

# A New Baseline Linear Combination Algorithm for Generating Urban Digital Elevation Models with Multitemporal InSAR Observations

Hui Luo, Zhenhong Li\*, *Senior Member, IEEE*, Zhen Dong\*, Peng Liu, Chisheng Wang, Jun Song

**Abstract**—The lack of high-resolution digital elevation model (DEM) data presents one major limitation for deformation mapping using synthetic aperture radar interferometry (InSAR) techniques with high-spatial-resolution radar imagery (e.g. TerraSAR-X). This paper presents a baseline linear combination (BLC) approach to generate interferograms with nearly-zero-baselines so as to minimize the effects of the uncertainties in the DEM used. It incorporates the baseline combination method with adjacent gradient networking to successfully unwrap the interferograms even in abruptly discontinuous areas, which in turn can be used to estimate a high-resolution DEM. The BLC approach does not require any deformation model; instead, it utilizes nearly-zero-baseline interferograms to assist with 3D phase unwrapping. Application of the BLC approach to the TerraSAR-X dataset in Shenzhen city, China shows that the BLC derived DEM agrees with the digital surface model (DSM) obtained from light detection and ranging (LiDAR) with a correlation coefficient of 0.998 and a root mean square error (RMSE) of 2.05m, demonstrating the effectiveness of the BLC approach. Note that the BLC approach is not only able to be employed in urban areas with high buildings, but also in mountain areas with steep slopes.

**Index Terms**—3D phase unwrapping, digital elevation model (DEM), multitemporal synthetic aperture radar interferometry (MTInSAR)

## I. INTRODUCTION

REPEAT-PASS interferometric synthetic aperture radar (InSAR) has been widely used to monitor surface displacements caused by natural hazards such as earthquakes, landslides and volcanoes, and by anthropogenic processes such as groundwater extraction, coal mining, and building load. Three-pass differential InSAR (DInSAR), without the need of topography information, has been proven a powerful approach for deformation monitoring [1]. More generally, an external digital elevation model (DEM) is required to remove the

topographic contributions in SAR interferograms (IFMs), particularly for those with a large spatial baseline (i.e. satellite separation). A range of DEMs such as the shuttle radar topography mission (SRTM) DEM, advanced spaceborne thermal emission and reflection radiometer global DEM (ASTER GDEM), and TanDEM-X DEM are freely available, but their spatial resolutions and accuracies are far from satisfactory for high-resolution InSAR applications, leading to the desperate demand for accurate DEMs with high resolution to precisely measure the surface displacements. Multitemporal InSAR (MTInSAR) has been developed to minimize the impacts of atmospheric disturbance and decorrelations, among which permanent scatterer InSAR (PS-InSAR) [2]-[4], small baseline subset (SBAS) [5], [6], and the combination of both techniques [7]-[9], have been proved powerful tools for DEM retrievals and deformation mapping. Major advances in MTInSAR have been made in recent years, including the development of SqueeSAR [10] and **persist scatterer pair** (PSP) [11]-[13].

Phase unwrapping is a vital step to resolve the interferometric phase ambiguity for both DEM generation and surface deformation mapping [14]. It can be straightforward when the Itoh smoothness condition is satisfied [15], i.e. the absolute value of the phase gradient is less than  $\pi$ . However, in the MTInSAR system, this condition is often violated due to the presence of big surface movements, strong atmospheric effects and/or large spatial gradients. A robust 3D phase unwrapping approach is needed for MTInSAR applications. Stanford method for PSs (StaMPS) [9], [16], [17] uses the phase evolution in the temporal dimension to guide unwrapping in the spatial dimension. It fully considers the temporal phase discontinuities caused by atmospheric effects and it has been successfully employed in many cases [18]-[20]. However, its

This work was supported by a Chinese Scholarship Council studentship awarded to Hui Luo and by National Natural Science Foundation of China (61771478). It was also supported by the UK NERC through the Centre for the Observation and Modelling of Earthquakes, Volcanoes and Tectonics (COMET, ref.: come30001) and the LICS and CEDRRIC projects (ref. NE/K010794/1 and NE/N012151/1, respectively), and by European Space Agency through the ESA-MOST DRAGON-4 project (ref. 32244).

H. Luo is with the college of electronic science of National University of Defense Technology, 410073 Changsha, China, and also with the School of Engineering, Newcastle University, NE1 7RU Newcastle, United Kingdom (email: luohui@nudt.edu.cn).

Z. Li is with the School of Engineering at Newcastle University, Newcastle NE1 7RU, United Kingdom (email: zhenhong.li@newcastle.ac.uk), and also

with the College of Geological Engineering and Geomatics, Chang'an University, Xian 710054, China.

Z. Dong is with the College of Electronic Science, National University of Defense Technology, 410073 Changsha, China (e-mail: dongzhen@vip.sina.com).

P. Liu is with the Department of Earth and Space Sciences, Southern University of Science and Technology, 518055 Shenzhen, China (email: liup3@sustc.edu.cn).

C. Wang is with the School of Architecture & Urban Planning, Shenzhen University, 518060 Shenzhen, China, (email: wangchisheng@szu.edu.cn)

J. Song is with the Integrated Geotechnical Investigation and Surveying Co., Ltd, 518040 Shenzhen, China (email: sdx1233@163.com).

\*Correspondence to: Z. Li (zhenhong.li@newcastle.ac.uk) and Z. Dong (dongzhen@vip.sina.com).

capability to reserve the height discontinuities is limited especially in urban areas where large phase gradients occur. In addition to StaMPS, a range of phase unwrapping and/or DEM retrieval approaches have been developed to handle the MT-InSAR data. Some of them are based on statistical estimation techniques, such as maximum likelihood (ML) [21]-[24] and maximum a posteriori (MAP) Bayesian estimation [25], [26]. For the ML approaches, exploiting both amplitude and phase information, multichannel phase and amplitude regularization (MCPAR) [27] reconstructs DEMs based on a new graph-cut optimization algorithm. It overcomes the heavy time consumption and memory occupation problems in the previous multichannel phase unwrapping (MCPU) [25], [28]. However, the assumptions of zero ground deformation and statistically independent IFMs, do not always hold true. MAP is characterized by good performances [25], [26]; MAP combined with Markov random fields [29] can recover topographic profiles affected by strong height discontinuities and noise can be rejected efficiently, yet is limited by the heavy computational burden.

Other techniques, such as minimum cost flow (MCF), Kalman filter (KF), and cluster analysis (CA), also have been developed and applied. The extended minimum cost flow (EMCF) and its improved versions [30], [31], make use of two Delaunay triangulations (in the temporal/spatial-perpendicular baseline plane and the azimuth/range plane) to successfully unwrap the phase in time and in space with MCF. However, it requires a deformation model and *a priori* information about topographic errors and velocities. The extended Kalman filter (EKF) approach has been adapted to work with multichannel interferometric stacks [32]-[34] but with difficulties handling the sharp height discontinuity issue because of its smooth surface hypothesis. [35] attempted to overcome the sharp discontinuity problem by implementing height discontinuity detection and postprocessing steps, but the non-deformation and distance related height gradient assumptions might not be consistent with the reality. The CA method is an approach for multibaseline phase unwrapping [36]-[38], which firstly clusters the pixels into groups and then finds for each group an ambiguity vector to unwrap the phases. It does not require any assumption on deformation model or phase gradients, but the  $N$ -dimensional-search to find optimal ambiguity vectors is computationally demanding and error-prone. Moreover, it requires a large number of pixels to ensure its accuracy.

Most recently, based on the work in [39], an integrated MTInSAR estimator [40] that combines the coherent point detection and phase unwrapping into a single step has been demonstrated. It can avoid the requirement of *a priori* deformation model, but its high computation burden brought about by the pixel-wise processing and freely connect network (FCN) checking is quite discouraging.

In this paper, for the first time, we present a non-model based approach to unwrap phase and to retrieve urban DEMs through the baseline linear combination (BLC) of IFMs with different spatial baselines. Baseline combination (BC) method was firstly proposed to reduce atmospheric effects on InSAR observations so as to generate high-resolution DEMs [41].

However, by summing up IFMs, the system is vulnerable to the increased noise for atmospheric estimation, and worse still it also requires unwrapped IFMs. The BC approach also shows potential for the estimation of ice motion in Antarctica, but still, accurate unwrapped IFMs are required and a constant displacement model is adopted [42]. The BLC method in this paper integrates the small baseline IFMs and the linear-combined ones with nearly-zero baselines to limit the phase errors caused by the inaccurate DEM. It avoids the use of inaccurate deformation model in the phase unwrapping steps and incorporates unwrapping and topography calculation with full resolution in a single operation.

The BLC approach starts with a phase unwrapping procedure, which is based on a two-step (1D temporal + 2D spatial) processing method, similar to the MCF phase unwrapping method. In particular, the key idea is to first carry out, for each arc connecting neighboring coherent pixels (CPs) or PS candidates (PSCs), a temporal phase unwrapping operation for temporally adjacent IFMs, based on the real and equivalent combined "IFMs" with very small baselines. The second step relies on the use of these results as a starting point for spatial phase unwrapping performed on each temporally adjacent IFM. In these steps, no *a priori* information is required. IFMs with very small baselines are generated using BLC to suppress the DEM-introduced topographic errors and the spatial differences between neighboring CPs and/or PSs are calculated to eliminate atmospheric errors. Furthermore, an outlier detection is implemented to make sure that all the remaining arcs and PS points are reliable. It involves the networking approach to define how to carry out the spatial differences. We propose an enhanced Delaunay triangular network (EDTN), with high reliability and acceptable computation burden. Finally, with the 3D unwrapped phase, the DEM errors can be easily retrieved. Comparing our results with digital surface model (DSM) obtained by light detection and ranging (LiDAR), a high correlation of 0.998 and a low RMSE of 2.07m are obtained.

The paper is organized as follows. Section II briefly explains the formulation of the proposed algorithm. Section III thoroughly describes the processing chain. Section IV is dedicated to the experimental results. In Section V, discussions are addressed. Finally, conclusions are presented in Section VI.

## II. ALGORITHM FORMULATION

### A. Interferometric phase and SBAS time series

We start here with the standard IFM formulation. After flattening the Earth and removing local topography, the unwrapped interferometric phase at pixel  $(x, r)$  computed from the SAR acquisitions at  $t_M$  (for the master image) and  $t_S$  (for the slave image), can be modeled as the summation of the phase contributions caused by deformation, topographic residual, atmospheric delay and other noise [1]. The topographic contribution can be written as follows

$$\phi_{topo}^{t_M, t_S}(x, r) = \frac{4\pi}{\lambda R \sin \theta} B_{t_M, t_S}^\perp \varepsilon(x, r) \quad (1)$$

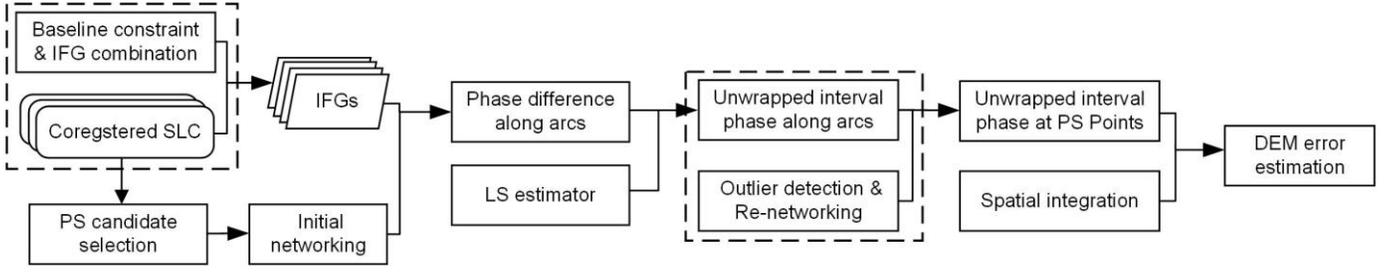


Fig. 1. Flowchart of the proposed baseline linear combination method. Dashed boxes imply the main steps characterizing the proposed method.

where  $\lambda$  is the radar wavelength,  $R$  is the sensor-target distance,  $\theta$  is the local incidence angle,  $B_{TM,TS}^\perp$  is the perpendicular baseline, and  $\varepsilon(x, r)$  is the topographic error present in DEM used in the interferometric processing. Note that the dependence on  $(x, r)$  variables is hereinafter not explicitly mentioned as the proposed technique implies a pixel-wise temporal analysis.

