Displacement Time Series Analysis Tutorial

Version 5.7.0

January 2023

Table of Contents

1. Introduction

The main goal of this tutorial is to give a basic description of the SARscape displacement time series analysis tools to obtain displacement classification maps and obtain different displacement parameters.

Three sets of tools for time series post postprocessing are available in SARscape. The first set of tools aim to filter time series reducing their noise. A second set of tools is intended for the extraction of displacement parameters such as periodicity, acceleration or coupling to an external measurement (i.e. temperature, rain, water level, etc) among others, and sequentially generate a classification map according to their behavior. The third set of tools generate a displacement classification via a clustering algorithm.

Time series filtering
This tool filters the deformation time series generated by PS or SBAS processes in raster or vectorial formats. It is intended to reduce the noise of the deformation time series preserving the signal as much as possible. For this, different filtering kernels/methods are available:

- Median: moving window median statistical filter.
- Mean: moving window mean statistical filter (square kernel convolution).
- Mode: moving window mode statistical filter.
- Triangular: triangular kernel convolution.
- Gaussian: gaussian kernel convolution.

Time series analytical and phenomenological fitting and classification
This set of tools fits raster or vectorial deformation time series using analytical and phenomenological models. It retrieves the best fitting coefficients and its RMS for each pixel in the time series. These coefficients represent different characteristics of the displacement such as speed, acceleration, oscillation amplitude and frequency.

The available models are linear, quadratic, periodical, or step function.

The linear model is defined by:

$$d = a_0 + a_1 t$$

With:

d the displacement [mm]

t time [year]

a_0 an initial displacement offset [mm]

a_1 the deformation speed [mm/year]

The quadratic model is defined by:

$$d = a_0 + a_1 t + a_2 t^2$$

With:

d the displacement [mm]

t time [year]

a_0 an initial displacement offset [mm]

a_1 the deformation speed [mm/year]

a_2 the deformation acceleration [mm/year ^{2}]

The periodical model is defined by:

$$d = a_0 + a_1 t + a_2 t^2 + a_3 \sin\left(\frac{a_4 + t}{a_5} 2\pi\right)$$

With:

d the displacement [mm]

t time [year]

a_0 an initial displacement offset [mm]

a_1 the deformation speed independent of the periodical trend [mm/year]*

a_2 the deformation acceleration [mm/year ^{2}]

a_3 the oscillation amplitude [mm]

a_4 a delay [days]

a_5 the oscillation period [days]

The step function model is defined by:

$$d = a_0 + a_1 t + a_3 H(a_2)$$

With:

d the displacement [mm]

t time [year]

a_0 an initial displacement offset [mm]

a_1 the deformation speed independent of the periodical trend [mm/year] *

a_2 the step date [days]

a_3 the step amplitude [mm]

The phenomenological model is defined by:

$$d = a_0 + a_1 t + a_2 M(t + a_3)$$

With:

d the displacement [mm]

t time

M external temporal phenomenological data [M unit]

a_0 an initial displacement offset [mm]

a_1 the deformation speed independent phenomenological component [mm/year] *

a_2 a scaling constant [mm/M unit] **

a_3 a delay between the phenomenological event and the displacement [days]

* This estimated speed corresponds to the deformation process not described by phenomenological or periodical process. For example, a structure may sink at constant speed and thermally dilate at the same time. This velocity corresponds to the sinking process.

** This scaling constant represents the coupling between the deformation and the external phenomenological magnitude. For example, in the case of external temperature measurements, this magnitude can be interpreted as the coefficient of thermal expansion of the target.

Once all the desired models are fitted, a classification processed can be performed by selecting the best fitting model for each time series. The best fitting model is selected according to one of the following criteria selectable by the user according to the lowest value of the following metrics:

- Akaike information criterion (AIC)

$$AIC = n \log\left(\frac{SSE}{n}\right) + 2k + \frac{2k(k+1)}{n-k-1}$$

- Bayesian information criterion (BIC)

$$BIC = n \log\left(\frac{SSE}{n}\right) + k \log(n)$$

With

$$SSE = \sum_{i=1}^{n}(d_i - D_i)^2$$

k the number of parameters
n the number of measurements

d the displacement obtained by the fitting model
D the measured displacement

Time series clustering

Another approach to classify time series is based on This tool is intended to cluster the deformation time series generated by PS or SBAS processes. It works by grouping nearby time series according to their similarity based on a modified DB Scan algorithm.

To quantify the similarity between time series two time series distances were introduced:

- Euclidean distance

$$D_{euc}(M,S) = \sqrt{\sum (T_m - T_s)^2}$$

- Cosine distance or Correlation

$$D_{cos}(M,S) = 1 - \frac{\sum T_m T_s}{\sqrt{\sum T_m{}^2 \sum T_s{}^2}}$$

Note that Euclidian metric is applied on normalized TS. The normalization is performed by subtracting the mean value and dividing it by its standard deviation. This helps matching TS with different amplitude but same behaviors.

Using a user provided time series distance threshold, the connections (in terms of similarities) of a given time series against others can be calculated. This is repeated for all time series in the scene. These connections define a *graph* in each *component* of it represents a cluster.

As ground deformations are usually localized, to improve the computational time, but more importantly avoiding an excess of connections which may lead to gigantic bad connected clusters where you can always find an indirect connection between two totally different times series, connections between time series are only allowed if they are inside a given radius.
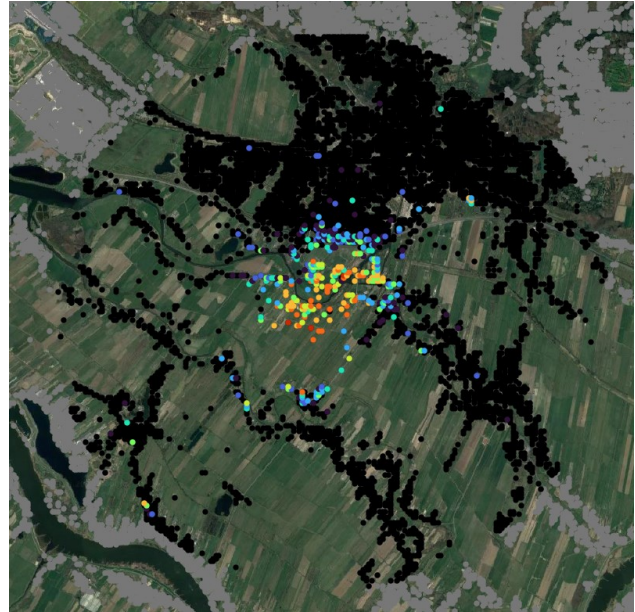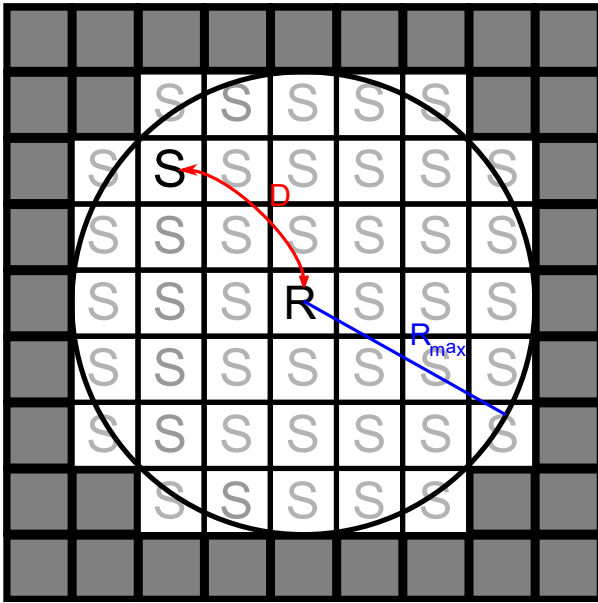
Figure 1 The length of this radius must be selected based on the dimension of the possible mechanisms present on the image. Increasing the radius not only has an impact on computing time and memory usage but also risk connecting pixels which may not be fiscally connected.
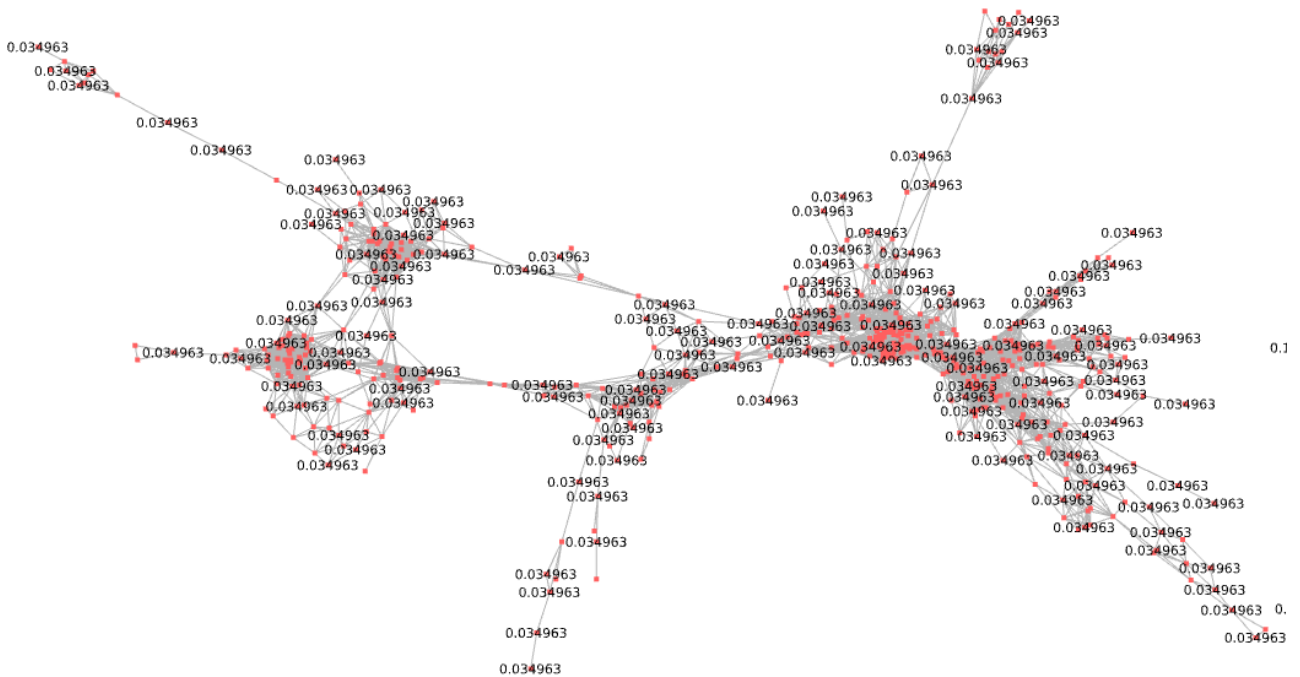

Figure 2 A badly connected graph in which to vertex (time series) may be connected through many intermediate time series in which the extremes of this connection may be totally different.

