A Nonlocal InSAR Filter for High-Resolution DEM Generation from TanDEM-X Interferograms

Gerald Baier, Cristian Rossi, Marie Lachaise, Xiao Xiang Zhu *Senior Member, IEEE*, Richard Bamler *Fellow, IEEE*

Abstract—This paper presents a nonlocal InSAR filter with the goal of generating digital elevation models of higher resolution and accuracy from bistatic TanDEM-X strip map interferograms than with the processing chain used in production. The currently employed boxcar multilooking filter naturally decreases the resolution and has inherent limitations on what level of noise reduction can be achieved. The proposed filter is specifically designed to account for the inherent diversity of natural terrain by setting several filtering parameters adaptively. In particular, it considers the local fringe frequency and scene heterogeneity, ensuring proper denoising of interferograms with considerable underlying topography as well as urban areas. A comparison using synthetic and TanDEM-X bistatic strip map datasets with existing InSAR filters shows the effectiveness of the proposed techniques, most of which could readily be integrated into existing nonlocal filters. The resulting digital elevation models outclass the ones produced with the existing global TanDEM-X DEM processing chain by effectively increasing the resolution from 12 m to 6 m and lowering the noise level by roughly a factor of two.

Index Terms—digital elevation model (DEM), interferometric synthetic aperture radar (InSAR), nonlocal filtering

I. INTRODUCTION

With the global availability of the digital elevation model (DEM) produced by the German Aerospace Center's (DLR) TanDEM-X mission, topographic data with so far nonexistent spatial resolution and height accuracy have become accessible on a global scale. The fact that a complete satellite mission was set in motion and executed for this sole purpose shows the demand and need for such data. Even more attention should be paid not to compromise resolution and accuracy after acquisition by imperfect processing steps.

Phase denoising is a mandatory step within any InSAR DEM production workflow. A more accurate phase estimate results not only in a less noisy DEM but also eases phase unwrapping.

Indiscriminate spatial averaging of the phase, also called boxcar multilooking, while being fast to compute and reducing the variance of the estimate, degrades resolution. To address this issue, more advanced filtering methods have been the topic of research for more than two decades. Lee's sigma filter and its later extensions [1–3] are examples of SAR and polarimetric SAR filters that include statistical tests for selecting pixels in the averaging process.

Nonlocal filters were first introduced for denoising optical images [4]. In recent years, the have become increasingly popular within the denoising community, due to their unsurpassed noise reduction and detail preservation. The foundation of their performance is a highly discriminate search for statistically homogeneous pixels,

G. Baier is with the Remote Sensing Technology Institute, German Aerospace Center, 82234 Wessling, Germany (e-mail: gerald.baier@dlr.de).

C. Rossi is with Satellite Applications Catapult, Harwell Campus, Didcot, OX11 0QR, United Kingdom.

M. Lachaise is with the Remote Sensing Technology Institute, German Aerospace Center, 82234 Wessling, Germany.

X. X. Zhu is with the Remote Sensing Technology Institute, German Aerospace Center, 82234 Wessling, Germany, and also with the Signal Processing for Earth Observation (SiPEO), Technical University of Munich, 80333 Munich, Germany.

R. Bamler is with the Remote Sensing Technology Institute, German Aerospace Center, 82234 Wessling, Germany, and also with the Institute for Remote Sensing Technology, Technical University of Munich, 80333 Munich, Germany.

somewhat akin to the sigma filter, within a large area during the filtering process. These features sparked research into adapting them to new domains, such as denoising regular SAR amplitude images [5–7], interferograms [8, 9], polarimetric SAR [10], and a unified approach for SAR amplitude images, interferograms and polarimetric SAR images [11]. Recent publications applied the nonlocal filtering paradigm to SAR stacks in the fields of differential SAR interferometry [12] and 3D reconstruction using SAR tomography [13]. The first nonlocal InSAR filter [8] piqued our interest to produce DEMs from bistatic TanDEM-X strip map interferograms with improved resolution and accuracy compared to boxcar multilooking, which is employed in DLR's processing chain for the global TanDEM-X DEM.

For the original operational processing chain [14–16], the need to cope with the data volume of the global DEM acquisition imposed severe design restrictions due to computational costs. Boxcar multilooking was finally chosen, as the resulting DEM fulfills the TanDEM-X accuracy requirements [17] and its computational costs are negligible compared to the other processing steps.

Our research was motivated by the need for even higher-resolution DEMs which led DLR to commence research on the high-resolution DEM (HDEM) product, with increased horizontal resolution and vertical accuracy over selected areas [18] compared to the default TanDEM-X DEM product. HDEMs rely on several new acquisitions with larger baselines resulting in smaller height errors from phase noise. For comparison, the heights of ambiguity for HDEM range from 10 m to 20 m whereas the values for the regular DEM start at 35 m and go up to 50 m. Thus, a boxcar averaging phase filter with a smaller spatial extent compared to the default processing toolchain suffices to fulfill the vertical accuracy goal and more of the original spatial resolution can be preserved. Table I gives the specifications of the two available DEM products from DLR.

Our goal was to create a DEM similar in accuracy to the HDEM specifications by reprocessing the acquisitions made for the global TanDEM-X DEM. The findings of our earlier investigation [19, 20] suggest that the qualities of nonlocal filters do indeed transfer to DEM generation. We were able to produce a RawDEM, the initial DEM product used for creating the final TanDEM-X DEM, with 6 m × 6 m resolution showing more details and less noise compared to the operational product with a resolution of 12 m × 12 m.