Referring to Equation (1), the DEM-introduced error is proportional to the perpendicular baseline and DEM error. To mitigate the spatial decorrelation phenomena, IFMs with small baselines are preferred to be used in time series analysis to minimize the effects of decorrelation and inaccuracies in the DEM used, i.e. SABS. Considering  $N+1$  single look complex images (SLCs) in the SAR stack and arranged chronologically as  $SLC_0, SLC_1, \dots, SLC_N, [t_0, t_1, \dots, t_N]^T$  the corresponding vector of SAR acquisition dates, we have  $M$  wrapped IFMs  $\boldsymbol{\varphi} = [\varphi_0, \varphi_1, \dots, \varphi_M]^T$ . With the unwrapped IFMs  $\boldsymbol{\Phi} = [\phi_0, \phi_1, \dots, \phi_M]^T$ , the components of the deformation and the DEM error can be jointly estimated in a least square (LS) sense [5]

$$[\mathbf{B}\mathbf{M}, \mathbf{b}] \mathbf{p}_c = \boldsymbol{\Phi} \quad (2)$$

where  $\mathbf{B}$  is an  $M \times N$  matrix linking the velocity vector and the IFMs stack,  $\mathbf{M}$  is the matrix models the deformation,  $\mathbf{b}^T = (4\pi/\lambda)[B_1^\perp/r \sin \vartheta, \dots, B_M^\perp/r \sin \vartheta]$ , and  $\mathbf{p}_c = [\mathbf{p}_v^T, \varepsilon]$  is a vector contains parameters accounting for the mean velocity and DEM error. The design matrix  $\mathbf{B}$  is with the generic  $(p, q)$  element  $\mathbf{B}(p, q) = t_{q+1} - t_q$  for  $IS_p \leq q \leq IE_p$ ,  $\forall p = 1, \dots, M$  and  $\mathbf{B}(p, q) = 0$  elsewhere, where  $IS_p$  and  $IE_p$  are the time-indexes of the slave and master images forming the  $p^{th}$  IFM, respectively.

### B. Baseline combined IFM

IFMs with small baselines are often exploited for the convenience of phase unwrapping and mitigation of decorrelation. Unfortunately, small baseline IFMs are not always available. Here we introduce a baseline linear combination (BLC) method to generate small baseline IFMs so as to suppress the DEM-introduced topographic errors. Besides the IFMs with small baselines, the ‘‘combined’’ or ‘‘equivalent’’ IFMs with small baselines are also considered. To illustrate it, we have the combined IFMs as

$$\tilde{\phi}_j = m\phi^k + n\phi^l \quad (3)$$

$$\tilde{\phi}_j = \frac{4\pi}{\lambda R \sin \theta} \tilde{B}_j^\perp \varepsilon + \frac{4\pi}{\lambda} \tilde{d}_j + \tilde{\phi}_{residual} \quad (4)$$

where  $\tilde{B}_j^\perp$  is the linear combination of the perpendicular baselines of two IFMs with  $\tilde{B}_j^\perp = mB_k^\perp + nB_l^\perp$  and  $\tilde{d}_j$  is the sum of the line of sight (LOS) displacements with  $\tilde{d}_j = md_k + nd_l$ . Theoretically,  $\tilde{B}_j^\perp$  can be infinitely approximated to 0 if the integer  $m/n$  is not restricted. In reality, to minimize the noise propagation, only the combinations with  $m, n = \pm 1$  are recommended, i.e. summing or subtracting two IFMs, if the data stack permitted; please see the detail on how to choose  $m$  and  $n$  in the last paragraph of this section. Without loss of generality, hereafter IFMs are referred to both types of IFMs, namely the original IFMs with small baselines, and the newly generated IFMs by BLC, if not specifically declared.  $\tilde{\phi}_{residual}$  mainly consists of atmospheric delays and thermal noise. Atmospheric delays can be largely canceled by differencing the phase between neighboring pixels [2], [3], and the differenced phase (phase along the arcs) can be expressed as

$$\Delta\phi = \Delta\phi_{def} + \Delta\phi_{topo} + \Delta\phi_{APS} + \Delta\phi_{noise} \quad (5)$$

When the two pixels are close enough (say  $d < 1$  km), the 3<sup>rd</sup> and 4<sup>th</sup> terms on the right hand of Equation (5) can be considered in a residual phase as

$$\Delta\phi = \Delta\phi_{def} + \Delta\phi_{topo} + \Delta\phi_{residual} \quad (6)$$

For each given arc, what we can observe are the  $M$  wrapped gradients of the interferometric sequence  $\Delta\boldsymbol{\varphi} = [\Delta\varphi_1, \Delta\varphi_2, \dots, \Delta\varphi_M]^T$ . If the wrapped phases are successfully unwrapped, Equation (6) can be easily obtained in a least square (LS) sense. Unfortunately, in urban areas, especially in big cities with high-rise buildings, DEM errors can be up to hundreds of meters, thus leading to dense phase fringes even with a relatively small baseline (say 100m), which makes phase unwrapping a big challenge. In the BLC case, the (equivalent) spatial baseline can always be small enough so that  $\Delta\phi$  can be considered within half a fringe, namely  $|\Delta\phi_{topo}| < \pi$ . Assuming that there is no earthquake, landslide or something similar during the whole time span,  $\Delta\phi_{def}$  can be considered small. For the above-mentioned reasons, we can simply consider  $\Delta\phi \triangleq \Delta\varphi$ . Note that  $\Delta\phi_{def}$  can exceed  $\pi$  in both original and combined IFMs if large displacements happen. To deal with this, an outlier detection is adopted to check the continuity, as will be shown in Section III-B. As we focus on urban applications in this paper, large displacements will then

not be considered in the rest of this paper.

For a dataset constituting of  $M$  IFMs (including  $M_1$  original IFMs and  $M_2$  combined IFMs) constructed from  $(N+1)$  acquisitions, the unknown temporally adjacent phase difference of each arc  $\Delta\theta = [\Delta\varphi_1 - \Delta\varphi_0, \Delta\varphi_2 - \Delta\varphi_1, \dots, \Delta\varphi_N - \Delta\varphi_{N-1}]^T$  and  $\Delta\varphi$  can be linked by a design matrix  $\mathbf{H}$ , as shown by

$$\mathbf{H}\Delta\theta = \Delta\varphi + \Delta\varphi_{noise} \quad (7)$$

where  $\mathbf{H}^T = [\mathbf{H}_1^T, \mathbf{H}_2^T]$  with  $\mathbf{H}_1$  and  $\mathbf{H}_2$  the design matrices describe the formulation of the original and the combined IFMs, respectively. The generic  $(p, q)$  element of  $\mathbf{H}_1$  is  $\mathbf{H}_1(p, q) = 1$  for  $IS_p \leq q \leq IE_p, \forall p = 1, \dots, M_1$ , and  $\mathbf{H}_1(p, q) = 0$  elsewhere.  $\mathbf{H}_2$  is regarded as the linear combination of 2 matrices  $\mathbf{H}_{21}$  and  $\mathbf{H}_{22}$  with the same structure as  $\mathbf{H}_1$ . The LS solution of the temporally adjacent phase difference can be obtained as

$$\Delta\hat{\theta} = (\mathbf{H}^T\mathbf{H})^{-1}(\mathbf{H}^T\Delta\varphi). \quad (8)$$

Equation (8) requires the full rank of  $\mathbf{H}$  to ensure the nonsingularity of the system. In Equation (3), theoretically,  $m/n$  can be any integer wanted, but  $\pm 1$  and  $\pm 2$  are recommended to minimize the noise propagation (please refer to the negative effects resulted from the large  $m$  and/or  $n$  values in Section V-C). Actually, the values of  $m/n$  greatly depend on one's own data stack, especially the baseline distribution, the number of SAR acquisitions, and the baseline span. Here, we give a brief guidance on how to carry out the combination with a given data stack (Table I). Firstly the combination is conducted under  $m/n=\pm 1$  only. If the nonsingularity requirement is not met,  $m/n$  is then enlarged until  $\mathbf{H}$  is nonsingular. With the nonsingular  $\mathbf{H}$ , combination optimization can be carried out by abandoning some unnecessary combinations in Step 4. It is easy to understand that once  $\mathbf{H}$  is nonsingular, there will be many redundant combinations. Some of them can be abandoned if their equivalent temporal baselines ( $\tilde{t} = |mt_k| + |nt_l|$ ) are too long. Of course, the abandonment could only be done when the nonsingularity and robustness of  $\mathbf{H}$  is ensured. The robustness means that with any single combination abandoned,  $\mathbf{H}$  is still nonsingular. The nonsingularities of  $\mathbf{H}$  can vary from one data stack to another, and the readers can refer to the simulated results in Appendix B, where the nonsingular probability of  $\mathbf{H}$  is given under different numbers of SAR acquisitions and baseline spans.

TABLE I  
PROCEDURE OF THE BASELINE COMBINATION

1.	Initialize $m/n=\pm 1$ .
2.	System formation. Do the baseline combination, choose the ones with $ \hat{B}_{j\perp}  < B$ , and get $\mathbf{H}$ .
3.	Singularity check. If $\mathbf{H}$ is singular, enlarge $m/n=\pm( m/n +1)$ , and jump to Step 2.
4.	Combination optimization. Eliminate redundancies according to one's own need.

### III. PROCESSING CHAIN

In this section, we summarize the main processing elements of the proposed baseline linear combination chain (Fig. 1).

#### A. Temporal coherent pixel (TCP) detection

Points of interest are those deemed coherent throughout all the IFMs used in the time series, including PSs and distributed scatterers (DSs). For simplicity, we only consider PSs in this paper as the dense PS distribution in the urban area is adequate for the processing. Of course, DSs can also be considered if there are not enough PSs, which is beyond the scope of this paper. The methods used for PS detection basically follow the amplitude dispersion index (ADI), namely standard deviation (SD) to mean ratio, proposed by Ferretti *et al.* [3]. Empirically a pixel is considered a PS candidate (PSC) only with an ADI less than 0.25. With these PSCs, a 'mask' map can be generated at this stage (see Fig. 5(b)).

