

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Prevention of Potential Catastrophes Depending on Interferometric Radar Technique and Artificial Intelligence

PhD Student: Lama Moualla - 36th Cycle

Supervisors: Dr. Alessio Rucci & Prof. Giampiero Naletto

Meeting - 14/12/2023

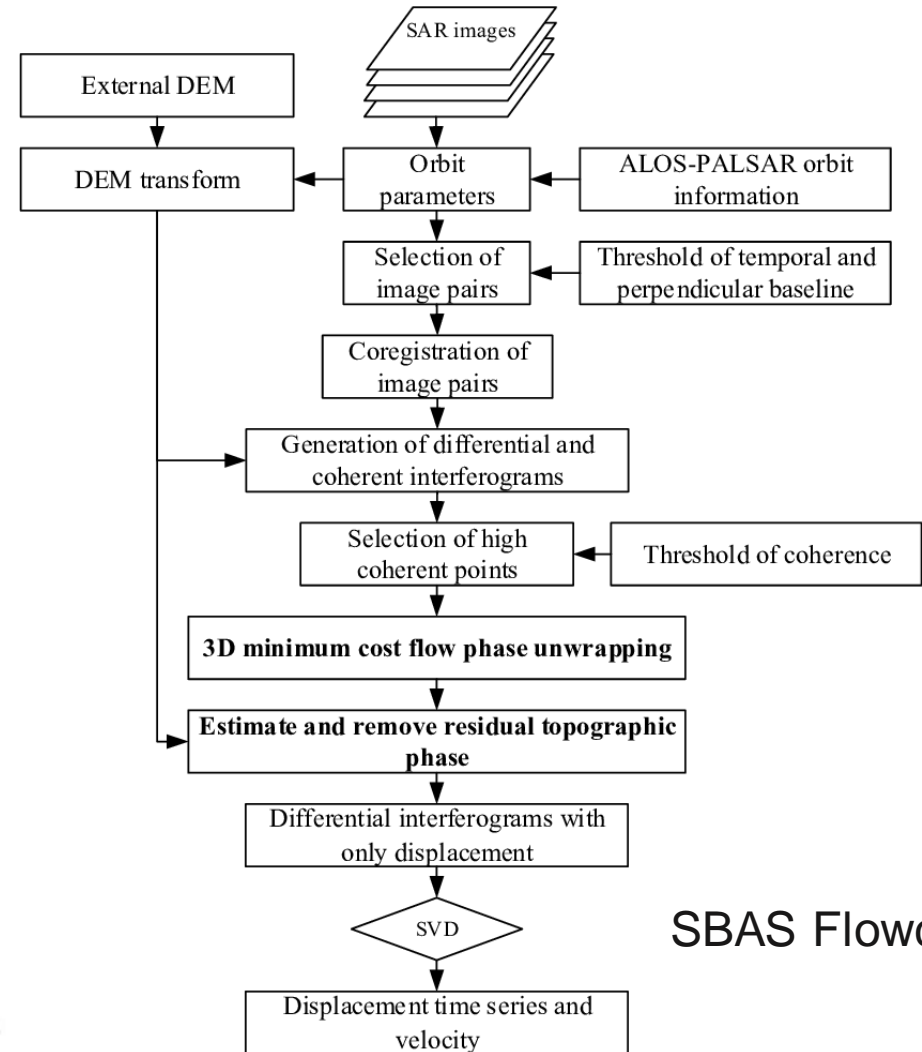
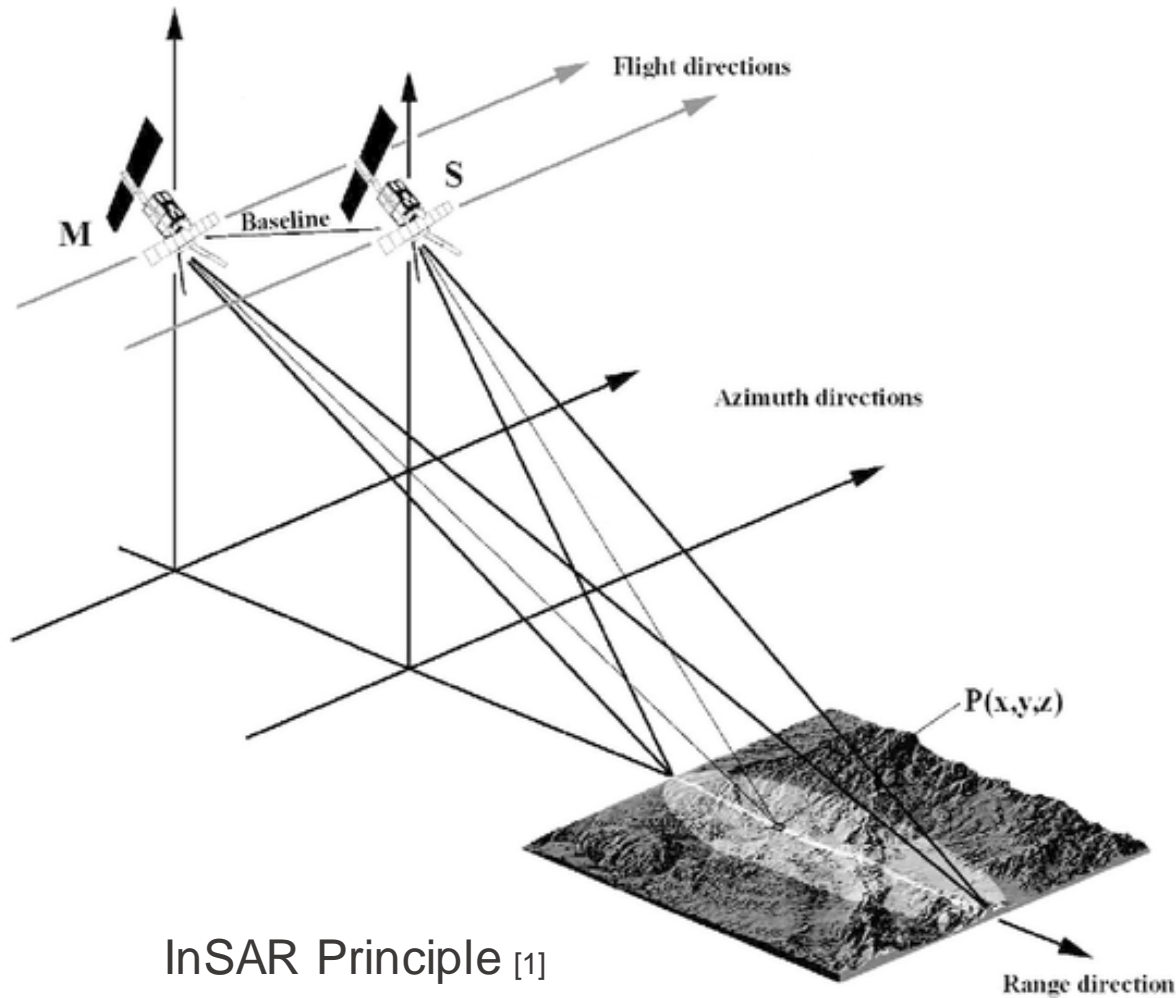


Table of Contents



- **Introduction**
- **Recap & Updated Results of the First Objective**
- **Results of The Second Objective**
- **Results of The Third Objective**
- **Conclusions**

Challenges in Deploying InSAR Technology



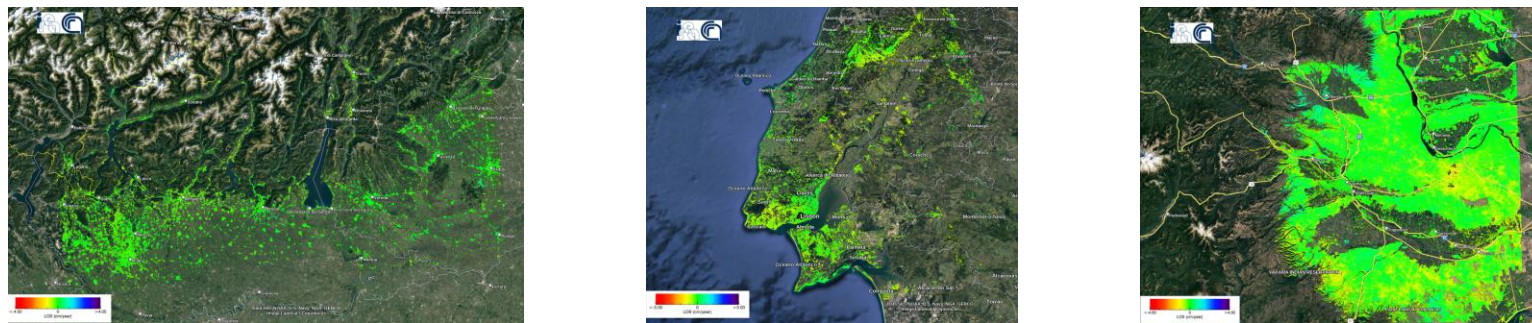


The First Objective

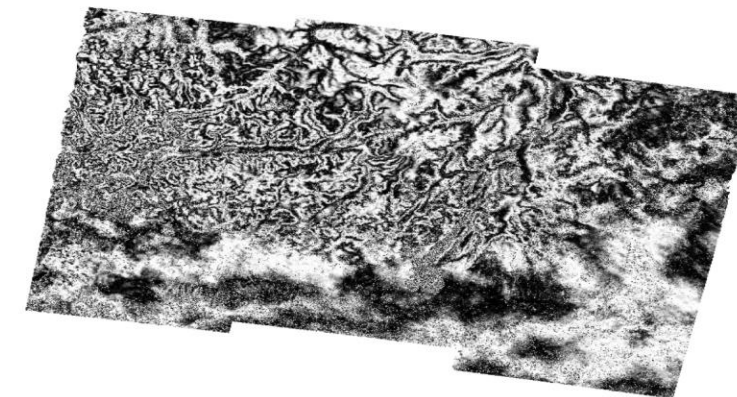


Using Intelligent algorithms in developing a methodology that can automatically analyze large InSar data packets and identify areas where infrastructures are at risk of displacement due to ground movement

