

A Data Mining Approach for Multivariate Outlier Detection in Postprocessing of Multitemporal InSAR Results

Matus Bakon, *Member, IEEE*, Irene Oliveira, Daniele Perissin, Joaquim Joao Sousa, and Juraj Papco

Abstract—Displacement maps from multitemporal InSAR (MTI) are usually noisy and fragmented. Thresholding on ensemble coherence is a common practice for identifying radar scatterers that are less affected by decorrelation noise. Thresholding on coherence might, however, cause loss of information over the areas undergoing more complex deformation scenarios. If the discrepancies in the areas of moderate coherence share similar behavior, it appears important to take into account their spatial correlation for correct inference. The information over low-coherent areas might then be used in a similar way the coherence is used in thematic mapping applications such as change detection. We propose an approach based on data mining and statistical procedures for mitigating the impact of outliers in MTI results. Our approach allows for minimization of outliers in final results while preserving spatial and statistical dependence among observations. Tests from monitoring slope failures and undermined areas performed in this work have shown that this is beneficial: 1) for better evaluation of low-coherent scatterers that are commonly discarded by the standard thresholding procedure, 2) for tackling outlying observations with extremes in any variable, 3) for improving spatial densities of standard persistent scatterers, 4) for the evaluation of areas undergoing more complex deformation scenarios, and 5) for the visualization purposes.

Index Terms—Data mining, InSAR, multivariate analysis, outlier detection.

I. INTRODUCTION

MULTITEMPORAL InSAR (MTI) techniques [1]–[7] are successfully applied in measuring of subtle deforma-

Manuscript received September 13, 2016; revised December 9, 2016, January 19, 2017, and February 21, 2017; accepted March 16, 2017. Date of publication April 11, 2017; date of current version July 17, 2017. This work was supported by the Slovak Grant Agency VEGA under projects 1/0714/15 and 1/0462/16 and by Portuguese FCT UID/AGR/04033/2013 and POCI-01-0145-FEDER-006958. TerraSAR-X data were provided by DLR under project ID LAN2833. Sentinel-1 data were provided by ESA under free, full, and open data policy adopted for the Copernicus program. Data have been processed by SARPROZ using Matlab, and Google Maps. (*Corresponding author: Matus Bakon.*)

M. Bakon and J. Papco are with the Department of Theoretical Geodesy, Slovak University of Technology, 810 05 Bratislava, Slovakia (e-mail: matusbakon@insar.sk; juraj.papco@stuba.sk).

I. Oliveira is with the Centre for the Research and Technology of Agro-Environmental and Biological Sciences, University of Tras-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal (e-mail: ioliveir@utad.pt).

D. Perissin is with the Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: perissin@purdue.edu).

J. J. Sousa is with the University of Tras-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal, and also with INESC TEC (formerly INESC Porto), 4200 Porto, Portugal (e-mail: jj.sousa@utad.pt).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2017.2686646

tions of both natural and man-made objects. The parameters of velocity, height, and others, sought as the ultimate MTI estimates, are commonly considered reliable when their ensemble coherence $\in [0, 1]$ exceeds a certain threshold of, e.g., 0.7 (see Fig. 1) and reaches the value of 1 [2], [4]. Loss of temporal coherence is commonly associated with the temporal and geometrical decorrelation. Noise from the signal delays caused by the atmospheric disturbances also prevents interferometric phase from being readable. Besides other reasons of inaccuracies such as subpixel positions, sidelobe observations, and orbit errors, there are difficulties in resolving nonuniform deformations. Possible scenarios include nonlinear movements such as high-phase gradients (like in the case of landslide activation process or earthquakes), seasonal patterns (e.g., thermal expansion due to temperature changes, dam oscillations related to the water level change) and other displacement-inducing effects, or a combination of more of them. Usually, only the eyes of InSAR experts are searching for the groups of scatterers that exhibit similar behavior, while evaluating their spatial relations and agreement of the estimated parameters within certain surroundings. New era of operational SAR satellites with frequent observations and enhanced swath coverage (Sentinel-1A), foreseen data boost from constellation missions (Sentinel-1B, TerraSAR Next Generation, Radarsat Constellation Mission, etc.), and nation-wide monitoring initiatives make this task more and more complicated. It is, therefore, of interest to reconsider, more closely, the practice of imposing simple threshold on ensemble coherence value and assess its full informative character, recognized in a range of thematic mapping applications. Although many advances have been achieved in exploiting low or partially coherent targets [8], [9], all efforts to evaluate higher order products often remain in the hands of end users, causing common concerns about the reliability of InSAR results by simply observing the locations of extreme velocities. To limit those concerns and possible misinterpretations, we would like to address the topic of a missing concept for finding statistically significant observations by removing those which appear outlying.

II. METHODOLOGY

Given a set of persistent scatterers (PSs) that have undergone standard MTI processing, we dispose of spatial data with multiple variables declared in every location. These variables define the velocity [see Fig. 1(a)], height or residual height,

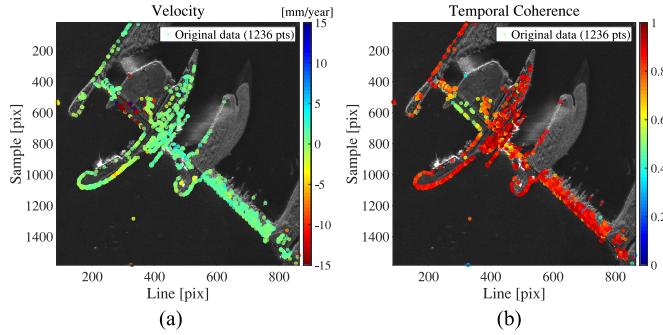


Fig. 1. (a) Velocity and (b) temporal coherence of the analyzed area.

their standard deviations, coherence [see Fig. 1(b)], and other parameters.

In our approach, spatial relations are analyzed among all variables employing multivariate analysis. It has to be noted that, for the current state of its development, we are neglecting the post-processing of deformation time series, thoroughly studied, for example, in [10] and [11]. Since a long history of equally sampled acquisitions is often required for the time-series analysis (which is usually not the case of former SAR missions like ERS, ENVISAT, etc.), our aim is to design a set of procedures that will, at the first level, eliminate multivariate outliers exploiting PSs' pointwise variables only.