#### B. Initial network formation and outlier detection

An initial Delaunay triangulation is computed, involving the arcs connecting the neighboring pixels of the selected PSCs in the azimuth/range plane. A first refinement is conducted by limiting the Euclidian distance of arcs within a given threshold (usually no more than 1km). Such a triangulation may involve the arcs relevant to SAR data pairs with large residuals (refer to Equation (9)). For each arc, the temporally adjacent phase of each arc can be obtained by Equation (8), and the residual phase becomes

$$\hat{\phi}_{res} = \Delta\varphi - \mathbf{H}\Delta\hat{\theta}. \quad (9)$$

Note that, Equation (8) holds true only when the assumption  $|\Delta\phi^j| < \pi$  is satisfied. If the residual is large,

$$\max\{|\hat{\phi}_{res}|\} > T_{\phi_{res}} \quad (10)$$

it is likely that the assumption is violated, and the corresponding arc should be discarded. Accordingly, to avoid these undesirable effects without losing the triangulation representation, a straightforward solution is to remove the arcs with large residuals [39] [43]. Here we prevent an unwrapping path with phase ambiguity instead of the path between singularities. The problem left here is to choose the residual phase threshold.

The design matrix  $\mathbf{D}$  linking SLCs and IFMs can be obtained as

$$\mathbf{D}^T = [\mathbf{D}_1^T, \mathbf{D}_{21}^T \pm \mathbf{D}_{22}^T] \quad (11)$$

where the generic element of  $\mathbf{D}_{1,21/22}(p, IS_p) = -1$ ,  $\mathbf{D}_{1,21/22}(p, IE_p) = 1$  with  $IS_p$  and  $IE_p$  the time-indexes of the slave and master images forming the original IFMs,  $\mathbf{D}(p, q) = 0$  elsewhere. Noted that  $\sum_p \mathbf{D}(p, q) = 0$ .

For PS points involved in the system, the noise covariance can be expressed as  $\mathbf{V}_{noise}^i = \sigma_{PS}^2 \mathbf{I}_{N+1}$ , if identical and uncorrelated noise is assumed, where  $\sigma_{PS}$  is the noise level of a single PS point and  $\mathbf{I}$  is the identity matrix. Then the noise

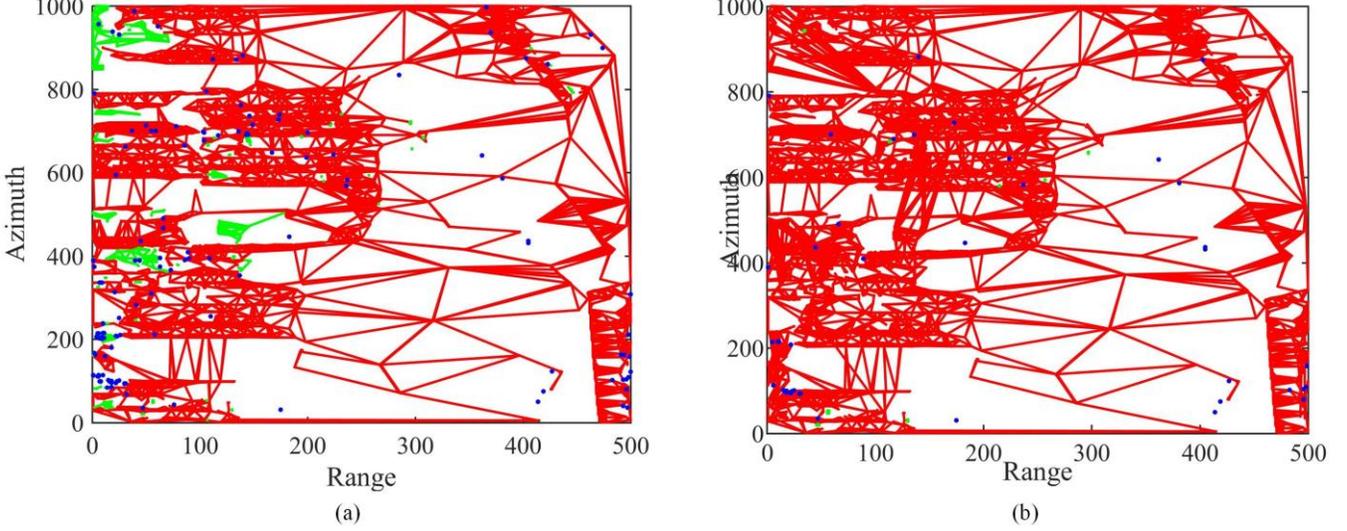


Fig. 2. Network illustration (a) DTN after outlier detection. (b) EDTN after re-network. Red lines and their vertices denote the arcs and points of the main network, the green lines and their vertices present the arcs and points of the subnetworks, and the blue points are the isolated points. Green and blue dots are the abandoned points. Note that only a small part of study scene is shown here for demonstration purposes.

variance of the  $M$  IFMs can be written as

$$\mathbf{V}_{PS}^{IFM} = \mathbf{D}\mathbf{V}_{noise}^i\mathbf{D}^T = \sigma_{PS}^2\mathbf{D}\mathbf{D}^T. \quad (12)$$

The noise variance of the arc doubles as  $\mathbf{V}(\Delta\varphi) = 2\sigma_{PS}^2\mathbf{D}\mathbf{D}^T$ . From Equations (7) and (8),  $\Delta\hat{\varphi} = \mathbf{H}\tilde{\mathbf{H}}\Delta\varphi$  can be obtained, where  $\tilde{\mathbf{H}} \triangleq (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$ , and its variance can be written as

$$\mathbf{V}(\Delta\hat{\varphi}) = 2\sigma_{PS}^2\mathbf{H}\tilde{\mathbf{H}}\mathbf{D}\mathbf{D}^T\tilde{\mathbf{H}}^T\mathbf{H}^T \quad (13)$$

Often the threshold  $T_{\phi_{res}}$  is

$$T_{\phi_{res}} = c\sqrt{\max(\mathbf{V}(\Delta\varphi))} + 2\sqrt{\max(\mathbf{V}(\Delta\hat{\varphi}))} \quad (14)$$

where constant  $c$  can be chosen as 3 or 4 [44], [45]. When Equation (14) is reached, the arc is regarded as an outlier at 95% confidence level.

### C. PS network formation: Enhanced Delaunay Triangular (EDTN)

Such an outlier detector described above may cause disconnected triangulation, bringing about either isolated points or isolated islands only internally connected (see the blue dots and green subnetworks in Fig. 2(a) respectively). To construct a fully connected network, here we introduce an enhanced Delaunay network (EDTN) method as follows.

After the outlier detection, the remaining arcs and points are denoted by  $\Sigma$  and  $\Pi$  separately, among which, there are 3 categories,  $\Sigma_0^M$  and  $\Pi_0^M$  of the main network (red lines and their vertexes in Fig. 2(a)),  $\Sigma_0^S = \{\Sigma_0^{S_i}\}_{i=1,\dots,K_0}$  and  $\Pi_0^S = \{\Pi_0^{S_i}\}_{i=1,\dots,K_0}$  of the secondary networks (green lines and their vertexes in Fig. 2(a)), and  $\Pi_0^I$  the isolated points (blue dots in Fig. 2(a)). Note that only one part of the studied scene is shown here for demonstration purposes.

Firstly, we connect the points in each  $\Pi_0^{S_i}$  to the nearest  $K$  points (for computational reasons,  $K < 100$ ) in  $\Pi_0^M$  followed by an outlier detection. If at least two points with arcs remaining

TABLE II  
COMPARISON BETWEEN DNT AND EDTN

Method		DTN	EDTN	
Before closure detection	Main network	Arcs Points	21,982 8,238	24,151 8,941
	Secondary network	Number Points	60 635	8 22
	Isolated points		134	44
	Abandoned points		769	66
	<b>Residues per interferogram</b>		<b>0.42</b>	<b>0.5</b>
After closure detection	Remained points		8238	8941
	Remained arcs		21,978	24,146

(to ensure the triangular), then all the points  $\Pi_0^{S_i}$  are connected to the main network by the arcs  $\Sigma_0^{S_i}$  and the newly added arcs  $\Sigma_0^{+1}$ . If all the newly connected arcs are denied as outliers, enlarge  $K$  as  $\alpha K$  ( $\alpha > 1$ ) and iterate until at least two arcs remain or no arc remains for a specified number of consecutive attempts. Secondly, repeat the same procedure for the points in  $\Pi_0^I$ . If successful, the corresponding points are connected to the main network by the newly added arcs  $\Sigma_0^{+2}$ . Let  $\Sigma_0^+ = \Sigma_0^{+1} + \Sigma_0^{+2}$ ,  $\Pi_0^+ = \Pi_0^{+1} + \Pi_0^{+2}$ . Then update  $\Sigma_1^M = \Sigma_0^+$ ,  $\Pi_1^M = \Pi_0^+$ ,  $\Sigma_1^S = \Sigma_0^S - \sum_p \Sigma_0^{S_p}$ ,  $\Pi_1^S = \Pi_0^S - \sum_p \Pi_0^{S_p}$ ,  $\Pi_1^I = \Pi_0^I - \Pi_0^{+2}$ , and iterate the connections until no more points can be added or a certain computation burden has been reached. The final network is  $\Sigma^M = \Sigma_0^M + \sum_j \Sigma_j^+$ ,  $\Pi^M = \Pi_0^M + \sum_j \Pi_j^+$ ; please refer to the red lines and their vertexes in Fig. 2(b). All the other points and arcs will be abandoned, i.e. the green lines and their vertexes for the remaining secondary networks, and the blue dots for the remaining isolated points.

Table II shows the quantitative comparison between the traditional DTN and the proposed EDTN. With DTN, we have to abandon 769 points out of 9,007 (8.5%) with DTN, while the number of abandoned points is 66 with EDTN (0.73%). **One**

further concern is that the EDTN may not be able to ensure more links which can meet the phase continuity assumption (i.e. the phase difference between two neighboring pixels cannot be greater than  $\pi$ ). Owing to the outlier detector introduced in Section III.B, almost all the remaining arcs can meet the phase continuity assumption, see Equations (9)~(14). To illustrate it, a statistical analysis has been carried out to compare the numbers of the residues (phase integration on each edge closure) of the DTN and EDTN (see Table II). The DTN shows a good consistency with very few residues (0.42 per interferogram), and the residues of the EDTN are slightly bigger (0.5 per interferogram). To further guarantee the reliability of the remaining arcs, a closure detection is implemented at triangular loops and a residual threshold of 1 radian is adopted [20]. After closure detection (without residues for both networks), there are 2168 more arcs and 703 more points reserved by the EDTN than the DTN (see Table II). Therefore, the EDTN makes phase unwrapping easier and more reliable because it offers a greater spatial coverage than the DTN. Note that the DTN can also be used in the BLC approach if there is no special requirement on the spatial coelevation.

