  | OPTIONAL INPUT. See Table 2 for information about what classification processes need this file. |
|  | File tag: **FI08**. |
| `FILE_INVENTORY_GEOLOGY_READ_MAP` | File describing how to find the information related to the kind of soils that must be identified during the classification process (saline or carbonate rock soils / quaternary unconsolidated sediments / potentially expansive lithologies). |
|  | CONDITIONED INPUT. The name of this file is mandatory when the geologic inventory is present (`FILE_INVENTORY_GEOLOGY`). Otherwise, it is not required. See sections 2.4.3 and 2.4.3.1 for a detailed description about this kind of files. |
|  | File tag: **FI09**. |
| `FILE_DISPLACEMENT_HORIZONTAL` | File containing the horizontal components of the groud movement measuread along the LOS. This file is the output of the LOS2HV tool (see [RD2] for details). |
|  | OPTIONAL INPUT. See Table 2 for information about what classification processes need this file. |
|  | File tag: **FI10**. |
| `FILE_ADA_CLASSIFIED` | The output file, result of the classification process. Its format is described in section 2.5). |
|  | MANDATORY OUTPUT. |
|  | File tag: **FI11**. |

*Table 4: The options controlling the behavior of the application: files.*

Figure 3 depicts a complete example of an ADAClassifier options file, where all the options described in Table 3 and Table 4 may be found.

```
# ============================================================================
#
# This is an ADACLASSIFIER options file.
#
# ----------------------------------------------------------------------------
#
```

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*
GA No: 777630

```
# THRESHOLDS
#
# -----------------------------------------------------------------------------
#
# Thresholds to decide when the percentage of intersections between ADAs and
# several inventories is enough for some purpose.
#
# ALL THRESHOLDS MUST BE WRITTEN AS PERCENTAGES but those relative to slopes,
# that must be in the range (0,90) degrees-
#


#
# Threshold used in the classification process finding whether an ADA may be
# considered a landslide. This threshold stands for the minimum
# percentage of intersection between the set of points in the ADA
# and the polygon in the landslides inventory that overlaps most.
#

THRESHOLD_INTERSECTION_LANDSLIDE      = 70.0


#
# Threshold used in the classification process finding whether an ADA may be
# considered a sinkhole. This threshold stands for the minimum
# percentage of intersection between the set of points in the ADA
# and the polygon in the sinkholes inventory that overlaps most.
#

THRESHOLD_INTERSECTION_SINKHOLE       = 70.0


#
# Threshold used in the classification process finding whether an ADA may be
# considered a land subsidence. This threshold stands for the minimum
# percentage of intersection between the set of points in the ADA
# and the polygon in the land subsidences inventory that overlaps most.
#

THRESHOLD_INTERSECTION_SUBSIDENCE     = 70.0


#
# Threshold used in the classification process finding whether an ADA may be
# considered a constructive / consolidation settlement. This threshold stands
# for the minimum percentage of intersection between the set of points in the
# ADA and the polygon in the infrastructures inventory that overlaps most.
#

THRESHOLD_INTERSECTION_IFS_SETTLEMENT  = 100.0


#
# The next threshold is a multiple one, including TWO values
# written as a list. All of them represent the percentage of
# intersection between the set of points in the ADA and the polygon
# in the geologic inventory that overlaps most.
#
# These thresholds are used in different classification processes.
#
# - First threshold in the list: saline or carbonate rock soils
#   (used when looking for sinkholes) and
# - second threshold: quaternary unconsolidated sediments (used
#   when looking for land subsidences).
#

THRESHOLD_INTERSECTION_GEOLOGY        = 70.0 70.0

# -----------------------------------------------------------------------------
#
# Thresholds for slopes. Given as a percentage. May exceed 100%
#

# Slope threshold used when looking for landslides.

THRESHOLD_SLOPE_LANDSLIDE             = 10.0

# Slope threshold used when looking for subsidences.

THRESHOLD_SLOPE_SUBSIDENCE           = 10.0

# -----------------------------------------------------------------------------
```