Relatively little space has been given to the utilization of data mining techniques for InSAR measurements [11]–[20]. Moreover, none of these techniques have dealt with our primary focus of multivariate outlier detection and removal, and most of them:

- 1) are based on the assessment of a single variable, i.e., velocity;
- 2) rely only on displacement time-series analysis, producing less reliable outputs for smaller or nonequally sampled datasets while avoiding spatial relations among the points;
- 3) lack robustness for performing analysis solely on a set of highly coherent points;
- 4) require manual handling of the data in different software and geographic information systems;
- 5) incorporate distortion to the spatial composition of point networks by involving interpolation or gridding;
- 6) are designed as a single-shot activity with exploitation of one of the data mining techniques (e.g., clustering);
- 7) do not support the analysis of more complex deformation scenarios over low-coherent areas.

Building upon six respective techniques, namely density-based spatial clustering of applications with noise (DBSCAN) [21], [22], principal component analysis (PCA) [23], graph theory grouping [24], Voronoi diagrams [25], median absolute deviation (MAD) [26], and Jaccard index [27], our implementation is applicable regardless of the specifics linked with every location in the world or purpose of its monitoring or could serve as a base for doing so. Selected mathematical tools are intentionally kept as simple as possible in order to emphasize the interconnectivity among three main steps [see Fig. 2(a)] for extraction of location-, data-, and application-driven outliers.

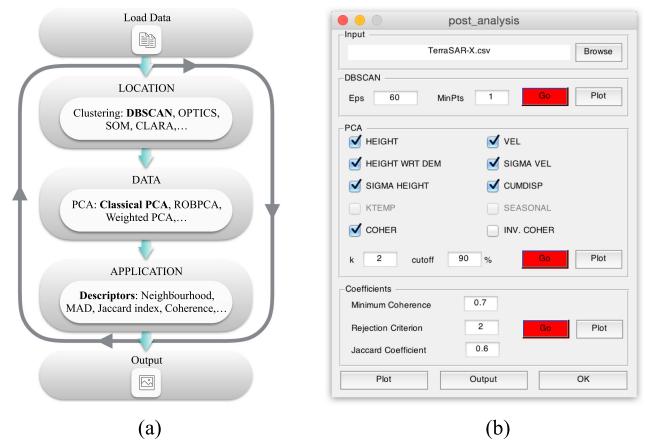


Fig. 2. (a) Flowchart and (b) GUI.

By iterating these key steps, while even involving advanced strategies to obtain clustered data structure and define outlier candidates (OCs) or changing descriptors to analyze local similarity or dissimilarity, it is possible to analyze different depths of the data in a seamless way and in various applications. A simple Matlab graphical user interface (GUI) [see Fig. 2(b)] has been built following these assumptions.

The tool is able to process comma-separated value files generated by SARPROZ software [28]; however, it is platform/application independent once the input is created following the simple structure of SARPROZ's output files. A description of the proposed methodology is accompanied by the results from monitoring of one of the biggest Europe's waterworks, Gabčíkovo–Nagymaros (part Cunovo) in Slovakia, through 52 TerraSAR-X images spanning years 2011–2013.

A. Location-Driven Outliers: Clustering Analysis

The estimation of PS parameters, such as velocities [see Fig. 1(a)], is performed within the system of equations utilizing several types of connection networks [29]. The connections (i.e., arcs), formed by all the scatterers in a study area, are often intentionally kept adjacent in order to decrease the impact of systematic errors [30]. In other words, systematic errors are causing the groups of PSs separated by a large distance to behave in a different manner from that of being anticipated in mathematical models amongst majority of points. One of many possible approaches to search for location-driven outliers is a clustering analysis.

For demonstration purposes, we adhere to a DBSCAN [21], [22] algorithm, mainly because of its mathematical simplicity and the ability to find clusters of arbitrary sizes and shapes together with detection of noise. As the whole process operates in 2-D space, image coordinates are used to define the location of points. Selection of the radius (Eps) for which the points are considered reachable is based on Euclidean distances computed between the pairs of PSs in the connection network. By plotting Euclidean distances in ascending order, it is possible to detect the knee of such a graph, expose deviating distances, and select the Eps parameter automatically [31] [see Fig. 3(a)].

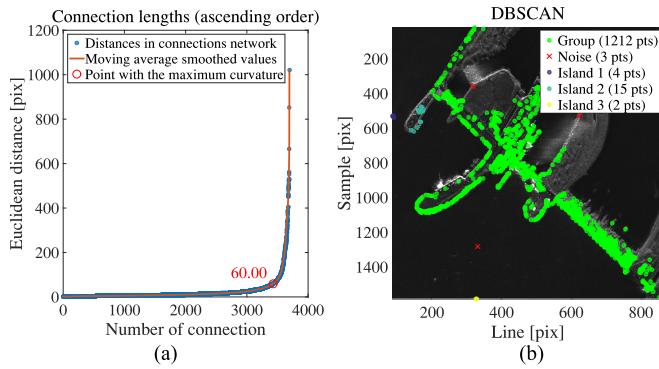


Fig. 3. (a) Selection of the Eps parameter and (b) clustered data structure obtained by DBSCAN.

The second input parameter required for a DBSCAN is a minimum number of neighboring points ($MinPts$) needed to form a cluster. The selection of the $MinPts$ parameter depends on the investigation strategy [31], [32] and preselection of the points for MTI processing. By DBSCAN, we can retrieve clusters of points classified as [see Fig. 3(b)]: GROUP, the core points of the dataset; ISLANDS, to be evaluated in the next steps; and NOISE, points to be discarded immediately when confirmed as outliers by PCA in the next step or kept when their coherence is greater than a selected minimum. Finally, we get a set of points with a clustered structure [see Fig. 3(b)], an advantageous one, provided that points allocated within the same cluster exhibit similar behavior, which will be analyzed further.