Components with *more edges* (connections) represent more reliable clusters avoiding sequential bridges where the first vertex may be totally different from the last. To rate the connectivity of each component the

mean path length (l_g) metric is introduced rangeing from 1 on a complete graph to 0 for not connected components.

$$l_g = \frac{e}{v(v-1)}$$
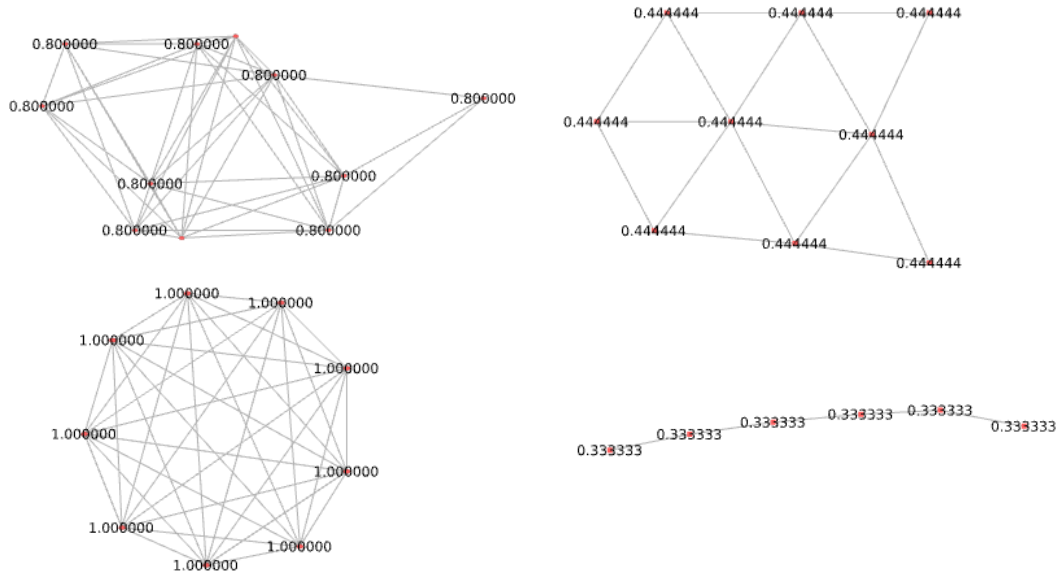
Figure 3 Examples of clusters with their mean path length

Figure 4 A graph is constructed with each vertex (red dot) representing a pixel and an edge (black line) representing each connection between pixels. The values on each vertex correspond to the mean path length of that component. This metric together with a min vertex number can be used to discard unreliable clusters.

An external thematic map, such as soil type, can be used to penalize cluster connections between map classes. The penalization constant is provided by the user.

2. Training Data sets description

The three exercises on this tutorial will be based on a single SBAS dataset in raster and vectorial format on Xiamen China although it could be performed with PS results as well.

In this urban scene different deformation processes are present. Some regions present linear and accelerated subsidence while some buildings present thermal dilatation.
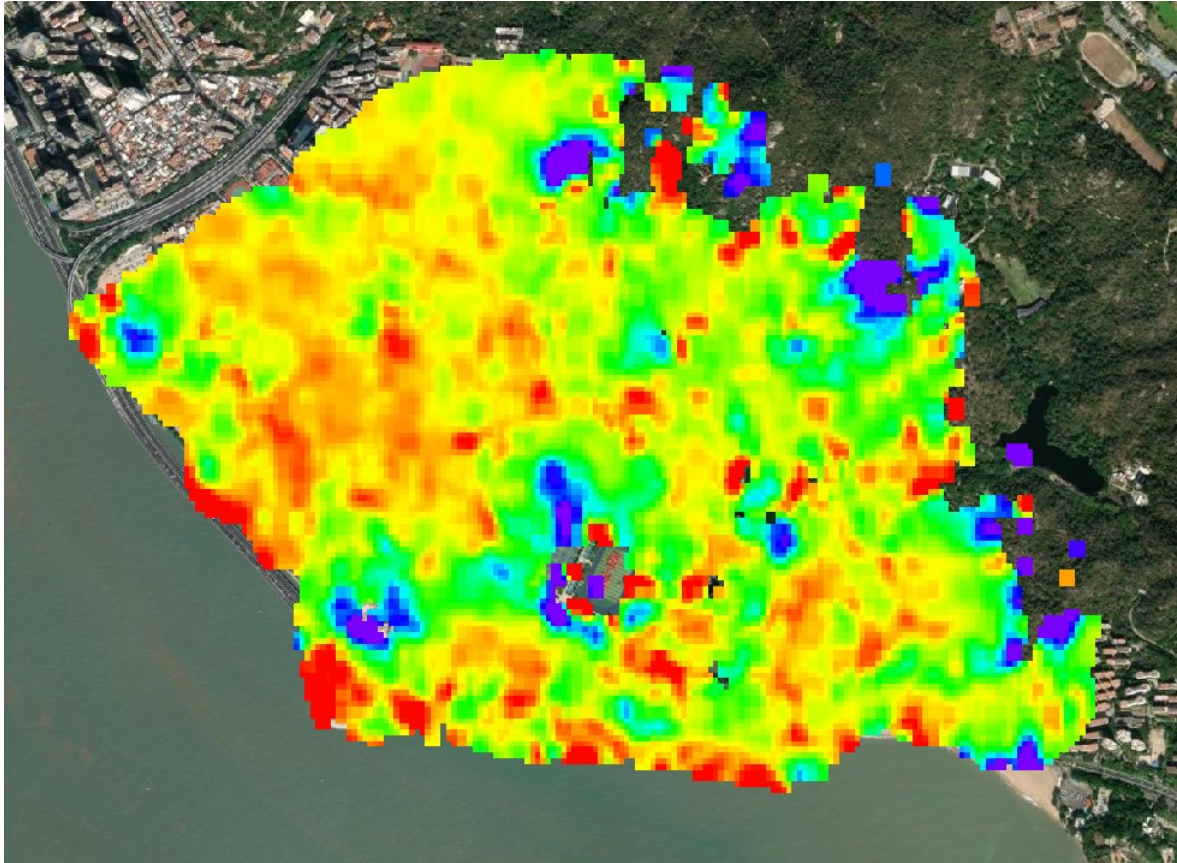
Figure 5 SBAS derived displacement velocity in raster format

Figure 6 SBAS derived displacement velocity in vector format

This results where generated using 95 images from 2015/07/06 to 2018/10/18 and are in line of sight of a single geometry even if this process can also be applied to decomposed (vertical-horizontal) data.

3. Time Series Filtering

The example of this section of the tutorial will be performed with vectorial, the raster version is left as exercise to the student. This step is suggested to be performed before clustering as it mitigates the signal noise aiding reducing the distance between time series.

The scope of this section is to get familiarized with the filtering functionalities of the time series filtering. To open the filtering panel go to *Interferometric Stacking/Post Processing Tools/Time Series Filtering*.

Input File List

Select the *SI-v8d0-h5d0-c0d2_0.shp* file (in general, more than one shp file can be selected) from SBAS\vector. In case of raster data select the single *SI_disp_geo_meta*

Principal Parameters

Filtering Method
The method used to filter the Time Series. Median, mean, mode, triangular, gaussian methods are available.

Window Size
The moving window size where the filtered method is applied. Larger windows will result in stronger filters.
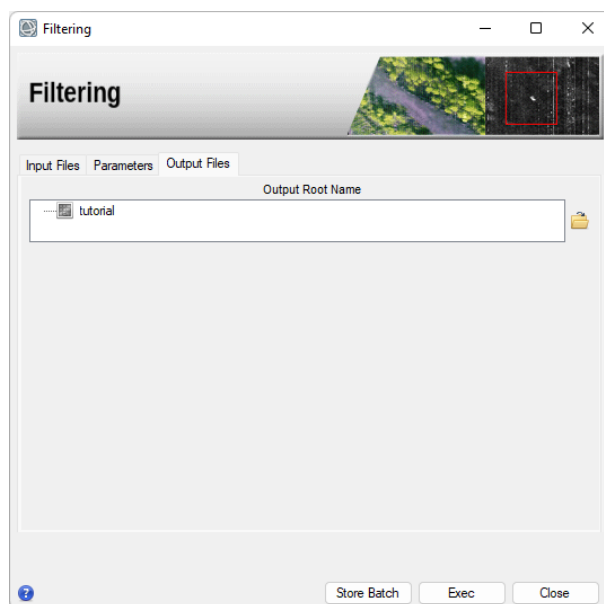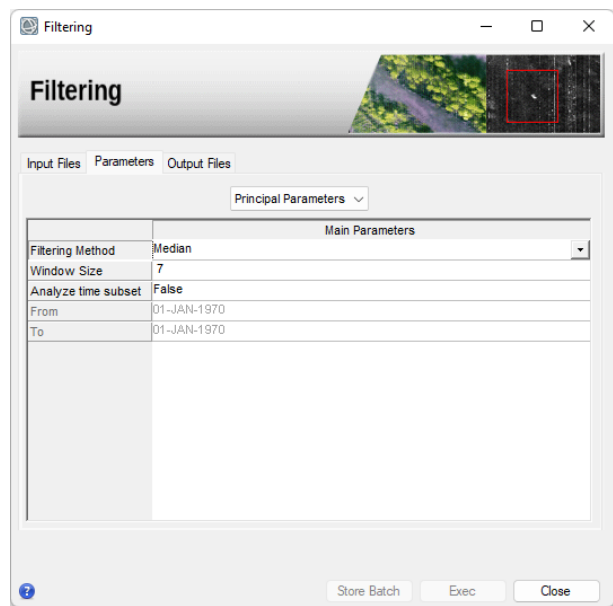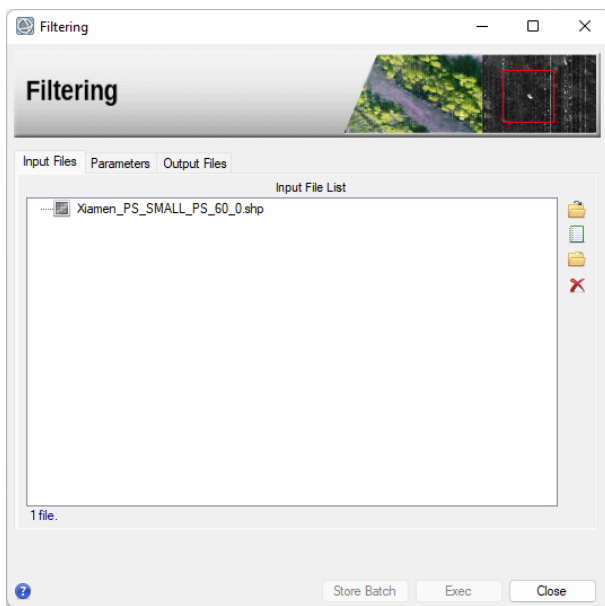
Analyze time subset
By setting this flag only the period defined by From and To will be analyzed.