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

```
#
# Thresholds to decide on the quality of curve fitting.
# The Pearson correlation coefficient value is used to accept or reject
# the fit. The values provided in the thresholds below must be in the
# range [-1, 1].
#

# Exponential function fit threshold. Used when looking for constructive or
# consolidation settlements.

THRESHOLD_EXPONENTIAL_FIT_SETTLEMENT    =   0.80

# ------------------------------------------------------------------------
#
# Thresholds to decide whether a displacement is big enough as
# to consider that a point is moving in the horizontal plane.
#

# Threshold used when looking for landslides.

THRESHOLD_DISP_HORIZONTAL_LANDSLIDE     =   5.01

# Threshold used when looking for sinkholes.

THRESHOLD_DISP_HORIZONTAL_SINKHOLE      =   5.02

# ------------------------------------------------------------------------
#
# FILES
#
# ------------------------------------------------------------------------
#
# Input files.
#

# Digital Terrain Model (DTM) header. Optional. Write a "dash" (-) when not available.

FILE_DTM                       = the_dtm_header.dtm

# ADA (Active Deformation Area) points.

FILE_ADA_POINTS                = ADA_points.shp

# Active Deformation Areas (ADAs, polygons)

FILE_ADA_POLYGONS              = ADA_polygons.shp

# Inventory: landslides. Optional. Write a "dash" when not available.

FILE_INVENTORY_LANDSLIDES      = inventory_landslides.shp

# Inventory: sinkholes. Optional. Write a "dash" when not available.

FILE_INVENTORY_SINKHOLES       = inventory_sinkholes.shp

# Inventory: subsidences. Optional. Write a "dash" when not available.

FILE_INVENTORY_SUBSIDENCES     = inventory_subsidences.shp

# Inventory: geology. Optional. Write a "dash" when not available.

FILE_INVENTORY_GEOLOGY         = inventory_geology.shp

# Geologic inventory read map. Mandatory when providing a geologic inventory.

FILE_INVENTORY_GEOLOGY_READ_MAP = inventory_geology_read_map.op

# Inventory: infrastructures. Optional. Write a "dash" when not available.

FILE_INVENTORY_INFRASTRUCTURES = inventory_infrastructures.shp

# Map of horizontal displacements. Optional. Write a "dash" when not available.

FILE_DISPLACEMENT_HORIZONTAL   = displacements_horizontal.shp

#
# Output files.
```

```
#

FILE_ADA_CLASSIFIED           = ADA_classified.shp

#
# End of file.

```

*Figure 3: Actual example of an ADAClassifier options (defaults) file..*

## 2.4   The (input) data files

ADAClassifier's input files are eleven, and may be grouped as follows:

- The ADA files (polygons and points),
- the DTM,
- the several inventories, that in turn should be grouped as
  - the geologic map and
  - all the remaining inventories, to finish with
- the file with the horizontal component of the ground movements.

The defaults (options) file is another input – stating, in fact, what are the actual inputs to the tools, but it has been thoroughly described in sections 2.3 so it is not included in this section. As stated before, full details on the syntax of option files may be found in [RD4].

### 2.4.1   The ADA (polygon and point) files

The files including ADA data must be the output of the ADAFinder tool. Please refer to [RD1] for a detailed description of these files.

### 2.4.2   The DTM

Only **grid-based** Digital Terrain Models stored as **binary** (as opposed to text) **files** are accepted by ADAClassifier. The format of these DTMs is as follows:

- These must contain *m* rows x *n* columns, where *m* and *n* are values greater than zero.
- The file is organized by rows. The first row corresponds to the northern boundary of the DTM. The last row, corresponds to the southern boundary of the DTM.  The first values in every row are part of the western boundary of the DTM. The last values in every row are part of the eastern boundary of the DTM. The previous statements imply that the first value in the first row corresponds to the upper left corner of the DTM, while the last value in the last row is located at its lower right corner.
- The distance between columns (in projected coordinates) is always the same. The distance between rows is always the same. The distances between rows and columns may be different.
- Each of the *m* rows contains *n* values.
- Each value must be stored as a floating-point value (4 bytes) or a double precision value (8 bytes).

In short, the kind of DTMs accepted by ADAClassifier are just data matrices. Such kind of organization leaves no room for metadata describing the characteristics of the file. This is the reason why DTM headers are required. Note that this format is the same one produced by the ENVI tool (however, see next section, the header describing DTMs does not adhere to the format used by ENVI).

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

### 2.4.2.1 *The DTM header file*

DTM header files are used to define the characterization parameters of the "raw" DTM files just described.

