AUTOMATED INSAR TIME-SERIES ANALYSIS TOOL FOR GEOLOGICAL INTERPRETATIONS IN NEAR-REAL TIME

Holly Hourston[1], Itahisa González Álvarez[1], Luke Bateson[1], Ekbal Hussain[1], Alessandro Novellino[1]

[1]British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

ABSTRACT

Large EO datasets are becoming increasingly challenging to analyse and interpret due to their size, a problem that will continue to worsen with longer satellite operating times. In this manuscript, we present an automatic tool to analyse interferometric synthetic aperture radar (InSAR)-derived deformation time series using a supervised machine learning regression algorithm. This processing tool at its current stage of development can map areas of anomalous ground motions, and different extents of seasonal behaviours. With the addition of British Geological Survey proprietary geological hazard susceptibility datasets, we can begin to make further interpretations of the ground movements. This algorithm will aid in the task of identifying areas of substantial long-term ground motions due to specific geological hazards. This will be applied specifically to vulnerable locations such as coastlines, and will be used to identify locations at risk of exacerbated seasonal ground motions due to the worsening effects of climate change.

Index Terms— Remote sensing, earth observation, time series analysis, machine learning, InSAR

1. INTRODUCTION

The ESA Sentinel-1 satellites have been operating together for 8 years, collecting interferometric synthetic aperture radar (InSAR) data of the Earth's surface to monitor ground deformations. InSAR data infers the movement of a planetary surface by measuring any phase differences between an outgoing and a reflected (or backscattered) radar signal- the backscattered signals are most coherent when reflecting off hard surfaces, such as roads, exposed outcrops, and buildings, making InSAR ideal for monitoring geological and anthropogenic ground motions, particularly in urban environments. The wealth of data obtained from these satellites, along with the high spatial resolution of InSAR observations (~15m), means that their time series can be used to monitor ground motions due to shallow geohazards around key areas such as infrastructure, high-risk fault lines, and homes. However, the quantity of Earth Observation (EO) data and its analysis is a time-consuming problem for remote sensing scientists, and automation is a necessary step for building near-real time understanding of how the ground surface is changing due to natural processes, anthropogenic pressures, and worsening climate change. Climate change in the UK is already causing drier summers and wetter winters- 7 out of the 10 hottest days on record in the UK have occurred in the 21st Century [1], and 7 out of the 10 wettest winters on record have been in the last 35 years [2].

Here, we present a Seasonal Trend Decomposition & Piecewise Linear Regression (STL-PLR) algorithm, which automates the analysis of InSAR deformation time series from Sentinel-1AB satellites and attaches a label to each InSAR measurement point based on its trend stability- or lack thereof. Firstly, the STL analysis separates out the general trend from any seasonality components and any residual components. Then, the PLR analysis examines the general trend pulled out from the STL and determines which measurement points are to be classed as "linear-stable", "linear-uplift", "linear-subsiding" and "non-linear" by identifying the number of breakpoints in each measurement point's deformation time series. Having a non-linear category will aid us in automatically identifying locations where there have been sudden changes, such as a landslide or new construction-related ground loading. PLR are regularly used in ground deformation analyses as they automatically indicate breakpoints in a time, which in our case signifies a change in deformation behaviour. PLR has been previously used to study the rates of acceleration or deceleration of landslides [3], and dam settlement subsidence monitoring [4]. Once we have the results of the trend behaviours belonging to each measurement point, these are then plotted to create a map of comparative movement through time.

The seasonal component extracted by the STL is also studied and compared to geological maps. A strong seasonal trend in this context is defined as any measurement point which has a greater variance in its seasonality component than variance in its residual component. The seasonality behaviour analysis has implications for shrink-swelling geologies, and other geological hazards which have seasonal behaviour, allowing us to identify the areas with ongoing shrink-swell activity and monitor these sites for potential changes in this behaviour.

We present the results of the algorithm for two test sites, one covering the Midlands region of England and demonstrating the effects of velocity thresholding and the initial outputs of the STL-PLR algorithm, and another covering the south coast of England and the Isle of Wight, where we demonstrate the seasonal analyses. The BGS's *GeoSure* [5] dataset consists of polygons of the most and least susceptible areas in Great Britain to 6 different geohazards: landslides, shrink-swell, running sands, compressible ground, collapsible deposits, and soluble rocks. *GeoSure* is based on the combination of BGS digital superficial and bedrock maps, the factors such as slope angles, and how significant these factors are thought to be at that location. The factors are combined to create an estimate of hazard susceptibility. By considering the susceptibility to various geohazards in our second test site, which contains known active geohazard locations, we validate this algorithm's ability to detect hazardous ground motion behaviours. Additionally, by cross-referencing with the hottest days of the years in our time series, we can begin to build a relationship between seasonality in InSAR, geology, and extreme heat. Our tool will aid in providing automatic interpretations of InSAR data in a geological context, and allow us to monitor new developments in ground motions in near-real time.

