MATTCH

Machine Learning Methods for SAR-derived Time Series Trend Change Detection

ESA/AO/1-9101/17/I-NB

Final Report

Deliverable D100.6

Version	1.1
Date	21/04/2020
Edit by	Marco Bianchi Christine Bischof Alessio Rucci Marco Basilico Francesco Lattari
Approved by	Marco Bianchi

TABLE OF CONTENTS

APPLICABLE DOCUMENTS

[RD1]	ESA/AO/1-9101 Documentation
[RD2]	MATTCH Proposal
[RD3]	MATTCH Contract

APPLICABLE DOCUMENTS

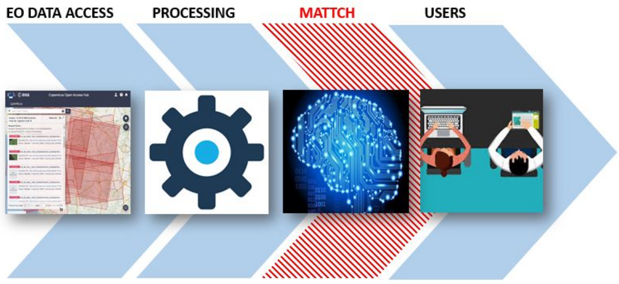
21/04/2020	Future Developments paragraph removed from this document. Ah hoc D100.7 created.

1. INTRODUCTION

This report summarizes the project as a whole, describing the scope, activity execution, the achieved results, and possible deviations from the baseline, comparison with planned objectives, and other information required by the contract.

2. SCOPE

The MATTCH project - Machine Learning methods for SAR-derived Time Series Trend Change Detection - aims to apply Machine Learning techniques to InSAR (Interferometric Synthetic Aperture Radar) derived surface deformation measurements, with the goal of identifying, among the huge number of measurement points (MP) identified by advanced InSAR algorithms, the ones exhibiting displacement time series characterized by a change in trend or, more generally, an “anomalous behavior”. This data screening step is extremely important to support the End Users Community in the exploitation of frequently updated (every few days) and highly populated (millions of MPs) information layers resulting from advanced InSAR analyses over large areas.



The main open problem is now the **data screening phase**: how to find, in this ocean of data, the information the final users are looking for?

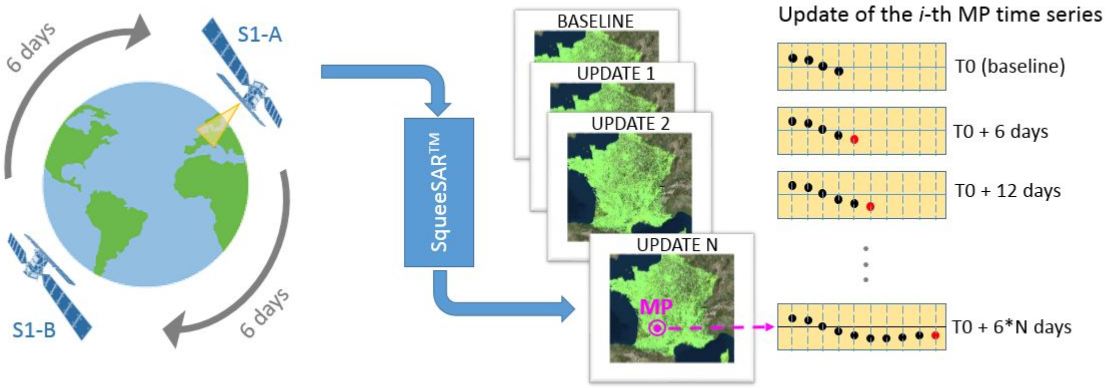


Figure 1: Schematisation of a SqueeSAR™ monitoring project using Sentinel-1 data. After a baseline processing at time T0, using all the images available from the archive, a new image is acquired every 6 days, triggering an update of the SqueeSAR™ results. The updated deformation data - each time consisting of hundreds of thousands/millions of measurement points (MP) - are delivered to the End User. Each MP is described by a displacement time series, with a new sample after each update (red dots in the diagrams on the right).

3. ACTIVITIES REVIEW

In this chapter, a recap of the activities in each WP is done.

3.1. WP100 - PROJECT MANAGEMENT

Code	WP100	**Company**	TRE ALTAMIRA	**Responsible**	Marco Bianchi
Start	KO	**End**	KO + 12 M		
The main purpose of the WP is to ensure reporting and coordination of the project, quality assurance, and risk management, through well-defined project management techniques. The goal will be pursued providing effective guidance to the project's activities, taking care of effective communication between all the participants, issuing appropriate documentation, and managing accounting duties during the entire project's lifecycle. In any case, where a discrepancy should rise between the project management activities here described and the activities deemed as necessary by the tender, the latter will prevail.					

This paragraph intends to describe the execution of the project in terms of fulfillment of the schedule.

Schedule

The MATTCH activities started with the KO meeting on **April 1st, 2019**.
Being the project duration equal to 12 months, the official activities closure date is **March 31st, 2020**.

The planned execution plan has been respected, with two exceptions:

- WP300: one additional month required, as agreed with the ESA Project Officer during the Mid-Term meeting on January 10th, 2020 (see minutes).
- Activity closure dates, one additional week required in order to cope with a slight delay in the information exchange with the User (UNIFI)

The bar chart of the project execution is therefore as follows:

Meetings

The following tables summarizes the list of meeting that took place during the projects execution:

Meeting	**Date**	**Premises**
KO	01/04/2019	Conference Call
Intermediate (MS1)	03/10/2019	Conference Call
Progress	03/01/2020	Conference Call
Final (MS2)	21/04/2020	Conference Call

Reporting

For the whole duration of the project, the following reporting has been provided to the ESA Project Officer:

KO Meeting	Minutes of the Meeting
Intermediate Meeting	Minutes of the Meeting
Progress Meeting	Minutes of the Meeting
Final Meeting	To be executed
Monthly update	Activity Reports

3.2. WP200 - REQUIREMENTS

Code	WP200	**Company**	TRE ALTAMIRA	**Responsible**	Christine Bischoff
Start	KO	**End**	KO + 3 M		
The main purposes of the WP is to gather requirements from the users and to assess the state of the art of machine learning applied to time series analysis. Requirements gathered from users will help the consortium in the implementation of algorithms to allow the final users to get maximum benefit from DInSAR products and services. The state of the art about machine learning applied to time series analysis will help the consortium to capitalize on past research, and to identify the more suitable approaches to be applied to DInSAR data.					

User Requirements

Thanks to a focus on the on-going TOSCANA project (see D200.1 Paragraph 2 for details), requirements have been collected through an interaction with the University of Florence, Earth Science department.