#### D. Temporally adjacent phase unwrapping and DEM retrieval

After removing the outliers, a spatial integration is performed to unwrap the temporally adjacent phase. Let  $P$  and  $L$  the numbers of PS points and the arcs involved in the final network respectively, and we have the relationship as

$$\mathbf{\Gamma}\boldsymbol{\theta} = \mathbf{\Delta}\hat{\boldsymbol{\theta}} \quad (15)$$

where  $\mathbf{\Gamma}$  is dimensioned  $L \times P$ , the design matrix linking the point temporally adjacent phase  $\boldsymbol{\theta}$  and the arc temporally adjacent phase  $\mathbf{\Delta}\hat{\boldsymbol{\theta}}$ . Obviously,  $\mathbf{\Gamma}$  is a sparse matrix with the generic element  $\mathbf{\Gamma}(l, i) = -1$ ,  $\mathbf{\Gamma}(l, j) = 1$  representing that the  $l^{\text{th}}$  arc starts with the  $i^{\text{th}}$  PS and ends with the  $j^{\text{th}}$  PS, and 0 otherwise. Equation (15) has a LS solution as

$$\hat{\boldsymbol{\theta}} = (\mathbf{\Gamma}^T \mathbf{\Gamma})^{-1} (\mathbf{\Gamma}^T \mathbf{\Delta}\hat{\boldsymbol{\theta}}). \quad (16)$$

The LS may bring about fringe non-congruency, but actually it is statistically optimal if there is no residue which is already guaranteed by the outlier detection. Like the other  $L^p$ -norm phase unwrapping methods ( $p > 0$ ), such as MCF, it is not congruent either as the necessity to suppress noise, which has been widely understood and accepted [14]. As described in Section II-A, the unwrapped phase comprises the topographic error, the deformation signal, atmospheric delays, and other thermal noise. The generic equation linking the design matrix and unknown variables is shown as

$$\mathbf{U}\mathbf{X} = \boldsymbol{\theta} \quad (17)$$

where  $\mathbf{U}$  is the design matrix with  $\mathbf{U} = [\mathbf{b}, f(\mathbf{T})]$ , and  $\mathbf{X}$  embodies the unknown variables with  $\mathbf{X} = [\varepsilon, f^{-1}(\mathbf{v})]^T$ .  $\mathbf{T} = 4\pi/\lambda[t_1 - t_0, t_2 - t_1, \dots, t_N - t_{N-1}]^T$  is the time temporally adjacent vector, and  $\mathbf{b}$  is the baseline vector with  $\mathbf{b}^T = (4\pi/\lambda)[B_1^\perp/r \sin \vartheta, \dots, B_M^\perp/r \sin \vartheta]$ .  $f(*)$  is dependent on different deformation models. We have the LS solution as

TABLE III  
PARAMETERS OF TERRASAR-X

Symbol	Description	Values
$R_0$	Sensor-to-target distance	645639 m
$f$	Operating frequency	9.65 GHz
$\theta$	Local incidence angle	39.5°

$$\mathbf{X} = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \boldsymbol{\theta}. \quad (18)$$

The dimension of  $\mathbf{U}$  is  $N \times (I+F)$ , where  $F$  is the number of parameters describing deformation model. System (17) is nonsingular if  $\mathbf{U}$  is column full rank, namely rank  $I+F$ . Note that  $F$  is typically small ( $F \ll N$ ), so this condition can always be met if the spatial and temporal baseline distributions are not fully coherent (it holds true almost in all cases).

As the main goal of this paper is to generate DEMs, we can leave the deformation in the residual as noise for later processing. Theoretically, after removing the contributions of the DEM error, followed by appropriate filtering [2], the LOS displacements and atmospheric artifacts can be estimated, however, it is out of scope in this paper and will be left for future research.

## IV. EXPERIMENTAL RESULTS

### A. Study area and data used in this study

Located in the southern coast of China, Shenzhen is one of

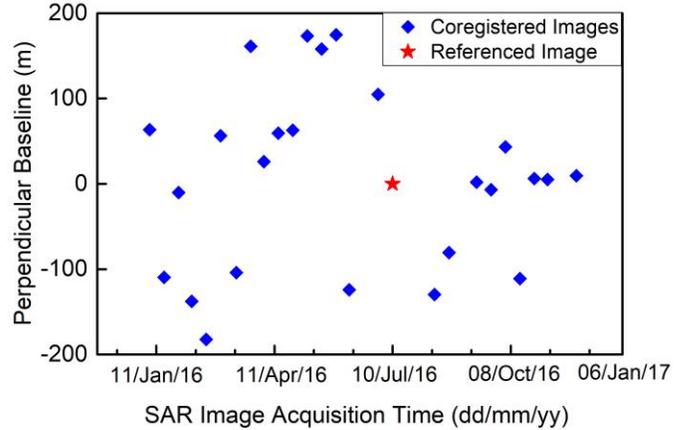


Fig. 3. The spatial/temporal baselines of the TerraSAR-X image used in this study

the fastest developing cities and is highly urbanized with many high-rise buildings and dense infrastructures [46], [47]. A dataset of 26 scenes 0.25m-resolution spotlight TerraSAR-X images (see the parameters in Table III) were acquired on the descending orbit over Shenzhen, China during the period from January to December 2016. To collaboratively process all the images, a referenced image (20160710, with format YYYYMMDD) was selected while the others were all coregistered to it. The spatial and temporal baseline distributions are shown in Fig. 3, where the referenced one is labeled as a red five star and the others denoted as blue diamonds.

### B. Unwrapped temporally adjacent phase

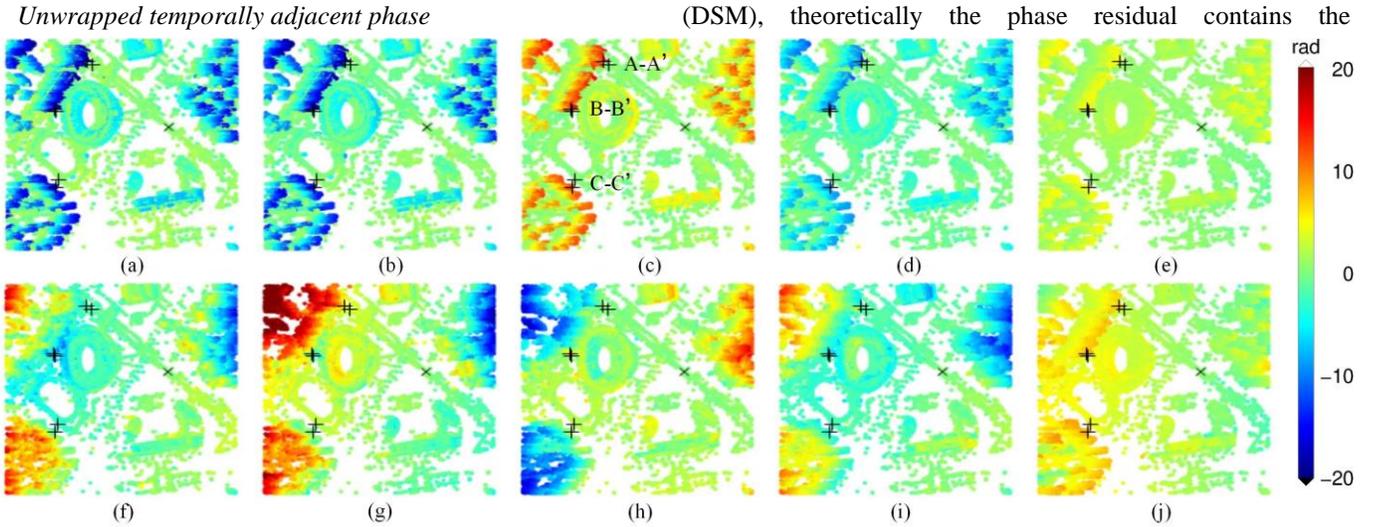


Fig. 4. Comparisons of 3D phase unwrapping results between the BLC approach and StaMPS. (a)~(e) Unwrapped phase obtained by the proposed BLC approach with baselines of -173m, -160m, 110m, -104m and 50m, respectively. (f)~(j) the corresponding unwrapped phase from StaMPS. Only one-fifth of the IFM coverage is shown, with color indicating the relative phase with respect to the cross marked point. The plus signs are the points pairs chosen from the building roof and the ground respectively for quantitative analysis in Table IV.

As mentioned earlier, one of the major challenges in InSAR-TS is 3D phase unwrapping. Following the triangulation steps described in Section III-C, we identify the sequence of InSAR IFMs to be generated for implementing the proposed BLC algorithm. For convenience, the “original” IFMs are from the temporal adjacent SLCs and temporal interval SLCs, i.e.

$$\phi^k = \{\Phi^{i+1} - \Phi^i, \Phi^{i+2} - \Phi^i\}. \quad (19)$$

In the Shenzhen case, we computed 108 IFMs with 15 original IFMs and 93 combined IFMs, with a maximum (equivalent) baseline of 10m (see the justification of the 10m threshold in Section V-B).

Here we compare the phase unwrapping results obtained from the proposed BLC approach with those from StaMPS [48], a popular InSAR time series package with 3D phase unwrapping capabilities. Fig. 3 (a)~(e) show the unwrapped phase derived from the proposed BLC approach whilst Fig. 3 (f)~(g) are the corresponding results derived from StaMPS, respectively. Note that only one-fifth of the IFM coverage is shown as space is limited, and the color indicates the unwrapped phase with respect to a point located at the road with a height of 0 m, denoted as black crosses. From left to right, the baselines are -173m, -160m, 110m, -104m and 50m.

By comparison, in areas without sharp discontinuities, these two methods showed similar patterns. However, where there are discontinuities, StaMPS failed to unwrap the phase, but the proposed BLC method still provided satisfactory results. To see the difference quantitatively, 3 pairs of points are chosen (see the plus signs in Fig. 4), denoted as A-A', B-B' and C-C' (marked in Fig. 4(c)). Each pair of points are located closely in the SAR images, one from the building roof and the other from the ground (sharply discontinuous). If we calculate the phase gradient for each point pair and then eliminate the topography contribution with a LiDAR-derived digital surface model

(DSM), theoretically the phase residual contains the

RMSs OF THE PHASE RESIDUAL OBTAINED BY BLC AND STAMPS

Method	Points pair			Overall RMS (rad)
	A-A' (rad)	B-B' (rad)	C-C' (rad)	
BLC	0.83	0.78	0.78	0.79
StaMPS	14.25	8.85	11.04	11.59

contributions from deformation, atmospheric noise, and other thermal noise. Because of the strong spatial correlation of these components, the residual is expected to be small if the phase is correctly unwrapped. Therefore, RMS (root mean square) of the phase residual is set as a quality index and a large RMS (e.g.  $>2\pi$ ) is definitely an indicator of incorrect phase unwrapping. RMSs derived from BLC and from StaMPS are listed in Table IV. The overall RMSs derived from BLC and from StaMPS are 0.79 rad and 11.59 rad respectively, implying that the BLC approach outperforms StaMPS when sharp discontinuities exist. As a result, the proposed BLC approach offers InSAR-TS, like SBAS, a robust 3D phase unwrapping approach.