2. METHODOLOGY

The algorithm developed for this study was first tested using a sample set of European Ground Motion Service (EGMS) data [6] for the Midlands region of England, UK, containing time series data for almost 290,000 InSAR measurement points (MP) (Figure 1). For each MP, the algorithm first uses a Seasonal-Trend decomposition algorithm that uses Locally Estimated Scatterplot Smoothing (LOESS) [7] to extract the trend and seasonality components of the signal. An automated piecewise linear regression (PLR) is then used to fit the trend with a variable number of linear segments, depending on the trends in the displacement time series. This technique allows us to expand beyond the uplifting, subsiding or stable indicators given in the EGMS, and define a new category of "non-linear" moving points. The variable number of linear segments used in the PLR is determined by gradually increasing the complexity of the model by testing it with the addition of one more segment and comparing the resulting adjusted R^2 coefficient of both models. If the increase in the new model's adjusted R^2 is larger than 3%, then the more complex model is accepted, and again a new segment is added. This process continues until this increase in adjusted R^2 becomes equal to or smaller than 3%, at which point we consider the algorithm to have converged.

MPs for which the PLR algorithm required only one segment to fit their trend are then classified as "Linear", while those needing two or more segments classed as "Non-linear". Linear trends are further grouped into "stable", "subsiding" and "uplifting". A velocity threshold, accounting for any bias and uncertainties in the ground deformation measurements is built into the algorithm, based on the 1-sigma standard deviation of the ground deformation rate in the chosen dataset. Any measurement points with linear velocities within the (±threshold) range are classed as stable points, as they cannot be reliably distinguished from noise within this threshold range.

In order to make geological interpretations beyond classification of the general behaviour, other behaviours and data are needed in the analysis. We re-run the STL-PLR algorithm on a new location, one with various known geohazards and extensively studied geology- the British south coast and Isle of Wight (Figure 2). We investigate the seasonal decomposition by using an unsupervised k-means clustering machine learning algorithm from Sci-kit Learn [8]. Seasonality strengths are generated for each MP based on their variance and cyclicity, the more cyclical and consistent each peak and trough of seasonality the stronger the seasonality. Seasonality strength is defined on a range of 0 to 1, with 1 being the strongest. At this point we now have a cluster of the most seasonally affected MPs in this test site, and by considering overlaps with the most susceptible locations for geological hazards, we can start to build an understanding of what deformation time series looks like for various geological processes and for meteorologically influenced locations.

3. RESULTS

3.1. Test Site 1: the Midlands

Figure 1 shows the original ground displacement time series for three locations in our sample dataset that present different behaviours (linear-stable in Fig. 1a, linear-subsiding in 1b and non-linear in 1c), according to a 1σ velocity threshold. Besides the ground displacement data, the plots also contain the trend component obtained from the STL decomposition (red lines), the results of the PLR fitting and break points (blue dashed lines), and the average ground deformation velocity obtained for each segment. The measurement points in panels (a) and (b) present linear trend behaviour (meaning that the PLR algorithm required only one segment to capture the characteristics of their trends). The measurement point in (a) is a linear-stable point, and in (b) is classes a linear-subsiding point. On the other hand, the deformation at the measurement point shown in panel (c) required three segments to fully capture the trend behaviour and is therefore assigned to the non-linear classification. The adjusted R^2 obtained for all three cases ranges from 0.68 to 0.99, highlighting the ability of our automatic processing tool to successfully capture the characteristics of, and changes in, the original InSAR signal. Overall, the preliminary results for our first dataset suggest that our algorithm successfully analyses the original signal in its constituent parts, and provides novel spatio-temporal patterns from the original ground displacement time series.