B. Data-Driven Outliers: PCA

One of the statistical tools capable of exposing multivariate outliers is PCA [23]. By mapping a high-dimensional space into a low-dimensional space, while retaining the maximum variability in terms of the variance–covariance structure, test limits could be applied in order to determine failing observations. The PCA technique has already been applied in a number of SAR/InSAR methodologies [11], [33]–[37]. However, the ambition here is different.

The goal is to detect outlying observations and, upon points classified in the previous DBSCAN step, track the behavior of all the points allocated within the same cluster in order to investigate why the points have become outliers. This could be the case of, for example, significant variance of residual height in comparison to the residual heights of majority of the points within the same cluster (ISLAND). Or the velocity could be extremely different from the core points (GROUP) of the dataset. The objective is to reveal multivariate outliers while keeping the statistical consistency of each cluster in any variable.

To distinguish between the regular observations and outliers for multivariate data, we construct a diagnostic plot [see Fig. 4(a)] according to [38]. For each observation, there are score distances and orthogonal distances to the PCA subspace.

To classify the observations, two cutoff lines are drawn, representing statistical confidence level (CL) of, e.g., 90%, for the points being OCs (see Fig. 5) when they exceed those limits.

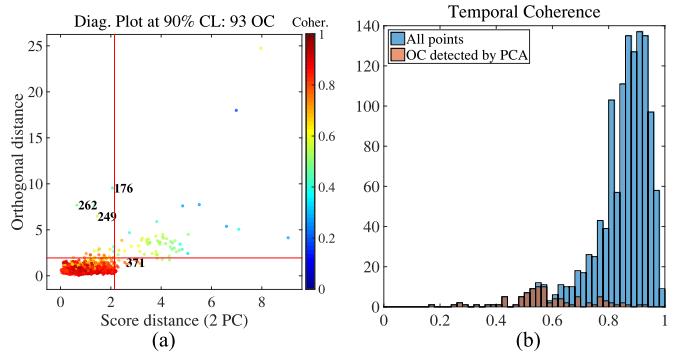


Fig. 4. (a) PCA analysis using diagnostic plot for retrieving OCs. (b) Histogram of the overall coherence and coherence covered by the OCs.

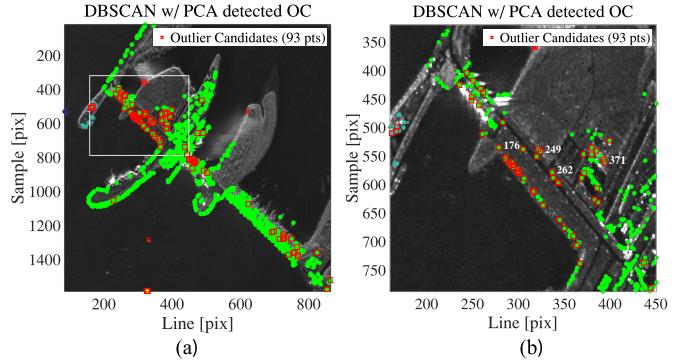


Fig. 5. (a) Clustered data structure subjected to PCA analysis. (b) Detail of the area depicted in white rectangle.

The observations in the diagnostic plot [see Fig. 4(a)] are colored proportionally to the temporal coherence and are accompanied by the histogram [see Fig. 4(b)] showing how much coherence has been assigned to points that have been identified as OCs. The whole procedure can be tuned by affecting the GUI [see Fig. 2(b)] and performed, even when only one variable is examined or different number k of principal components (PCs) is chosen to be retained.

C. Application-Driven Outliers: Decision-Making Process Using Descriptors for Quality Control and Local Similarity Analysis

Prior to the final analysis, the graph theory [24] is applied to find the connected components of undirected graph and separate the OC detected by PCA (only). Those outliers that have not been excluded as noise in the first round of DBSCAN and PCA are separated in the following way: 1) outliers without any outlier in the neighborhood are considered isolated; and 2) outliers with the presence of other outliers in the neighborhood form an outlier group (see Fig. 6).

Only the points that share Voronoi adjacency cells [25] and are within Eps radius are taken as neighboring. This way, though unfortunately computationally less efficient, the points could have different amount of “natural” neighbors in close surrounding to the limit of the distance (Eps) that implies boundary of noise.

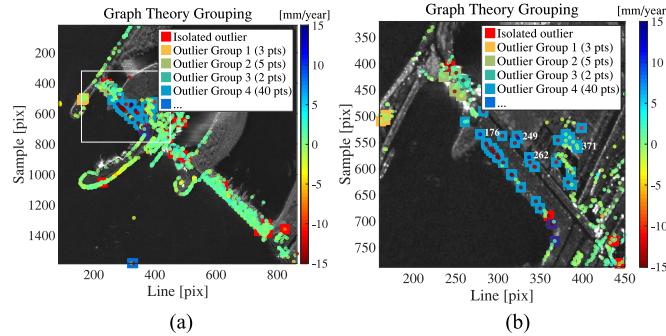


Fig. 6. (a) Outliers group creation with a detail of the area (b) within the white rectangle.

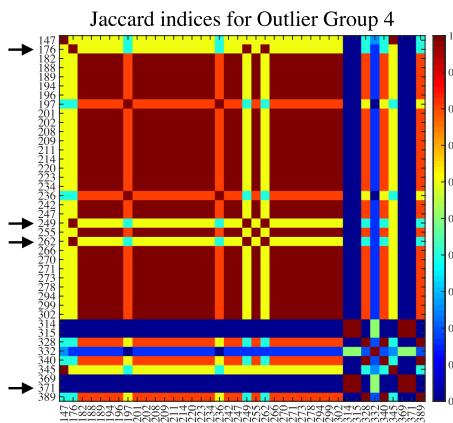


Fig. 7. Matrix of Jaccard indices.

TABLE I
 SAMPLE SET OF POINTS WITH VARIABLES FLAGGED 0 OR 1 IN ACCORDANCE
 WITH THE REJECTION CRITERION IMPOSED ON THEIR MAD

Point ID	Vel	σ Vel	Height	Res. height	σ Height	Cum. Disp.	Coher.
176	0	0	1	1	1	1	1
249	0	0	1	1	1	1	1
262	0	0	1	1	1	1	1
371	1	0	0	0	0	0	0
Majority	0	0	1	0	1	0	1

The abbreviations (Vel., Coher., etc.) stands for variables defined in each point (Velocity, Coherence, etc.).