Code	Name	Description
UR1	EFFECTIVE	An effective way to retrieve relevant information, namely changes in the trend of time series, from the millions of measurement points provided at each update
UR2	AUTOMATIC	The ‘trend change’ layer has to be created automatically to ensure the timely delivery of this information
UR3	TUNABLE	The trend change detection needs to be calibrated according to the end users needs, i.e. according to the magnitude or time frame of the trend change
UR4	GIS-READY	The trend change information needs to be delivered in a practical format, such as a GIS layer

Suggestions for further developments are also provided:

False positives/ negatives

Occasionally, points affected by seasonal deformation are highlighted by the trend change algorithm, giving false positives – this could be improved. The occurrence of false negatives is difficult to confirm since so far, there has been no site where ground monitoring and DInSAR are clearly in disagreement.

Characterization of the trend change

It would be useful to have a clearer indication of the time period over which the change in trend occurred, for example, whether it was an abrupt step or a slower change. Visualizing the date at the center of time period affected by the trend change would be useful. A better characterization of the trend change would also include a more detailed analysis of the spatial dimension of a ground deformation anomaly detected in the DInSAR data. Furthermore, it could be useful to provide confidence in the occurrence of a trend change.

Automatically linking trend changes to the most likely cause

Ideally, a detected ‘anomalous’ deformation pattern, both in time and space, could be checked against a list of known ‘typical’ deformation patterns identified for phenomena such as different landslide types. This serves as complementary information to identify and assess the ground deformation phenomena that are causing the anomaly.

Past Experience

Several initiatives (funded projects, workshop, conferences) triggered and funded by ESA and other public Entities have run in the past years aiming at making the EO community closer to the Users and Stakeholders needs. Many of these projects are based on the collection of Requirements by the Users, collected in order to make the EO community efforts effective and well addressed.
The goal of the usability of EO derived data is once again stressed as critical.

Santorini Report

Considering the topic exploited in the MATTCH projects, we consider as relevant all those comments aimed at recommending the full exploitation of the whole EO data capacity, with the highest update rate possible:

- Regularly updated landslide maps (susceptibility, hazard and risk maps)and landslide inventories, including location, type, area, volume, intensity, state and style of activity of observed phenomena; updated distribution of landslide-affected areas to help understanding of ongoing and future instability.
- Long-term monitoring of areas at higher risk, with regular and consistent observations, to improve understanding of landslide kinematics and facilitate assessment of their future evolution; site-specific information on the instability conditions is needed to associate the identified motions with causative factors and triggers, and to analyse zones with different susceptibility to landslides;
- Post-event motion and damage assessment, mapping of affected areas and identification of safe zones for relocation of assets at risk; residual hazard and risk zonation.
- Landslide vulnerability assessment and modelling; forecasting and early warning.

The Geo-Hazard Exploitation Platform

Among the requirements for the execution of the GEP (Geohazard Exploitation Platform):

- Support provision of systematic PSI maps using repeat Sentinel-1 observations
- Further expand the tools and a portal for data discovery and to visualize results
- Support dissemination of available results, historical INSAR or PSinSAR products including results generated using the TEP and results provided by the user community.

The EU-GMS Initiative (EEA)

The EU-GMS is expected to provide motion products updated every 12 months or incrementally, i.e. by processing new images after each new acquisition over the whole European territory.
This is to stress how the whole community is heading to a massive exploitation of the Sentinel-1 data: millions of measurement points, updated at least every 12 Months, to be distributed to the stakeholders.
We foresee a relevant application of the MATTCH methodology, meeting the need of an additional layer for data usability.

Machine Learning Knowledge Base

As reported in D200.2, the collected bibliography demonstrates how recurrent neural networks are able to model the temporal correlation among samples in time series. In specific, these kinds of architectures have been successfully employed to detect anomalous behaviours during the evolution of particular phenomena. We consider the methods based on recurrent networks, in particular LSTM networks, good candidates for detecting trend changes in InSAR-derived time series.
Similarly to what is proposed in literature, the idea is to use a recurrent model to detect whether an observed measurement point corresponds to a transition point between two trends (change point) or not. However, differently from the time series observed in literature, InSAR sequences are characterized by a short temporal extent with a variable noise distribution, which make the application of unsupervised techniques even more difficult. Additionally, in InSAR time series trend can change multiple times in the same sequence while the examples observed in the bibliography are characterized by large periods of normal behaviour, which are used to train the network. The learned model is then evaluated on subsequent periods to detect anomalous patterns.
In conclusion, we think that it is worth investigating both unsupervised and supervised techniques to solve the considered change point detection task. However, we expect both the approaches to require much more effort to be successfully applied in the case of InSAR time series. We further expect supervised approaches to have better performance with respect to unsupervised ones, due to the difficulty of the latter to generalize in the case of time series with variable length, variable noise distribution and multiple trend changes.

3.3. WP300 - MACHINE LEARNING, APPROACH DEVELOPMENT

<table>
<tr><td>Code</td><td>WP300</td><td>Company</td><td>POLIMI</td><td>Responsible</td><td>Matteo Matteucci</td></tr>
<tr><td>Start</td><td>KO + 2 M</td><td>End</td><td>KO +10M (1 extra)</td><td></td><td></td></tr>
<tr><td colspan="6">The main objectives of the WP are:

● Choosing and selecting suitable machine learning approaches and algorithms, with a particular emphasis to deep learning ones, to be applied to time series from DInSAR data.
● Develop, implement, and train the machine learning algorithms coming from the selected approaches on selected datasets.
● Test the performance of the resulting algorithms on real data.</td></tr>
</table>

Approach and Algorithm Description Report (D300.1)

The collected bibliography about state-of-the-art methods for the analysis of time series allowed to identify a set of data-driven approaches extendable to the considered task of detecting trend changes in InSAR-derived time series. The family of considered methods lies in the field of Deep Learning (DL) approaches based on recurrent models known as Recurrent Neural Networks (RNNs).
The designed algorithms are based on Long Short-Term Memory (LSTM) networks, which are recurrent architectures composed by LSTM units. Each unit employs two main components: the memory and a set of learnable gates which allow to update the content of the memory.

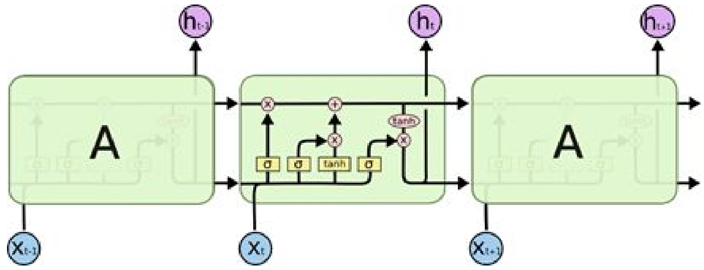


Figure 2. Long Short-Term Memory (LSTM) cell.

Multiple LSTM layers are stacked one on top of the other and the output of the previous recurrent layer is given as input to the following one. This allows capturing lower and high-level features describing the original input sequence. The features extracted by the last LSTM layer are finally given as input to an additional fully connected layer which provides the final prediction.