These adopt the form of an options file, that is, a series of labels and values. The syntax of options files (and therefore, for this particular case, of the DTM header file) is fully defined in [RD4]. Table 5 lists and describes the different parameters found in DTM header files.

| Label | Description | Comments |
|-------|-------------|----------|
| DATA_FILE | (Optionally full path and) name of the file containing the actual DTM data. When no full path precedes the file name, the DTM data file must be located in the same folder than the DTM header file (this is the recommended option). | |
| TOP_LEFT_CORNER_X | X coordinate of the top left corner of the bounding box of the DTM. | Projected coodinates. |
| TOP_LEFT_CORNER_Y | X coordinate of the top left corner of the bounding box of the DTM. | Projected coordinates. |
| DELTA_X | Distance between consecutive columns in the data grid (X spacing). | Greater than zero. |
| DELTA_Y | Distance between consecutive rows in the data grid (Y spacing). | Greater than zero. |
| N_ROWS | Number of rows in the data grid. | Greater than zero. |
| N_COLUMNS | Number of columns in the data grid. | Greater than zero. |
| SAMPLE_TYPE | Type of the data used to store the elevations (each single value). Set it to one to state that 32-bit, floating point values are used. When double precision, 64-bit values are selected, set this parameter to 2. | Integer value, either 1 or 2. |

*Table 5: The options in a DTM header file.*

An actual example of a DTM header file is depicted in Figure 4.

```
# =====================================================================
#
# This a DTM header file, defining a Digital Terrain Model organized as a
# regular grid.
# ---------------------------------------------------------------------


# Name of the file containing the actual data (the grid).


DATA_FILE        = illes_canaries_mar_0


#
# X & Y coordinates of the top left corner of the DTM. Projected coordinates
# only!!!
```

```
TOP_LEFT_CORNER_X = 262597.500
TOP_LEFT_CORNER_Y = 3165002.500


# Spacing between samples along the x & y axis.


DELTA_X          = 5.0
DELTA_Y          = 5.0


# Number of rows and columns in the grid.


N_ROWS           = 20902
N_COLUMNS        = 40802


# Type of sample: either 1 (32-bit floating point number) or
# 2 (64-bit double precision number).


SAMPLE_TYPE      = 1


# End of file.
```

*Figure 4: Actual example of an ADAClassifier DTM header file.*

### 2.4.3   The inventories

Up to five different inventories may be used by ADAClassifier to carry out its tasks. These are:

1. the landslides,
2. sinkholes,
3. land subsidences,
4. infrastructures and
5. geologic inventories.

All these inventories must be implemented as Esri type 5 (polygon) shapefiles.

The .shp files for these inventories are all alike, thus they contain the polygons defining the areas where the features of interest (landslides, sinkholes, etc.) lie. These files (.shp) are fully standardized so no specific comments need to be made to describe them.

Concerning the attribute files (.dbf), these do not matter in the case of inventories 1 to 4 in the list above (landslides to infrastructures) since they are not used by ADAClassifier to perform its task. The case of the attributes file of the geologic inventory is, however, different. See section 2.4.3.1 for details.

#### 2.4.3.1   *The attribute (.dbf) file in the geologic inventory. The read map file*

The .dbf file is crucial in the case of geologic inventories, since it is necessary to ascertain what is the kind of soil where a feature is lying. In the case of ADAClassifier, it is necessary to tell apart carbonate or saline soils and quaternary unconsolidated sediments. Such information, that is, the kind

**MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats**
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

of lithology, must be coded in one of the attributes of the .dbf file.

Geologic maps originating from different sources may code such information (the kind of lithology) in a different column or attribute of the .dbf file. For example, one inventory might include these codes in the third column of the .dbf file while another one could do the same using the first column. Moreover, the codes themselves used to describe the different lithologies may not be the same – that is, the code for carbonate or saline rock soils may be 23 in some geologic inventory and 532 in another one. Still worse, there may not exist a code to describe, precisely this kind of soils; on the contrary, it could be that the aggregation of other soil subclasses would constitute the one ADAClassifier is looking for. For instance, a geologic map could include two different lithologies, i.e., carbonate rock soils with the code 765 plus another one, saline rock soils with the code 799. Then, it would be the aggregation of these two classes what would constitute the "carbonate or saline rock soils" sought by ADAClassifier.

To manage all these differences, ADAClassifier includes a read map for geologic maps ,where all these particularities are described, thus making possible to read a wide range of geologic inventories avoiding the need to modify the software.

Two conditions, however, must be satisfied:

1. There must exist a column in the .dbf (attributes) file where the kind of lithology is coded.
2. The kinds of lithologies sought by ADAClassifier must be present (that is, a code for these exists) either directly (direct correspondence between the lithology sought and the one present in the inventory) or via aggregation of classes (several kinds of lithologies in the inventory may be considered as components making the one sought).