### C. Estimated DEM Error

On the basis of 3D phase unwrapping, the DEM errors can be estimated and then added to the DEM used in the interferometric processing step. Let's focus on the stadium located in the central left of the scene (the red box and the zoomed image in Fig. 5(a)), between whose inner and outer ring, very low heights were detected, not consistent with the truth of 40m-around-tall stadium. This is mainly for two reasons. Firstly, the side-looking satellite makes some seats detectable in far range due to the specific ring structure, while the nearly orthographic LiDAR only measures the roof height. Secondly, unlike the residential or office buildings, the polytetrafluoroethylene membrane employed in the stadium roof [49] is much thinner and easier to be penetrated for X-band TerraSAR-X radar signals. Instead, the airborne LiDAR carries

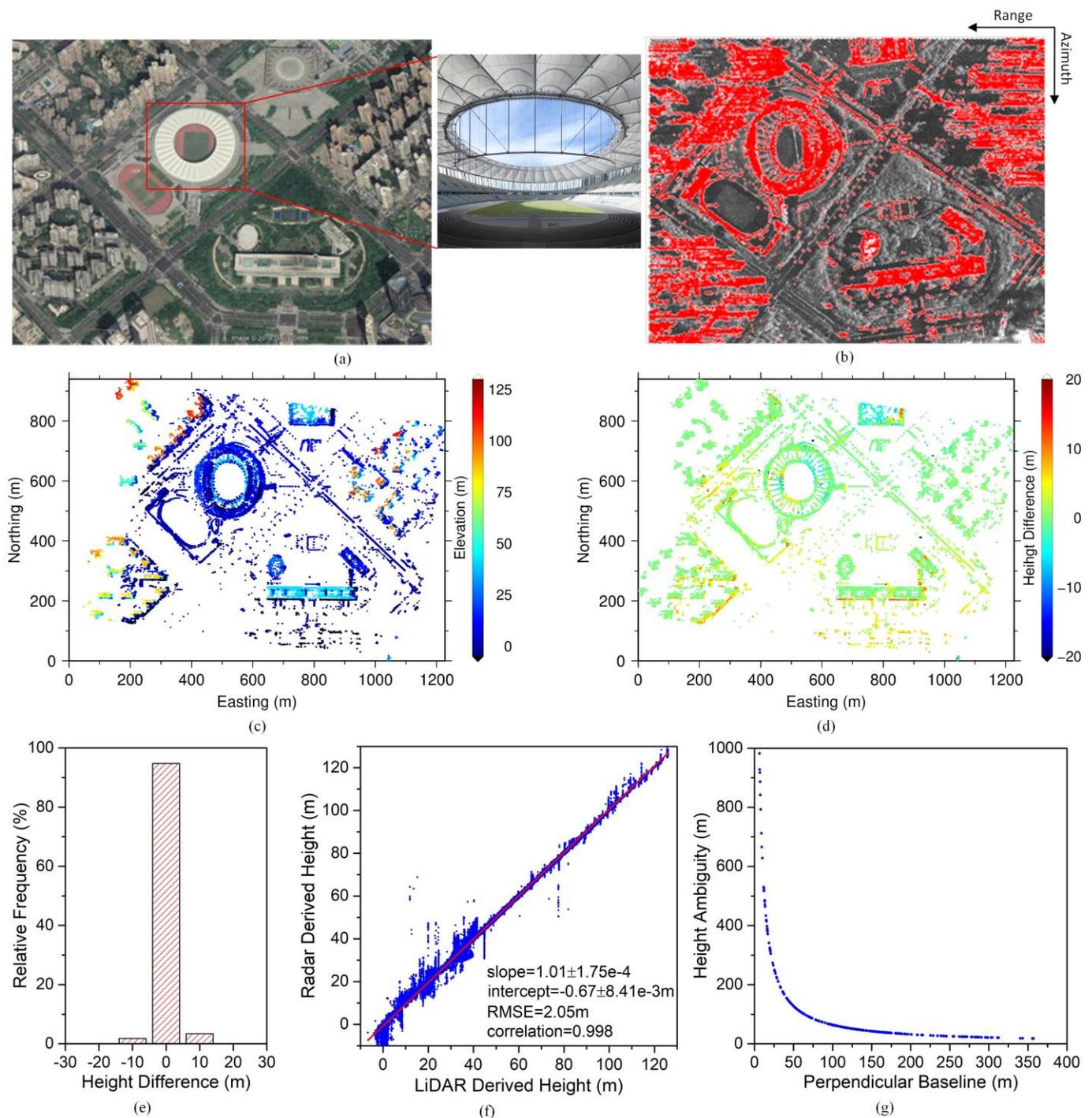


Fig. 5. Validation of the BLC-derived DSM using a LiDAR DSM. (a) Optical image from Google Earth with the zoomed view of the stadium. (b) PSC (in red) superimposed on the mean SAR amplitude image. PSCs are mainly on infrastructure especially on buildings, and scarcely on vegetation. (c) Geocoded results of the superposition of estimated height on SRTM. (d) Height differences between LiDAR and the estimated heights. (e) Histogram of (d). (f) The correlation between the LiDAR and radar derived heights with a total correlation of 0.998 and an RMSE of 2.05m. The red line represents a regression line that best fits the plotted set of points with a slope of 1.01 and an intercept of -0.67. The points in red rectangles denoting the building profiles with geocoding error. (g) Height ambiguity of the data stack with a highest height ambiguity of 17.8m.

near infrared (micrometer) signals, which hardly has a penetrating capability. According to the zoomed image in Figs. 4(a) and (b), it is inferred that the satellite signal penetrates or side-looks the stadium roof and the points detected with lower heights are the seats. The lower height ranges from 30m to 15m, well matching the heights of the seats, providing strong

supporting evidence for the conclusion about the penetration of the radar signals.

After eliminating the pixels with penetrating effects, we compared the BLC-derived heights with those derived from LiDAR. The accuracy of the estimated DEM can be found out by the comparative analysis of setting LiDAR-derived DSM as

a reference, see the difference map in Fig. 5(d). A total of 127,346 PS points were detected in the study area with a density of 231,500/km<sup>2</sup>, among which 6665 points (5.2%) were observed with height differences over 5m, 59 points (0.046 %) with height differences over 15m, and the rest points (94.8%) with height differences within [-5m, 5m] (see the histogram in Fig. 5(e)).

The consistency between LiDAR and radar-derived heights was also assessed using a correlation analysis as shown in Fig. 5 (f), where the red line represents the regression line with a slope of  $1.01 \pm 1.75e-4$ , and an intercept of  $-0.67 \pm 8.41e-3$  m. A high correlation coefficient, 0.998, was obtained between the two data sets with an RMS of 2.05m. Note that the radar images used here have a very limited baseline range of 357m (Fig. 3), suggesting a low height sensitivity. Using the height ambiguity equation  $H_a = \lambda R \sin \theta / 2B_{\perp}$  [50], the system has the highest sensitivity with an  $H_a$  of 17.8m (Fig. 5(g)). According to the height error equation in [50],  $H_{error}$  is derived as

$$H_{error} = H_a \cdot \sigma_{\phi} / 2\pi \quad (20)$$

with  $\sigma_{\phi}$  being the phase standard deviation of the IFM, which can be described as follows [51]

$$\sigma_{\phi}^2 = \frac{1}{2N_L} \frac{1-\gamma^2}{\gamma^2} \quad (21)$$

where  $N_L$  is the number of looks ( $N_L = 1$  in this study) and  $\gamma$  is the IFM coherence. As all the pixels studied here are PSs and/or TCPs,  $\gamma = 0.85$  is a reasonable value (refer to [52]). Employing the error propagation theory, the autocovariance of the DEM error can be obtained as

$$\text{var}(\text{DEM}_{error}) = \mathbf{A} \text{var}(\mathbf{H}_{error}) \mathbf{A}^T \quad (22)$$

where  $\text{var}(\mathbf{H}_{error})$  is the autocovariance of the error height for the original IFMs, and  $\mathbf{A}$  is the design matrix linking the original IFMs and the combined IFMs. Finally, an error height of 2.66m is obtained, which is close to the 2.05m RMS. Another explanation is that according to [3], small values of the amplitude dispersion index  $D_A$  is a good estimation of the phase stand deviation, i.e.  $\sigma_{\phi} = D_A = 0.25$ . Using this approximation,  $\text{DEM}_{error} = 1.52m$ , also agrees with the RMS well.

The validation experiment conducted in this paper is only a coarse result as only one track of TerraSAR-X data is exploited. Surely, results with even higher accuracy can be obtained if more tracks of data are available.

Besides the stadium, the height differences were also found caused by geocoding process, especially near building profile or edges (the Y-axis parallel lines aligned by blue points in Fig. 5(f)), which will be analyzed in detail in Section V-C. Except for some noise (few isolated points with big height differences in Fig. 5(f)), the error caused by geocoding, and the limitation of the essential resolution, the BLC-derived results match the reality well.

## V. DISCUSSIONS

### A. FCN or DTN

There are two commonly used networking methods, namely DTN and freely connected network (FCN). The latter is preferable in recent literature because of its stronger connection and higher reliability and accuracy [53]-[55]. But when processing big scenes, the computational demand is too high to be acceptable. Let  $P$  the number of points involved. The FCN approach has to connect all the PSCs by  $C_P^2 = P \times (P - 1) / 2$  arcs (if there is no limit on the distance) and then checks all the arcs, while the number of arcs required to be connected and checked by DTN approach is only  $O(P \log P)$  [54]. Therefore, the computational complexity of FCN is  $O(P^{1-1/n}/2)$  times greater than that of DTN. Here for the first time, we propose EDTN with a high reliability and an acceptable computation burden.

The light computational demand is one of the reasons to choose DTN. More importantly, with the EDTN described in Section III-C, most of the PSCs are kept in the network and little information is lost with an acceptable computation burden. To illustrate it, we take a small area from the study scene, the same area demonstrated in Section III-C (Fig. 3). Originally 9,007 PSCs were selected. EDTN kept 8975 PS points and 25,060 arcs, while FCN kept 8984 PS points (only 0.1% improvement) at the cost of 28,228,098 arcs (1000 times greater) for checking. Using a desktop with 4 cores at 3.5GHz and 64GB RAM, it cost 7.76 and 709.86 seconds for EDTN and FCN processing respectively. The comparison between FCN and DTN/EDTN is listed in Table V. From the theoretical respect, FCN has quadratic complexity, while DTN just has quasilinear complexity. Consequently, EDTN is a good compromise in terms of both computation burden and reliability.

### B. BLC IFMs or SBAS

As described in Section IV-B, we only consider one and/or two IFMs for combined IFMs to minimize the propagated noises,

$$\begin{aligned} \tilde{\phi}_k &= \Phi_{k_1} - \Phi_{k_2} + p(\Phi_{k_3} - \Phi_{k_4}), k_{1,2,3,4} \in 0, 1, 2, \dots, N \\ k_1 &\neq k_2, k_3 \neq k_4, k_2 \neq k_3, k_1 \neq k_4, p \in \{0, 1\} \end{aligned} \quad (23)$$

For a dataset with  $N+1$  images, if no baseline constraint is applied, the number of IFMs is  $C_N^2 = N(N+1)/2$  and the number of combined IFMs can be up to  $C_{N+1}^2 \times C_{N-1}^2 + C_{N+1}^2 \times C_N^2 = N^2(N^2-1)/4$ , with an increase of nearly  $N^2/2$  times. For the dataset used in this study, with baseline constraints of 10m and 0.2y, the numbers of IFMs for SBAS and BLC are 15 and 497 respectively. Even if we are more tolerant of the baseline separation, say  $|B_{\perp}| < 40m$  and  $|t| < 0.4y$ , only 65 IFMs are selected for SBAS case and 1 SLC has to be abandoned because none of the IFMs composed by the abandoned one meets the baseline requirements. Therefore, the combined IFMs have many more redundancies than the conventional SBAS approach.