The outlier structure created upon such principles (see Fig. 6) is then passed to the algorithm, and its performance is tested within allocated DBSCAN clusters. The variables for isolated outliers and grouped outliers are then flagged 0 or 1 (see Table I), according to MAD [26]:

$$\text{MAD} = b \cdot M_i(|x_i - M_j(x_j)|) \quad (1)$$

where x_i is a value of the variable in the inspected point, x_j is the $j = 1, 2, \dots, n$ observation of a nonoutlying neighbor or nonoutlying point remaining from the whole cluster, respectively. M is the median of the series and $b = 1.4826$ is a constant linked to the assumption of normality of the data [39]. For isolated outliers, if discrepancy in terms of exceeding the rejection

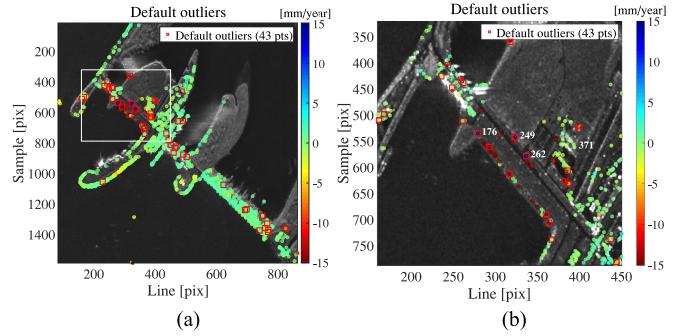


Fig. 8. (a) Propagation to default outliers with a detail (b) over white rectangle area.

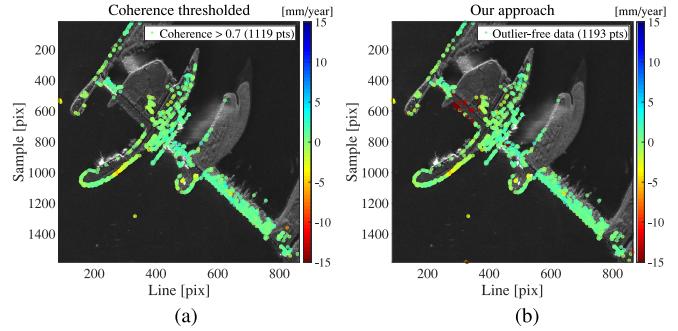


Fig. 9. (a) Velocities after imposing the threshold of 0.7 on temporal coherence value and (b) final postanalysis result.

criterion of 3 (very conservative), 2.5 (moderately conservative), or 2 (poorly conservative) [26]:

$$\frac{x_i - M}{\text{MAD}} > |\pm 3| \quad (2)$$

in any variable has not been found and the coherence of the inspected point is greater than a selected minimum, the point is kept in the dataset. Otherwise, it is recognized as the default outlier (see Fig. 8).

For OCs in groups, the Jaccard similarity coefficient [27] for each pair inside the group is computed (see Fig. 7) as

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (3)$$

evaluating final sample sets—the vectors A and B of zeros and ones for each variable in every point (see Table I), 1 as a flag for variable exceeding rejection criterion again, meaning that the point is breaking the rules of the allocated DBSCAN cluster in a given variable.

Point pairs with the Jaccard index (see Fig. 7) lower than a certain threshold (e.g., 0.6, where 0 stands for complete dissimilarity and 1 for complete similarity) are seized for the key step of the whole procedure: the minimum coherence value having a final word in deciding whether the point has an insufficient amount of coherence (lower than a chosen minimum, e.g., points no. 176, 249, and 262) and would be given away from the dataset, or the coherence is sufficient enough (higher than a minimum, e.g., point no. 371) for the point to be excluded—such point will be kept in a dataset as prone to be problematic and/or

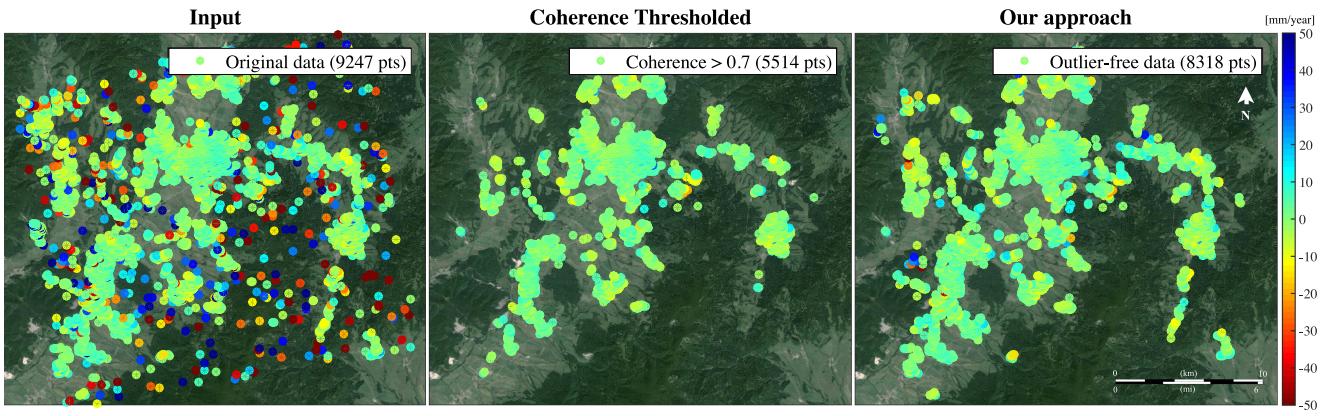


Fig. 10. Deformation map over area of active landslides in Prievidza, Slovakia.