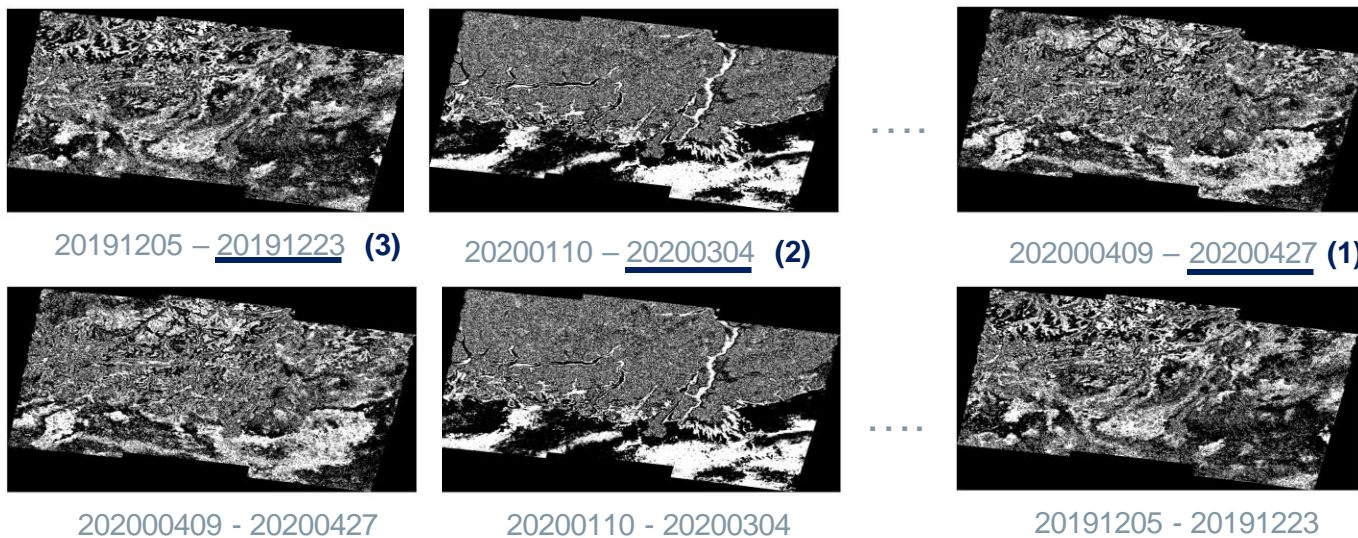
Recap of The First Objective Results



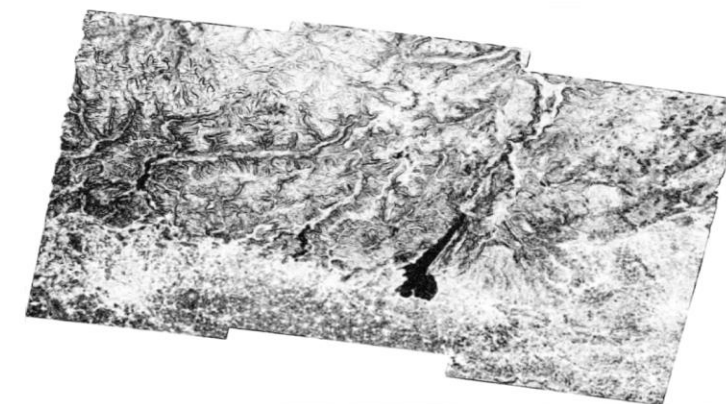
a)- Parallel - Small BAseline Subset Displacements Maps (Lombardy – Lisbon - Washington)



Wrapped Interferogram (Before Applying The High-Pass Filter)



b)- The chronological sorting of the interferograms before creating the dataset and inputting the training samples into the model

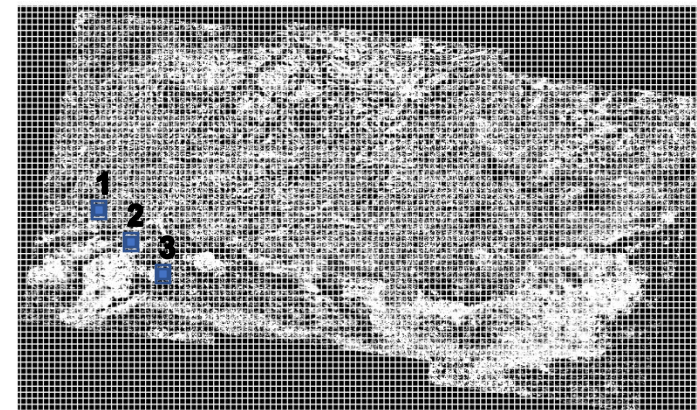
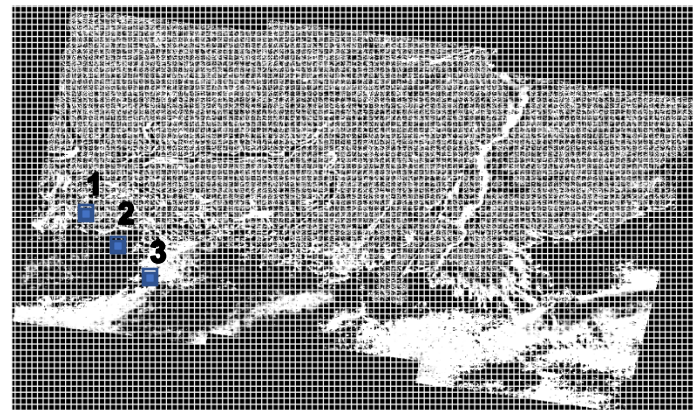
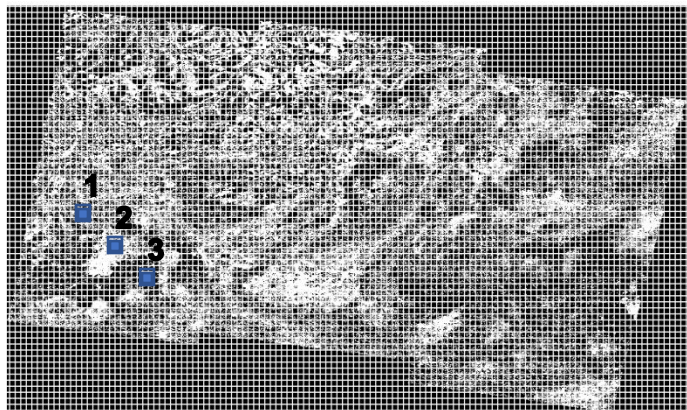


Wrapped Interferogram (After Applying The High-Pass Filter)

c)- Considering that the main power of the low frequency signal comes from the atmospheric artifacts , it was necessary to apply a high pass filter



Recap of The First Objective Results



1st Interferogram

2nd Interferogram

Last Interferogram VEL Label (-0.7 cm/year)

1	1.15
2	-2.10
3	-1.75

3.10
2.55
-1.15

2.05	-1.5	Fast Negative
-3.09	0.9	Undefined
-1.68	-0.1	Undefined

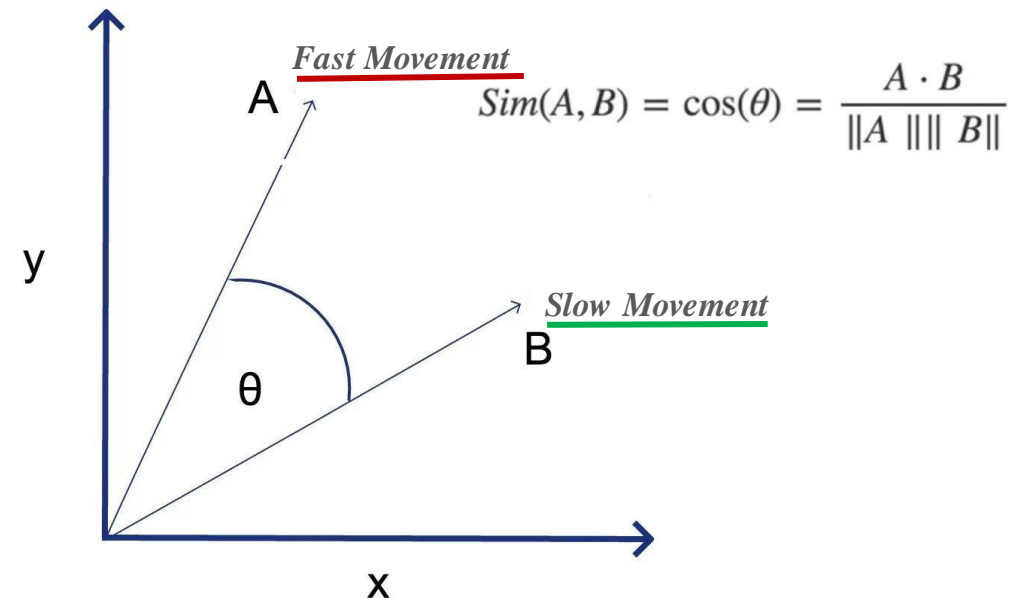
Training Examples
Phase Values (Radian)

Labels ??
The Velocity Values of The
Measurement Point Inside The Pixel
(cm/year)



The Best Trained Classifier: Cosine-KNN

- Handling of Angular Data
- Dimensionality and Feature Relationships
- Noise and Small Variations
- Suitability for Complex Patterns
- Generalization Capability

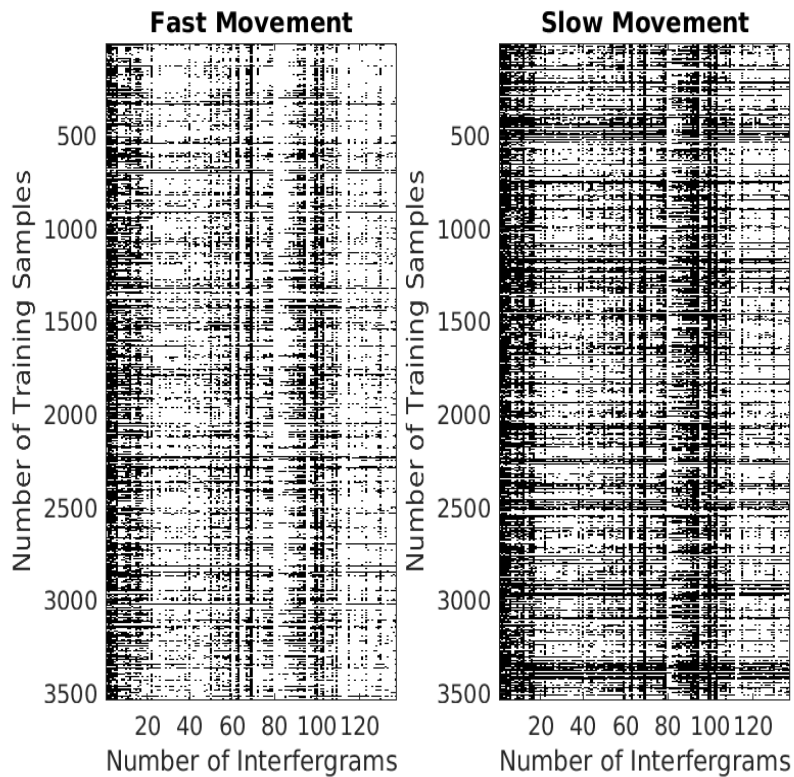


Trained Classifiers:

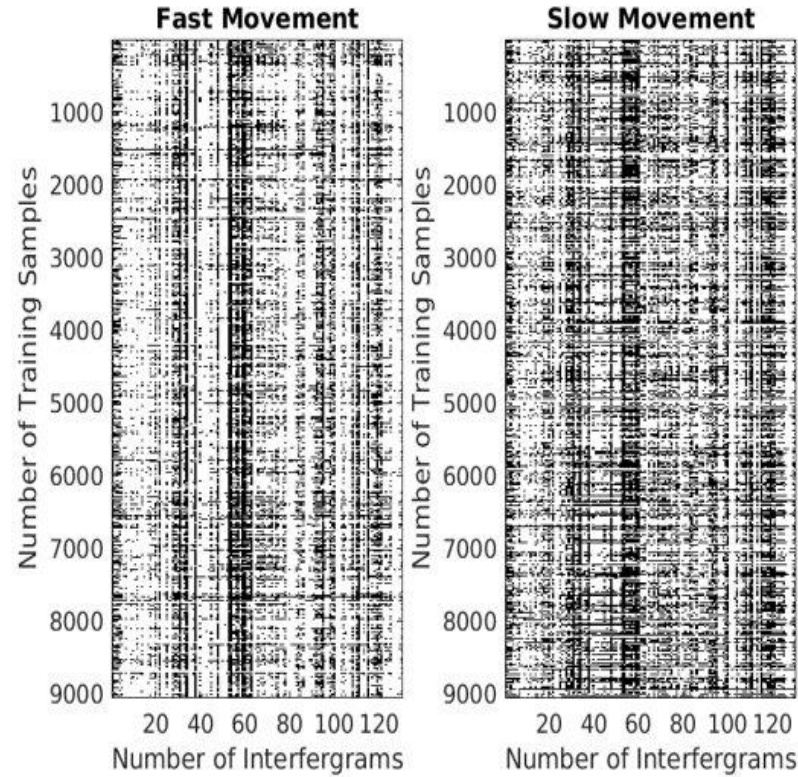
Cosine KNN, Subspace KNN (Ensemble), Medium Neural Network, Logistic Regression, Cubic SVM, Medium Tree, Fine Tree, Bagged Tree (Ensemble), Quadratic Discriminant, 2D Convolution Neural Network and Long short-term memory (LSTM)



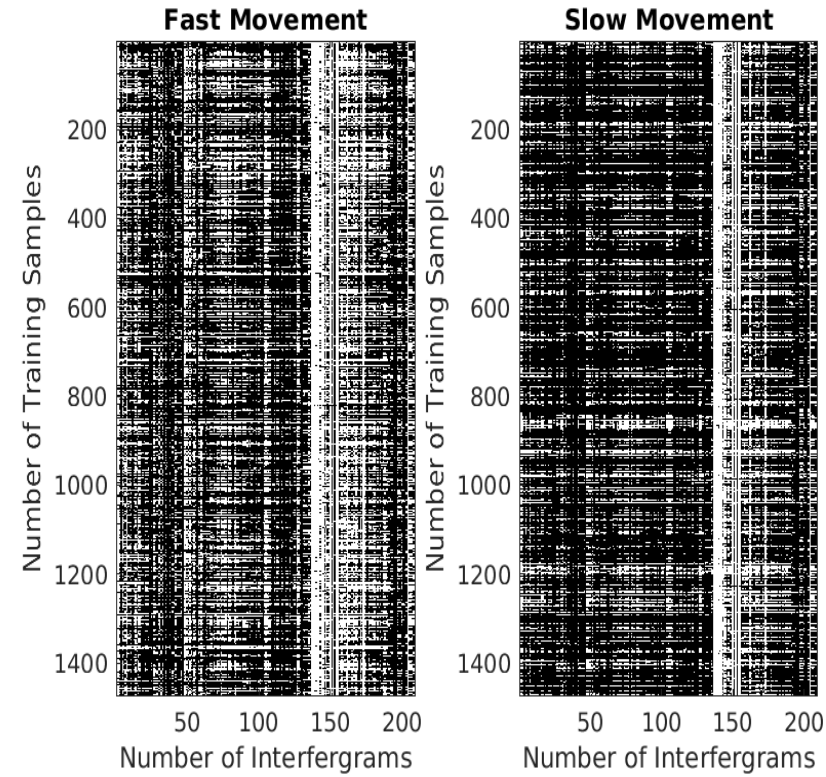
The Matrices Representing Slow and Fast Motion



Lombardy Dataset



Lisbon Dataset



Washington Dataset

The black color represents the magnitude values larger than 0.9 radian
The white color represents the magnitude values smaller than 0.9 radian



Pseudo-Labeling

A model is trained and tested on the labeled data



The trained model is used to make predictions on the pseudo-labelled data



A new training set is created by combining the testing samples with a prediction accuracy of 0.9 or higher with the original labeled data



The pseudo-labeled training samples are removed from the original test set of unlabeled training samples



The model is retrained using the combined labeled and pseudo-labeled data



Pseudo-Labeling Results

	Cosine K-NN	1st PS	2nd PS
Number of training samples	952	12334	29474
Number of testing samples	238	3084	7368
Accuracy of validation	72.8%	97.8%	98.7%
Accuracy of the test set	70.6%	97.9%	98.8%

Results of Pseudo-Labeling PS on Lombardy Dataset
(Fast Positive/Undefined)

	Cosine K-NN	1st PS	2nd PS
Number of training samples	4884	6102	6684
Number of testing samples	1220	1526	1672
Accuracy of validation	81.2%	85.1%	86.1%
Accuracy of the test set	82.3%	84.3%	86.8%

Results of Pseudo-Labeling PS on Washington Dataset
(Fast Negative/Undefined)

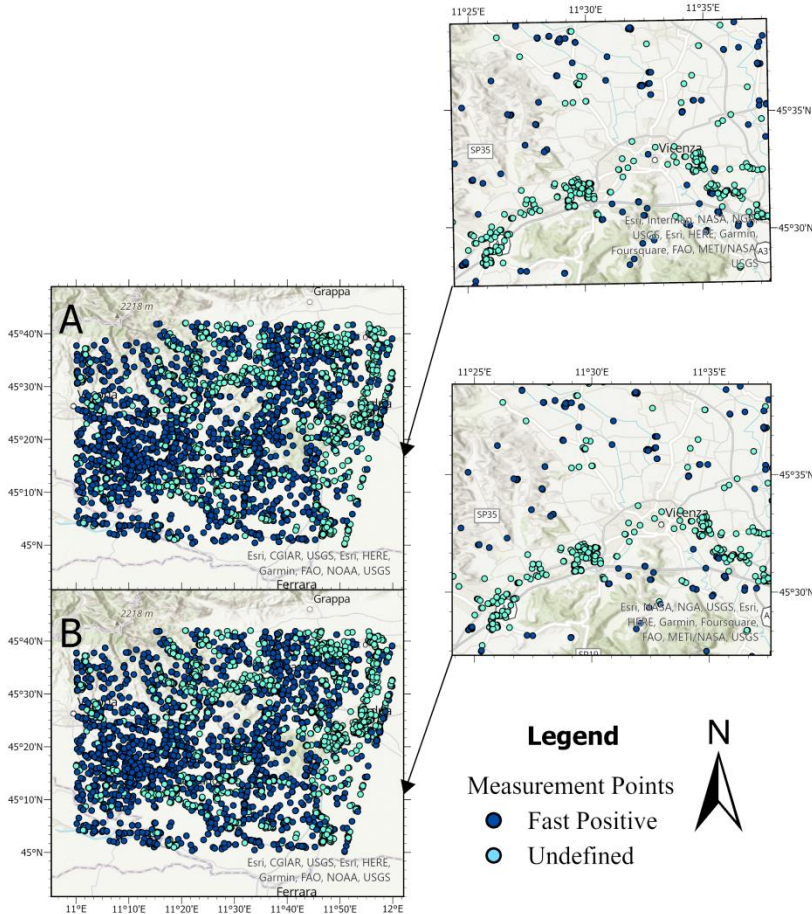
	Cosine K-NN	1st PS	2nd PS	3rd PS	4th PS
Number of training samples	14596	17812	20130	21576	22736
Number of testing samples	3650	4454	5032	5394	5684
Accuracy of validation	79.4%	83.4%	85.5%	86.6%	87.2%
Accuracy of the test set	79.4%	82.8%	85.7%	87.1%	87.1%

Results of Pseudo-Labeling PS on Lisbon Dataset
(Positive/Negative)