Then, the read map for geologic inventories file may be used to define such particularities. This file adheres to the syntax of the simple options files defined in [RD4]. The parameters to define are described in Table 1.

| Label | Description | Comments |
|---|---|---|
| POSITION_FOR_CODE | Position of the column in the .dbf file where the codes stating the different kinds of lithologies present in the geologic inventory are stored. | This is a ZERO-based index. That is, the first column in the .dbf file gets the index 0; the second one gets 1 and so on. Integer value, greater than or equal to zero. |

| Label | Description | Comments |
|-------|-------------|----------|
| CODE_LIST_FOR_SALINE_OR_CARBONATE_ROCK_SOIL | Value or list of values containing the codes representing the lithologies in the geologic inventory that may be considered exactly as or constituting carbonate or saline rock soils. | Either a single value or a list of these (see [RD4] for a description about how to write lists of values). In all cases, values must be integers. |
| CODE_LIST_FOR_QUATERNARY_UNCONSOLIDATED_SEDIMENTS | Value or list of values containing the codes representing the lithologies in the geologic inventory that may be considered exactly as or constituting quaternary unconsolidated sediments. | Either a single value or a list of these (see [RD4] for a description about how to write lists of values). In all cases, values must be integers. |

*Table 6: The parameters in a read map file for geologic inventories.*

An actual example of a read map for geologic inventories used by ADAClassifier may be found in Figure 5.

```
# =============================================================================
#
# This is the ADAClassifier's file defining how the fields in the geology
# inventory must be read.
#
# READING shapefiles containing a geology map --------------------------------
#
# ADAClassifier needs a minimal set of attributes present in shapefiles
# containing geology inventories to work properly. These attributes are,
# for each element in the inventory,
#
# - Code stating what kind of element it is (as for instance, water body,
#   quaternary sediment, saline soil, etc.)
#
# To be able to read a wide range of shapefiles, ADAClassifier includes a
# mechanism to define WHERE to find the needed fields listed above.
# By providing such information, ADAClassifier will be able to read many
# different shapefiles containing geology inventories with no changes in the
# software itself.
#
# The way to do this is to assign, to each field, the POSITION in the
# shapefile where the field sought may be found. For instance, assuming a
# shapefile whose attribute table is defined as follows:
#
# Identifier Perimeter Area Code Description ...
#
# and assuming that the code sought is the field "Code" above, its position
# would be 3 (first column uses index 0). Should it be needed, the position
# of the field named "Description" would be 4. An so on.
#
# Additionally, it is necessary to provide the list of codes defining
# each geological phenomenon sought. This version of the file supports
# the following phenomena:
#
# - saline or carbonate rock soils,
# - quaternary unconsolidated sediments.
#
```

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

```
# To define such positions and the lists of codes for each geological
# phenomenon sought, a set of pairs of labels / values must be defined below.
# The labels defining positions start with POSITION_FOR_, followed by
# the specific name of the field sought. The labels for the list of
# codes begin with the text CODE_LIST_FOR_, followed by a sentence
# stating the kind of phenomenon for which the list is provided.
#
# Change these values to meet the needs of the specific shapefile to read.
#
# Note that the list of codes for geological phenomena may be long. This means
# that the list of values may become pretty long consequently. It is possible
# to break the label into multiple lines using the backslash at the end of
# each line. The last "sub-line" must NOT include a backslash, indicating thus
# that the list of values is over. See the example below:
#
# THE_LABEL =  VALUE1 VALUE2 VALUE3 \
#              VALUE4 VALUE5 VALUE6 \
#              ...... ...... ...... \
#              VALUEX VALUEY VALUEZ
#
#
# GENERAL COMMENTS  -------------------------------------------------------
#
# Note that ALL the pairs label / value listed below MUST exist. Values
# may be changed, but the pairs label / value themselves may not be
# removed.
#
# Work with caution!
#
# =========================================================================
#
# -------------------------------------------------------------------------
#
# POSITIONS for:
# - Code
#
# -------------------------------------------------------------------------

POSITION_FOR_CODE =  3

# -------------------------------------------------------------------------
#
# LISTS OF CODES for the diverse geological phenomena.
#
# -------------------------------------------------------------------------

CODE_LIST_FOR_SALINE_OR_CARBONATE_ROCK_SOIL        = 9 90 91 92 93 94 \
                                                     95 96 97 98 99

CODE_LIST_FOR_QUATERNARY_UNCONSOLIDATED_SEDIMENTS = 8 80 81 82 83 84 \
                                                     85 86 87 88 89
# End of file.
```

*Figure 5: Actual example of an read map file for geologic inventories..*

### 2.4.4  The file with the horizontal component of ground movements

This file, containing the values of the horizontal component of the ground movement along the LOS is described in [RD2]. Note that ADAClassifier makes use of the **polygon** (not point) version of these files.