For typical SAR sensors, taking TerraSAR-X for example, whose parameters are shown in Table III, with which the topographic phase error is  $\phi_e = 3.13\pi B_{\perp} \varepsilon \times 10^{-4}$ . As the buildings with an absolute height of  $>300m$  are defined as

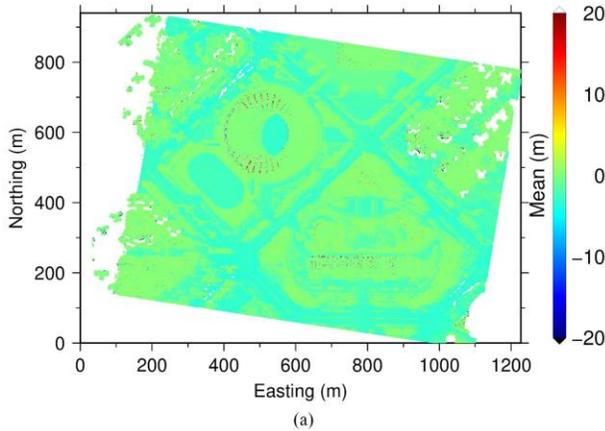
supertall skyscrapers (no more than 100 worldwide), and the low-resolution DEM used can only eliminate the topography of the ground, the maximum value of the residual height  $\varepsilon$  equals to or is just less than the absolute height. Without loss of generality,  $\varepsilon$  can be upper limited to 300m. To guarantee  $|\phi_\varepsilon| < \pi$ ,  $|B_\perp|$  should be limited to 10m, which is the reason for the baseline threshold in Section IV-B.

Let's focus on the spatial difference between neighboring pixels. Adopting very small and normal baselines, say 10m and 150m respectively, the corresponding maximum gradient residual height  $\Delta\varepsilon$  are 300m and 20m to keep  $|\Delta\phi_\varepsilon| < \pi$ . In the normal baseline case, most of the arcs with a height difference greater than 20m will be dropped at the outlier detection step described in Section III-B, leading to a very sparse network. Comparatively, in the very small baseline case, the deleted arcs are mostly the ones with big stratified distributed atmospheric artifacts. Namely, the dropped arcs in the 10m-baseline IFMs are only affected by the atmospheric effects, while in the 150m-baseline case, they are deleted due to big residual topographic contributions and large atmospheric residuals. This is also the reason why in the previous work [40], [54]-[56], normal baseline IFMs adopt FCN at the cost of heavy computation burden to ensure that enough arcs remain after the outlier detection. The BLC method utilizes the light computationally burdened EDTN to ensure that there are sufficient arcs for further analysis (see Table V).

It is worth mentioning that there is a special scenario with all the baselines being regularly distributed, which rarely happens in real life. The generic  $i^{\text{th}}$  SLC has a baseline of  $B_i^\perp = B_0^\perp + (i-1)\bar{b}$  with  $\bar{b}$  the even baseline separation, and the baseline of the combined IFM can be written as  $\tilde{B}_k^\perp = (k_1 - k_2 + p(k_3 - k_4))\bar{b}$ . If  $(k_1 - k_2 + p(k_3 - k_4)) = 0$ , i.e.  $\tilde{B}_k^\perp = 0$ , the topographic phase is totally canceled; in other words, the combined IFM won't be influenced by the DEM error anymore, whether in continuous or sharply discontinuous case.

### C. Noise Analysis

In the outlier detection process, it is important to decide the residual threshold, which has already been addressed and described from a statistical perspective in Section III-B. As a



Method	FCN	DTN/EDTN
Theoretical computation <sup>a</sup>	$(P \times (P - 1))/2$	$O(P \times \log P)$
Original PS points	9,007	9,007
Arcs before outlier detection	40,558,521	26,928
PS points after outlier detection	8,984	8,975
Arcs after outlier detection	28,228,09	25,060
Time consumed in total <sup>b</sup>	709.9 s	7.8 s

<sup>a</sup> $P$  is the PS points in total, big  $O(n)$  notation means the algorithm has an order of  $n$  time complexity; <sup>b</sup>s = second

matter of fact, a simple decision can be made according to the PS noise. Assuming that noise is identical and uncorrelated, the most disadvantageous case is the combination of 2 IFMs sharing one SLC,

$$\tilde{\phi}_k = \Phi_{k_1} - \Phi_{k_2} + (\Phi_{k_1} - \Phi_{k_3}) = 2\Phi_{k_1} - \Phi_{k_2} - \Phi_{k_3} \quad (24)$$

whose noise is  $6\sigma_{ps}^2 ((2\sigma_{ps})^2 + \sigma_{ps}^2 + \sigma_{ps}^2)$ , while only  $4\sigma_{ps}^2$  for the other combined IFMs,

$$\tilde{\phi}_k = \Phi_{k_1} - \Phi_{k_2} + (\Phi_{k_3} - \Phi_{k_4}) \quad (25)$$

and  $2\sigma_{ps}^2$  for the original IFMs (or taken as one SLC canceled)

$$\tilde{\phi}_k = \Phi_{k_1} - \Phi_{k_2} + (\Phi_{k_3} - \Phi_{k_1}) = \Phi_{k_3} - \Phi_{k_2} \quad (26)$$

For convenience, only the most disadvantageous case is considered, i.e. each arc shares a noise level of  $12\sigma_{ps}^2$ . If we expect an accuracy of 5m (smaller than the height sensitivity in Section IV-C) for the estimated DEM (for TerraSAR-X satellite  $\sigma_{DEM} = 0.049rad$  provided that baseline is limited to 10m), as a result  $\sigma_{arc} = 0.17rad$  and the threshold  $T_{\phi_{res}}$  in Equation (14) can be set as 0.17 rad.

One should note that, except for infecting  $T_{\phi_{res}}$  value, the increased noise also deteriorates the system. The general linear combination case is then investigated. Similar to Equation (24),

$$\begin{aligned} \tilde{\phi}_k &= m(\Phi_{k_1} - \Phi_{k_2}) + n(\Phi_{k_1} - \Phi_{k_3}) \\ &= (m+n)\Phi_{k_1} - m\Phi_{k_2} - n\Phi_{k_3} \end{aligned} \quad (27)$$

whose noise  $\sigma_{\tilde{\phi}_k} \triangleq \beta\sigma_\phi$ , where  $\sigma_\phi$  is the noise standard deviation of the original interferogram and  $\beta$  is the noise

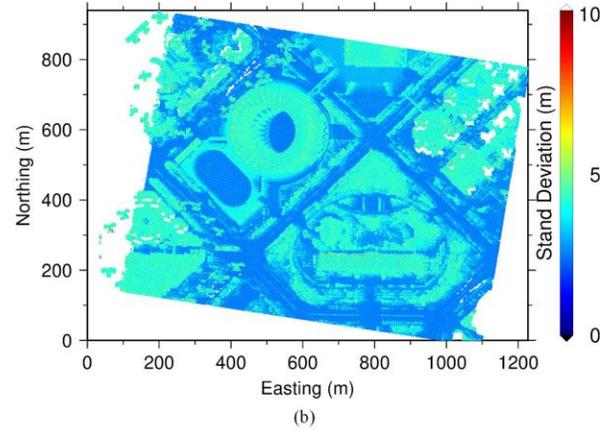


Fig. 6. Statistical results of the geocoding errors caused by the uncertainties in DSM, realized by adding Gaussian noise with 0 mean and 5m standard deviation, processed by 1000 times of Monte Carlo simulations. (a) Mean value. (b) Standard deviation.

amplification factor with  $\beta = \sqrt{m^2 + n^2 + mn}$  (here  $mn \geq 0$ ). Similar to Equations (25) and (26), the general cases are

$$\tilde{\phi}_k = m(\Phi_{k_1} - \Phi_{k_2}) + n(\Phi_{k_3} - \Phi_{k_4}) \quad (28)$$

$$\begin{aligned} \tilde{\phi}_k &= m(\Phi_{k_1} - \Phi_{k_2}) + n(\Phi_{k_3} - \Phi_{k_1}) \\ &= (m - n)\Phi_{k_1} + m\Phi_{k_3} - n\Phi_{k_2} \end{aligned} \quad (29)$$

whose noise amplification factors are  $\beta = \sqrt{m^2 + n^2}$  and  $\beta = \sqrt{m^2 + n^2 - mn}$ , respectively.

From the above analysis, compared with the conventional InSAR techniques, the combined interferograms are interfered by heavier noise (up to  $\beta = \sqrt{m^2 + n^2 + mn}$ ). For example, for  $m/n = \pm 1$ ,  $\beta = \sqrt{3}$ , and for  $m/n = \pm 2$ ,  $\beta = 2\sqrt{3}$ . This is the reason why we would like to limit  $m$  and  $n$  to the smallest values in the premise that the system is nonsingular.

Due to the inaccurate orbit knowledge, the geocoding error should also be investigated for the estimated DSMs. Generally, for flat areas, if the deviation is small and the offset can be controlled within some pixels, the geocoding error can be neglectable. Unfortunately, in the case with large discontinuities over the layover area, it is very sensitive to the geocoding error. To illustrate this, we performed simulations on the LiDAR data used in Section IV-C. Firstly, the LiDAR DSM was back-geocoded to the radar coordinate and Gaussian noises with 0 mean and 5m standard deviation were attached. Note that the 5m standard deviation is our targeted accuracy. Then the noise-added DSM was forward-geocoded to the geodetic coordinate. After 1000 times of Monte Carlo simulation, the statistical results of the geocoding error relative to noise-free geocoded results are shown in Fig. 6, with (a) and (b) the mean and standard deviation respectively. It is clear in Fig. 5 that, in the flat areas, the mean error and standard deviation are mostly

around 0 and 5m, while the mean error is over 20m and the standard deviation can be over 10m near the building edges where sharp discontinuities exist, corresponding to the high height differences in Fig. 5(d). As the building façades constitute quite a large proportion of the detected PS, it is sensible that the ratio of PSs with big height difference in Fig. 5(d) exceeds that of presented in Fig. 6. More detailed factors affecting geocoding errors are depicted in Appendix A.

## VI. CONCLUSIONS

A baseline linear combination (BLC) method has been demonstrated in this paper. By combining two IFMs, IFMs with very small baselines (nearly zero) can be constructed, with which 3D phase unwrapping can be easily done. The application of BLC approach to the TerraSAR-X data over Shenzhen city suggests a high correlation of 0.998 between LiDAR and TerraSAR-X derived heights with a low RMSE of 2.05m, demonstrating the effectiveness of the proposed method.

In addition, the approach proposed is extremely suitable for urban areas with high-rise buildings which will encounter phase unwrapping errors by conventional InSAR methods. Nevertheless, when applied to small towns just with sparse infrastructures, its superiority cannot be fully made use of. Worse still, it can hardly be applied to heavily vegetated areas as this method is TCP-processed based and there would not be enough TCPs in those areas. Also, the orbit tightly controlled TerraSAR-X, with a very lower elevation resolution, may hinder the validation of the method. It will be helpful if more data stacks with a larger baseline span are available.