Output Root Name

Output Root Name_Filtering Method_disp_geo_meta: Filtered deformation meta and its corresponding files.
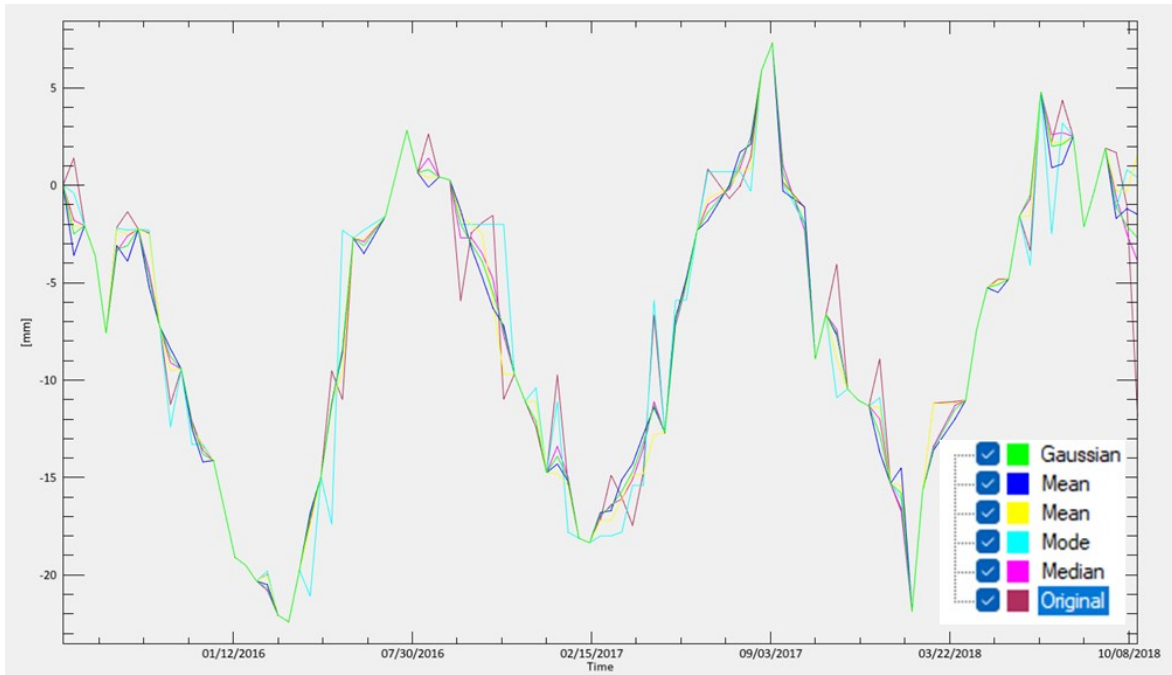*Input Name_fil.shp*: Filtered deformation shapefiles.
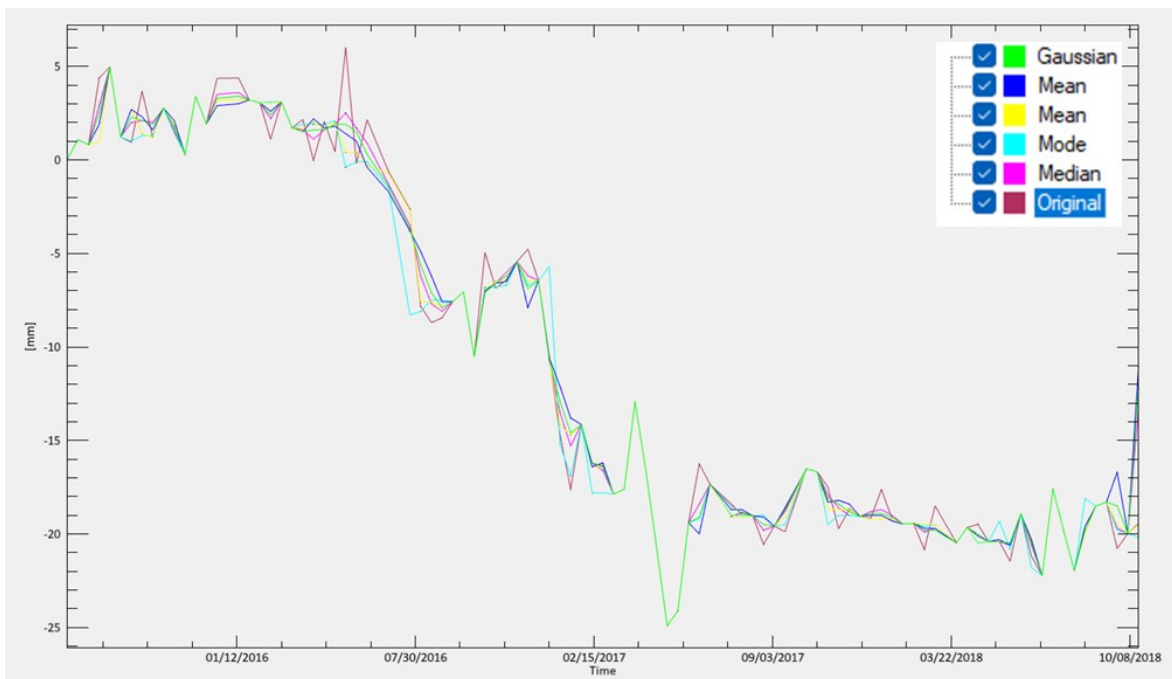
Figure 7 Example of different filtering methods

Figure 8 Example of different filtering methods

You can compare the resulting time series with the original one. Typically, we suggest using the median filter as it seems to reduce the noise but keeping steps opposed to the mean filter where these steps seem to be rounded.

4. Time Series Classification

The example of this section of the tutorial will be performed with raster, the vector version is left as exercise to the student. This process is based on three steps, of which one is optional, which must be performed sequentially.

The first step uses four analytical models to fit the time series. The second step fits the time series against an external value and the third step selects the best fitting model.

The scope of this process is not only to classify the scene in different behaving areas but to extract physical parameter of the displacement such as accelerations, oscillating frequencies, coupling constant and delay between displacement and an external value (i.e. dilatation constant and thermal delay if the external value is temperature).

Parametrical fitting

To open this panel, go to *Interferometric Stacking/Post Processing Tools/Time Series Classification/1B - Parametrical Analysis.*

Input Files

Select the SI_disp_geo_meta file from SBAS\geocoded_disp. In case of raster data select the single SI-v8d0-h5d0-c0d2_0.shp (in general, more than one shp file can be selected).

Parameters - Principal Parameters

Analyze time subset

By setting this flag only the period defined by From and To will be analyzed.

Generate fitted TS

By setting this flag the fitted time series for each model will be saved. This may be useful to analyze different model performance.

Set the *Generate fitted TS* as true. It will write the fitted deformation files which will be useful to understand the results.

Linear fit

By setting this flag the linear model fitting will be performed.

Quadratic fit

By setting this flag the quadratic model fitting will be performed.

Periodical fit

By setting this flag the periodical model fitting will be performed.

Heavyside fit

By setting this flag the Heaviside model fitting will be performed.

SHP TS prefix String

String preceding the date characters in the shapefile dbf. (i.e. D_ for D_20230101)

Parameters - Other Parameters

Limits for the possible fitting parameters can be introduced.

Output Files

A directory is created for each executed model combining the Output Root Name and the model's name (I.e.: projectXXX_Linear, projectXXX_Periodical, projectXXX_Quadratic).

In the case of raster input inside each of these directories you will find:

*modelName*_aN

Images with fitting parameter of order N.

*modelName*_daic

Image with the delta Akaike Information criterion to compare the goodness off fitting among models.

*modelName*_rms

Image with the fitting RMS error.

*modelName*_meta

Meta file of fitting parameters and the fitted time series (if selected).
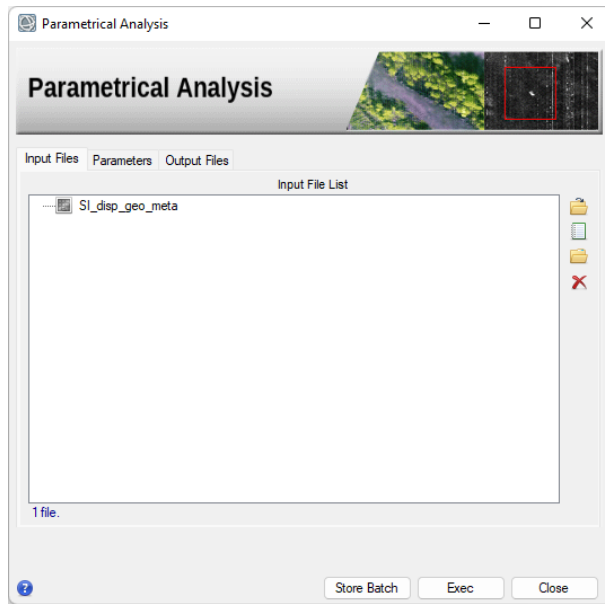
modelName_input_TS

Fitted time series images (if selected).

In the case of vector input inside each of these directories you will find:

modelName_imputfile.shp

A shapefile containing the fitting constants and results.