## 2.5  The output file

The ADAClassifier tools (either the GUI-based or command-line-based versions) create a single output file.

It is a type 5 (polygons) Esri shapefile, containing the ADAs and the result of the classification

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*
GA No: 777630

process.

In fact, the output file of ADAClassifier simply *extends* the attribute table (.dbf file) of the input file with ADA polygons (see section 2.4.1) which, in fact, is one of the outputs of ADAFinder – called "The ADA file" by this application.

Therefore, most of the attributes included in ADAClassifier output file are those present in the input one. These common attributes are described in [RD1]. However, Table 7, where the full structure of the output file is described, the attributes originating (and being copied) directly from the input file are included for the sake of completeness. Note that all these attributes are highlighted using an *italic* font. Please, refer to [RD1] for any clarification needed about these. The last six attributes, highlighted in **boldface**, are the contribution of this tool, ADAClassifier, and state the results of the (up to) four classification subprocesses.

| Field | Description |
|---|---|
| *ADA_ID* | *ADA identifier. This is a sequential number starting at 1 for the first ADA found. Used to link points in the output points file with the ADAs in this one.* |
| *N_APS* | *Number of active points that are included in the ADA.* |
| *X_MEAN* | *Mean X coordinate for the ADA. Computed as the mean of the X coordinates of the active points included in it.* |
| *Y_MEAN* | *Mean Y coordinate for the ADA. Computed as the mean of the Y coordinates of the active points included in it.* |
| *VEL_MIN* | *Absolute value of the minimum velocity of the active points included in the ADA.* |
| *VEL_MEAN* | *Absolute value of the mean velocity of the active points included in the ADA.* |
| *VEL_MAX* | *Absolute value of the maximum velocity of the active points included in the ADA.* |
| *VEL_CLASS* | *Velocity class.* |
| *DEF_MEAN* | *Mean of the last n values of the deformation time series of the active points included in the ADA.* |
| *TNI_CLASS* | *Temporal Noise Index Class. The number (and labels) of the classes depend on the number of temporal noise thresholds set by label* TNITABLE_THRESHOLDS *in the application's defaults file.* |
| *SNI_CLASS* | *Spatial Noise Index Class. The number (and labels) of the classes depend on the number of spatial noise thresholds set by label* SNITABLE_THRESHOLDS *in the application's defaults file.* |
| *QI_CLASS* | *Quality Index Class. The values of the classes are defined by label* QITABLE_CLASSES *in the tool's defaults file.* |
| **LANDSLIDE** | Flag stating the result of the classification subprocess looking for landslides. The results are coded as follows: -1: not classified. 0: not a landslide. 1: Potential landslide. 2: Landslide. |
| **SETTLEMENT** | Flag stating the result of the classification subprocess looking for constructive or consolidation settlements. The results are coded as follows: -1: not classified. 0: not a settlement. 1: Potential settlement. 2: Settlement. |
| **SINKHOLE** | Flag stating the result of the classification subprocess looking for sinkholes. The results are coded as follows: -1: not classified. 0: not a sinkhole. 1: Potential sinkhole. 2: sinkhole. |

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*

GA No: 777630

| Field | Description |
|---|---|
| `SUBSIDENCE` | Flag stating the result of the classification subprocess looking for subsidences. The results are coded as follows: -1: not classified. 0: not a subsidence. 1: Potential subsidence. 2: subsidence. |
| `PI_GEO_01` | Percentage of intersection between the points in the ADA and the polygons in the geologic inventory that cover saline or carbonate soils. |
| `PI_GEO_02` | Percentage of intersection between the points in the ADA and the polygons in the geologic inventory that cover quaternary unconsolidated sediments. |
| `PI_INFRA` | Percentage of intersection between the points in the ADA and the polygons in the infrastructures inventory. |
| `PI_LAND` | Percentage of intersection between the points in the ADA and the polygons in the landslides inventory. |
| `PI_SINKHOL` | Percentage of intersection between the points in the ADA and the polygons in the sinkholes inventory. |
| `PI_SUBS` | Percentage of intersection between the points in the ADA and the polygons in the land subsidences inventory. |
| `VEL_HOR` | Horizontal component of the mean of velocity of the movement undergone by the ADA (mm/yr) |
| `SLOPE` | Slope of the ADA (percentage). |
| `EXP_R2` | $r^2$ coefficient measuring the goodness of the fit of the averaged time series of the points included in the ADA and an inverse negative exponential function of the form $y = Be^{Cx}$ |
| `EXP_RMS` | RMS measuring the goodness of the fit of the averaged time series of the points included in the ADA and an inverse negative exponential function of the form $y = Be^{Cx}$ |
| `EXP_B` | $B$ coefficient of the inverse negative exponential function of the form $y = Be^{Cx}$ matching the averaged time series of the points included in the ADA. |
| `EXP_C` | $C$ coefficient of the inverse negative exponential function of the form $y = Be^{Cx}$ matching the averaged time series of the points included in the ADA. |