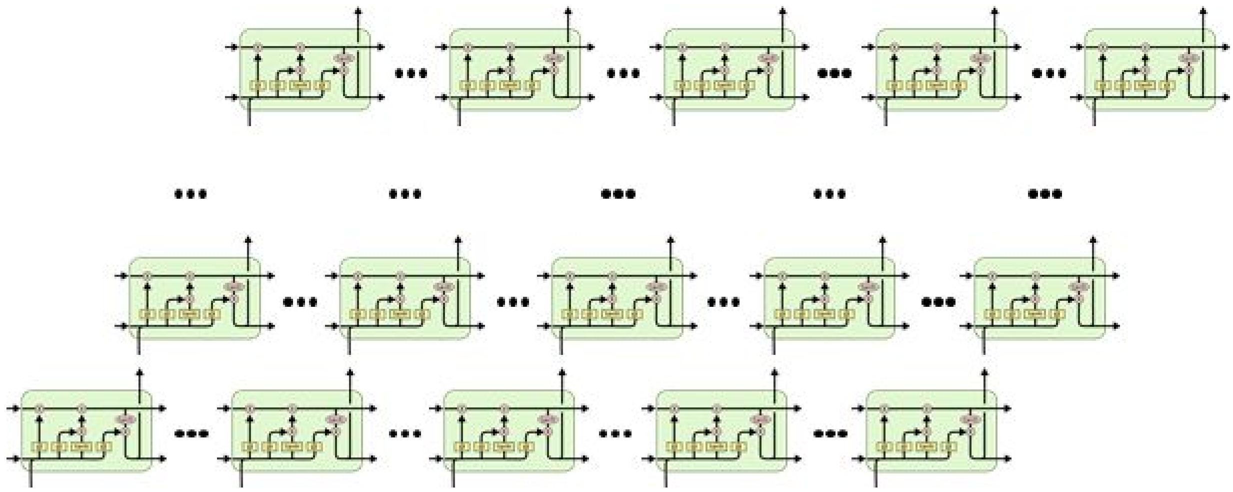


Figure 3. Stacked LSTM layers. The output of previous layers are fed in input to the following ones.

Specific approaches are considered analyzing the issues of the Unsupervised Learning, Supervised Learning and Non-Uniform Time Series Sampling.

Training Sets Report (D300.2)

The data collected for training the proposed models differs between two consecutive phases of the project. In particular, we can divide the conducted experiments in: training on real time series and training on simulated time series. The details of the respective training dataset are reported in D300.2.

The change points can be classified into two main classes based on the correspondent trend change:

Real data - Step change point

In correspondence of this kind of change point no velocity change is observable between the two trends but there is an abrupt displacement.

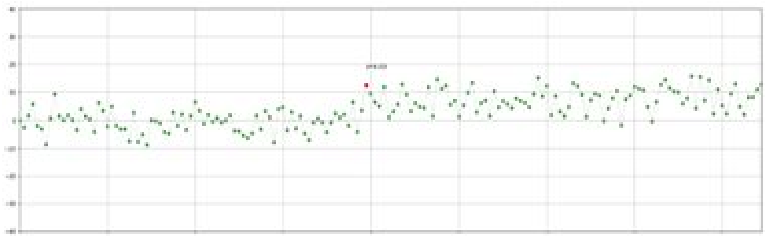


Figure 4. Example of 'step' change point (red dot).

Real data - Velocity change point

In correspondence of this kind of change point no abrupt displacement is observable but there is a variation in the velocity.

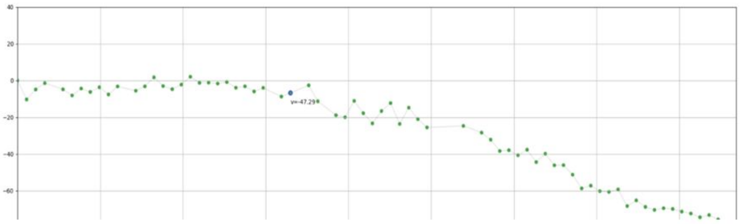


Figure 5. Example of 'velocity' change point (blue dot).

Real data - Step+Velocity change points

In correspondence of this kind of change point both an abrupt displacement and a change in the velocity are observable.

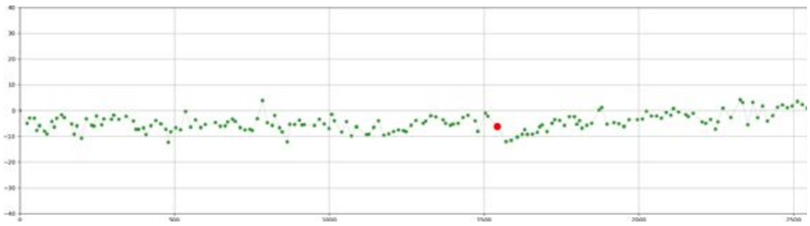


Figure 6. Example of 'step+velocity' change point (red dot).

Dataset of simulated time series

The goal of the simulated dataset is to improve the generalization capabilities of the proposed models. Previous training data, composed by real time series, gave the opportunity to explore and test different models and to understand the problems of the considered task. However, some problems related to the training data were encountered during the tests:

- targets for training were given by the statistical algorithm of the company. This involves overfitting on the reference algorithm by limiting the capability of the deep learning method of improving the detection accuracy.

- strictly related to the one highlighted above is the problem of the absence of reliable targets. In particular, given that the targets are provided by another algorithm, they were not suitable for a fair comparison between different methods.

- even if the number of collected real data was very large, the learned models had some difficulties to generalize on those test time series with a dynamic very different from the one observed during training.

Report on Algorithms Training (D300.3)

In order to train and evaluate the proposed model a large suite of experiments has been designed. The implementation of both the training and evaluation stages is based on Python programming language and the learning process has been accomplished through the PyTorch deep learning framework, which allows to pre-process data, run the computation on GPU, optimize model parameters through automatic differentiation, evaluate results with quantitative metrics, and visualized learning with the support of the Tensorboard interface.

All the experiments have been performed using Adam optimization algorithm with a learning rate LR = 0.001 and a batch size of 128 time series, i.e., the number of samples used for the optimization at each iteration. As already stated, we started by considering directly a supervised learning strategy with a bidirectional model.

In order to evaluate the generalization performance of each model during training, we created 4 different validation datasets, each one of 12800 time series: one composed by only normal sequences, one composed by time series with only step change points, one composed by time series with only velocity change points, one composed by time series with all kinds of change points. The metrics used in the evaluation are:

- *True Positives (TP)*: number of correctly predicted change points
- *False Positives (FP)*: number of wrong detections
- *False Negatives (FN)*: number of missed detections
- *Precision*: ratio of the correctly predicted change points to the total of predicted change points
- *Recall:* ratio of the correctly predicted change points to the total of actual change points
- *F1 score*: weighted average of precision and recall

In D300.3, the history of the experiments, which can be divided into simulated time series with uniform sampling rate and simulated time series with non-uniform sampling rate, are described in detail.

Report on Algorithms Performances (D300.4)

In this task the performances achieved during tests are reported. The document is divided into two parts based on the kind of data used for training: real or simulated time series.

Real data time series experiments are also divided into Unsupervised and Supervised learning tests.

Some details about the computation time evaluated on a given test dataset of 631616 time series.