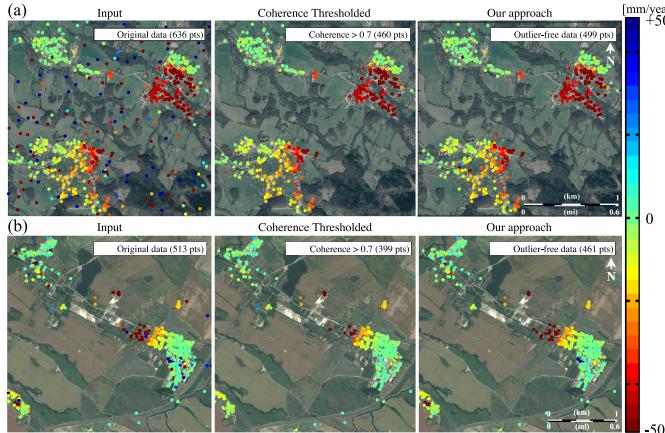


Fig. 11. Problematic areas of (a) Velka Lehotka and Hradec that rest exclusively upon active landslides and undermined area of (b) Kos municipality in the central part of Slovakia.

ambiguously integrated to the corresponding outlier group (see Fig. 8).

One could easily grasp the whole concept by following the behavior of the labeled points in Figs. 4–8 and Table I. Points that have been identified as default outliers (176, 249, and 262) show an uplifting motion unlike the majority of subsiding points within the same outliers group.

There are extreme cases when whole ISLANDS are indicated as outliers. Then, the points classified as GROUP are used for the computation of MAD and Jaccard index statistics for each inspected point. Likewise, a minimum coherence rule is applied in order to confirm the statistical consistency of parameters within the cluster in comparison to the majority of the points inside the dataset. Except for those extreme cases (e.g., when there are more outliers in the cluster than nonoutlying points and MAD statistics is biased, or the number of points to keep is lower than a minimum number of neighboring points needed to form a cluster *MinPts*, etc.), when whole clusters are indicated as default outliers, this key process is responsible for preserving the group of scatterers with similar statistical nature (see Fig. 9), even though their coherence is weakened and by the rules of standard thresholding procedure they would be discarded.

Observing the original deformation map [see Fig. 1(a)] and the results obtained by thresholding on coherence [see Fig. 9(a)],

there are 1119 out of 1236 points retrieved by the standard thresholding procedure. By applying the procedures described in this research, we have identified 1193 points [see Fig. 9(b)] with 74 new points belonging also to the problematic part of the dam. These points undergo a strong subsiding motion at the rates of -10 mm/year to -15 mm/year and have much larger deformation velocity, and their ensemble coherence is around 0.5. However, now they preserve consistent information in comparison to the results obtained by the thresholding procedure, where the total loss of information over this area occurs. The problematic part of the dam corresponds to the original riverbed of the Danube river, and it is under investigation due to persistent seasonal movements that are difficult to observe by using standard linear mathematical model assumption for persistent scatterer InSAR (PSInSAR) deformation estimates.

III. EXPERIMENTAL RESULTS

Relying on the C-band observations of Sentinel-1A, we would like to present the performance of our algorithm by the results obtained over the area of active slope failures in Prievidza, Slovakia. Performing PSInSAR processing (see Fig. 10) on 32 images from interferometric wide swath mode acquired along ascending track 175, we have identified 9247 scatterers. By imposing the standard threshold of 0.7 on ensemble coherence value, this amount decreased dramatically to 5514 PS points. However, by applying postprocessing analysis following the procedures proposed within this research, we have got 8318 scatterers, more than half of the total amount of standard PS points (see Fig. 10), that exhibit spatial and statistical dependence among themselves, as described in Section II. Thanks to this, the problematic areas could be assessed in more detail, as the deformation phenomena of these localities tend to be diminished by the standard thresholding procedure (see Fig. 10).

Let us now consider the performance of our algorithm over small areas showing deformation time series as well. The problematic areas of Velka Lehotka, Hradec, and Kos municipality processed by 38 images along Sentinel-1A's descending track 51 are depicted in detailed views (see Fig. 11).

Velka Lehotka and Hradec, the suburbs of Prievidza, rest exclusively upon active landslides [see Fig. 11(a)]. The area of Kos municipality [see Fig. 11(b)] is heavily undermined.