You can limit or fix a fitting parameter going to Other Parameters.
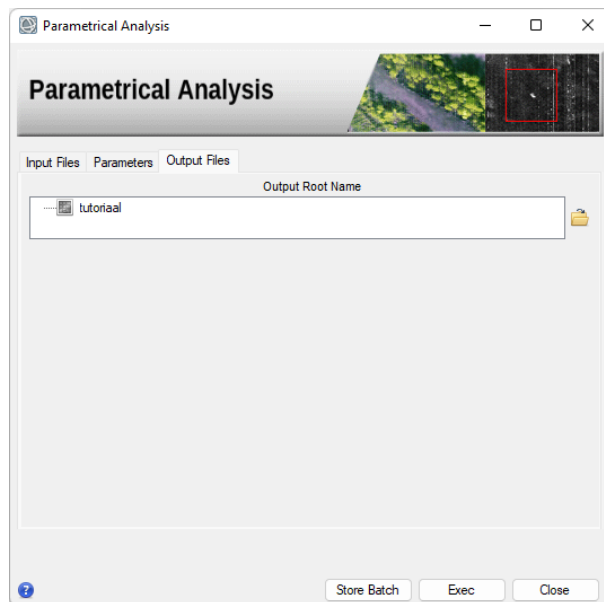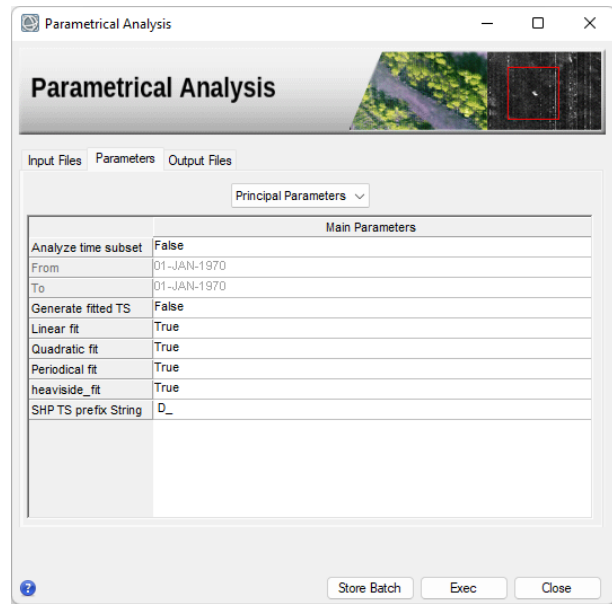
As an example, open the a_3 and a_5 parameter files from the fitted periodical model folder. The a_3 parameter corresponds to the oscillation amplitude while a_5 corresponds to the period.
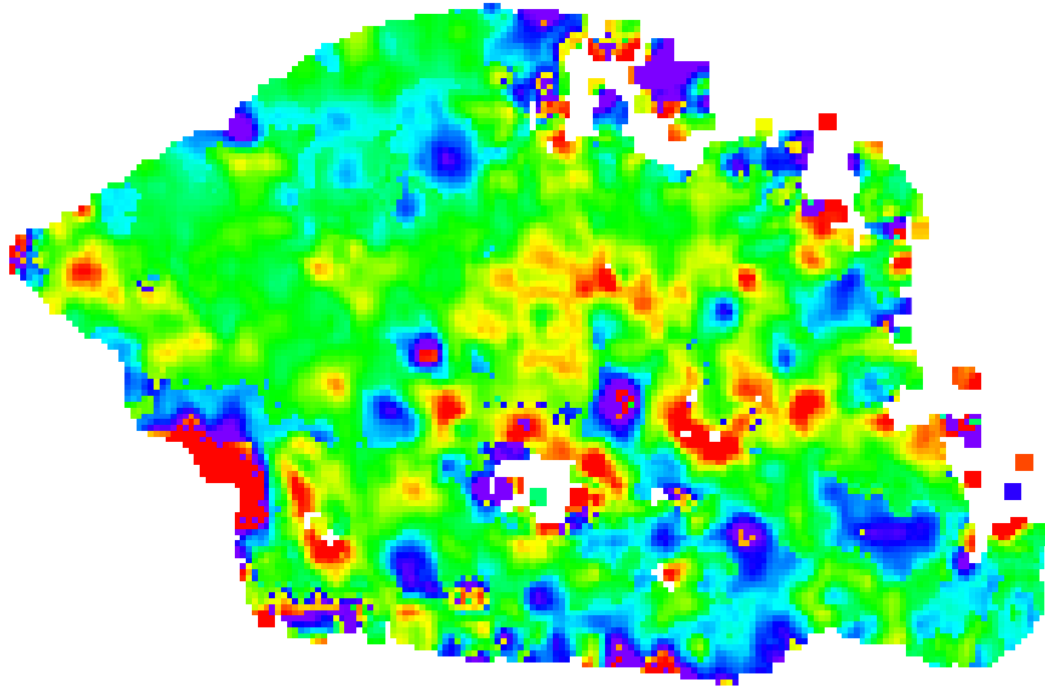
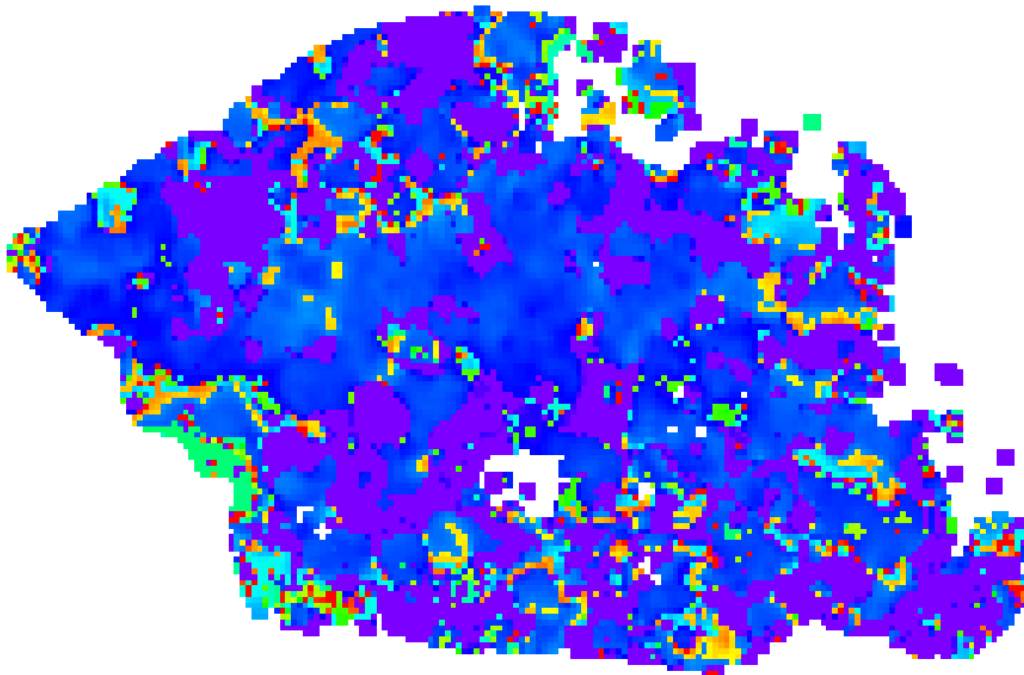Figure 9 a_3 parameter from the periodical model representing the amplitude of the oscillation.

Figure 10 a_5 parameter from the periodical model representing the delay (or phase) of the oscillation.

We found some buildings that have around 7mm oscillation amplitude and a 365-day period. It is to the student to open other files to find accelerations.

With hints we can suspect that temperature may be causing some dilation on the buildings, this motivates us to try fitting the displacement against an external temperature measurement.

Phenomenological fitting

To open this panel, go to */Interferometric Stacking/Post Processing Tools/Time Series Classification/1A - Phenomenological Analysis.*

As seen in the previous step, temperature seems to be related to the dilation of some buildings, so we intend to fit the displacement against temperature. To do this we first need to prepare the historical temperature of the area.

You can provide it via a csv file with the corresponding format or using the era5 download tool which creates a series o temperature maps corresponding to the timeseries.

If you want to provide a csv file with the external quantity like the following example

```
date,temp
01-01-2014,14.9
02-01-2014,16.5
03-01-2014,17.1
04-01-2014,14.7
05-01-2014,13.4
06-01-2014,15.3
07-01-2014,16.9
08-01-2014,17.2
09-01-2014,11.1
10-01-2014,12.8
11-01-2014,15.3
…
```

Note that if you provide a csv file there will not be any spatial information, the same value is going to be applied all over the image. This format can be useful for example in the case of damns water levels where a map has no sense.

After generating an Era5 account, if you prefer to download the temperature data from ECMWF Era5 you can use the seascape tool found on *Import Data / Download / ECMWF Era5 Download*.

Input Files

As Input file select the S1_disp_geo_meta you are working with. This will populate the input list with all the displacement images which are going to be used to define the geographic area and time interal.

DEM/Cartographic System

Output Projection

Select the cartographic system you are working with.

Parameter

Definition of the product type. In this case we need the temperature so we will leave the default value which 2 metre temperature.
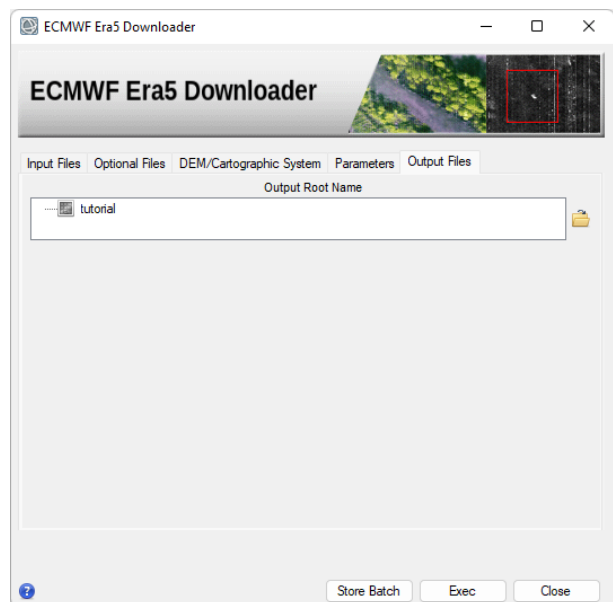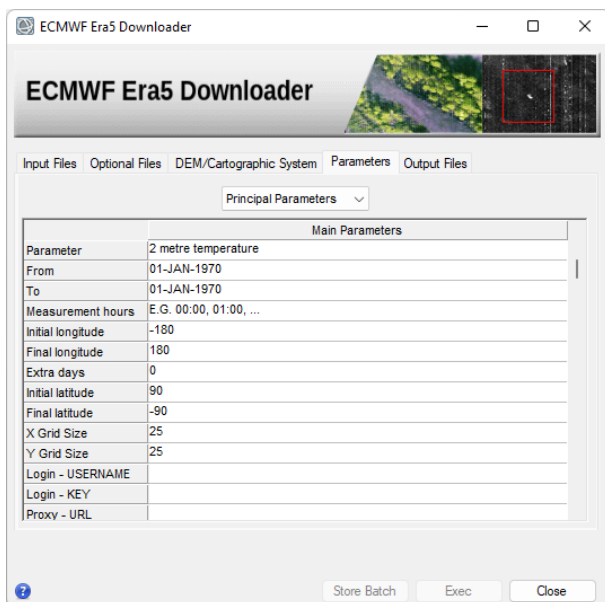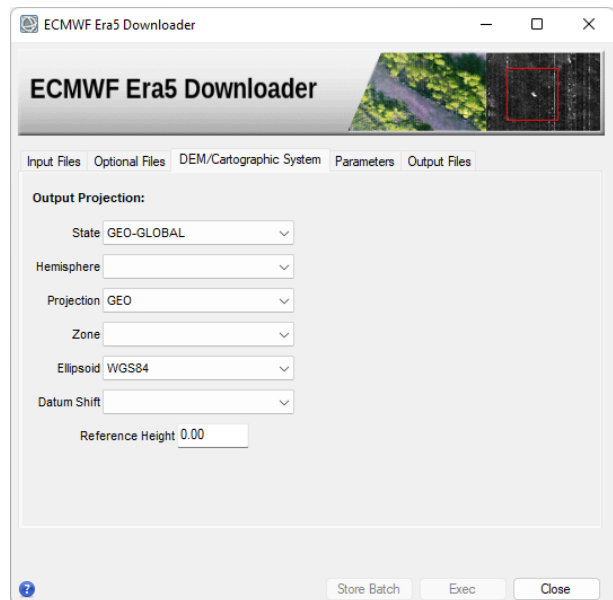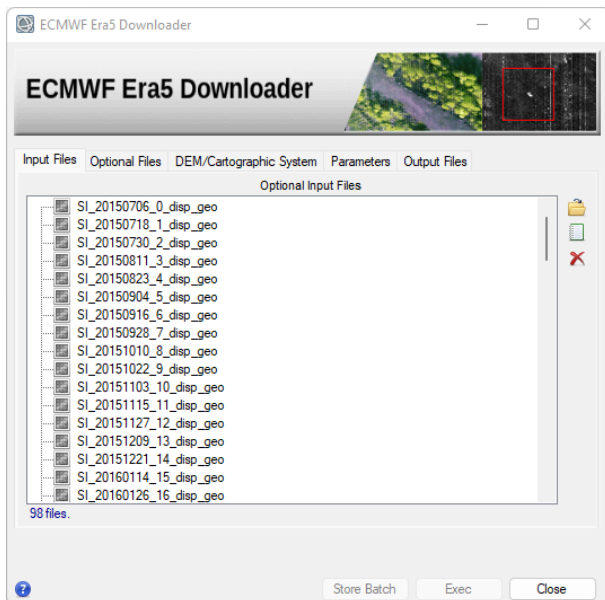
Login - USERNAME

Username (UID) of ECMWF ERA5 Web-API login credentials.

Login - KEY

Web-API key of ECMWF ERA5 Web-API login credentials.