- The proposed deep learning approach takes about **15 minutes** to process all the time series on a single GeForce GTX 1080 Ti NVIDIA's GPU.
- The baseline statistical approach takes about **3 hours and 15 minutes** for processing on a machine with 2 x Xeon 8-Core E5-2640v3 2.6 Ghz 25MB.

3.4. WP400 - SOLUTION ENGINEERING AND VALIDATION

Code	WP400	**Company**	TRE ALTAMIRA	**Responsible**	Alessio Rucci
Start	KO + 5 M	**End**	KO + 12 M		
The main purpose of the WP is the engineering and integration of the algorithmic approaches selected in previous work packages into a processing chain, ready to be used in a production environment.					

As per the proposal document [RD2], WP400 activities are grouped into 3 tasks:

- ICT Solution Implementation
- Algorithm Engineering
- Algorithm Validation and Fine Tuning

ICT SOLUTION IMPLEMENTATION

ICT solution selection

Since the whole processing chain in TREA is moving towards cloud-based solutions, it has been decided to assess cloud services options for the purposes of the MATTCH project.
This action must be considered in the framework of the on-going activities for the use of a processing chain in the cloud, therefore the constraints in the selection of the proper cloud service is not only due to the requirements linked to MATTCH.
While the WP300 activities are running, an investigation on Machine Learning algorithms on Amazon's cloud, exploiting the AWS's P3 instances family, is performed.
In order to exploit the potentialities of the GPU-based instances the solution is implemented adopting the AWS optimized AMI for GPUs, Machine Learning optimized images that includes the latest NVIDIA GPU-acceleration through pre-configured CUDA and cuDNN drivers (more details about AMI in D400.1, paragraph 2.1).

Prototype integration into the SqueeSAR chain

The MATTCH prototype is expected to run on a large scale project, as the Toscana Region project actually is.
To this goal, the ICT implementation must consider the data volume.
The first step is the execution of the MATTCH prototype on local CPUs: in this case, no GPUs Graphical Processing Units) are available and therefore no enhancement in the computational time is expected.
A second step, two possible approaches are considered to apply the MATTCH prototype algorithm on SqueeSAR data processed on Amazon Cloud: an AWS BATCH cloud-managed cluster has been used.

ALGORITHM ENGINEERING

In order to reach a production-ready environment, some actions are taken about the code optimization, the threshold set-up, the output data format.

For continuous monitoring projects, such as the one over the Tuscany region, we did identified as output of the TS-change algorithms two additional fields of information to add to the usual shapefile which are: estimated step estimated velocity variations occurred in the last 150 days for each measurement point.

- variation in average displacement rate ("DVEL" field in the out shapefile);
- abrupt variations in the average value of the time-series ("DSTEP").

This impacted with some modification to the standard SqueeSAR results format (D400.1, Annex B).
After the first delivery, thanks to the User's, a tuning on the algorithm parameters has been performed in order to reduce false positives.

ALGORITHM VALIDATION AND FINE-TUNING

The TOSCANA project

TRE ALTAMIRA has a running project project with University of Florence - Earth Science Department to monitor the territory of Regione Toscana using Sentinel-1 data on both ascending and descending geometries.

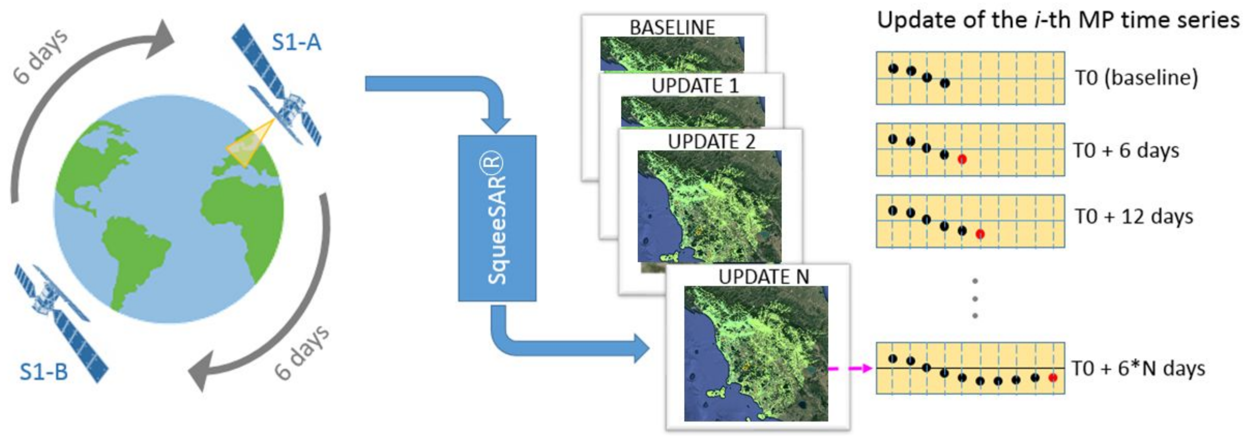


Figure 7: Schematisation of a SqueeSAR® monitoring project using Sentinel-1 data over Tuscany.

The Sentinel-1 datasets

The Toscana projects relies on the analysis of a number of swaths available over the territory, as represented by the following table:

Geometry	Track (relative orbit)	Deliverable labels
Ascending	15	ELBA TOSCANA OVEST
Ascending	117	TOSCANA EST GIGLIO
Descending	168	ELBA GIGLIO OVEST
Descending	95	EST