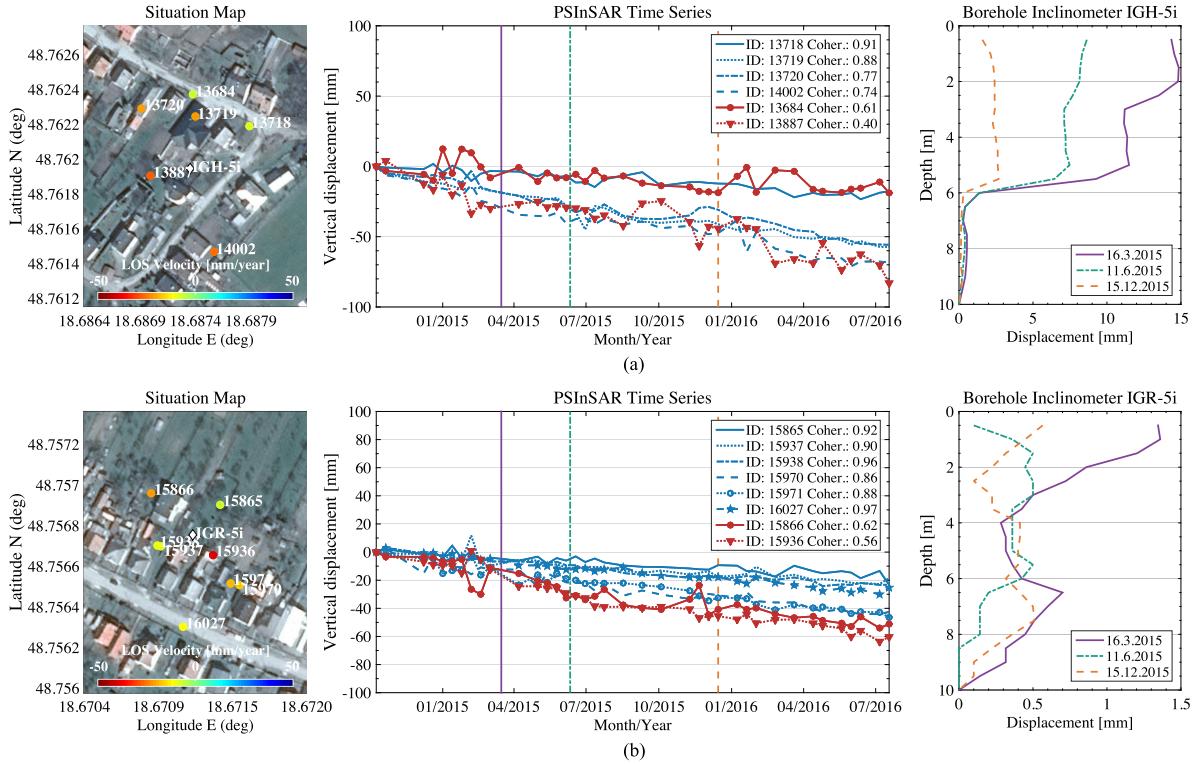


Fig. 12. Time series of low-coherent points (red) that have been preserved by applying the postprocessing procedure. Deformation signal of points with coherence greater than the standard threshold of 0.7 is shown in blue, and precise inclinometer measurements are in the separated graphs on the right.

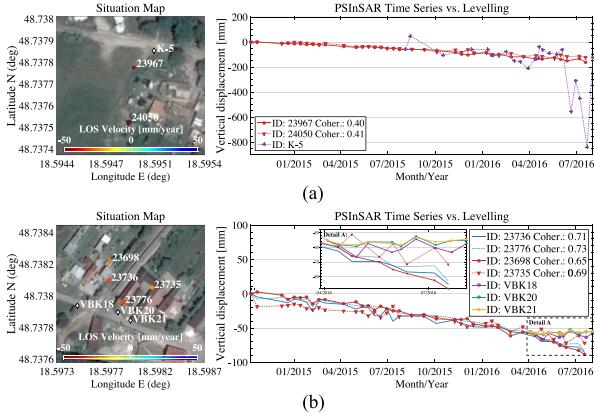


Fig. 13. Time series of low-coherent points (red) that have been preserved in final results after application of the postprocessing procedure in comparison to leveling measurements. Deformation signal of points with coherence greater than the standard threshold of 0.7 is shown in blue.

The 2012–2013 reactivation of slope deformations in Hradec and Velka Lehotka caused serious damage to the local infrastructure and buildings. The assessment of the physical activity in these locations has become crucial since the emergency conditions in June 2013. Current monitoring techniques focus on observation of the groundwater level and precise inclinometer measurements. While the method of precise inclinometer provides information on the deformation evolution directly on the shear zone, the InSAR observations enable to assess the deformation phenomena fully across the sliding areas. In the area of 4.6 km^2 , 636 radar scattering targets [see Fig. 11(a)] have been identified. Standard threshold of 0.7 on ensemble coherence has decreased this amount to 460 PS points. The approach



Fig. 14. Photodocumentation of structural deformation in monitored areas. (a) Disrupted structure in Hradec located close to the point ID 13684 [see Fig. 12(a)]. Photos courtesy of Google Street View. (b) Example of a damage to the infrastructure in Kos municipality over the subsiding part, where most of the points have been preserved by the algorithm [see Fig. 11(b)]. (c) Building crack in Kos municipality observed near the location of retained point ID 23735 [see Fig. 13(b)].

introduced in this study enables us to identify 499 scatterers [see Fig. 11(a)]. This has increased point densities in low-coherent areas for better evaluation of the complex deformation scenario.

The reliability of scatterers that have been preserved by the algorithm is proven by the inclinometer measurements (see Fig. 12). As the data from PSInSAR technique and inclinometer measurements are of different nature and the direct comparison is not possible, the separated graphs have been used instead. The graphs in Fig. 12 are showing the stage of the deformation for respective acquisition dates. Although the deformation time series for low-coherent point ID 13684, 13887 [see Fig. 12(a)] and 15866, 15936 [see Fig. 12(b)] are notably more unstable (red), they show a deformation trend similar to the points that

overpass the standard threshold of 0.7 (blue) [see Fig. 12(a)]. The inclinometer measurements prove that the areas are affected by movements. For the point 13887, the closest to the location of borehole inclinometer IGH-5i in Fig. 12(a), a more disruptive deformation signal is observed after the last inclinometer measurements. Several structures around it would have been left without any information if the standard thresholding procedure had been used, similarly to other 38 newly preserved points that might be crucial for identification of damaged structures. This is exactly the case of the point ID 13684 with time series shown in Fig. 12(a) and damage evidenced on nearby structures [see Fig. 14(a)].

The unfavorable conditions of slope stability in Kos municipality [see Fig. 11(b)] have been impaired by underground mining of brown coal. Mining has caused acceleration of mass movements and subsidence of the territory above mined-out underground spaces. Although the coal mining has been gradually ceasing out, the subsidence still represents an ongoing process. The area is regularly monitored by leveling measurements that have been used for the comparison with InSAR deformation time series in this study. As observable in Fig. 11(b), 513 radar scattering targets have been identified in the area, from which only 399 points can be regarded as PS points, thanks to their ensemble coherence higher than 0.7. Again, the outlier removal procedure has been applied preserving 461 out of 513 targets. As a result, there are 62 low-coherent targets that would be discarded by the standard thresholding procedure. The vast majority of these points is of extreme subsiding motion [see Fig. 11(b)].

Observing the deformation time series of these points as compared to the leveling measurements, the areas undergo rapid deformation changes that have been identified successfully again. Deformation rates of leveling point ID K-5 [see Fig. 13(a)] close to the end of monitoring period are undetectable by C-band observations using standard PSInSAR methodology with a linear model assumption as they exceed the interferometric phase detectable in one pixel (half of the radar wavelength per revisit period, i.e., 28 mm/12 days). Damages to structures and local infrastructure caused by sudden changes are documented by photos from the respective area [see Fig. 14(b)]. These changes are presumably the reason for noise incorporated in the deformation time series of point ID 23967 and 24050 [see Fig. 13(a)]. Another evidence of structures affected by strong subsiding motion is shown in Fig. 13(b). Not only the time series of points 23735 and 23698 correspond to the disrupted structures [see Fig. 14(c)], they began to be monitored by leveling measurements (VBK18, VPK20, and VPK21) only recently, in April 2016, due to persistent problems. In summary, low-coherent points that are preserved by our algorithm are valuable for detecting deformation changes.