Output Root Name

Name of the output root. The data files with its meta will be created in this directory.

This process can take some minutes, do not close the panel.

Once the download is completed, we can proceed to fit the displacements against the temperature. To do this open the panel found on *Interferometric Stacking / Time Series Classification / Phenomenological Analysis*.

Input File

The displacement meta file (*_meta) . Both decomposed or LOS meta files are supported.

Phenomenological Data File

A list of external measurements in csv and/or meta files format such as the ones obtained from the ERA5 or ECMWF download tools.

Parameters - Principal Parameters

Analyze time subset

By setting this flag only the period defined by **From** and **To** will be analyzed.

Generate fitted TS

By setting this flag the fitted time series for each model will be saved. This may be useful to analyze different model performance.

Output Files

A directory is created combining the ***Output Root Name*** and the model's name from the Phenomenological Data File name (I.e.: ***projectXXX_temperature***). In that directory the files are:

In the case of raster input inside each of these directories you will find:

*modelName*_aN

Images with fitting parameter of order N.

*modelName*_daic

Image with the delta Akaike Information criterion to compare the goodness off fitting among models.

*modelName*_rms

Image with the fitting RMS error.

modelName_meta

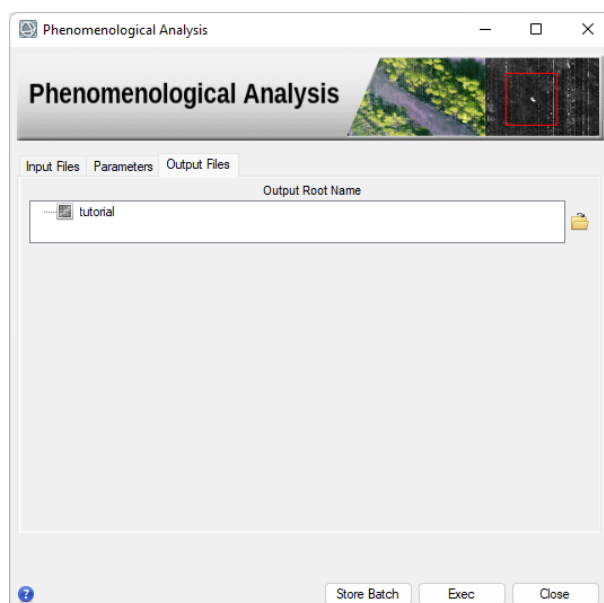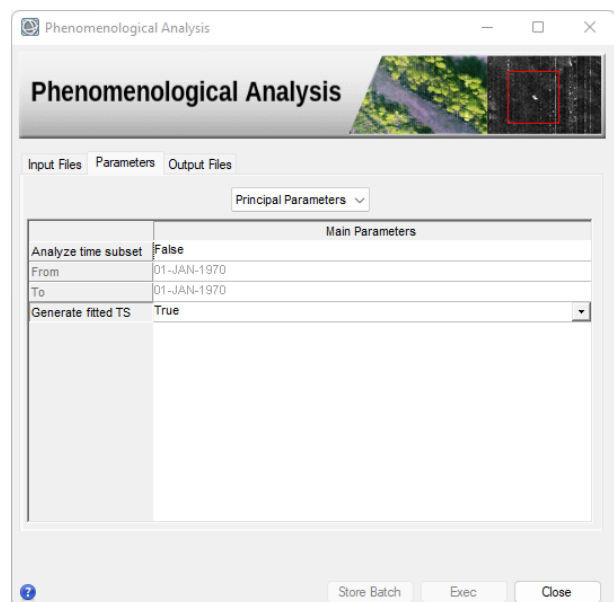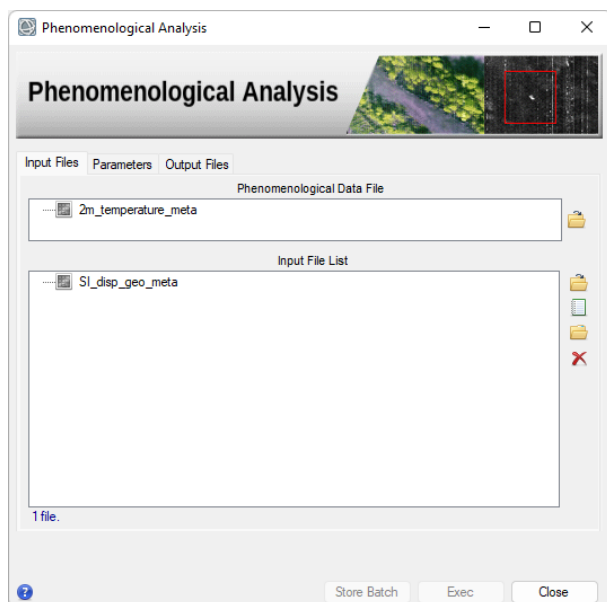Meta file of fitting parameters and the fitted time series (if selected).

modelName_input_TS

Fitted time series images (if selected).

In the case of vector input inside each of these directories you will find:

modelName_imputfile.shp

A shapefile containing the fitting constants and results.

Using the time series analysis tools found on *General Tools / Time Series Analyzer / Raster*, we can look at the results of the fitting comparing the temperature and periodical model for the a point over the oscillating buildings.
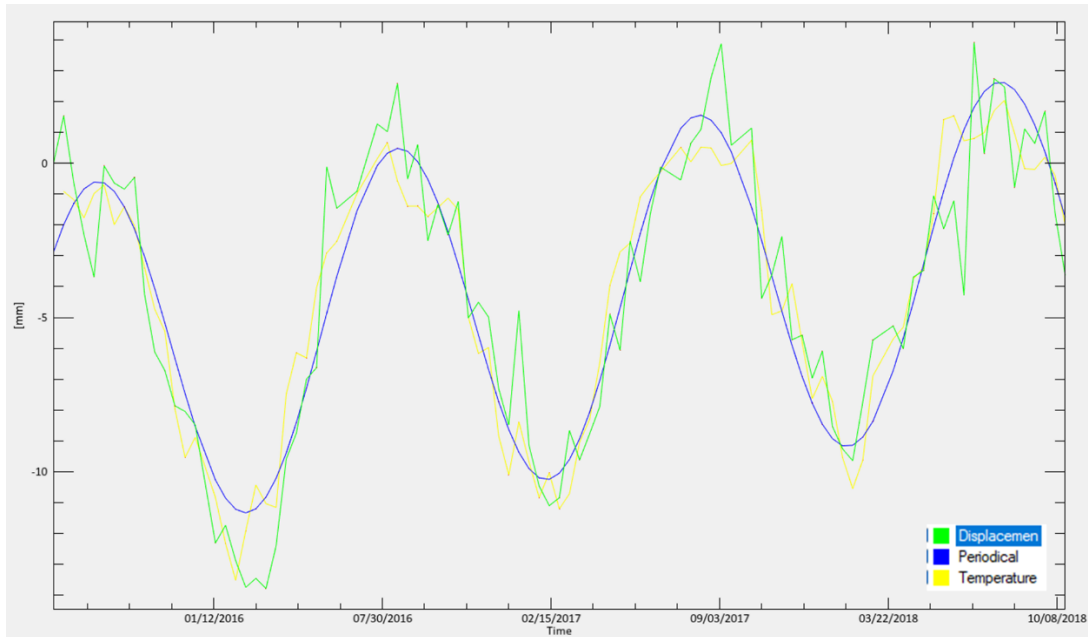

Figure 11 Comparison of periodical and temperature fitting of the displacement.

Again, we can look at the coupling coefficient (a_2) and delay (a_3) of the fitting finding areas that seem to couple with the temperature variation.
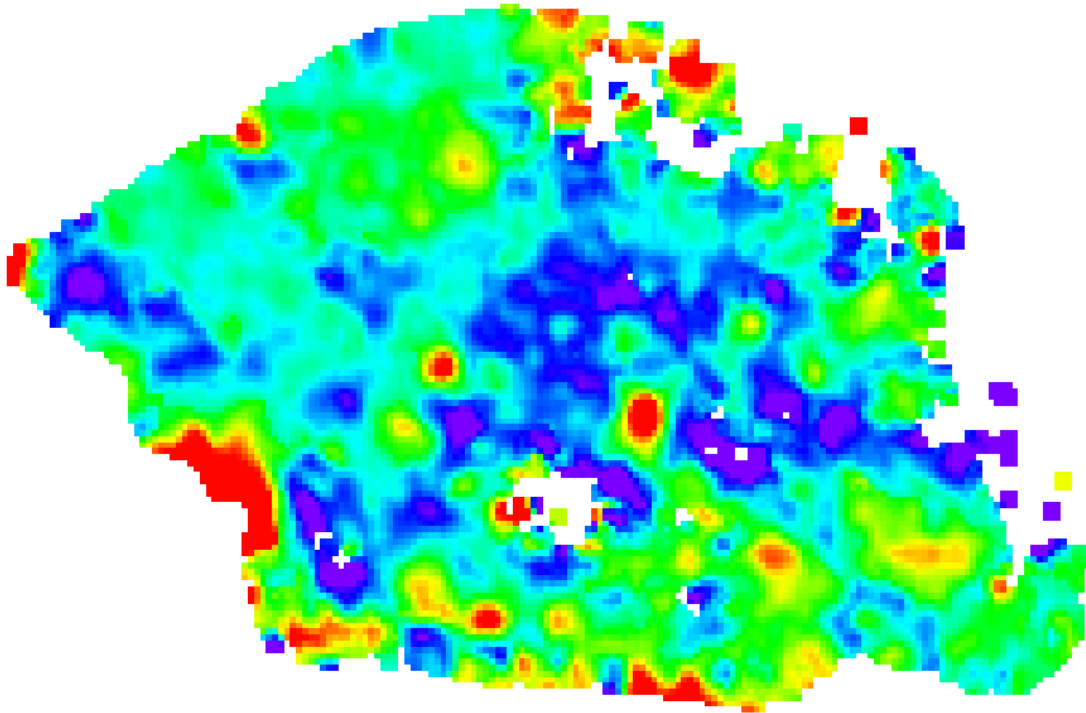
Figure 12 Coefficient constant between deformation and temperature. Higher values show higher relation between temperature and deformation.
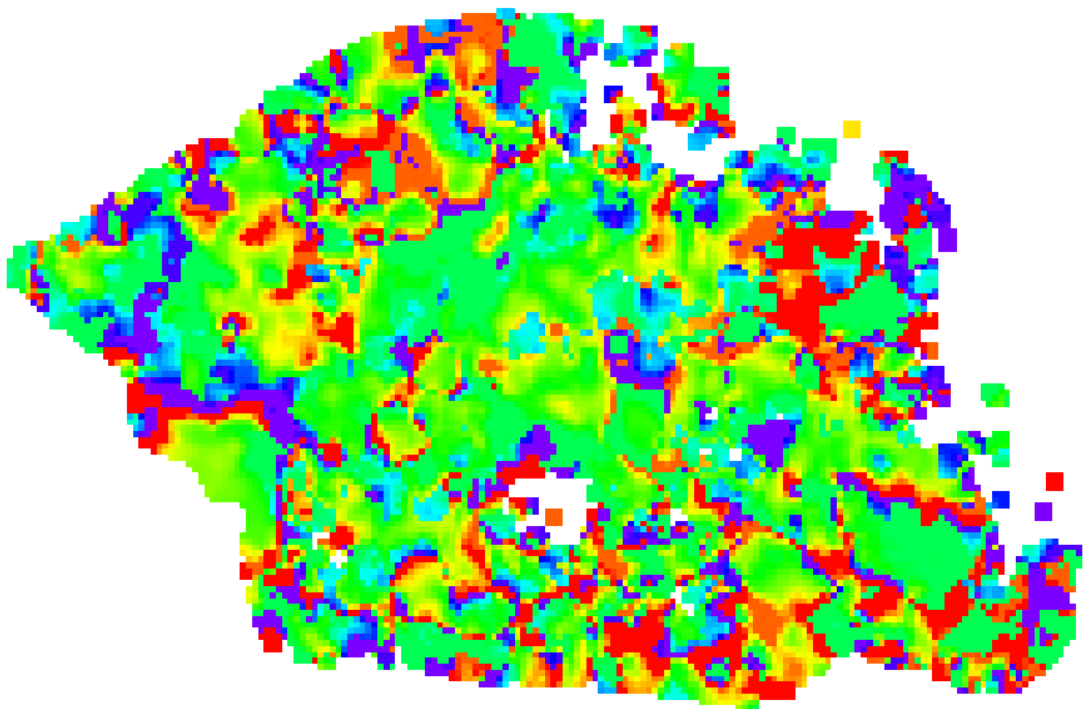

Figure 13 Delay between deformation and temperature change.