Table 1: Sentinel 1 swath used in the Toscana project

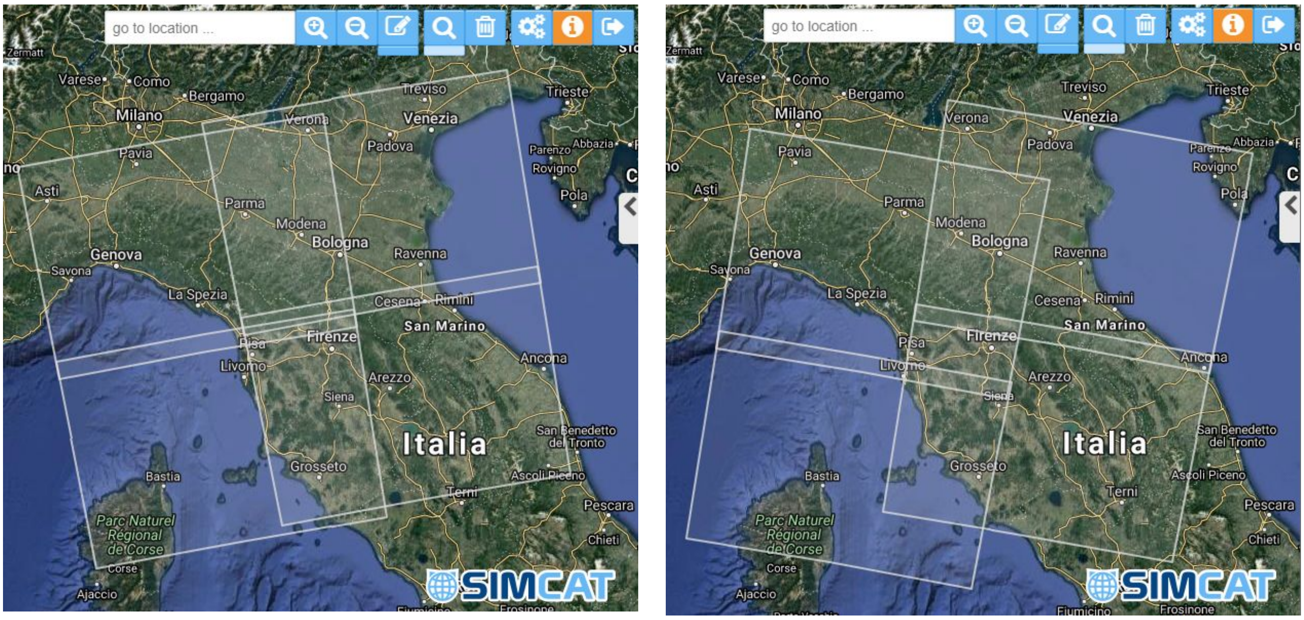


Figure 8: Sentinel-1 coverage over Tuscany, Ascending (left) and Descending swaths.
SIMCAT is the TRE ALTAMIRA tool to browse all the satellite catalogs at once.

MATTCH delivery 1

After the conclusion of activities of WP300 and the first two tasks of WP400, the main target of the project is now to apply the MATTCH prototype, integrated in the TRE ALTAMIRA production environment, to the project delivery of the TOSCANA project.

This took place in parallel with the “standard” delivery on 03/03/2020, whene the following products have been generated using the standard processing chain (SqueeSAR® + Statistical approach) and - along with them - the MATTCH results have been produced and delivered too to the User (UNIFI) using the TREmaps® web interface (standard mean of delivery for all projects in TRE ALTAMIRA).

Results are visible on TREmaps® using the access credential provided to ESA.

Results labels and updates date are reflected in the following table:

File	Group Label	Track	Geom.
TOSCANA_ELBA_SNT_IW_A_T15_REG_TOSC_SC81364_81364_300 TOSCANA_ELBA_SNT_IW_A_T15_REG_TOSC_SC81364_81364_300-trend_variation	TOSCANA ELBA 2020-02-25	15	A
TOSCANA_ELBA_SNT_IW_D_T168_REG_TOSC_SC81362_81362_296 TOSCANA_ELBA_SNT_IW_D_T168_REG_TOSC_SC81362_81362_296-trend_variation	TOSCANA ELBA 2020-02-24	168	D
TOSCANA_EST_SNT_IW_A_T117_REG_TOSC_SC81164_81164_315 TOSCANA_EST_SNT_IW_A_T117_REG_TOSC_SC81164_81164_315-trend_variation	TOSCANA EST 2020-02-20	117	A
TOSCANA_EST_SNT_IW_D_T95_REG_TOSC_SC81167_81167_315 TOSCANA_EST_SNT_IW_D_T95_REG_TOSC_SC81167_81167_315-trend_variation	TOSCANA EST 2020-02-19	95	D
TOSCANA_GIGLIO_SNT_IW_A_T117_REG_TOSC_SC81365_81365_299 TOSCANA_GIGLIO_SNT_IW_A_T117_REG_TOSC_SC81365_81365_299-trend_variation	TOSCANA GIGLIO 2020-02-20	117	A
TOSCANA_GIGLIO_SNT_IW_D_T168_REG_TOSC_SC81366_81366_296 TOSCANA_GIGLIO_SNT_IW_D_T168_REG_TOSC_SC81366_81366_296-trend_variation	TOSCANA GIGLIO 2020-02-24	168	A
TOSCANA_OVEST_SNT_IW_A_T15_REG_TOSC_SC81163_81163_311 TOSCANA_OVEST_SNT_IW_A_T15_REG_TOSC_SC81163_81163_311-trend_variation	TOSCANA OVEST 2020-02-25	15	A
TOSCANA_OVEST_SNT_IW_D_T168_REG_TOSC_SC81170_81170_304 TOSCANA_OVEST_SNT_IW_D_T168_REG_TOSC_SC81170_81170_304-trend_variation	TOSCANA OVEST 2020-02-24	168	D

SqueeSAR® results are provided in both geometries, and, along with them, hotspots identified with the MATTCH algorithm are delivered.

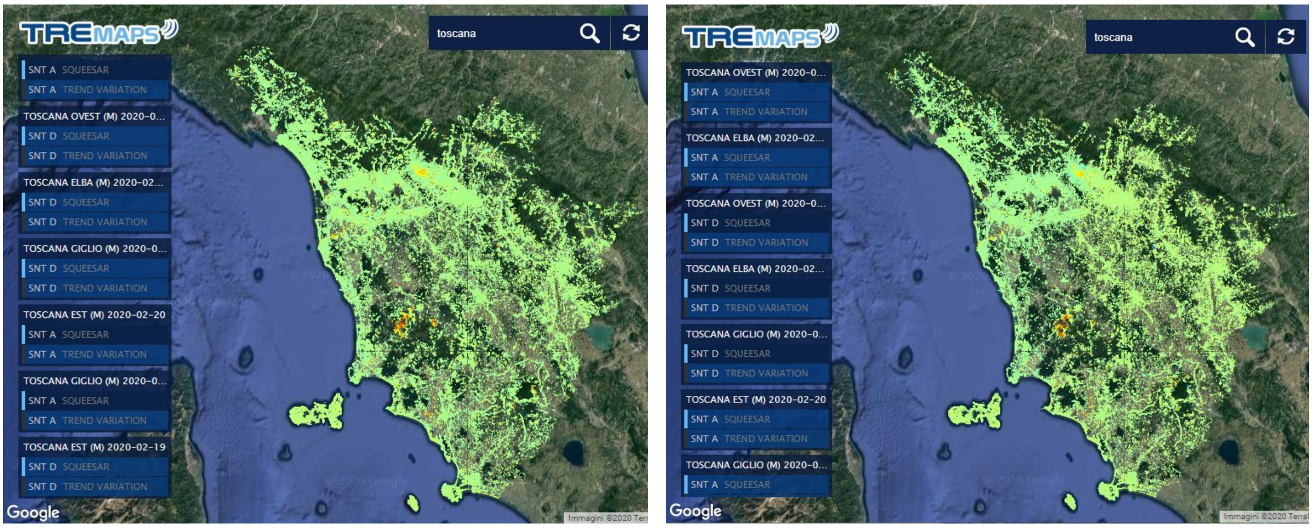


Figure 9: MATTCH results computed on the 03/03/2020 delivery of the TOSCANA project: ascending (right) and descending (left)