It is worth mentioning that the BLC method is originally designed for urban areas, but actually it is perfectly applicable whenever big height differences exist (e.g. layover and/or

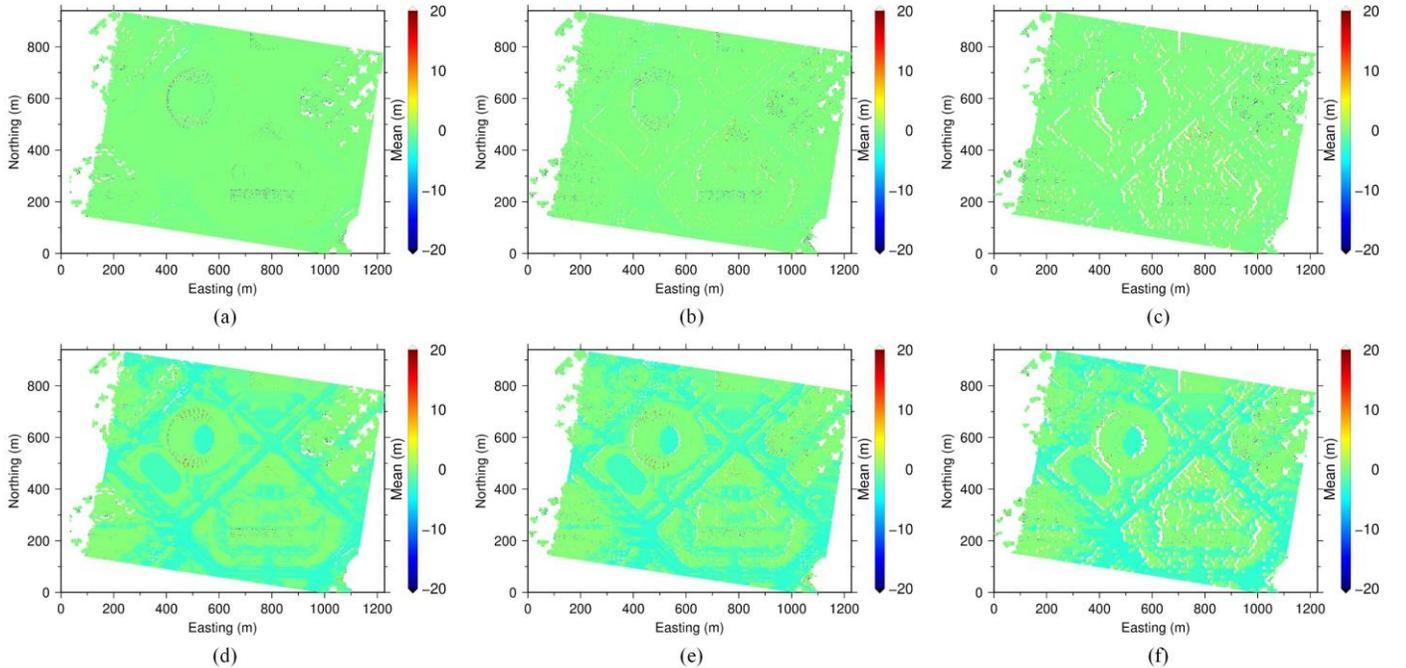


Fig. A1. Mean value of the statistical results of geocoding error caused by DSM error, realized by adding Gaussian noise with 0 mean, processed by 1000 times of Monte Carlo simulations by 1m, 5m and 10m LiDAR from left to right. The upper row is with noise of a standard deviation 1m, while the bottom row is with noise of a standard deviation of 5m.

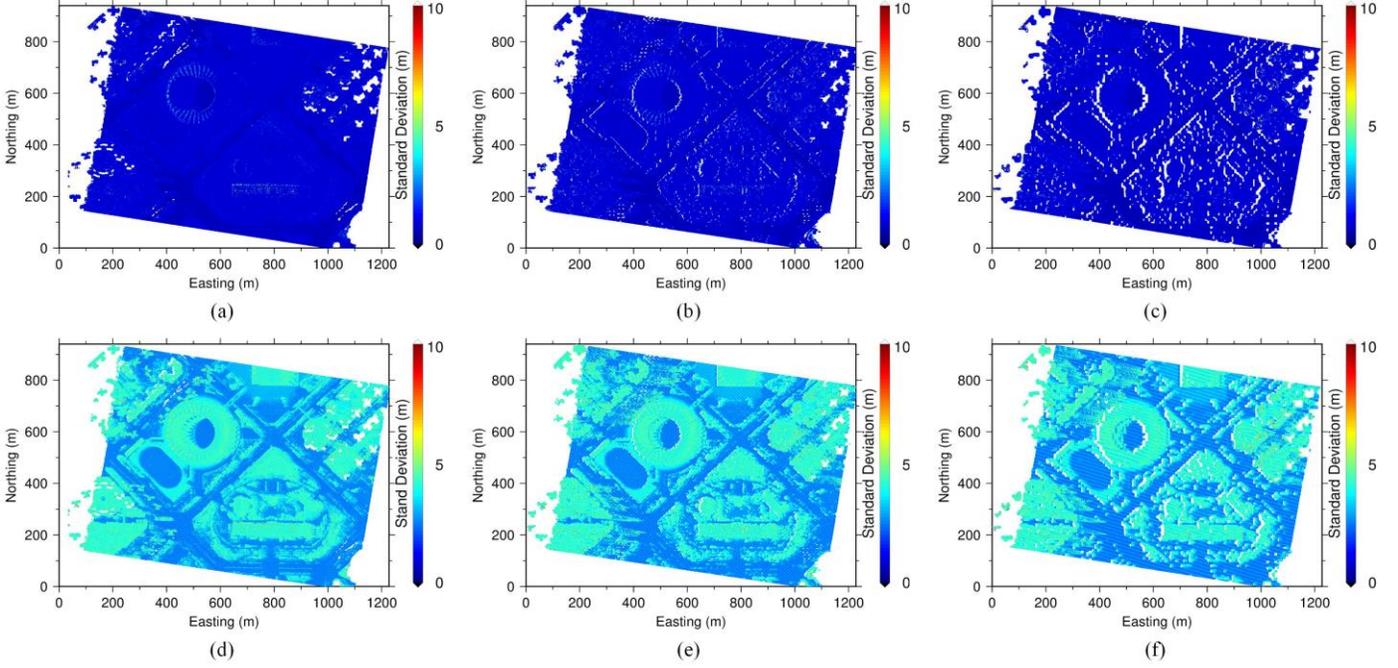


Fig. A2. The corresponding standard deviations of Fig. A1.

shadow areas) and wherever TCP exists. One possible application is for steep mountains or cliffs, especially with rocks and not heavily vegetated ones. Also, this can be applied in certain geological hazards, like the monitoring of volcano craters and faults whose activity is closely related to earthquakes.

#### APPENDIX A

This appendix illustrates the geocoding errors caused by DSM with different spatial resolutions and at different noise levels, as a supplement of Fig. 6. The corresponding mean and standard deviation results are shown in Figs. A1 and A2 respectively, with the same experimental condition but with different noise levels and spatial resolutions. (a)~(c) are with 1m noise standard deviation and (d)~(f) are with 5m noise standard deviation. (a) and (d) are with 1m DSM resolution, (b) and (e) are with 5m DSM resolution and (c) and (f) are with 10m DSM resolution, respectively. From the figures, a conclusion can be drawn that the higher the spatial resolution and the lower the noise, the better performance. When the resolution of the DSM is low, say lower than 10m, the geocoding performance dramatically deteriorates. Similarly, high noise can severely deteriorates the geocoding performance.

#### APPENDIX B

This appendix aims to validate that the combination method is feasible by Monte Carlo simulation.

It is assumed that the position of satellite distributes randomly in the space, and with a specific perpendicular baseline span  $B_{\perp}$ , whose typical value is 400m for TerraSAR-X and 2000m for CosmoSky-Med. We can assume that the perpendicular baseline obeys the normalized distribution  $b \sim N(0, \sigma_b^2)$  (relative to the master acquisition). According to

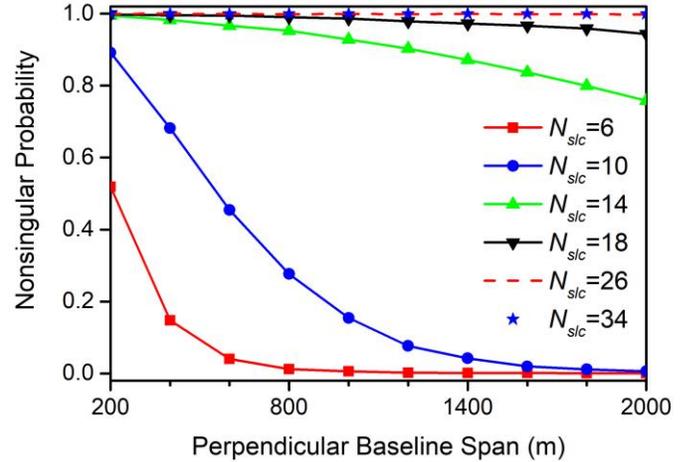


Fig. B1. Perpendicular baseline span versus nonsingular characteristic under different numbers of SAR acquisitions with  $m/n = \pm 1$ .

Pauta criterion ( $3\sigma$  criterion), the perpendicular baseline  $b$  distributes within  $(-3\sigma_b, 3\sigma_b)$  with a probability of 99.7%.  $\sigma_b$  can be set as  $\frac{1}{6}B_{\perp}$  so as to make  $b$  distribute within  $(-\frac{B_{\perp}}{2}, \frac{B_{\perp}}{2})$  with a high probability of 99.7%. If the design matrix  $\mathbf{H}$  is full rank (i.e. a nonsingular system) with a probability equaling to or greater than 99.7%, the system can be considered nonsingular. Although the analytical solution of the nonsingular probability of  $\mathbf{H}$  is difficult to realize, it can be evaluated by Monte Carlo simulations. Based on 10,000 MC simulations, the baseline span with different  $N_{slc}$  versus nonsingular probability are plotted in Fig. B1 (as here we only set  $m/n = \pm 1$ , and  $\mathbf{H}$  only contains elements  $0, \pm 1, \pm 2$ , its rank can be obtained quickly and accurately). From Fig. B1, it is clear that the systematic nonsingularity increases with the number of SAR acquisitions  $N_{slc}$  while decreases with the perpendicular baseline  $B_{\perp}$ . When

$N_{slc} \geq 26$ , the system is always nonsingular when  $B_{\perp} \in [200, 2000]$ . For our case of  $B_{\perp} = 400\text{m}$  and  $N_{slc} = 26$ , the system nonsingularity can be ensured.

Note that, the above-discussed nonsingularity is based on the assumption  $b \sim N(0, \sigma_b^2)$ . However, it is possible that one acquisition stays further away from the others (we call it the isolated acquisition). Thus  $b \sim N(0, \sigma_b^2)$  does not hold, and neither the lines in Fig. B1. Following the guidance in the last paragraph in Section II-B,  $m/n$  can be enlarged to  $\pm 2$ . And if after the enlargement, the system is still singular (though rarely happens), one can simply abandon the isolated acquisition.

#### ACKNOWLEDGEMENT

The authors would like to thank Julian Austin for his valuable comments and suggestions that improved this article. The authors would also like to thank the anonymous reviewers for their constructive comments and valuable suggestions that have helped improve the quality of this manuscript.