We can notice on the left side of the image an area that seem to respond stronly to the temperature with a coupling value of about 0.7 mm/deg and a delay of a 5 days with respect to the temperature.

This step can be performed sequentialy with as many phenomenological measurements as desired. For example, a part of the image may deforming bya the act of the temperature and an other part coupled with precipitations.

Classification

Once we have fitted all the desired models, we can classify the image according to the best fitting model according to the Akaike Information Criterion or the Bayesian Information Criterion. This will expose areas with different behaviours. Some attention must be taken in the selection of models avoiding competition among them. For example, in the previous case we saw that periodical and temperature models seem to fit very well in some parts. In this case is better to leave one of the two models out of the classification.

To open the classification tool, go to */Interferometric Stacking/Post Processing Tools/Time Series Classification/2 - Classification*.

Input File

Modeling meta files (*_meta) created the Phenomenological Modeling and/or the Analytical Modeling panels.

Select all the _meta files generated by the fitting processes with the exception of the periodical fitting.

Parameters

Generate Classified TS

By setting this flag a composed fitted time series will be generated with the fitted time series of the best fitting model.

Select Generate Classified TS true

Majority filter

By setting this flag a majority filter will be performed with a filter window defined by **MF window size**.

Select majority filter true

MF window size

Dimension of the sliding window used to perform the majority filter.

L1 threshold

The L1 norm (normalized by the number of dates in the time series) displacement threshold to consider a pixel as a stable one. Pixels under this threshold will be considered as no displacement points in the classification map.

RMSE threshold

The maximum root mean square of allowed for a model to be considered. If all models yield an RMSE over this threshold this pixel will be displayed as not classified on the classification map.

In the case of raster input inside each of these directories you will find:

Output Files

A directory is created for each executed model combining the **Output Root Name** and the **"_classification"** string (I.e.: **projectXXX_classification**). Inside each of these directories you will find:

classification

A classification image showing the best fitting model for each pixel according to the lowest delta Akaike Information Criterion, non-classified areas, and stable areas.

classification_filtered

classification image filtered with a majority filter (if selected)

classification_daic

Delta Akaike Information Criterion value of the best fitting model for each pixel. *

classification_rms

RMS error of the best fitting model for each pixel. *
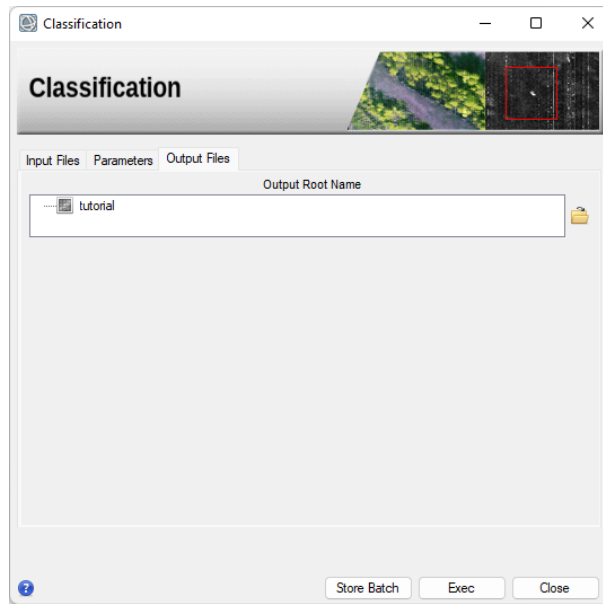
classification_input_TS

Best fitted time series images (if selected). This TS is a combination of the individual TS of each model using the TS of the best fitting model for each pixel. *

In the case of vector input inside each of these directories you will find:
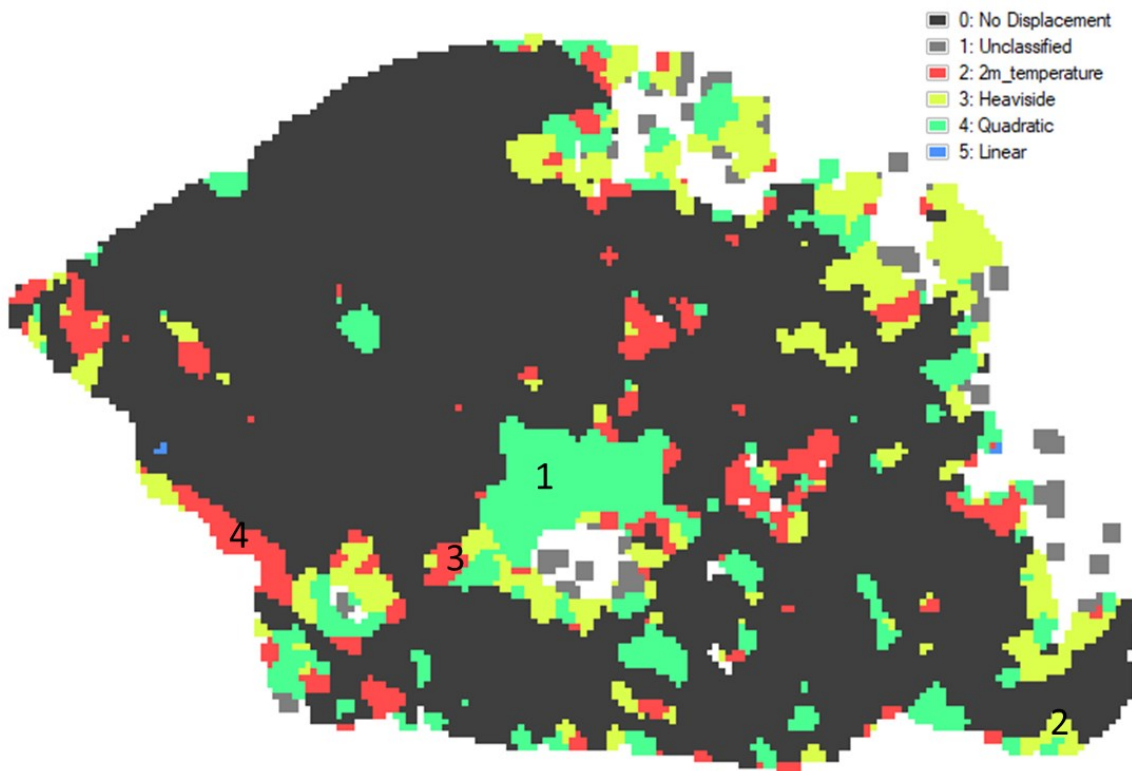
classification_imputfile.shp

If the classification majority filter is performed, these files correspond with this filtered classification.
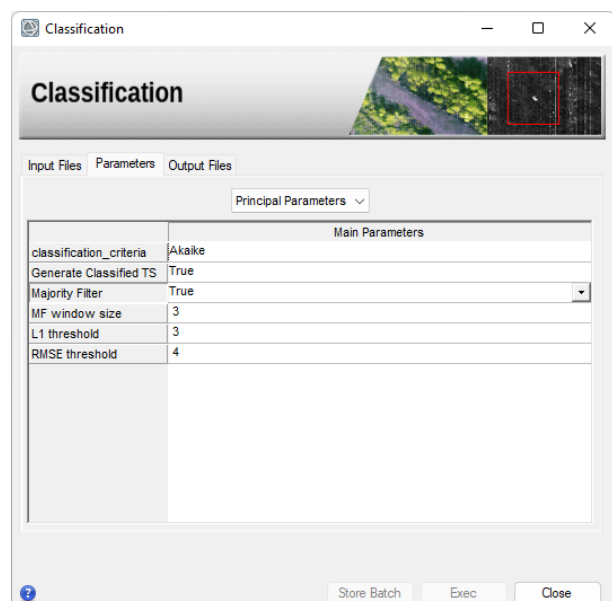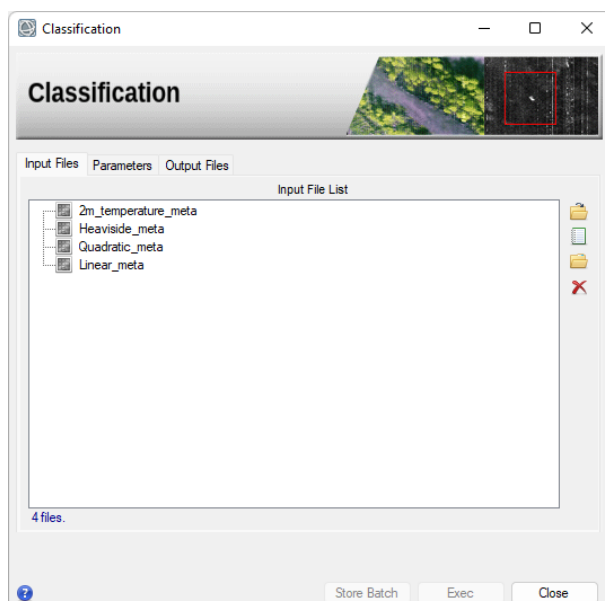
Open the filtered *classification_filtered* file in ENVI

You will find a classification map with the best fitting models, the *unclassified* areas due to high *rms* result of the fitting and the *no displacement* areas (stable) according to the *L1 Threshold*.