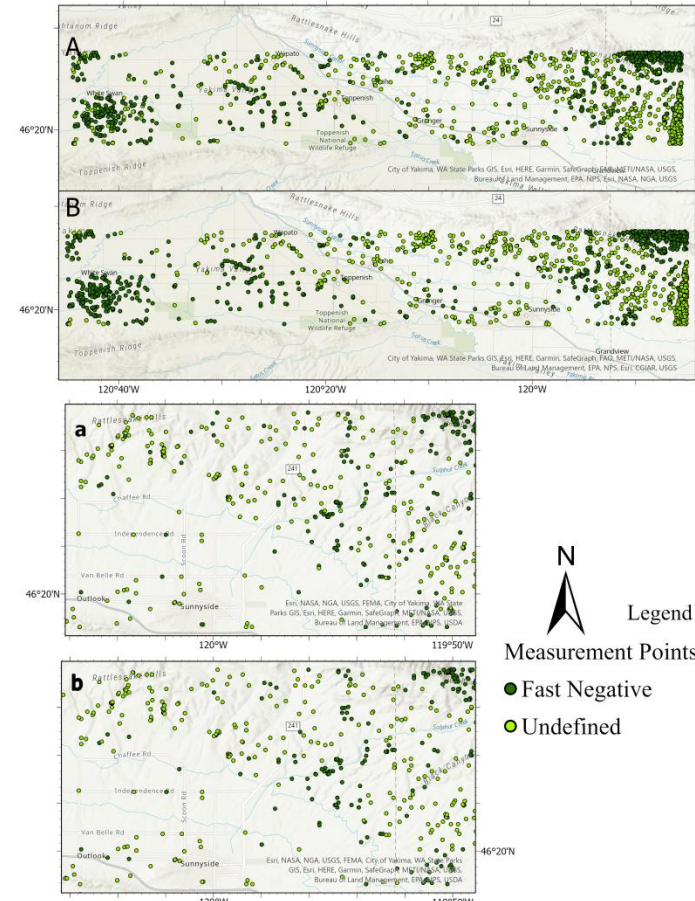


The Ground-truth Test Set vs Predictions

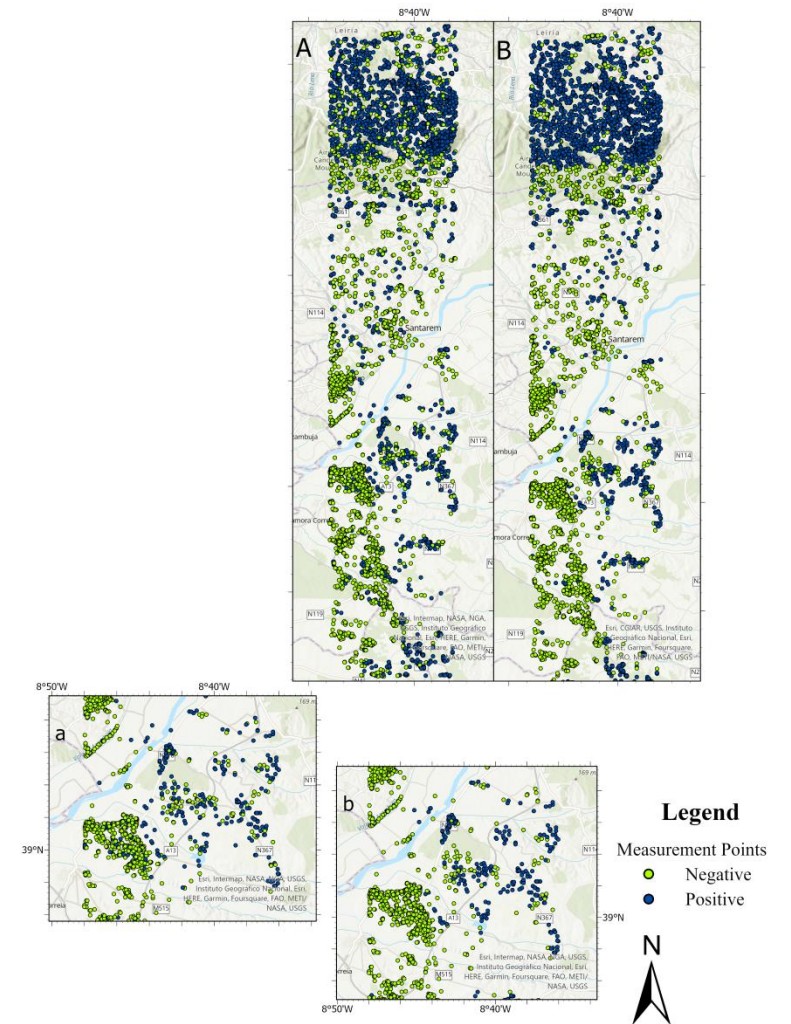
Lombardy Dataset (Fast Positive/ Undefined Movement Model)



Washington Dataset (Fast Negative/ Undefined Movement Model)



Lisbon Dataset (Positive / Negative Movement Model)

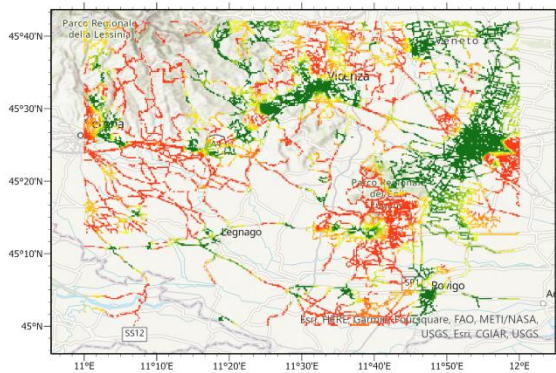
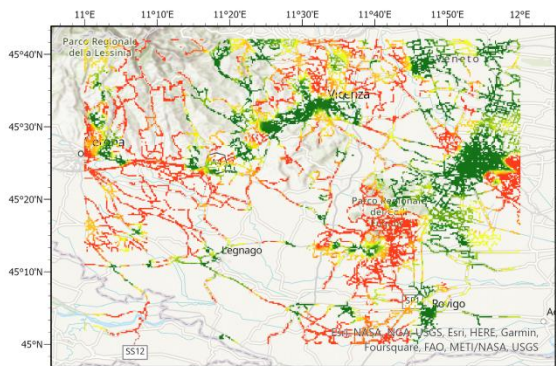


Figures a display the ground truth of the test dataset. Figures b display the predictions of the test dataset.

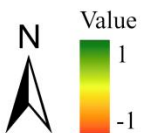


Masked Roads Network

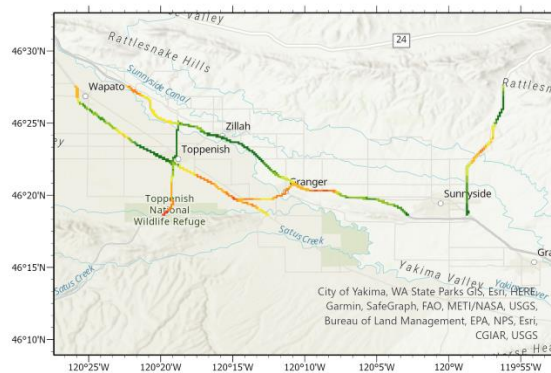
Fast Positive Movement Roads Indication (Milan Dataset)



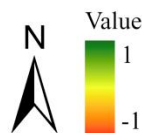
Legend



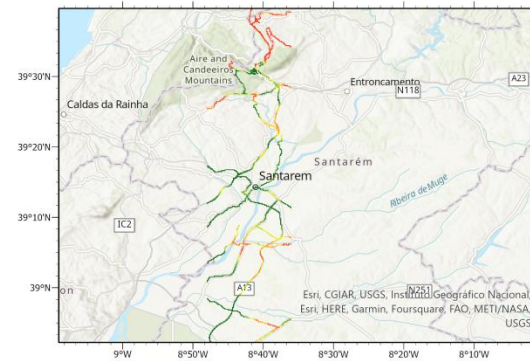
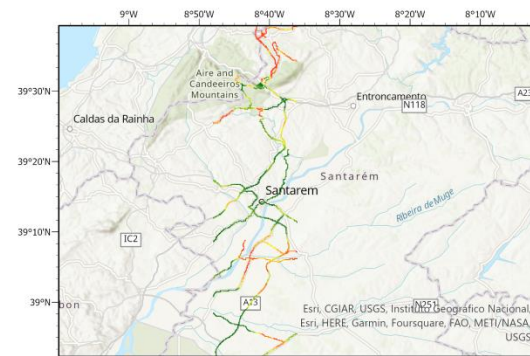
Fast Negative Movement Roads Indication (Washington Dataset)



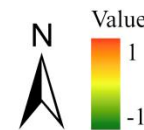
Legend



Positive/Negative Movement Roads Indication (Lisbon Dataset)



Legend



The top figures represent the masked roads of the ground truth test set; while the bottom figures represent the masked roads of the predicted test sets
The value of 1 expresses the positive movement while the value -1 expresses the negative movement



The Second Objective

Establishing a predictive model for the displacements of the infrastructure studied in the research, based on the methodology that we will try to develop



Data Preprocessing

Before Resampling

Name of the case study	Number of the time steps	Temporal difference between the time steps (days)
Milan	50	12
Lisbon	48	6, 12, 18, 24
Washington	75	12, 24, 48, 222

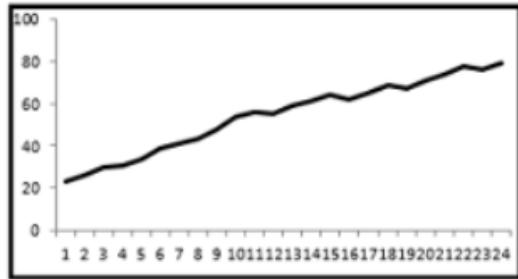


Data Preprocessing

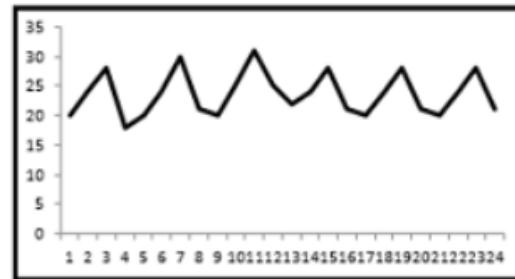
- **Missing Values Imputation (Backward Filling)**
- **Feature Engineering (Embedding Time as a Second Feature in Time Series Data)**