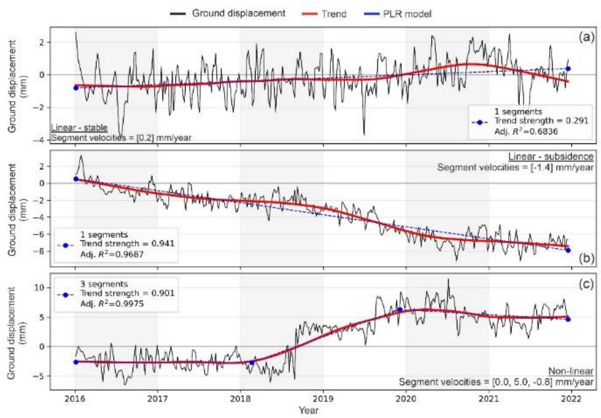

Figure 1: Three examples of measurement points that present different trend behaviour in test site 1- (a) is a linear-stable point, (b) is a linear-subsiding point, and (c) a non-linear point showing a significant uplift period between 2018-2020. The original ground displacement InSAR time series are the black lines. Red lines correspond to the trend component obtained in each case from the STL decomposition, while blue dashed lines represent the segments of average deformations following similar trends, determined by the PLR. The number of segments required by the algorithm to fit the trend are shown in legend boxes. Additional text inserts in each panel also show the trend behaviour label assigned to each location when applying a 1-σ velocity threshold (1.3 mm/year) and the ground velocity obtained for each segment of the PLR fit.

3.2. Test Site 2: the English Southern Coast and Isle of Wight

Our second test focusses on the seasonality component of extracted signals. A visual representation of trend label results is shown in Figure 2. The seasonal decomposition results of one measurement point are shown in Figure 3, which shows a break point in the late summer of 2021. When visually comparing the seasonal signal through time, there is an inconsistency in the seasonal component occurring in early 2022, and an unusually high seasonal signal is observed during the summer of 2022, coinciding with the heatwave seen across the UK. This point falls into the middle of the seasonality clustering groups identified by the SciKit-Learn k-means clustering. This is a good example of a location which has had considerable ground motion changes due to changing seasonal influence with extreme high heat.

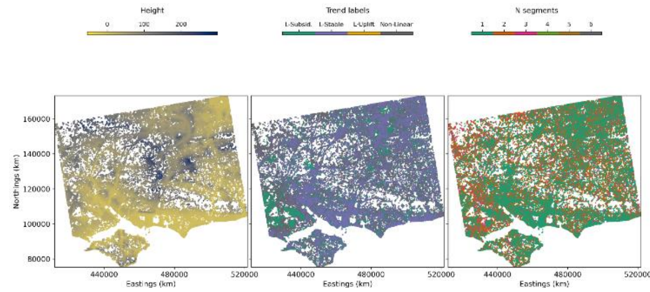

Figure 2: From left to right, an elevation plot, trend labels plot and number of segments plot for the InSAR measurement points observed from the South coast of England and Isle of Wight. The STL-PLR algorithm was run for a velocity threshold of 1.9 mm/year. As described in the text above, the non-linear points are any measurement point whose deformation time series requires more than 1 segment to describe the overall trend behaviour- this can be seen when comparing the "trend labels" and "N segments" subplots.

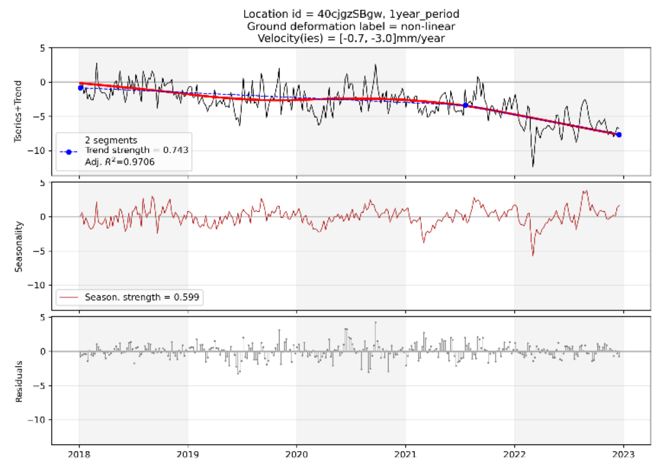

Figure 3: The deformation trend (top, red) with seasonality (middle) and residual (bottom) data subtracted via STL.

Next, we cluster together groups of InSAR measurement points to spatially map out regions of similar seasonality signals. Using the elbow method, we determined the optimal number of clusters to be three. We applied a k-means clustering algorithm to cluster on the strength of seasonality through time. This produced three clusters of similarly behaving seasonal trends, with the most seasonal (or "strongest seasonal strength") cluster including points with a seasonality strength in the range 0.73-0.99. In the most seasonal cluster 16% of the MPs correspond to linearly subsiding areas and 14% to non-linear deformation points, with the remaining points showing long-term stable deformation. This indicates that although the ground motion itself is considered stable at present for most points, these areas are at risk of future changes due to climate since their current state of stability is reliant on the current seasonality.