You can now verify the classification comparing the selected fitted to the original displacement using the generated meta files and the time series analysis tools.
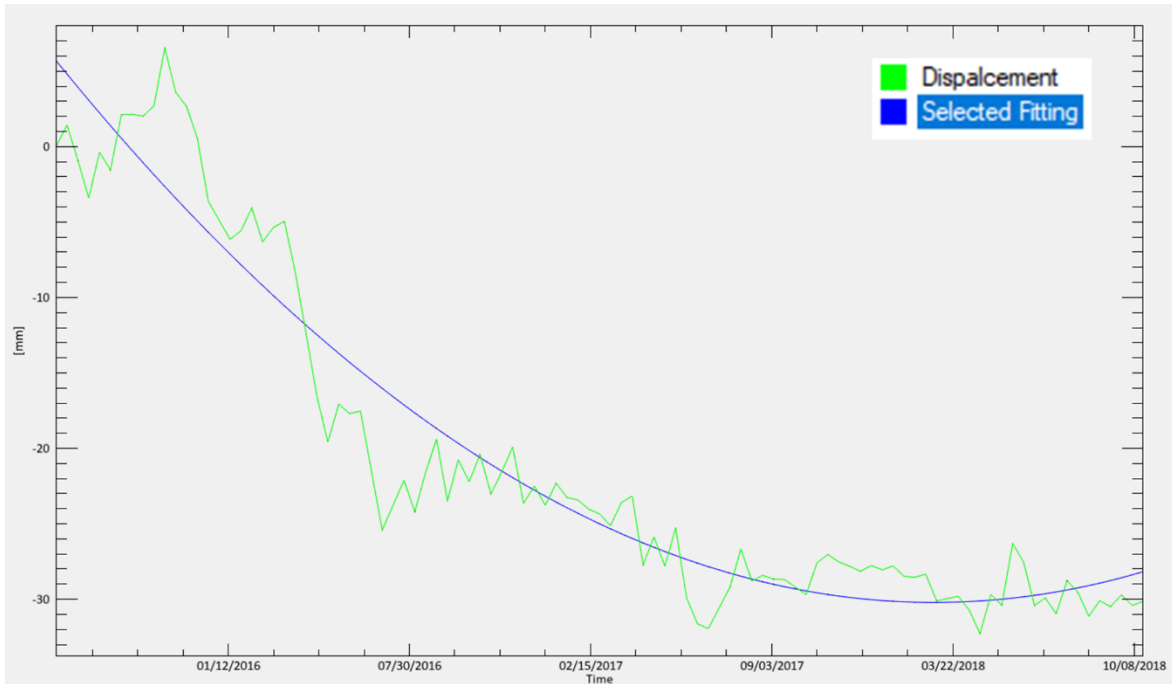
Some examples

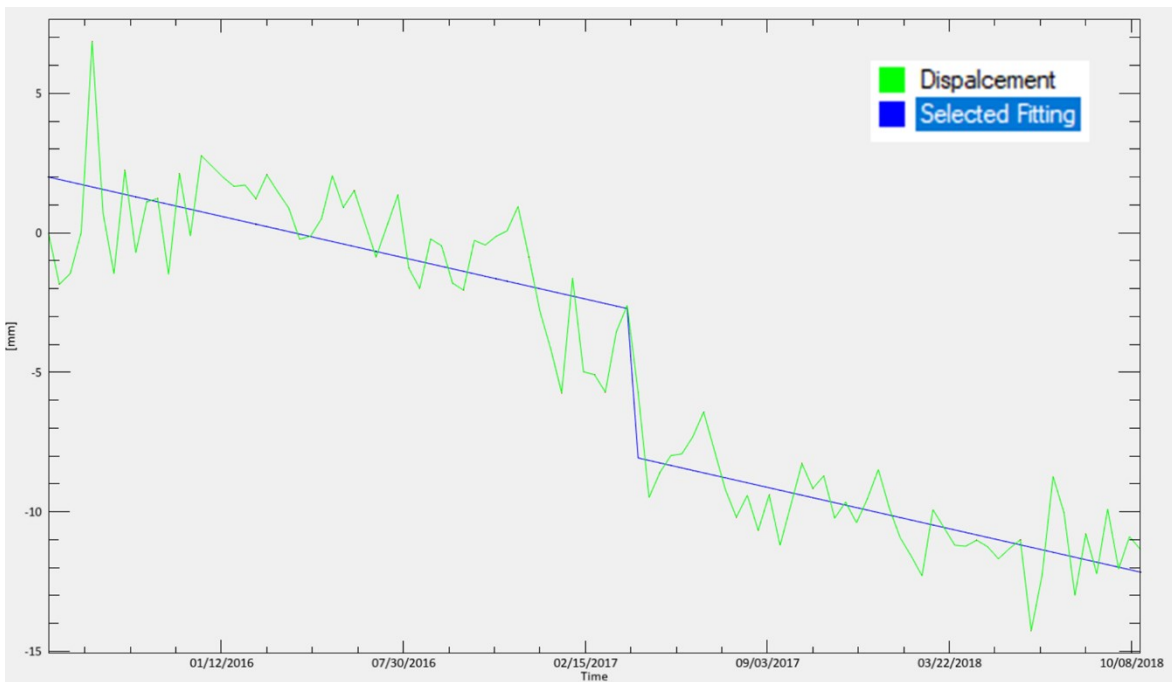Figure 14 Decelerated displacement in zone 1 classified as quadratic.

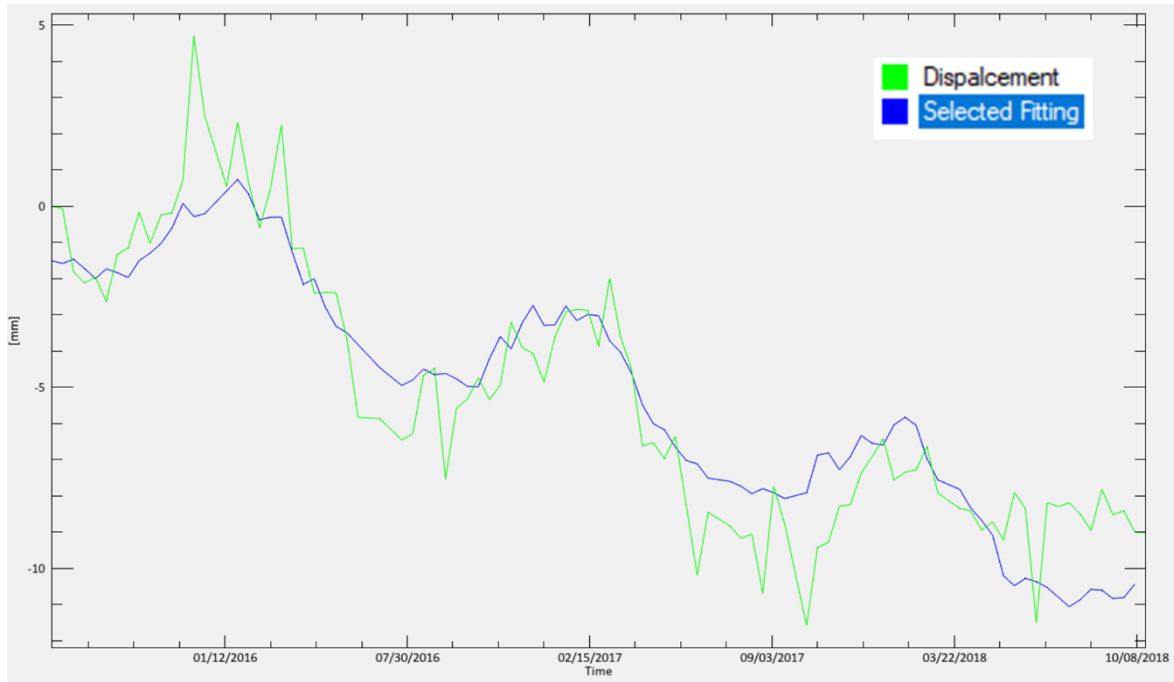Figure 15 A change of trend classified as Heaviside in zone 2.

Figure 16 An oscillated area classification as temperature in zone 3. Considering that the delay in this case is much bigger than the buildings examples of the previous section, we could suspect that the triggering element may be rain instead of temperature as they have the same periodicity but different delay (or phase). Its left for the student adding the rain model and limiting the possible delays.
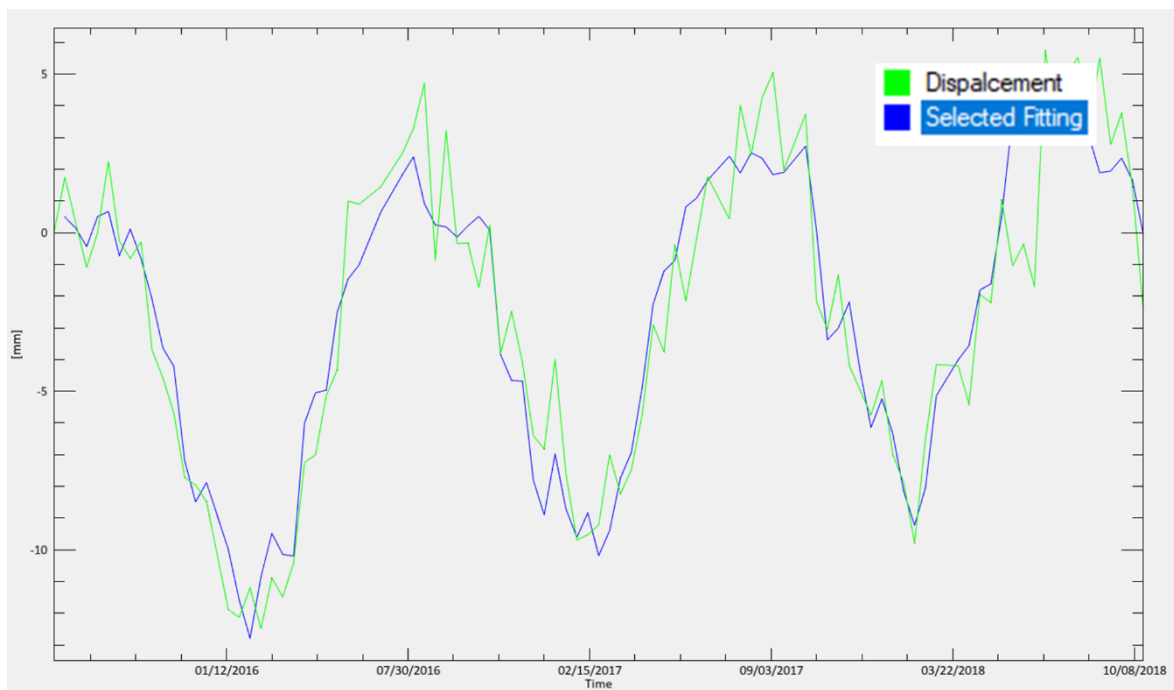

Figure 17 The buildings from the previous section classified as temperature triggered displacement. Zone 1.

Clustering

Another approach for classifying deformation time series without any external information is based on DB Scan Clustering algorithm. Compare with the previous examples, this technique has the advantage of differentiating different kinds of movements even if both have the same behaviours (linear, quadratic, etc) as long as they have different evolutions. For example, if two areas coupling to the temperature but different delay and the clustering metric is set to Euclidean, they will not match. The disadvantage of this classification is that is strongly dependant on user defined thresholds so some trial-and-error procedure may be needed to achieve good results.

This example is going to be performed with vectorial data. To avoid nonmatching time series due to noise it is strongly recommended to filter the data as shown in the filtering section prior to this step. In this example, our selection is the median filter.

Open the panel found at *Interferometric Stacking/Post Processing Tools/Time Series Classification/Clustering*.

Input File List

Single displacement meta file (_meta) or multiple displacement vector files (.shp). Both decomposed and LOS meta files are supported.

Thematic Map

Optional external thematic map.

TS Distance Metric

Metric used to evaluate the similarity of the time series. Euclidean or Cosine available.