IV. CONCLUSION AND FUTURE WORK

This paper presents a novel workflow for detecting outliers in postprocessing of MTI results. Tested upon Sentinel-1 data, this approach has shown its potential in increasing point densities of standard PS point networks. While preserving spatial dependence among low-coherent areas, the main benefits of this methodology include enhanced details visible in deforma-

tion maps, highlighted zones of scatterers that would require deeper investigation in terms of systematic errors mitigation, replacement of the frequency band, correction of the processing procedures, and others. The platform will help to interpret higher order MTI products by removing statistically insignificant observations, conserving the full informative character of the whole range of an ensemble coherence value. The multidisciplinary character of the proposed approach allows for modifying the procedures in order to operate with any heterogeneous 2-D point clouds of arbitrary high-dimensional variables. Along with the statistical analysis of time-series data, the state of the art of this approach will focus on making these procedures fully automatic and capable of predictions in order to assist the mitigation of inaccuracies and support the interpretation of higher order products such as displacement maps. Thanks to the availability of both sensing geometries over the studied area of Prievidza, the possibilities of joint postprocessing of ascending and descending geometries with decomposition of line-of-sight vectors will be investigated in future work.

REFERENCES

- [1] A. Ferretti, C. Prati, and F. Rocca, "Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2202–2212, Sep. 2000.
- [2] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.
- [3] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2375–2383, Nov. 2002.
- [4] B. M. Kampes, *Radar Interferometry: Persistent Scatterer Technique*. Dordrecht, The Netherlands: Springer, 2006.
- [5] C. Werner, U. Wegmüller, T. Strozzi, and A. Wiesmann, "Interferometric point target analysis for deformation mapping," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2003, vol. 7, pp. 4362–4364.
- [6] A. Hooper, P. Segall, and H. Zebker, "Persistent scatterer interferometric synthetic aperture radar for crustal deformation analysis, with application to Volcán Alcedo, Galápagos," *J. Geophys. Res.: Solid Earth*, vol. 112, no. B7, pp. 1978–2012, 2007.
- [7] A. Hooper, "A multi-temporal InSAR method incorporating both persistent scatterer and small baseline approaches," *Geophys. Res. Lett.*, vol. 35, no. 16, pp. 1–5, Aug. 2008.
- [8] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci, "A new algorithm for processing interferometric Data-Stacks: SqueezSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3460–3470, Sep. 2011.
- [9] D. Perissin and T. Wang, "Repeat-pass SAR interferometry with partially coherent targets," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 271–280, Jan. 2012.
- [10] L. Chang and R. Hanssen, "A probabilistic approach for InSAR Time-series postprocessing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 421–430, Jan. 2016.
- [11] R. Boni, G. Pilla, and C. Meisina, "Methodology for detection and interpretation of ground motion areas with the A-DInSAR time series analysis," *Remote Sens.*, vol. 8, no. 8, 2016, Art. no. 686.
- [12] C. Meisina *et al.*, "Geological interpretation of PSInSAR data at regional scale," *Sensors*, vol. 8, no. 11, pp. 7469–7492, 2008.
- [13] V. B. H. Ketelaar, *Satellite radar interferometry subsidence monitoring techniques*. Dordrecht, The Netherlands: Springer, 2009. [Online]. Available: <http://dx.doi.org/10.1007/978-1-4020-9428-6>
- [14] A. Spata *et al.*, "A clustering analysis of three dimensional ground deformation map estimated by integrating DInSAR and GPS dataset," in *Proc. 4th Int. Conf. Phys. Control*, Sep. 2009. [Online]. Available: <http://lib.physcon.ru/doc?id=7360a6f0112f>
- [15] G. Milone and G. Scipi, "A clustering approach for studying ground deformation trends in Campania region through PS-InSAR™ time series analysis," *J. Appl. Sci.*, vol. 11, pp. 610–620, 2011.