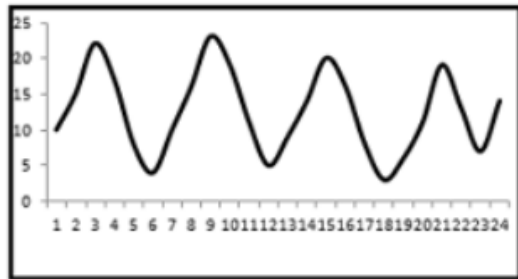
Components of Time Series



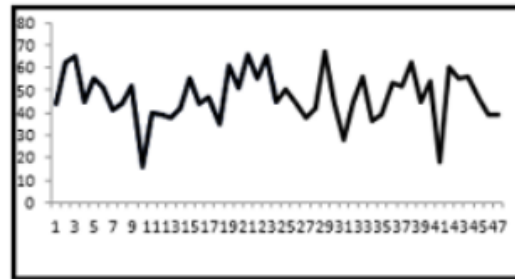
(a) Trend



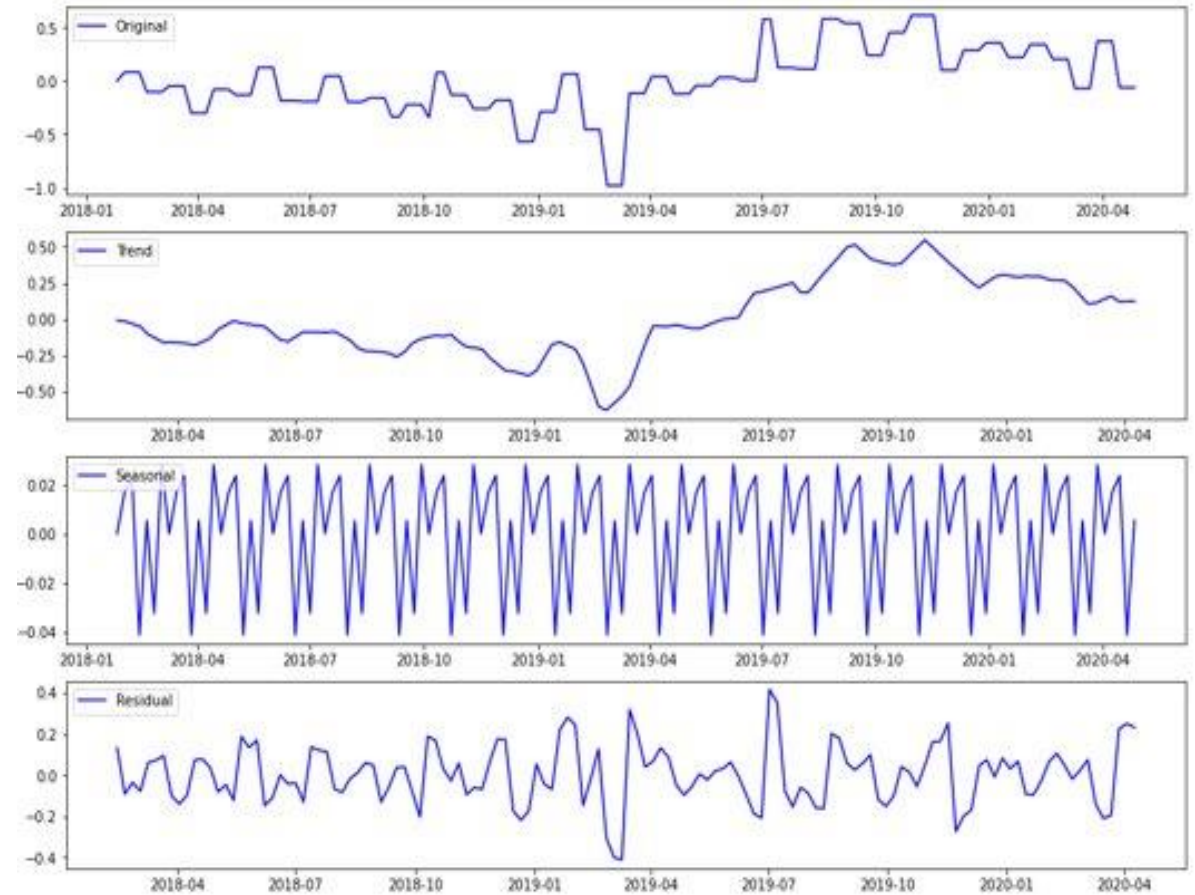
(b) Seasonality



(c) Cyclicity



(d) Irregular



Lisbon Dataset

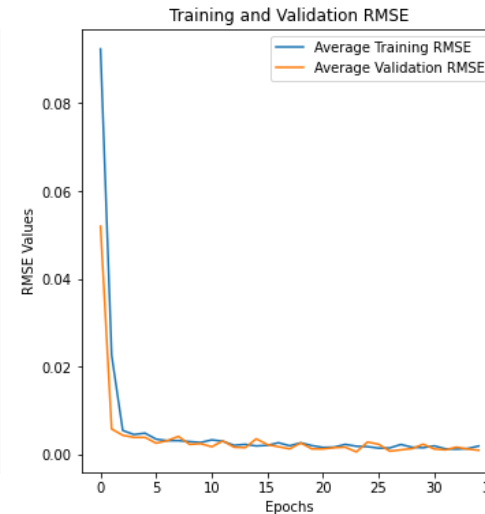
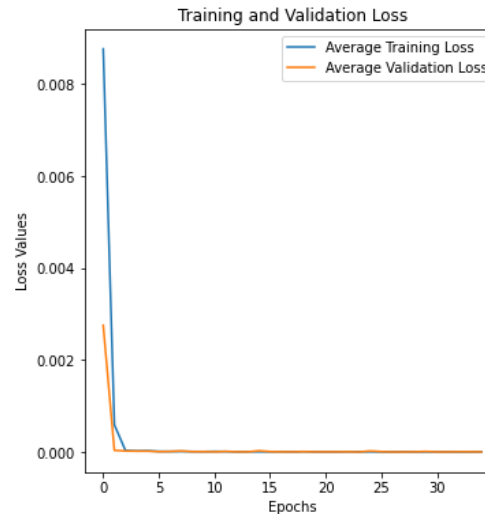
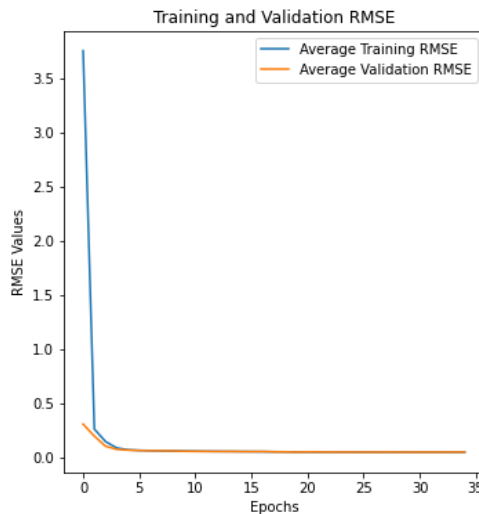
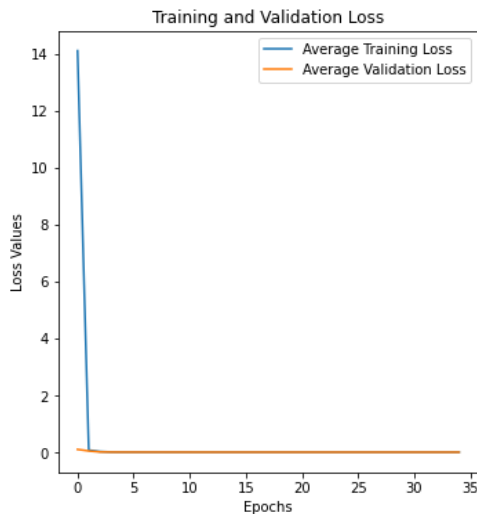
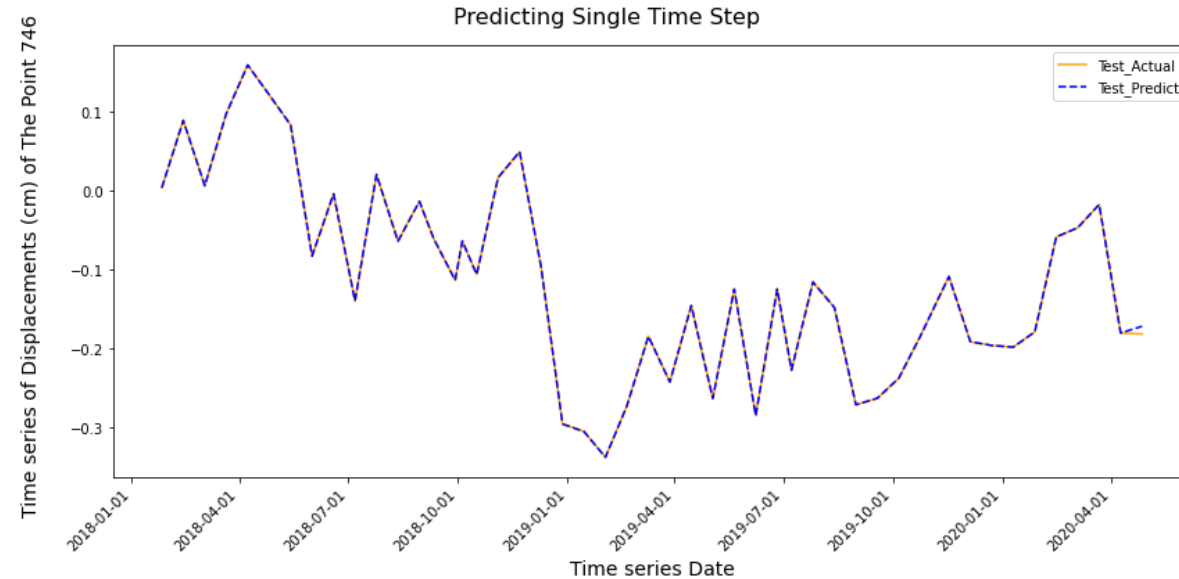
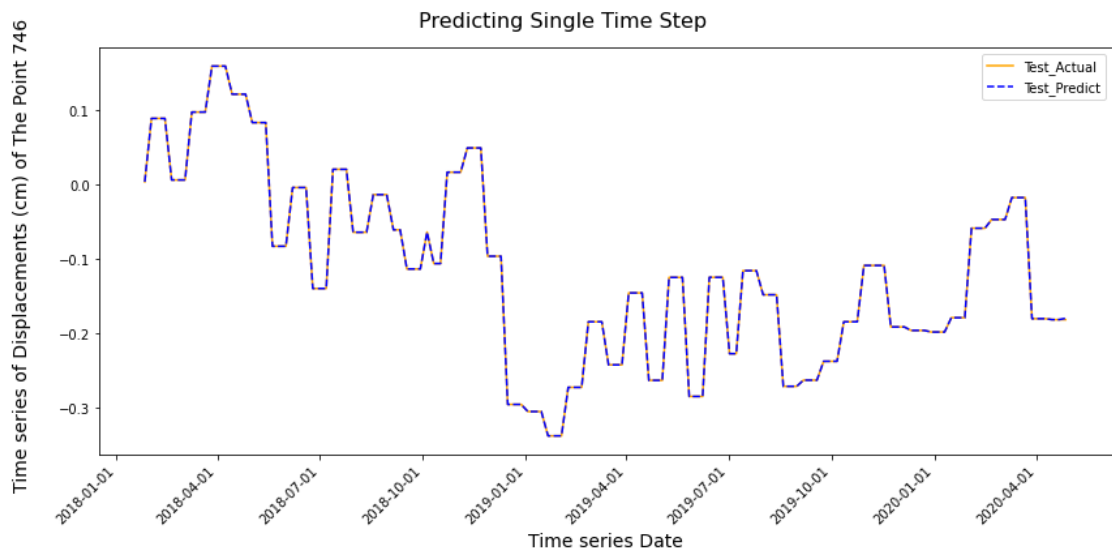


Why LSTM ?