Our clusters are shown in Figure 4, which identify regions (in red) that are the most influenced by seasonal changes. These are areas that may be at-risk for future ground instabilities as our climate becomes drier/wetter. There is a clear red band along the southern coast, which indicates that the MPs in this zone have highly seasonal ground motions. Cross-referencing with *GeoSure* we found that of the two most susceptible classes for shrink-swell hazard, 29% of InSAR measurement points that overlap with these areas move non-linearly. For the two most landslide-susceptible classes, 41% of all the InSAR measurements which overlap with these areas are moving non-linearly. Additionally, this region contains clay and chalk-rich bedrock, both of which can become less stable with heavy rainfall or can shrink in high heat/dry weather conditions.

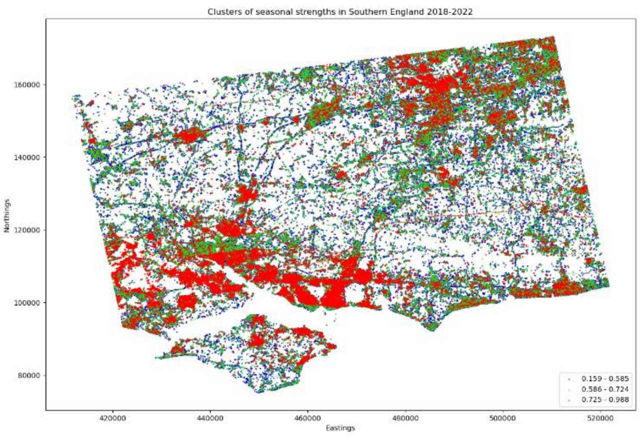

Figure 4: The overlay of three clusters of seasonal strengths in test site 2.

4. CONCLUSIONS

We have outlined a process-independent data analysis tool for interpreting spatio-temporal patterns in InSAR-derived deformation time series. The algorithm contains no geological or process-related information, nor does it assume any kind of spatial correlation. We have demonstrated the capability to determine the most seasonal ground deformations for any location covered by the EGMS, which enables determination of places at risk of seasonal instabilities due to future climate.

The output of this research will be an automated tool for the analysis of InSAR ground deformation time series, resulting in classification of subsiding, uplifting, stable, and non-linearly measurement points over the UK. The aim is to automate interpretation models for investigating and monitoring processes. Next steps include further refining the “non-linear” category, by building a database of characteristic deformation profile models for specific geohazards from real-world and AI-generated examples, and further development to understand misclassifications. In the future, this work will aid in creating an automatically updated dynamic model of the UK. We anticipate that this tool will continue to be developed along with BGS geological maps, with uses for geohazard monitoring, identification of at-risk areas at risk of exacerbated seasonal ground motions due to the worsening effects of climate change.

5. REFERENCES

[1] Met Office Press Office, “Record high temperatures verified”, accessed via webpage on 20th May 2024 via https://www.metoffice.gov.uk/about-us/news-and-media/media-centre/weather-and-climate-news/2022/record-high-temperatures-verified. 2022.

[2] Carbon Brief, “Analysis: How UK winters are getting warmer and wetter”, accessed via webpage on 20th May 2024 via https://www.carbonbrief.org/analysis-how-uk-winters-are-getting-warmer-and-wetter/. 2024.

[3] A. Urgilez Vinueza, A.L. Handwerger, M. Bakker, et al. “A new method to detect changes in displacement rates of slow-moving landslides using InSAR time series.” *Landslides.* Springer, Volume 19, pp. 2233-2247, doi: 10.1007/s10346-022-01913-8, 2022.

[4] M. Evers, A. Thiele, H. Hammer, E. Cadario, K. Schulz, S. Hinz. “Concept to Analyze the Displacement Time Series of Individual Persistent Scatterers.” *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, pp. 147–154, doi: 10.5194/isprs-archives-XLIII-B3-2021-147-2021, 2021.

[5] K A Lee, D Diaz Doce. “User Guide for the British Geological Survey GeoSure dataset: Version 8”. British Geological Survey Open Report, OR/17/050. pp. 18., 2017.

[6] M. Costantini et al., "European Ground Motion Service (EGMS)," *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, Brussels, Belgium, 2021, pp. 3293-3296, doi: 10.1109/IGARSS47720.2021.9553562. 2021.

[7] S. Skipper, and J. Perktold. “statsmodels: Econometric and statistical modeling with python.” *Proceedings of the 9th Python in Science Conference.* Austin, Texas, 2010.

[8] Pedregosa et al., “Scikit-learn: Machine Learning in Python”, JMLR 12, pp. 2825-2830, 2011