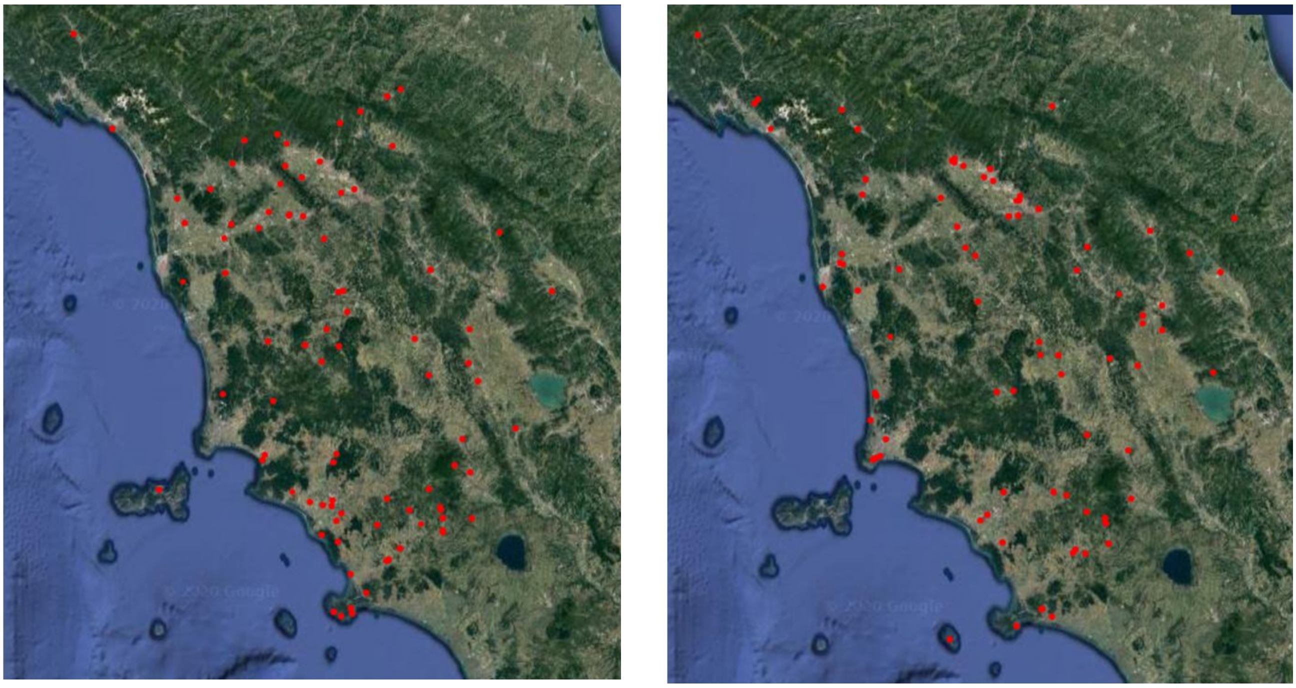


Figure 10: MATTCH results computed on the 03/03/2020 delivery of the TOSCANA project: ascending (right) and descending (left)

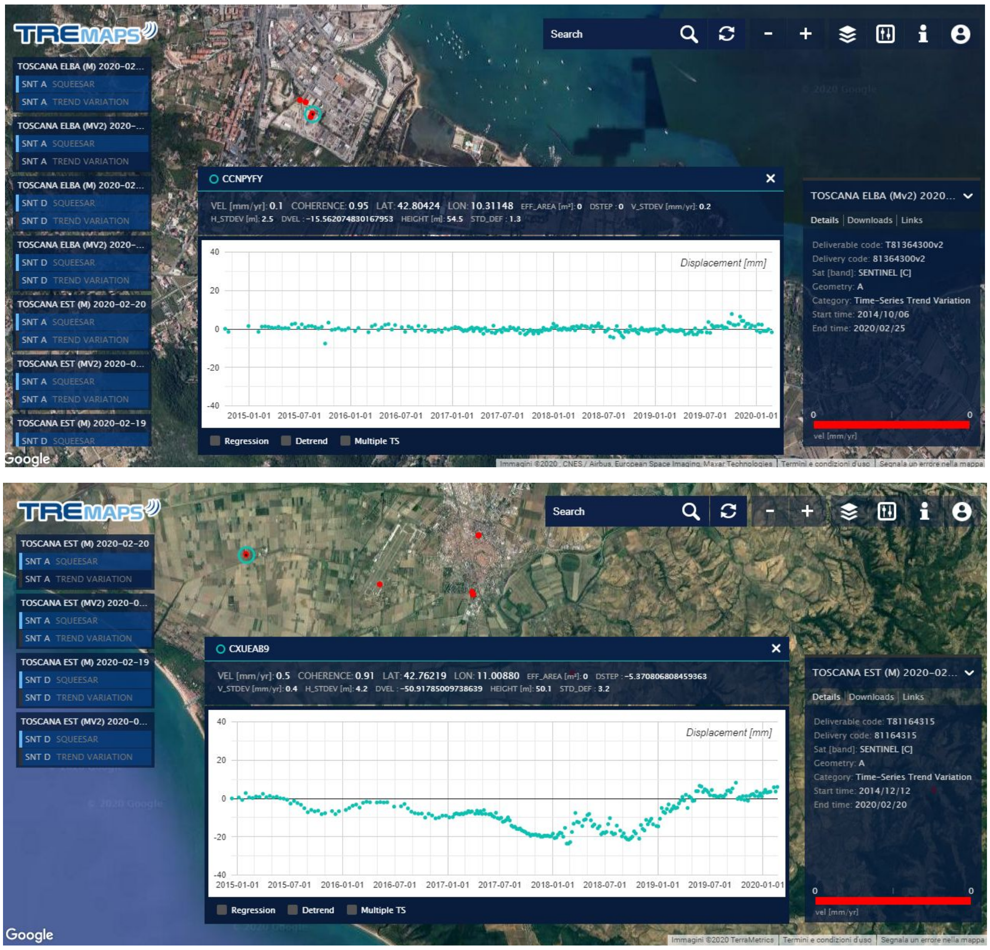


Figure 11: Examples of time series showing a trend change, delivered to UNIFI in the 03/03/2020 delivery of the Toscana project.

User's feedback

The User has provided a report after comparing the 03/03/2020 standard delivery against the MATTCH one. The full report is available in D400.1 - ANNEX A.

The User's comments can be summarized around 3 main pillars:

1. The MATTCH prototypes seems to be sensitive to time series showing in the very last part a coherent and well defined trend made of a few pixels: in this case, it is intended as a trend variation (see Figure A3 in the ANNEX A)
2. Temporal window of 150 days: in case of weak/small changes both the statistical and the MATTCH approach have difficulties in finding the right time sample where the change is located. When the change is close to the temporal window limitation, it is easy that one approach assigns the first sample inside the 150 days temporal window (hotspot detected) while the other approach assigns the first sample outside of it (hotspot not detected). See Figure A8 in the ANNEX A;
3. Seasonality: the statistical approach is applied to displacement time series after the de-trend of seasonal components. The MATTCH one operates on the displacement time series as they are (the training is made under this assumption, to avoid the need of a seasonal model that could bias the assessment (see D300.4). The MATTCH prototype seems to be more sensitive than the static one on seasonal time series. See Figure A7 in the Annex A.

Fine tuning and MATTCH delivery 2

As foreseen in the description of WP400, the Users' feedback referred to the first delivery have been assessed by POLIMI and TREA jointly, in order to evaluate possible correction to the scripts.