#### REFERENCES

- [1] H. Yu, H. Lee, T. Yuan and N. Cao, "A Novel Method for Deformation Estimation Based on Multibaseline InSAR Phase Unwrapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5231-5243, Sept. 2018.
- [2] 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.
- [3] 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.
- [4] C. Colesanti, A. Ferretti, F. Novali, C. Prati, and F. Rocca, "SAR monitoring of progressive and seasonal ground deformation using the permanent scatterers technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 7, pp. 1685-1701, Jul. 2003.
- [5] 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.
- [6] R. Lanari, O. Mora, M. Manunta, J. J. Mallorqui, P. Berardino and E. Sansosti, "A small-baseline approach for investigating deformations on full-resolution differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1377-1386, Jul. 2004.
- [7] Y. Yan, M.-P. Doin, P. Lopez-Quiroz, F. Tupin, B. Fruneau, V. Pinel and E. Trouve, "Mexico city subsidence measured by InSAR time series: joint analysis using PS and SBAS approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 5, no. 4, pp. 1312-1326, Aug. 2012.
- [8] B. Bayer, D. Schmidt and A. Simoni, "The influence of external digital elevation models on PS-InSAR and SBAS Results: implications for the analysis of deformation signals caused by slow moving landslides in the northern Apennines (Italy)," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2618-2631, May 2017.
- [9] A. Hooper, "A multi-temporal InSAR method incorporating both persistent scatterer and small baseline approaches," *Geophys. Res. Lett.*, vol. 35, no. 16, Aug. 2008.
- [10] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca and A. Rucci, "A new algorithm for processing interferometric data-stacks: SqueeSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3460-3470, Sep. 2011.
- [11] M. Costantini, F. Minati, F. Trillo and F. Vecchioli, "Enhanced PSP SAR interferometry for analysis of weak scatterers and high definition monitoring of deformations over structures and natural terrains." in *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS)*, Melbourne, Australia, Jul. 2013, pp. 876-879.
- [12] M. Costantini T. Chen, Y. Xu, et al., "High resolution ground deformations monitoring by COSMO-SkyMed PSP SAR interferometry: accuracy analysis and validation," in *Proc. ESA FRINGE*, Frascati, Italy, Sep. 2011.
- [13] M. Costantini, S. Falco, F. Malvarosa, F. Minati, F. Trillo and F. Vecchioli, "Persistent scatterer pair interferometry: approach and application to COSMO-SkyMed SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 7, no. 7, pp. 2869-2879, Jul. 2014.
- [14] H. Yu, Y. Lan, Z. Yuan, J. Xu and H. Lee, "Phase Unwrapping in InSAR : A Review," *IEEE Geosci. Remote Sens. Magazine*, vol. 7, no. 1, pp. 40-58, March 2019.
- [15] K. Itoh, "Analysis of the phase unwrapping algorithm," *Appl. Opt.*, vol. 21, no. 14, pp. 2470, Jul 1982.
- [16] A. Hooper, and H. A. Zebker, "Phase unwrapping in three dimensions with application to InSAR time series," *J. Opt. Soc. Amer. A*, vol. 24, no. 9, pp. 2737-2747, Sep. 2007.
- [17] A. Hooper, "A statistical-cost approach to unwrapping the phase of InSAR time series," in *Proc. ESA FRINGE Workshop*, Frascati, Italy., Nov. 30-Dec. 4, 2009, p. 6.
- [18] P. Shanker F. Casu H. Zebker and R. Lanari, "Comparison of persistent scatterers and small baseline time-series InSAR results: a case study of the San Francisco Bay area," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 592-596, July 2011.
- [19] Sadeghi, Z., M.J. Valadan Zoej and M. Dehghani, "An improved persistent scatterer interferometry for subsidence monitoring in the Tehran Basin," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 6, no. 3, pp. 1571-1577, Jun. 2013.
- [20] H. Hussain, A. Hooper, T. J. Wright, R. J. Walters, and D. P. S. Bekaert, "Interseismic strain accumulation across the central North Anatolian Fault from iteratively unwrapped InSAR measurements," *J. of Geophysical Research: Solid Earth*, vol. 121, no. 12, pp. 9000-9019, Dec. 2016.
- [21] Y. Dong, H. Jiang, L. Zhang and M. Liao, "An efficient maximum likelihood estimation approach of multi-baseline SAR interferometry for refined topographic mapping in mountainous areas," *Remote Sens.*, vol. 10, no. 3, Mar. 2018.
- [22] V. Pascazio and G. Schirinzi, "Multifrequency InSAR height reconstruction through maximum likelihood estimation of local planes parameters," *IEEE Trans. Image Process.*, vol. 11, no. 12, pp. 1478-1489, Dec. 2002.
- [23] G. Fornaro, A. Monti Guarnieri, A. Pauciuolo and F. De-Zan, "Maximum likelihood multi-baseline SAR interferometry," *IEE Proc. Radar Sonar and Navig.*, vol. 153, no. 3, pp. 279-288, June 2006.
- [24] M. Eineder and N. Adam, "A maximum-likelihood estimator to simultaneously unwrap, geocode, and fuse SAR interferograms from different viewing geometries into one digital elevation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 24-36, Jan. 2005.
- [25] G. Ferraiuolo, V. Pascazio, and G. Schirinzi, "Maximum a posteriori estimation of height profiles in InSAR imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 1, no. 2, pp. 66-70, Apr. 2004.
- [26] G. Ferraiuolo, F. Meglio, V. Pascazio and G. Schirinzi, "DEM Reconstruction Accuracy in Multichannel SAR Interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 191-201, Jan. 2009.
- [27] A. Shabou, F. Baselice, and G. Ferraioli, "Urban digital elevation model reconstruction using very high resolution multichannel InSAR Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4748-4758, Nov. 2012.
- [28] G. Ferraioli, A. Shabou, F. Tupin, and V. Pascazio, "Multichannel phase unwrapping with graph-cuts," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 562-566, Jul. 2009.
- [29] Shabou, A. and F. Tupin, "A Markovian Approach for DEM Estimation From Multiple InSAR Data With Atmospheric Contributions," *IEEE Geosci. and Remote Sens. Letters*, vol. 9, no. 4, pp. 764-768, July 2012
- [30] A. Pepe, and R. Lanari, "On the extension of the minimum cost flow algorithm for phase unwrapping of multitemporal differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2374-2383, Sep. 2006.
- [31] A. Pepe, Y. Yang, M. Manzo and R. Lanari, "Improved EMCF-SBAS processing chain based on advanced techniques for the noise-filtering and selection of small baseline multi-look DInSAR Interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4394-4417, Aug. 2015.
- [32] D. Chirico and G. Schirinzi, "Multichannel interferometric SAR phase unwrapping using extended Kalman smoother," *Int. J. Microw. Wireless Technol.*, vol. 5, pp. 429-436, Jun. 2013.
- [33] X. Xianming and P. Yiming, "Multi-baseline phase unwrapping algorithm based on the unscented Kalman filter," *IET Radar, Sonar Navigat.*, vol. 5,

- no. 3, pp. 296–304, 2011.
- [34] B. Osmanoglu, T. H. Dixon and S. Wdowinski, “Three-dimensional phase unwrapping for satellite radar interferometry, I: DEM generation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1059–1075, Feb. 2014.
- [35] R. Ambrosino, F. Baselice, G. Ferraioli and G. Schirinzi, “Extended Kalman Filter for Multichannel InSAR Height Reconstruction,” *IEEE Trans. on Geos. and Remote Sens.*, vol. 55, no. 10, pp. 5854–5863, Oct. 2017.
- [36] H. Yu, Z. Li and Z. Bao, “A cluster-analysis-based efficient multibaseline phase-unwrapping algorithm,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 478–487, Jan. 2011
- [37] H. Liu, M. Xing and Z. Bao, “A cluster-analysis-based noise-robust phase-unwrapping algorithm for multibaseline interferograms,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 494–504, Jan. 2015.
- [38] Z. Jiang, J. Wang, Q. Song and Z. Zhou, “A refined cluster-analysis-based multibaseline phase-unwrapping algorithm,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 9, pp. 1565–1569, Sep. 2017.
- [39] L. Zhang, X. Ding and Z. Lu, “Modeling PSInSAR time series without phase unwrapping,” *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 547–556, Jan. 2011.
- [40] S. Wu, L. Zhang, X. Ding and D. Perissin, “Pixel-Wise MTInSAR Estimator for Integration of Coherent Point Selection and Unwrapped Phase Vector Recovery,” *IEEE Trans on Geosci. and Remote Sens.*, pp. 1–10, 2018.
- [41] Dominique Derauw and A. Orban, “Baseline combination for InSAR DEM atimetric resolution enhancement,” in *Proc. ESA FRINGE*, Frascati, Italy, Dec. 2003.
- [42] Y. Zhou, C. Zhou, E. D. and Z. Wang, “A baseline-combination method for precise estimation of ice motion in Antarctica,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5790–5797, Sep. 2014.
- [43] K. Koch, *Parameter Estimation and Hypothesis Testing in Linear models*, 2nd ed. Berlin, Germany: Springer-Verlag, 1999
- [44] P. Jia, and Z. Zhu, *Optical Estimation and Applications*, Beijing: Science Press, 1984.
- [45] Z. Wang, D. Yi, X. Duan, J. Yao and D. Gu, *Measurement Data Modeling and Parameter Estimation*, Boca Raton: CRC Press, 2012.
- [46] Y. Du, G. Feng, Z. Li, X. Peng, J. Zhu, and Z. Ren, “Effects of external digital elevation model inaccuracy on StaMPS-PS Processing: a case study in Shenzhen, China,” *Remote Sens.*, vol. 9, no. 11, Nov. 2017.
- [47] B. Xu, G. Feng, Z. Li, Q. Wang, C. Wang, and R. Xie, “Coastal subsidence monitoring associated with land reclamation using the point target based SBAS-InSAR method: a case study of Shenzhen, China,” *Remote Sens.*, vol. 8, no. 8, Aug. 2016.
- [48] A. Hooper, D. Bekaert, K. Spaans and M. Arikian, “Recent advances in SAR interferometry time series analysis for measuring crustal deformation,” *Tectonophysics*, vol. 514–517, pp.1–13, 2012
- [49] D. Knight, R. Whitefield, E. Nhieu, F. Tahmasebinia, P. Ansourian and F. Alonso-Marroquin, “Transient dynamic analysis of the Bao’an Stadium,” *AIP Conference Proceedings*, 2016. 1762(1): p. 020001.
- [50] C. Choi and D. Kim, “Optimum Baseline of a single-pass In-SAR system to generate the best DEM in tidal flats,” *IEEE Trans. Geosci. Remote Sens.*, vol. 11, no. 3, pp. 919–929, Mar. 2018
- [51] P. Rosen, S. Hensley, I. R. Joughin, F. K. Li, S. Madsen, E. Rodriguez and R. Goldstein, “Synthetic aperture radar interferometry,” *Proc. IEEE*, vol. 88, no. 3, pp. 333–382, Mar. 2000.
- [52] M. Martone, B. Brautigam, P. Rizzoli, C. Gonzalez, M. Bachmann and G. Krieger, “Coherence evaluation of TanDEM-X interferometric data,” *ISPRS J. Photogram. Remote Sens.*, vol. 73, pp. 21–29, 2012
- [53] T. Li, G. Liu, H. Lin, H. Jia, R. Zhang, B. Yu and Q. Luo, “A hierarchical multi-temporal InSAR method for increasing the spatial density of deformation measurements,” *Remote Sen.*, vol. 6, no. 4, pp. 3349, Apr. 2014.
- [54] G. Liu, S. M. Buckley, X. Ding, Q. Chen and X. Luo, “Estimating spatiotemporal ground deformation with improved persistent-scatterer radar interferometry,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 9, pp. 3209–3219, Sep. 2009.
- [55] G. Liu, H. Jia, R. Zhang, H. Zhang, H. Jia, B. Yu, and M. Sang, “Exploration of subsidence estimation by persistent scatterer InSAR on time series of high resolution TerraSAR-X images,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens. (JSTARS)*, vol. 4, no. 1, pp. 159–170, Mar. 2011.
- [56] S. Rebay, “Efficient unstructured mesh generation by means of Delaunay Triangulation and Bowyer-Watson algorithm,” *J. Comput. Phys.*, vol. 106, no. 1, pp. 125–138, May 1993.