- **Handling Temporal Dependencies**
- **Spatial Correlations**
- **Noise and Anomaly Tolerance**
- **Multivariate Time Series Capability**
- **Adaptability**
- **Real-Time Prediction**



Results of One Time Step Predictions (Lisbon Dataset)



Learning Curves Influenced by the Imputation of Missing Values

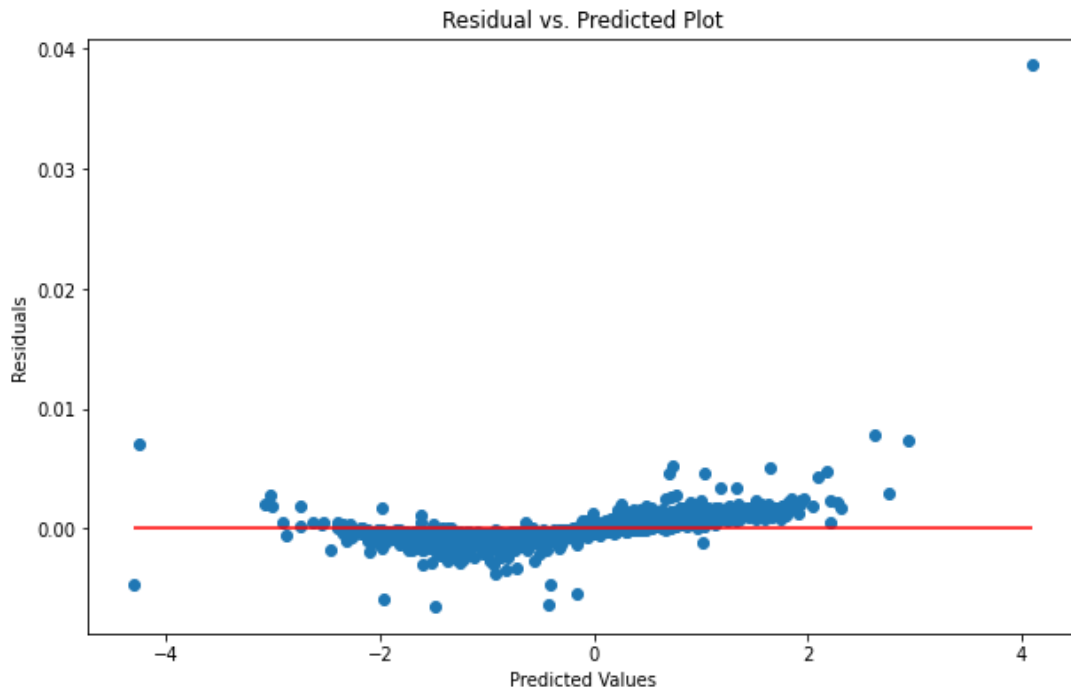
Learning Curves Influenced by the Time Embedding



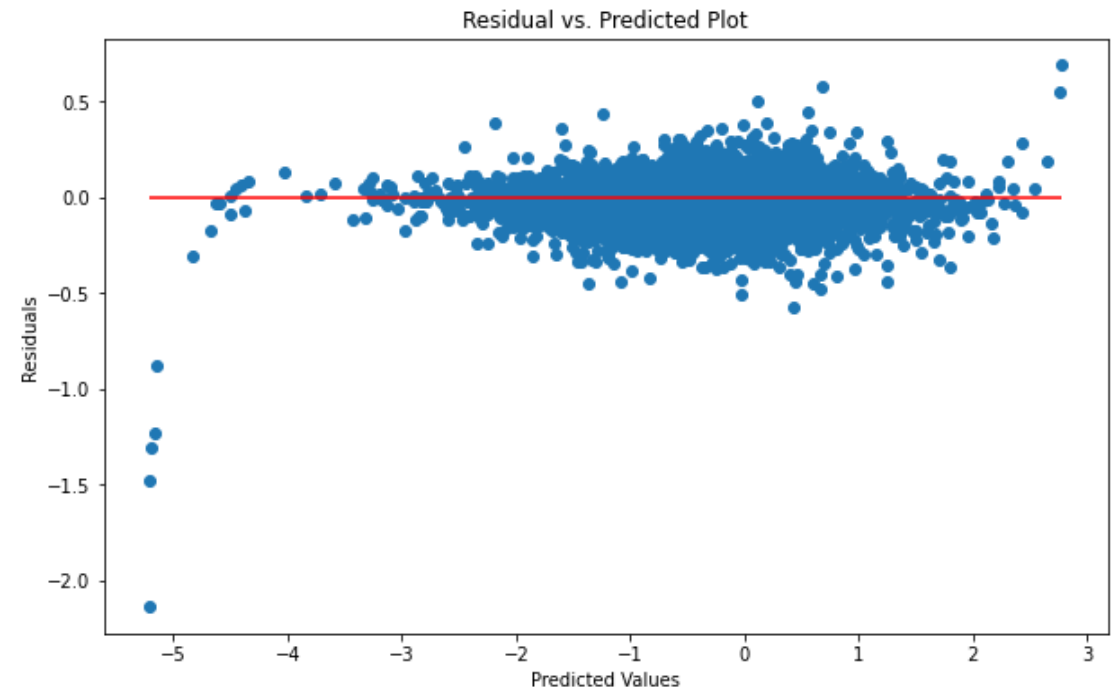
Results of One Time Step Predictions (Lisbon Dataset)

Are the residuals exhibiting constant variance across different levels of predicted values?

- **Homoscedasticity** is a desirable property in a regression model, as it suggests that the model appropriately captures the variance in the data across all levels of the independent variable(s).
- **Heteroscedasticity** often suggests that the model is missing important features, a misspecification in the model, or that a transformation of variables might be necessary.



Homoscedasticity as Affected by the Imputation of Missing Values



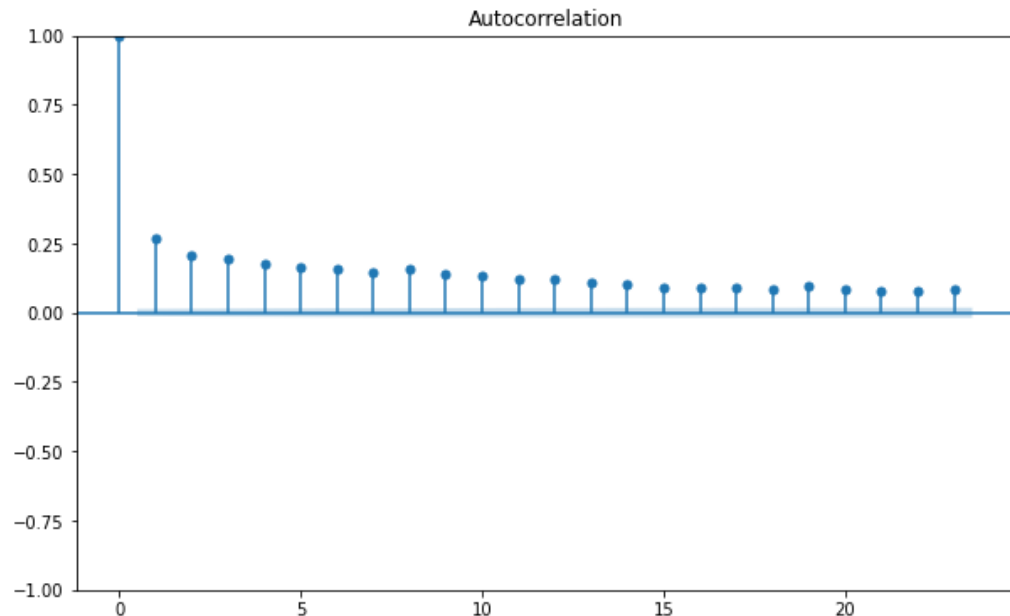
Homoscedasticity as Affected by the Time Embedding



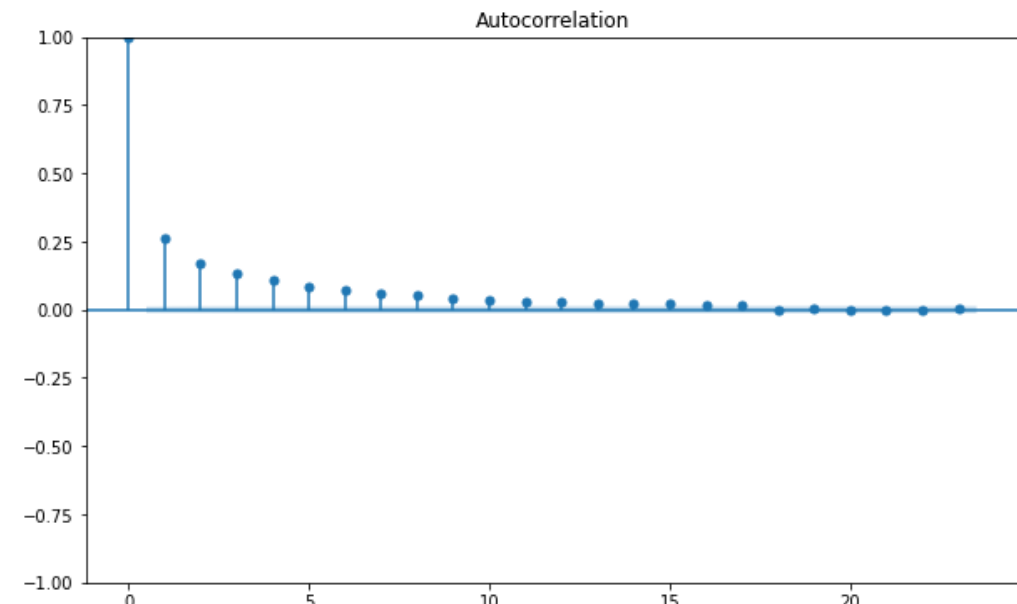
Results of One Time Step Predictions (Lisbon Dataset)

The Autocorrelation Function (ACF) is indispensable in the analysis of time series data as it reveals any remaining correlation in the residuals of the model's predictions.

- Significant autocorrelation at any lag could suggest that the model has not fully captured the predictive structure within the data, indicating room for improvement.
- The absence of such correlation, on the other hand, would affirm that the model's predictions are not systematically biased by overlooked temporal dependencies.



ACF Charts Depending on the Imputation of Missing Values



ACF Charts Depending on the Time Embedding



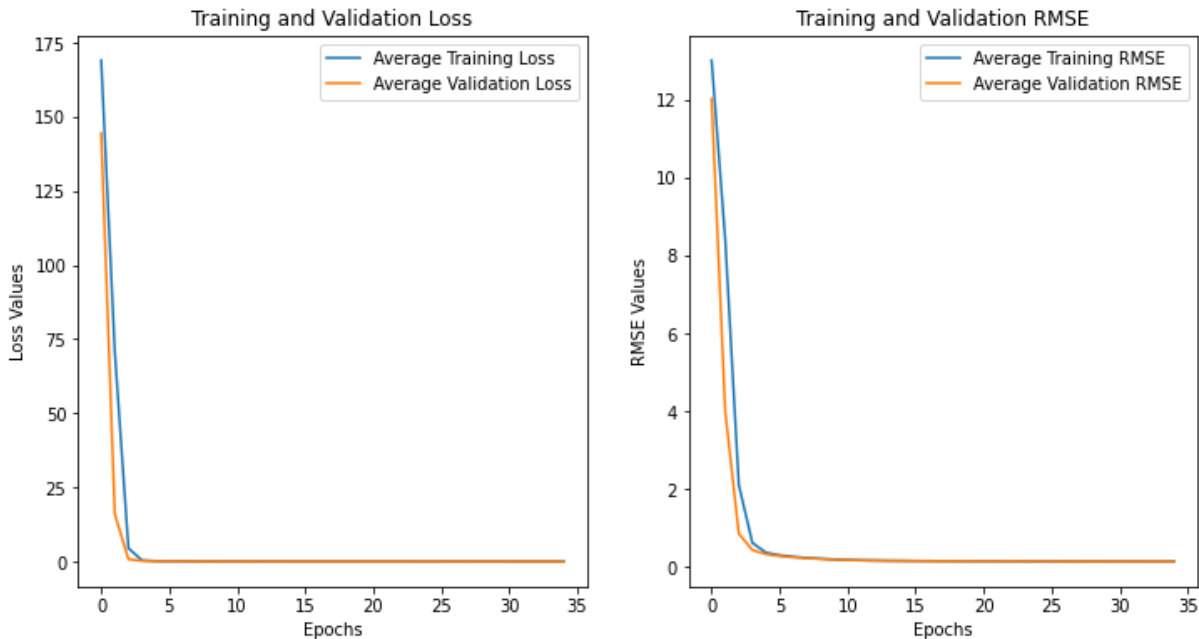
Results of Multiple Time Steps Predictions