In order to test the outcome of the refined algorithm, the script has been applied to the same results over the Toscana region, producing overall a significant reduction of false-positive hotspots.

In summary:

Layer	Total MP	Statistic Trend Change	MATTCH version 1	MATTCH version 2
TOSCANA ELBA 2020-02-25 (A)	25046	0	6	6
TOSCANA ELBA 2020-02-24 (D)	23349	0	0	0
TOSCANA EST 2020-02-20 (A)	631616	80	248	171
TOSCANA EST 2020-02-19 (D)	390478	40	178	122
TOSCANA GIGLIO 2020-02-20 (A)	4438	0	0	0
TOSCANA GIGLIO 2020-02-24 (D)	4442	0	3	0
TOSCANA OVEST 2020-02-25 (A)	197386	9	19	0
TOSCANA OVEST 2020-02-24 (D)	308668	9	100	33

Table 4: Number of detected hotspots executing the MATTCH script - version 1 and the version 2 (refined after User's feedback)

The table shows a general reduction in the number of detected hotspots: this is mainly related to points showing a seasonal trend. This impacts most of the issues highlighted by the User.

3.5. WP500 - DISSEMINATION

Code	WP500	Company	POLIMI	Responsible	Francesco Lattari
Start	KO + 8 M	End	KO + 12 M		
The main purpose of the WP is to create a comprehensive dissemination strategy and exploitation plan; they will contain an overview of the overall strategy, target audiences, key messages as well as detailed communication and dissemination activities for the duration of the project.					

Conferences

Abstracts have been submitted to attend the following conferences.

In the proposal, a reference to the Living Planet Symposium 2019 (Milano) was made, but given the MATTCH project KO date (01/04/2019), it was not possible to produce any relevant results on time.

Event	Earth Observation Phi-Week
Venue	ESA-ESRIN Frascati (Rome)
Title	Recurrent Neural Networks for Trend Change Detection in InSAR Time Series
Date	9-13/09/2019
Session	AI4EO (6) - Wed 11/09/2019
Speaker	Francesco Lattari, Politecnico di Milano, Milano, Italy
Authors	Matteo Matteucci (1), Alessio Rucci (2), Christine Bischoff (2), Marco Basilico (2), Emanuele Passera (2) || (1) Politecnico di Milano, Milano, Italy, (2) TRE ALTAMIRA s.r.l., Milano, Italy

Event	ASAR 2019: A workshop on Synthetic Aperture Radar
Venue	Saint-Hubert, Quebec (Canada)
Title	Machine Learning Methods for SAR-derived Time Series Trend Change Detection
Date	Wed 02/10/2019
Speaker	Giacomo Falorni, TRE ALTAMIRA Inc., Vancouver BC (Canada)
Authors	Francesco Lattari (1), Emanuelle Passera (2), Alessio Rucci (2), Christine Bischoff (2), Marco Basilico (2), Andrea Bonarini (1), Matteo Matteucci (1) || (1) Politecnico di Milano, Milano, Italy, (2) TRE ALTAMIRA s.r.l., Milano, Italy

Event	FRINGE 2020
Venue	Delft
Title	Machine Learning methods for SAR-derived Time Series Trend Change Detection (MATTCH)
Date	CANCELLED[1]
Speaker	Francesco Lattari, Politecnico di Milano, Milano, Italy
Authors	Matteo Matteucci (1), Alessio Rucci (2), Christine Bischoff (2), Marco Bianchi (2) || (1) Politecnico di Milano, Milano, Italy, (2) TRE ALTAMIRA s.r.l., Milano, Italy

It is also planned to participate in the ESA EO Phi Week 2020, scheduled for the end of September at ESRIN in Frascati.

Publications

The goal of both POLIMI and TRE ALTAMIRA is to produce a publication in a peer-reviewed journal during 2020.

Website

A dedicated web page has been designed reporting the description of the main steps of the project.

The web page address is:

https://site.tre-altamira.com/mattch/

Social media

Posts on the projects results are being published using the LinedIn account of TRE ALTAMIRA

[1] (02/04/2020) Due to COVID-19 outbreak and the guidance from the relevant authorities, ESA and the Organising Committee of FRINGE 2020 has decided to **cancel** the workshop.

https://www.linkedin.com/company/tre-altamira/
https://twitter.com/tre_altamira

4. CONCLUSIONS

In order to summarize the main achievements of the project, we propose the following conclusion divided into conclusion about the algorithm implementation and conclusion about the User's Requirements description. Subsequently, a paragraph talking about future developments is provided.

4.1. THE ALGORITHM

Considering the state-of-the-art from literature (D200.2), the temporal dependencies between measurement points (MPs) provided by the SqueeSAR™ analysis make the Recurrent Neural Network architectures well suitable for analysing displacement time series. RNN represents the starting point.
Mono and bi-directional Multiple Long Short-Time Memories layers, both in supervised and unsupervised modes, have been implemented.

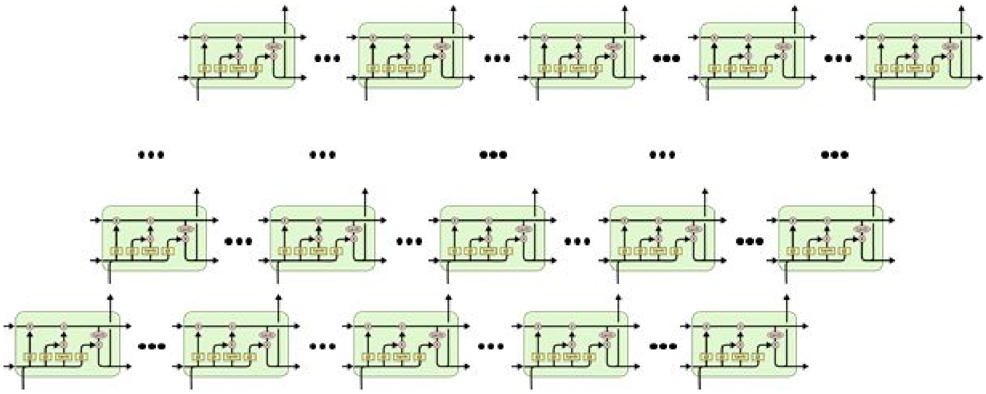


Figure 12: Stacked LSTM layers. The output of previous layers are fed in input to the following ones.

One peculiarity, in the InSAR domain, is the irregular sampling of the time series: due due missing acquisition, it may happen to have a non uniform temporal distance between two subsequent samples. Three methodologie are tested. Details are presented in D300.1.