*Table 7: The attributes in the output, classified ADAs file.*

## 2.6 Installing the tools

To install the tool, please refer to [RD5].

## 2.7 Using the tool

The tool has two incarnations that must be used in different ways.

To use **the GUI version** it is enough to start the application by whatever means at hand (start menu, desktop shortcut…), fill the data in the dialog shown by the tool (see section 2.2, Figure 1 and Figure 2) and click on the "Go" button. Once the process is over, a message box warning about the result (success, failure and its reason) will show up. The process may be interrupted at any moment by closing the GUI window.

Note that once all the data in the GUI has been typed, it is possible to save it as an options file keeping these values for later use. To do it, use the button labelled "Save options file". Any errors related to data (as invalid values, out of range, etc.) will be reported. The options file will only be saved once

that no errors are detected.

Conversely, an existing options file may be loaded using the button "Load options file", thus saving the effort of retyping all the required information.

To use the **batch version,** a command line must be opened. To start it, type the following command:

```
adaclassifier_cmd <options_file>
```

where `<options_file>` stands for the full path and name of the defaults (options) file. See section 2.3 and [RD4] for more details about ADAClassifier's option files.

Note that in the case of the command line version, the option file does not need to stay in the application's directory. This is so because options files are the means to control the application and it is much more convenient let the users manage these files at his / her will.

Note also that if the tool is not located in the current (windows) path, then it will be necessary to prepend the path to the directory where the tool is stored.

*Dissemination Level: CO =* Confidential

MOMIT - Multi-scale Observation and Monitoring of railway Infrastructure Threats
*Project funded by S2R JU under the European Union's H2020 R&I programme*
GA No: 777630

# 3 Conclusions and Recommendations

The family of ADAClassifier tool completely automates the process of classifying the ADAs identified by ADAFinder, according to the decision table shown in Table 1. It is a fast tool (a between two or three minutes in a regular computer: i5, 8Mb RAM, this depending on the amount of data to process and the number of classification processes to run). This makes ADAClassifier a truly production tool. It may be integrated as a toolbox (ArcGIS) or plugin (QGIS). Moreover, the kind of inputs and outputs managed may be obtained from GIS tools as well, being therefore no obstacle in a "production" workflow based on GIS tools. The same may be said about the outputs.

The application is configurable, this meaning that it may be adapted to read different kind of input geologic inventories and a (small set of) DTMs, providing these include the minimum set of attributes needed to operate.

# Parent Documents

[PD1]    Shift2Rail Joint Undertaking (S2R JU) – Multi-Annual Action Plan (MAAP) – Rev. 3 – 26/11/2015

[PD2]    Shift2Rail Joint Undertaking (S2R JU) – Annual Work Plan 2017 – Version 1.1– 23/12/2016

[PD3]    MOMIT – Description of Action (DoA) – GA 777630

[PD4]    MOMIT – Grant Agreement (GA) – GA 777630 – 22/11/2017

[PD5]    MOMIT – Consortium Agreement (CA) – Version 1.0 – 01/11/2017

*Dissemination Level: CO* = Confidential

# Reference Documents

[RD1]   Navarro, J.; Cuevas, M. MOMIT technical note – ADAFinder - A tool to automate the detection of Active Deformation Areas.

[RD2]   Navarro, J.; Cuevas, M. MOMIT technical note: LOS2HV - Compute horizontal & vertical components of movement

[RD3]   Esri.       Esri       shapefile       description.       1999. https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf. Last visited: March 13th, 2018.

[RD4]   Navarro, J.; Cuevas, M. MOMIT technical note – Simple options file: technical description.

[RD5]   Navarro, J.; Cuevas, M. MOMIT technical note – Installing the software.

# End of the document