Layer (type)	Output Shape	Param #
rnn (RNN)	(None, 38, 225)	206100
rnn_1 (RNN)	(None, 75)	90600
dense (Dense)	(None, 20)	1520
reshape (Reshape)	(None, 10, 2)	0

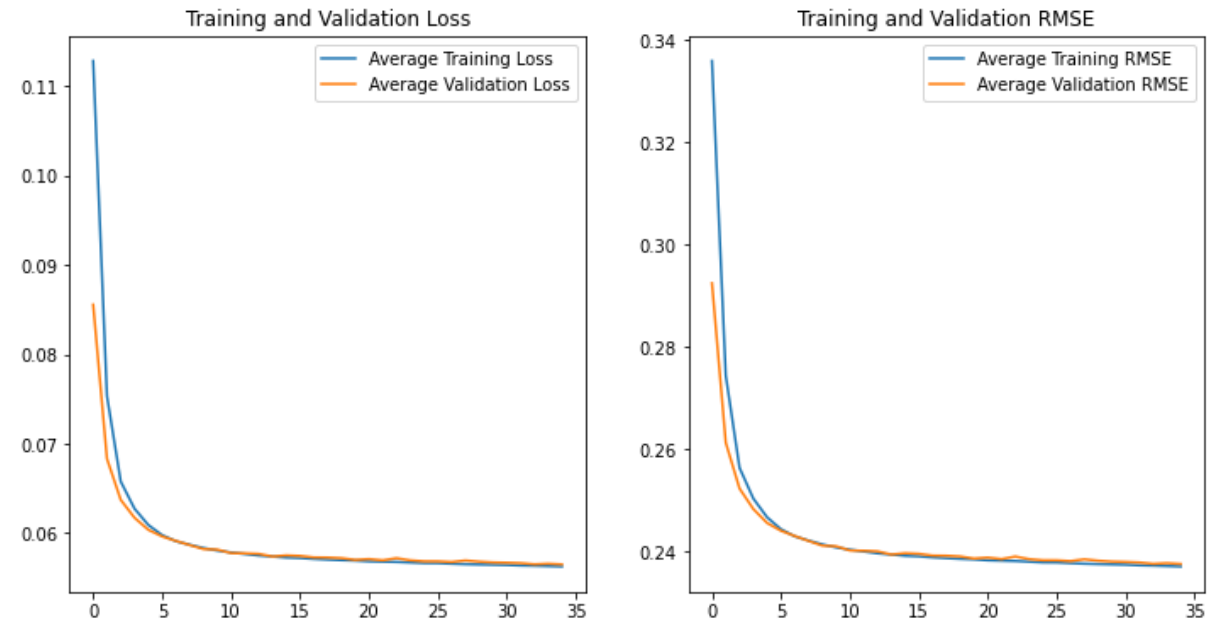
Total params: 298,220
Trainable params: 298,220
Non-trainable params: 0

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 110, 100)	40800
lstm_1 (LSTM)	(None, 50)	30200
dense (Dense)	(None, 28)	1428

Total params: 72,428
Trainable params: 72,428
Non-trainable params: 0



Learning Curves of training Time Gated LSTM model (10 Steps)



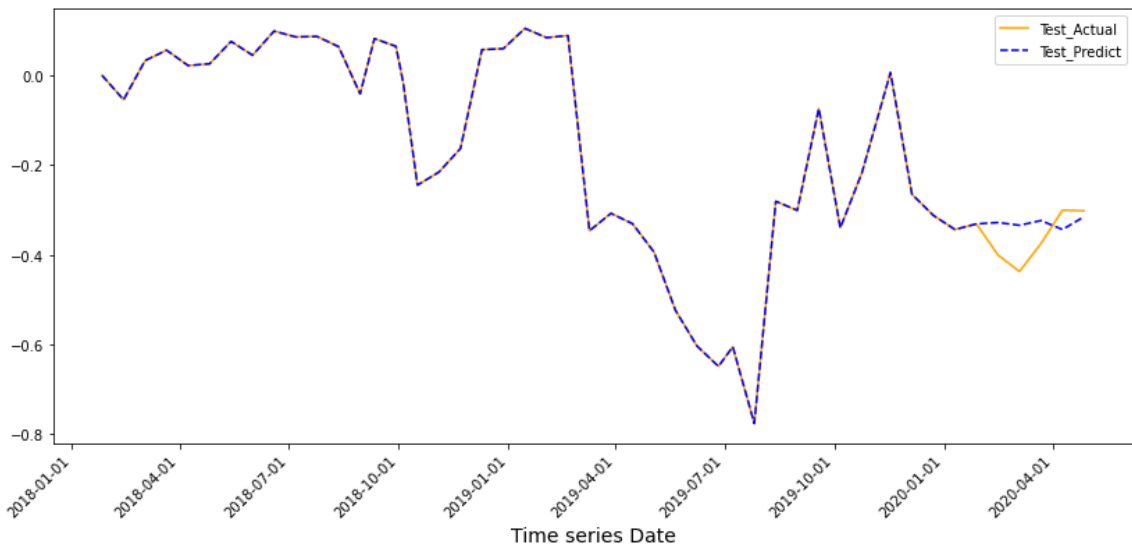
Learning Curves of training standard LSTM model (28 Steps)



Results of Multiple Time Steps Predictions

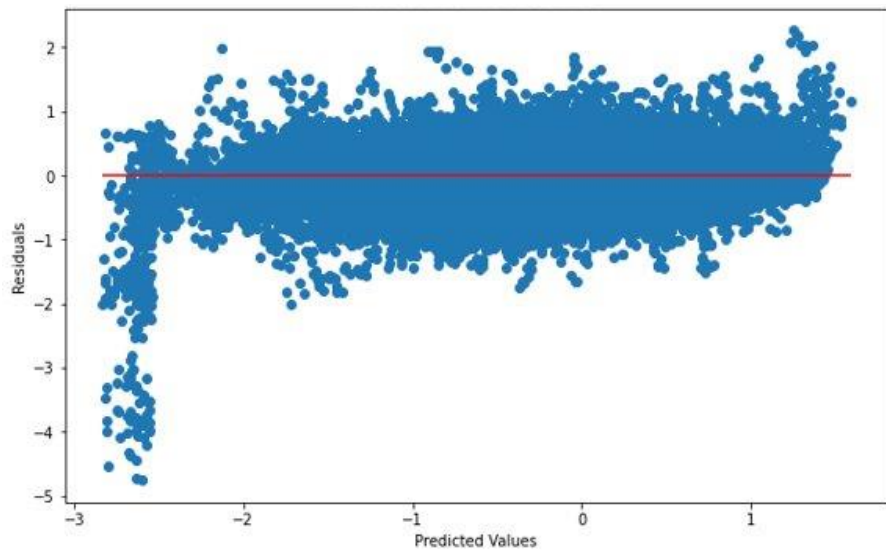
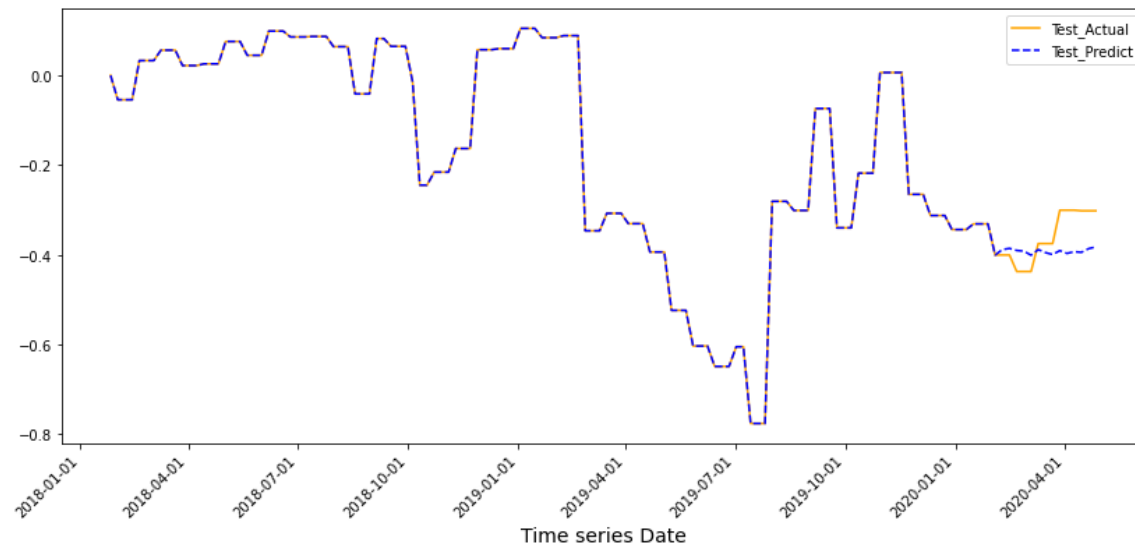
Time series of Displacements (cm) of The Point 37

Predicting Five Time Steps

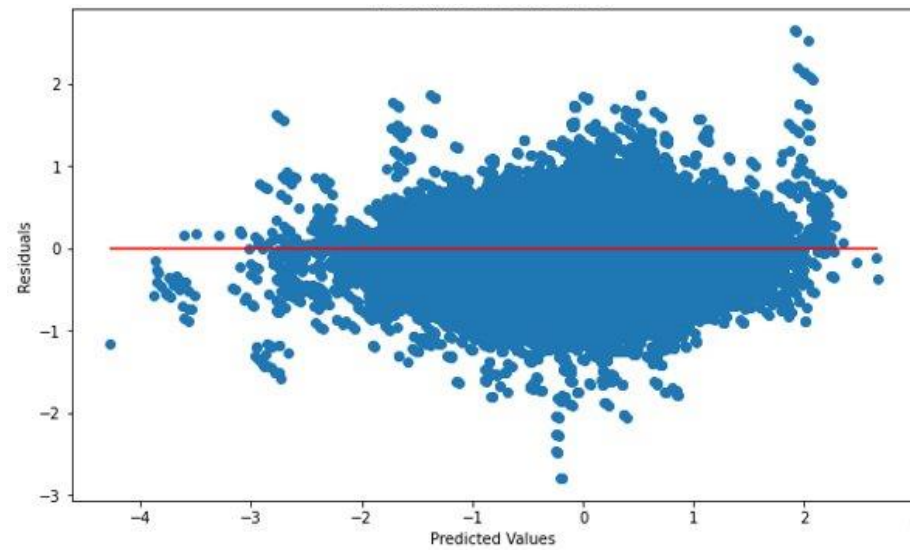


Time series of Displacements (cm) of The Point 37

Predicting Fourteen Time Steps



Homoscedasticity of training Time Gated LSTM model



Homoscedasticity of training Standard LSTM model



The Third Objective

Determine the functions of the Geographic Information Systems (GIS) toolbox that will be developed to integrate the final work results within the ArcGIS Pro environment