Maximum Distance

Distance threshold to consider two series similar expressed as the mean value for each series date after series normalization (if corresponds). Lower values may produce many small clusters while high values may generate few big clusters

Deformation Correlation Length

Maximum distance in meters between two points to be considered as neighbors in the TS comparison. This value should be in the order of the dimension of the deformation process present in the scene. Increasing this value will increase the complexity of the process as more points must be analyzed also allowing the creation of giant clusters with low connectivity.

L1 threshold

Threshold to consider stable points, those point are not considered for the classification.

Minimum Cluster Points

Minimum number of points on a cluster to consider it valid. Clusters with less elements will be mark as unclassified.

Min cluster connectivity

Number between 1 and 0 of the minimum mean path length of a cluster to consider it valid.

Normalize Time Series

Flag to normalize the time series prior to metric application.

Inter-Class Connectivity

Penalization factor for inter thematic region connections. Zero means no inter region connection, 1 no penalization.

Analyze time subset

By setting this flag only the period defined by **From** and **To** will be analyzed.

Output Root Name

Name of the output root. This file is mandatory.

• *Output Root Name_ TS Distance Metric_***clustering:** Classification map in raster or vectorial format (depending on the input type). Class 0 corresponds to stable points (under the L1 threshold). Class 1 correspond to non-classified points.

• *Output Root Name***.dot:** Clustering graph in graphviz format which can be displayed using software such as Tulip.

• *Output Root Name_ TS Distance Metric***.shp:** Classification vector map. Class 0 corresponds to stable points (under the L1 threshold). Class 1 correspond to non-classified points.

• *Output Root Name***.dot:** Clustering graph in graphviz format which can be displayed using software such as Tulip.
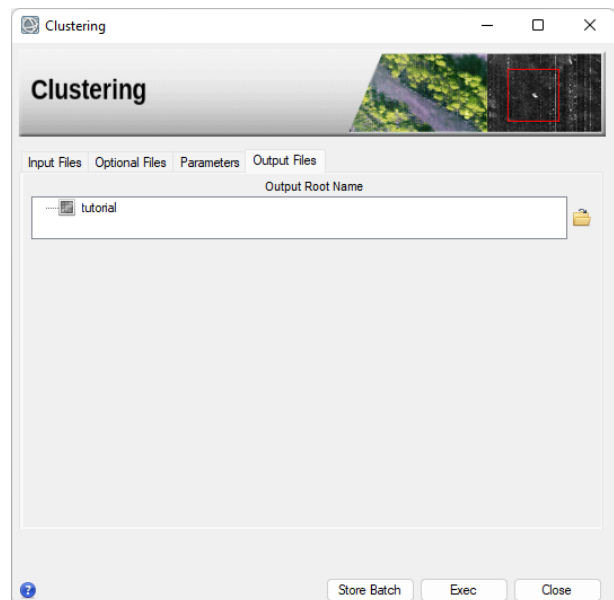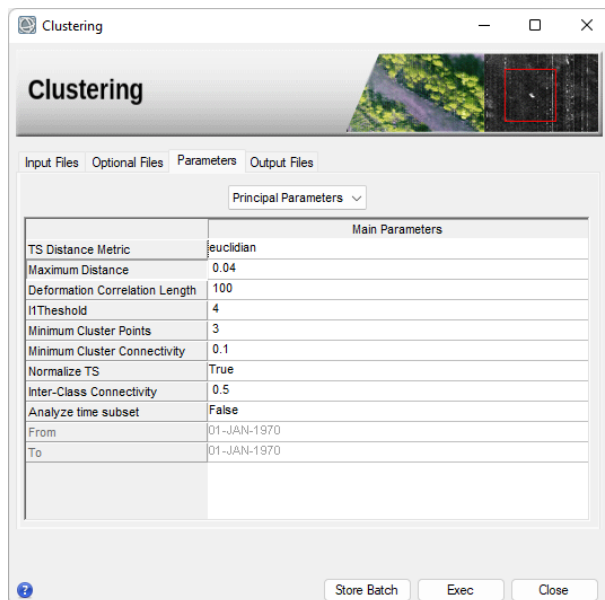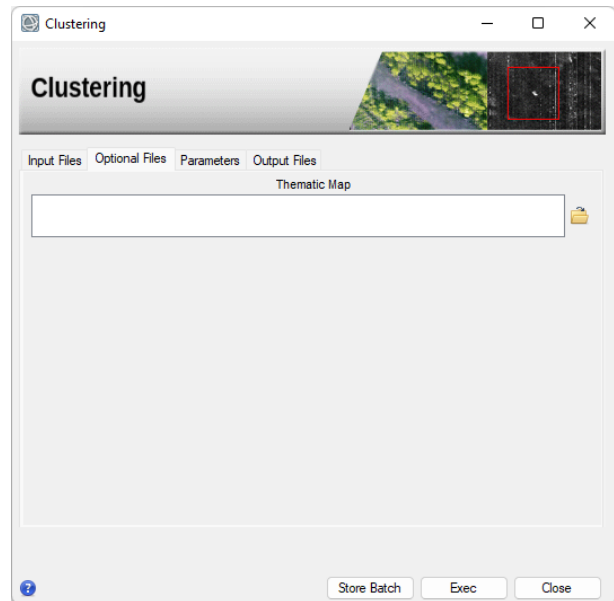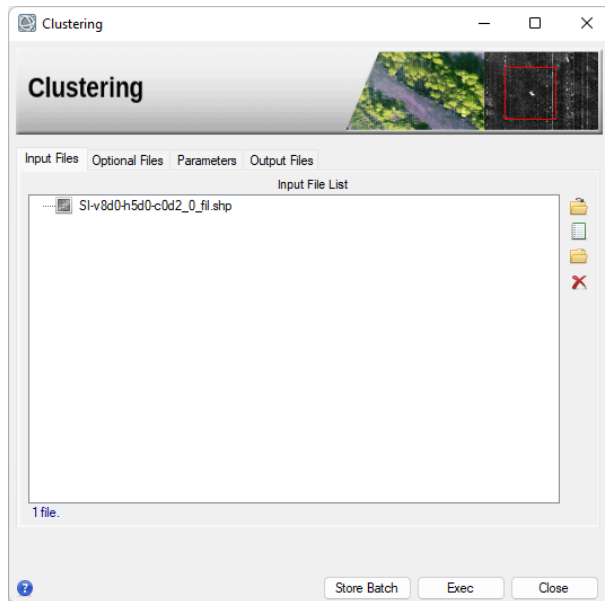
Open the resulting shp file and the Vector Time Series Analyser. Select the dbf attribute to display bay going to *Options / Change Attribute Field* from the *TS Vector Analyzer* and select ClusterId.
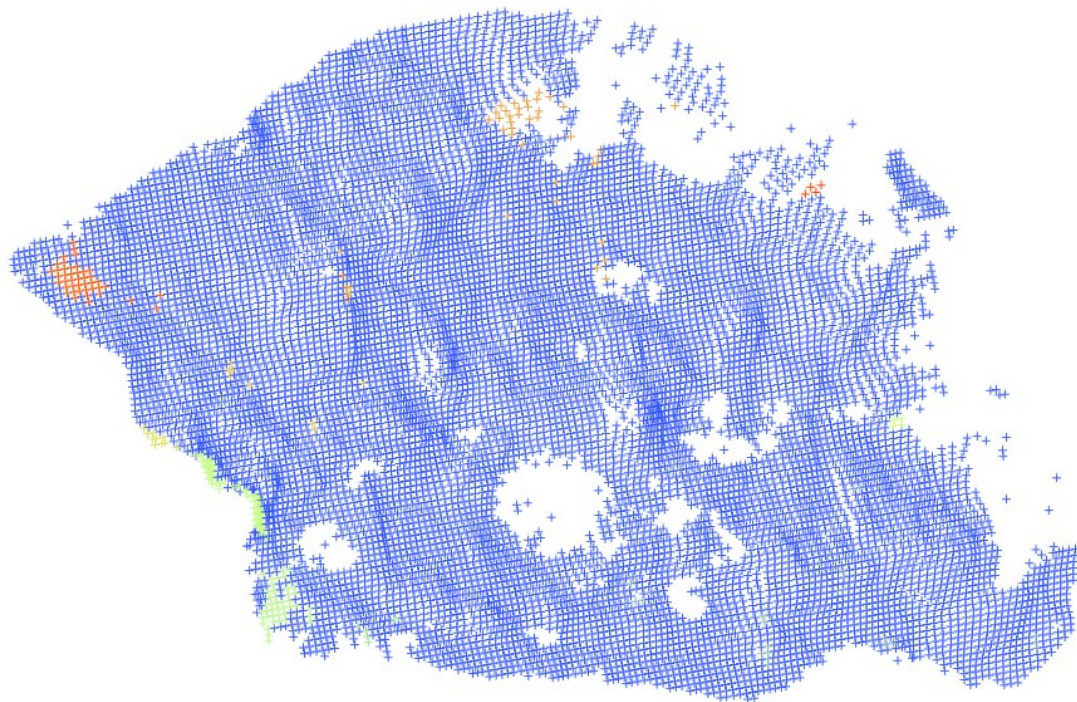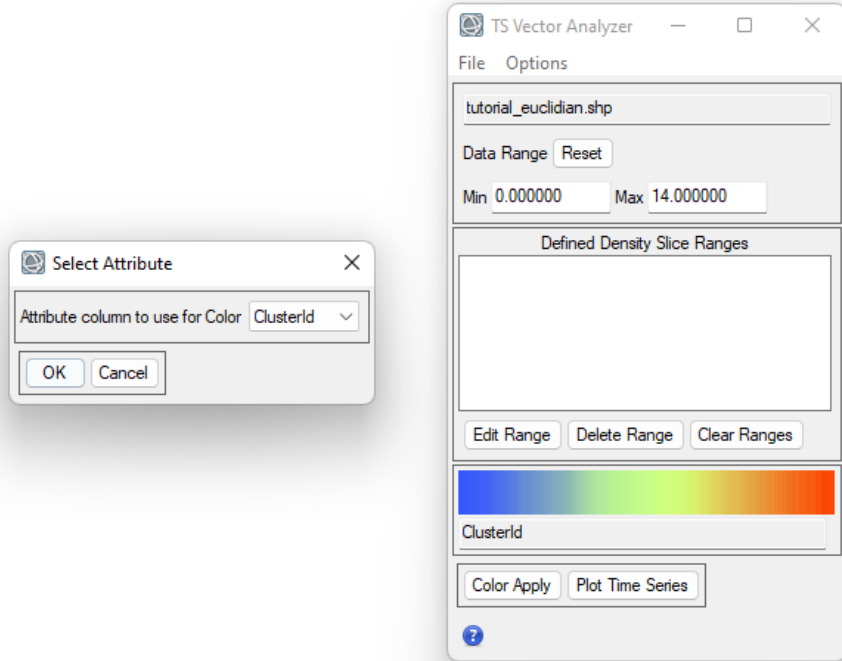
Figure 18 In blue stable regions (under L1 threshold), light blue unclassified areas. In other colors, the resulting clusters

Cluster IDs 0 represent stable areas, 1 represent areas which could not be cauterized because they have no connections with other time series. IDs from 2 onwards represent the different clusters.