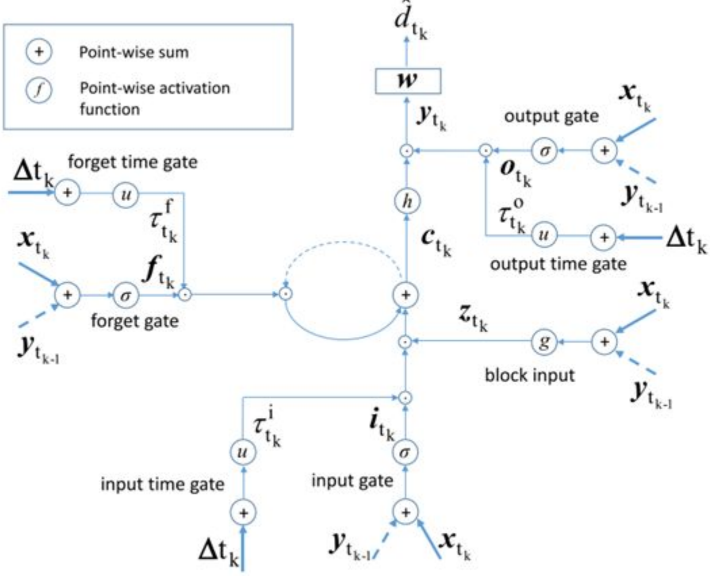


Figure 13. Time-Gate Long Short-Term Memory (TGLSTM) cell.

The implemented methodology relies on the choice of the proper training sets. Two families of experiments have been considered: real-data training sets and simulated data training sets. After an analysis of the results provided in the two cases, the simulated training sets have been chosen, to ensure more generalizability of the results (see D300.2).
As a conclusion, thanks to the experiments described in D300.3 and D300.2, the prototype has been implemented considering:

- TG (Time Gated) LSTM to cope with the non-uniform temporal sampling
- Use of simulated data training seat

In terms of algorithm performances, it also proved that the proposed solution is faster than the statistical approach.

Additionally, we report also some details about the computation time evaluated on a given test dataset of 631616 time series.

- The proposed deep learning approach takes about 15 minutes to process all the time series on a single GeForce GTX 1080 Ti NVIDIA's GPU.
- The baseline statistical approach takes about 3 hours and 15 minutes for processing on a machine with 2 x Xeon 8-Core E5-2640v3 2.6 Ghz 25MB.

The prototype has been engineered and included in the TRE ALTAMIRA processing environment (D400.1, paragraphs 2 and 3).

4.2. THE USER

Thanks to the involvement of the User it is possible to compare the requirements identified in D.200.1 against the comments received after the MATTCH results delivery in WP400.

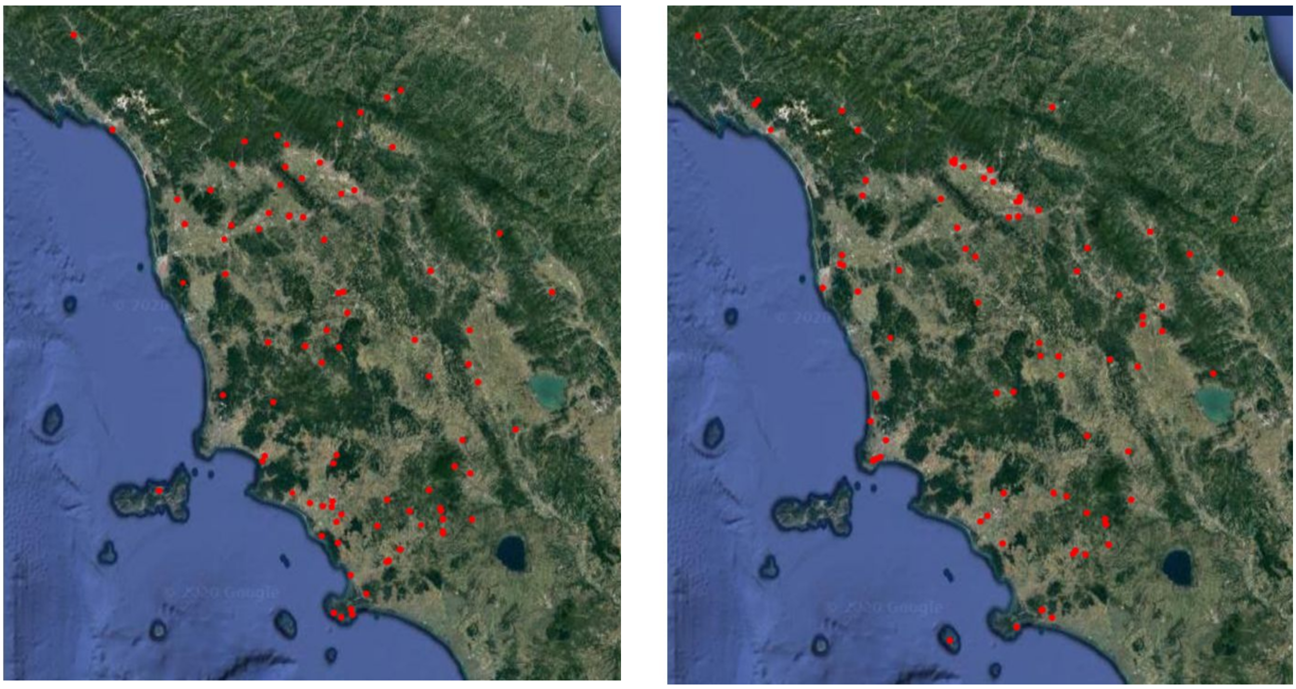


Figure 14: MATTCH results computed on the 03/03/2020 delivery of the TOSCANA project: ascending (right) and descending (left)

Recalling the D200.1 outcome, in the table below, it is agreed that the approach fulfils the needs expressed by the User:

Code	Name	Description	Achieved
UR1	EFFECTIVE	An effective way to retrieve relevant information, namely changes in the trend of time series, from the millions of measurement points provided at each update	YES
UR2	AUTOMATIC	The ‘trend change’ layer has to be created automatically to ensure the timely delivery of this information	YES
UR3	TUNABLE	The trend change detection needs to be calibrated according to the end users needs, i.e. according to the magnitude or time frame of the trend change	YES
UR4	GIS-READY	The trend change information needs to be delivered in a practical format, such as a GIS layer	YES

Table 3: User Requirement described in D200.1 vs achievement status

More in detail the activities in WP400 have shown that:

UR1 - The MATTCH approach consists in a Machine-Learning based algorithm to be applied to the measurement points. All the displacement time series are processed and changes in the trend are provided. The algorithm, included in the SqueeSAR® processing environment, is tailored to run as a tested and robust step of the processing chain.

UR2 - The MATTCH approach has been tested inside the Toscana Project scheme. It has been shown that the results are timely produced together with the standard delivery to the client (who is the Final User in this project) and uploaded on the TREmaps® web-based application.

UR3 - The MATTCH approach has shown to be efficient also when applied in the 150 days window as required by the User. Thanks to an iteration with the User’s, it was also demonstrated that it is possible to tune the algorithm parameters accordingly with the User’s feedback after the first delivery.

UR4 - All the MATTCH deliverables are produced in .SHP format (vector layer). With respect to the standard SqueeSAR® deliverable (see Annex B), two fields are added to all the measurement points analyzed:

a. DVEL [mm] (difference in the average deformation rate)
b. DSTEP [mm] (step in time series)

The results are visible on TREmaps® and can be downloaded following the proper link. The .SHP format can be easily ingested in all GIS systems.