- [16] P. Lu, N. Casagli, F. Catani, and V. Tofani, "Persistent scatterers interferometry hotspot and cluster analysis (PSI-HCA) for detection of extremely slow-moving landslides," *Int. J. Remote Sens.*, vol. 33, no. 2, pp. 466–489, 2012.
- [17] M. Berti, A. Corsini, S. Franceschini, and J. Iannaccone, "Automated classification of persistent scatterers interferometry time series," *Natural Hazards Earth Syst. Sci.*, vol. 13, no. 8, pp. 1945–1958, 2013.
- [18] P. Lu, S. Bai, and N. Casagli, "Investigating spatial patterns of persistent scatterer interferometry point targets and landslide occurrences in the Arno river basin," *Remote Sens.*, vol. 6, no. 8, 2014, Art. no. 6817. [Online]. Available: <http://www.mdpi.com/2072-4292/6/8/6817>
- [19] D. Peduto, L. Cascini, L. Arena, S. Ferlisi, G. Fornaro, and D. Reale, "A general framework and related procedures for multiscale analyses of DInSAR data in subsiding urban areas," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 186–210, 2015.
- [20] D. Notti *et al.*, "A user-oriented methodology for DInSAR time series analysis and interpretation: Landslides and subsidence case studies," *Pure Appl. Geophys.*, vol. 172, no. 11, pp. 3081–3105, 2015.
- [21] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Discovery Data Mining*, 1996, pp. 226–231.
- [22] M. Daszykowski, B. Walczak, and D. Massart, "Looking for natural patterns in data—Part 1. Density-based approach," *Chemometrics Intell. Lab. Syst.*, vol. 56, no. 2, pp. 83–92, 2001.
- [23] I. T. Jolliffe, *Principal Component Analysis*. New York, NY, USA: Springer, 1986.
- [24] R. Tarjan, "Depth-first search and linear graph algorithms," *SIAM J. Comput.*, vol. 1, no. 2, pp. 146–160, 1972.
- [25] F. Aurenhammer, "Voronoi diagrams—A survey of a fundamental geometric data structure," *ACM Comput. Surveys*, vol. 23 no. 3, 1991, Art. no. 345405.
- [26] C. Leyls, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Social Psychol.*, vol. 49, no. 4, pp. 764–766, Jul. 2013. [Online]. Available: <http://dx.doi.org/10.1016/j.jesp.2013.03.013>
- [27] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Upper Saddle River, NJ, USA: Pearson, 2005.
- [28] D. Perissin, SARproZ software. Official Product Web Page, 2016. [Online]. Available: <http://www.sarproz.com/>
- [29] R. F. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis*. Dordrecht, The Netherlands: Kluwer, 2001.
- [30] R. Zhang *et al.*, "A hierarchical approach to persistent scatterer network construction and deformation time series estimation," *Remote Sens.*, vol. 7, no. 1, 2014, p. 211. [Online]. Available: <http://www.mdpi.com/2072-4292/7/1/211>
- [31] M. Elbatta and W. M. Ashour, "A dynamic method for discovering density varied clusters," *Int. J. Signal Process., Image Process. Pattern Recognit.*, vol. 6, no. 1, p. 123, 2013.
- [32] P. Sharma and Y. Rathi, *Efficient Density-Based Clustering Using Automatic Parameter Detection*. Singapore: Springer, 2016, pp. 433–441.
- [33] L. Ferro-Famil, A. Reigber, E. Pottier, and W. Boerner, "Multi-baseline polarimetric sar data classification using the complex wishart distribution and principal component analysis," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2001, vol. 6, pp. 2712–2714.
- [34] D. Singh, V. Chamundeeswari, K. Singh, and W. Wiesbeck, "Monitoring and change detection of natural disaster (like subsidence) using synthetic aperture radar (SAR) data," in *Proc. Int. Conf. Recent Adv. Microw. Theory Appl.*, Nov. 2008, pp. 419–421.
- [35] Y. August, D. Blumberg, and S. Rotman, "Identifying low reflection amplitude and low level phase noise points for permanent scatterer (PS) interferometry," in *Proc. IEEE Int. Conf. Microw., Commun., Antennas Electron. Syst.*, Nov. 2011, pp. 1–4.
- [36] H. W. Yun, J. R. Kim, C. Y. Soo, and H. S. Yoon, "The application of InSAR time series for landcover classification," in *Proc. Asia-Pacific Conf. Synth. Aperture Radar*, Sep. 2013, pp. 308–311.
- [37] G. Fornaro, S. Verde, D. Reale, and A. Pauciullo, "Caesar: An approach based on covariance matrix decomposition to improve multibaseline-multitemporal interferometric SAR processing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2050–2065, Apr. 2015.
- [38] M. Hubert, P. J. Rousseeuw, and K. Vandendriessche, "ROBPCA: A new approach to robust principal component analysis," *Technometrics*, vol. 47, no. 1, pp. 64–79, Feb. 2005. [Online]. Available: <http://www.tandfonline.com/doi/abs/10.1198/004017004000000563>
- [39] P. J. Rousseeuw and C. Croux, "Alternatives to the median absolute deviation," *J. Amer. Statist. Assoc.*, vol. 88, no. 424, pp. 1273–1283, 1993.



Matus Bakon (M'15) is working toward the Ph.D. degree with the Department of Theoretical Geodesy, Slovak University of Technology, Bratislava, Slovakia, with research interests primarily in deformation monitoring via satellite radar interferometry (interferometric synthetic aperture radar).

After gaining experience at Czech Technical University in Prague, Leibniz University Hannover, and Chinese University of Hong Kong, he worked with the University of Tras-os-Montes and Alto Douro (Portugal) where he devoted his attention to the development of alert and monitoring systems for physical structures. His research interests include data mining, principal component analysis, feature extraction, and pattern recognition in interferometric synthetic aperture radar.



Irene Oliveira received the Doctorate degree in statistics from the University of Aberdeen, Aberdeen, U.K., in 2004.

She is an Assistant Professor with the University of Tras-os-Montes and Alto Douro, Vila Real, Portugal, where she is also a Researcher with the Centre for the Research and Technology of Agro-Environmental and Biological Sciences. Her current research interests include multivariate statistics, more specifically, methodologies and extensions of principal component analysis and multivariate methods for environmental and ecological studies.



Daniele Perissin received the Ph.D. degree in information technology from the Politecnico di Milano, Milan, Italy, in 2006.

He is an Assistant Professor at the School of Civil Engineering, Purdue University, West Lafayette, IN, USA. He has authored a patent on the use of urban dihedral reflectors for combining multisensor interferometric synthetic aperture radar (InSAR) data and has published about 100 research works in journals and conference proceedings. He is the developer of SARPROZ software for processing multitemporal InSAR data.

Dr. Perissin received the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING Best Paper Award in 2012.



Joaquim Joao Sousa received the Ph.D. degree in surveying engineering from the University of Porto, Porto, Portugal, and Delft University of Technology, Delft, The Netherlands, in 2010.

He is an Assistant Professor with the Department of Engineering of Science and Technology, University of Tras-os-Montes and Alto Douro (UTAD), Vila Real, Portugal. He has authored national publications on geographic information systems (GIS) and topography. He is an Integrated Researcher with the Technology and Science Associate Laboratory (INESC TEC), Porto, and a member of UTAD's Urban Eco-Efficiency Unit. Recently, he has focused on the use of unmanned aerial vehicles for agroforestry applications. His main research interests include multitemporal interferometric synthetic aperture radar techniques for deformation monitoring and GIS.



Juraj Papco received the Ph.D. degree in geodesy and cartography from Slovak University of Technology, Bratislava, Slovakia, in 2010.

Since 2004, he has been a Researcher and Lecturer in global geodesy, physical geodesy, and remote sensing with Slovak University of Technology. He is also a Researcher at the National Centre of Earth's Surface Deformation Diagnostic in the area of Slovakia. As the principal investigator, he has an experience in managing five satellite-radar-interferometry-related projects, unique for Slovakia territory. His main research interests include satellite geodesy (global navigation satellite systems), remote sensing, satellite radar interferometry (interferometric synthetic aperture radar), deformation monitoring, gravity field modeling, gravimetry, land surveying, and geographic information systems.