Displacement Predictions Toolbox

A toolbox to interactively predict single or multiple steps of displacements time series

Geoprocessing
Displacement Predictions

Parameters Environments

- * Dataset File
- CNR P-SBAS Dataset
- * Directory
- * min. Latitude
- * max. Latitude
- * min. Longitude
- * max. Longitude
- * Number of Time Steps
- Nodes Number of The LSTM Layer (Optional)
- Learning Rate (Optional)
- Number of Epochs (Optional)
- Plot Learning Curves (Optional)

Run



Displacement Predictions Toolbox

Insert the path for the dataset csv file and check the box if it generated by CNR P-SBAS service

Insert the path for the output file where the results will be saved

Enter the geographic coordinates of the boundary for the studied area

Insert the number of time steps you want to predict

Change the default values of the following hyperparameters in case you are not satisfied with the RMSE Result

Check the final box if it is important to be sure of the learning process efficacy

The image shows a software interface for 'Displacement Predictions' with two panels. The left panel is a simplified version of the form, and the right panel is the full application window. Arrows point from the text instructions to the corresponding input fields in both panels.

Left Panel (Simplified Form):

- * Dataset File: [Text Field] CNR P-SBAS Dataset
- * Directory: [Text Field]
- * min. Latitude: [Text Field]
- * max. Latitude: [Text Field]
- * min. Longitude: [Text Field]
- * max. Longitude: [Text Field]
- * Number of Time Steps: [Text Field]
- Nodes Number of The LSTM Layer (Optional): [Text Field]
- Learning Rate (Optional): [Text Field]
- Number of Epochs (Optional): [Text Field]
- Plot Learning Curves (Optional)

Right Panel (Full Application Window):

Geoprocessing Displacement Predictions

Parameters Environments

- * Dataset File: [Text Field] CNR P-SBAS Dataset
- * Directory: [Text Field]
- * min. Latitude: [Text Field]
- * max. Latitude: [Text Field]
- * min. Longitude: [Text Field]
- * max. Longitude: [Text Field]
- * Number of Time Steps: [Text Field]
- Nodes Number of The LSTM Layer (Optional): [Text Field]
- Learning Rate (Optional): [Text Field]
- Number of Epochs (Optional): [Text Field]
- Plot Learning Curves (Optional)

Run



Toolbox Improper Execution Error Messages

The screenshot shows the ArcGIS Desktop interface with the Geoprocessing tool running. The error message is as follows:

Displacement Predictions (MyProject6atbx)

Started: Today at 6:07:46 PM
Completed: Today at 6:08:48 PM
Elapsed Time: 1 Minute 2 Seconds

The number of predicted time steps is not valid. Refill the field of this parameter with a valid value.

Parameters Environments Messages (3)

The number of predicted time steps is not valid. Refill the field of this parameter with a valid value.

Failed script (null)...

Failed to execute (TSPREDICTION).

Failed at Wednesday, November 22, 2023 6:08:48 PM (Elapsed Time: 1 minutes 1 seconds)

The Geoprocessing tool parameters are:

- Dataset File: C:\SecondObTbx\DTSLOS_Jama_20200107_20210E
- CNR P-SBAS Dataset
- Directory: C:\SecondObTbx
- min. Latitude: 45
- max. Latitude: 45.5
- min. Longitude: 10
- max. Longitude: 10.5
- Number of Time Steps: -1** (highlighted in red)
- Nodes Number of The LSTM Layer (Optional):
- Learning Rate (Optional):
- Number of Epochs (Optional):
- Plot Learning Curves (Optional)

The error message also indicates that the 'Number of Time Steps' parameter is invalid and should be set to a valid value.



Toolbox Improper Execution Error Messages

The screenshot displays the ArcGIS Desktop interface. The main map area shows a dark, textured area representing displacement predictions. A toolbox window titled 'Displacement Predictions (MyProject6atbx)' is open, showing the following error messages:

Displacement Predictions (MyProject6atbx)

Started: Today at 6:03:20 PM
Completed: Today at 6:06:17 PM
Elapsed Time: 2 Minutes 57 Seconds

The number of displacemnets samples is not allowed to exceed 100000. Change the extent of the area or recheck your dataset.

Parameters Environments Messages (3)

The number of displacemnets samples is not allowed to exceed 100000. Change the extent of the area or recheck your dataset.

Failed script (null)...

Failed to execute (TSPREDICTION).

Failed at Wednesday, November 22, 2023 6:06:17 PM (Elapsed Time: 2 minutes 56 seconds)

The toolbox window also shows parameters for the tool, including Dataset File, Directory, min. Latitude, max. Latitude, min. Longitude, max. Longitude, Number of Time Steps, Nodes Number of The LSTM Layer (Optional), Learning Rate (Optional), Number of Epochs (Optional), and Plot Learning Curves (Optional).



Results of Running The Toolbox (First Part)

The screenshot displays a GIS application window with three main panels:

- Contents Panel (Left):** Shows a search bar and a drawing order list. The 'Map' layer is expanded, showing 'XYpointsShP' as the active layer. Other layers include 'World Topographic Map' and 'World Hillshade', both of which are unchecked.
- Map Panel (Center):** Displays a map view of a point cloud, where individual points are represented as small green and brown dots. A dialog box is overlaid on the map, containing the text: "Window title", "The dataset has been trained and tested.", "The Root Mean Squared Error is 0.099", and an "OK" button.
- Geoprocessing Panel (Right):** Shows the configuration for the 'Displacement Predictions' tool. The 'Parameters' tab is active, with the following settings:
 - Dataset File: C:\SecondObTbx\DTSLOS_lama_20200107_202106
 - CNR P-SBAS Dataset
 - Directory: C:\SecondObTbx
 - min. Latitude: 45
 - max. Latitude: 45.5
 - min. Longitude: 10
 - max. Longitude: 10.5
 - Number of Time Steps: 1
 - Nodes Number of The LSTM Layer (Optional):
 - Learning Rate (Optional):
 - Number of Epochs (Optional):
 - Plot Learning Curves (Optional)A 'Run' button is visible at the bottom of the panel. Below the tool name, a progress bar indicates the tool is 'Running...'. At the very bottom of the interface, the status bar shows a scale of 1:195,444, coordinates of 10.3209565°E 45.1984312°N, and 'Selected Features: 0'.



Results of Running The Toolbox (Second Part)

The screenshot displays the ArcGIS Pro interface. The main map area shows a scatter plot of points. A window titled "Figure 1" is overlaid on the map, showing a time series plot. The plot's y-axis is labeled "Time series Displacements of The Point 37053" and ranges from -0.4 to 0.8. The x-axis is labeled "Time series Date" and shows dates from 2020-01-01 to 2021-10-01. The plot contains two lines: a solid red line for "Actual Values" and a dashed blue line for "Predicted Values". The predicted values closely follow the actual values until late 2021, where they begin to diverge. A small dialog box titled "Another Point" with an "Exit" button is also visible.

The right-hand side of the interface shows the "Geoprocessing" pane with the "Displacement Predictions" tool running. The parameters are as follows:

Parameter	Value
Dataset File	C:\SecondObTbx\DTSLOS_lama_20200107_202106
Directory	C:\SecondObTbx
min. Latitude	45
max. Latitude	45.5
min. Longitude	10
max. Longitude	10.5
Number of Time Steps	5
Nodes Number of The LSTM Layer (Optional)	
Learning Rate (Optional)	
Number of Epochs (Optional)	
Plot Learning Curves (Optional)	<input type="checkbox"/>

The status bar at the bottom indicates the map scale is 1:195,444 and the coordinates are 10.1885651°E 45.5291775°N.



Conclusions

1st Obj.

- The methodology developed in this research produced meaningful datasets to identify ground displacements using machine learning, with the Cosine K-Nearest Neighbor algorithm demonstrating superior performance, especially in adjacent areas
- The application of pseudo-labeling notably improved validation accuracy, marking a significant advancement in geospatial analysis techniques
- The trained models worked consistently across three different geographical datasets, although further validation is needed
- Testing the Persistent Scatterer Interferometry technique using this workflow for further evaluation is recommended

2nd Obj.

- LSTM models are proficient in forecasting single or multiple steps in InSAR displacement time series, particularly for regular sequences.
- For irregular time series, employing advanced LSTM models that are sensitive to time, such as the Time Gated LSTM, is advisable despite the increased computational costs involved.

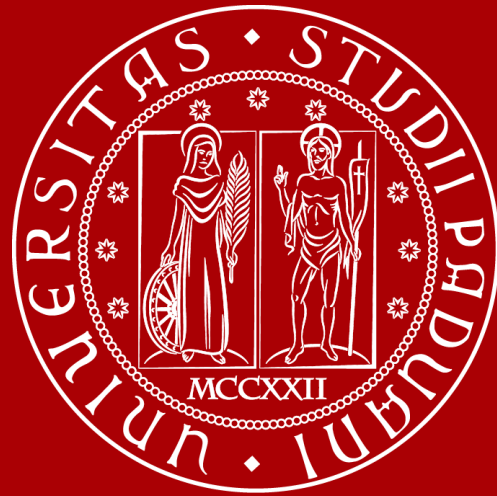
3rd Obj.

- The fusion of AI and InSAR within the GIS framework remains in its initial stages, and the developed toolbox has underscored the significance of such an interdisciplinary amalgamation



- [1] Crosetto, M., Castillo, M., & Arbiol, R. (2003). Urban subsidence monitoring using radar interferometry. *Photogrammetric Engineering & remote sensing*, 69(7), 775-783.
- [2] He, L., Wu, L., Liu, S., Wang, Z., Su, C., & Liu, S. N. (2015). Mapping two-dimensional deformation field time-series of large slope by coupling DInSAR-SBAS with MAI-SBAS. *Remote Sensing*, 7(9), 12440-12458.

Thanks for the attention



UNIVERSITÀ
DEGLI STUDI
DI PADOVA