Yet our straightforward application of NL-InSAR, the nonlocal filter introduced in [8], led to undesired terrace-like artifacts in the final DEM. We also found, that the more recently published NL-SAR filter [11] was unsuitable for DEM generation as it showed a tendency for oversmoothing.

This paper further elaborates on the issues we encountered when applying the nonlocal filtering paradigm to InSAR denoising and proposes a new nonlocal InSAR filter that takes these into consideration. A key feature is its compensation of the deterministic, topographic phase component, which hampers the search for statistically homogeneous pixels in mountainous terrain. It further factors in the diversity of natural terrain by using a local scene heterogeneity measure to select key filtering parameters instead of relying on a global, fixed set. These techniques can readily be integrated into existing nonlocal InSAR filters to also bolster their

Table I
RESOLUTION AND ACCURACY REQUIREMENTS OF THE STANDARD GLOBAL TANDEM-X DEM AND THE LOCALLY AVAILABLE HDEM [21].

	Independent pixel spacing	Absolute horizontal and vertical accuracies (90 %)	Relative vertical accuracy (90 % linear point-to-point)
(global) TanDEM-X DEM	12 m (0.4″ at equator)	10 m	2 m (slope ≤ 20 %) 4 m (slope > 20 %)
(local) TanDEM-X HDEM	6 m (0.2″ at equator)	10 m	goal: 0.8 m (90 % random height error)
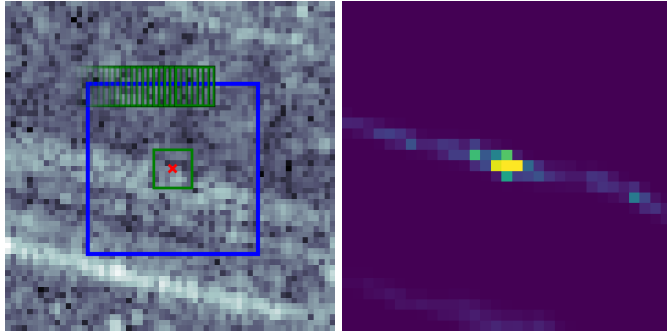
(a) Search window (blue square) and patches (green squares)   (b) Similarity map

Figure 1. The nonlocal filtering process: Inside the search window (blue square) centered at the pixel that is to be filtered (a), all pixels are checked for their similarity by comparing their surrounding patches to the center patch (green squares). The corresponding similarity map (b) shows that similar pixels are located along the edge.

performance. A comparison with a LiDAR DEM gives an impression of and quantifies the level of improvement that can be achieved by employing nonlocal filters instead of conventional filters to real data. Concerning the vastly increased computational cost, with the advances in semiconductor manufacturing processes and computing architecture, especially graphics processing units (GPUs), large-scale nonlocal filtering of SAR interferograms is nowadays feasible [22].

The paper is structured as follows. Section II briefly introduces the nonlocal filtering concept with respect to SAR interferometry. The design decisions of the proposed filter are described in Section III and are backed up by the experiments in Section IV. We discuss the impact of the new filter in Section V and conclude together with an outlook in Section VI.

II. NONLOCAL INSAR FILTERING

What sets nonlocal filters apart from other filters is the large area they operate over for denoising each pixel. This area, called the search window, is inspected for similar pixels. Their absolute position does not influence the later filtering process, true to their name "nonlocal", unlike with many conventional neighborhood filters. For detecting similar pixels, nonlocal filters do not only rely on comparing the pixel value alone, but also take their surrounding areas, henceforth referred to as patches, into account. By doing so, textures, structures and features help identifying similar pixels and influence the filtering results to a far larger degree than with conventional filters.

Fig. 1 illustrates this filtering process, where, in order to denoise the pixel marked by the red cross, all pixels inside the search window (blue square) are considered by comparing their surrounding patches to the center pixel's patch (all as green squares). The resulting similarity map is depicted on the right and shows that the most similar pixels are located along the edge.

In the original version of the nonlocal filter, the Euclidean distance between patches was used as a measure of similarity. This measure is

the least square estimate for additive white Gaussian noise, a common and practical model for optical images. As the noise characteristics of SAR profoundly differ, the earlier referenced filters for SAR, InSAR and polarimetric SAR all define similarity criteria depending on the statistics of the observed quantities: the speckle noise for SAR amplitude images, the interferometric phase for InSAR, or the covariance matrix for (Pol)(In)SAR.

The patch dissimilarities Δ in the search window are mapped into weights w by a kernel. In most cases, an exponential kernel or a slight adaption thereof is used

$$ w = e^{-\frac{\Delta}{h}} \, , \tag{1} $$

where h sets the trade-off between filtering strength and detail preservation. In the following, we assume that the weights are normalized to sum to one. The estimate of an image z, in our case the interferogram

$$ z = u_1 \bar{u}_2 = A_1 A_2 e^{j\varphi} = |z| e^{j\varphi} \tag{2} $$

of the master and slave images, at the pixel location $\mathbf{x}$ is computed as the weighted mean over the corresponding search window $\partial_{\mathbf{x}}$

$$ \hat{z}_{\mathbf{x}} = \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} z_{\mathbf{y}} \, . \tag{3} $$

The argument $\hat{\varphi} = \angle \hat{z}$ of $\hat{z}$ is the estimate of the true interferometric phase θ. In a similar fashion, estimates of the intensity

$$ \hat{I}_{\mathbf{x}} = \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \frac{|u_{1,\mathbf{y}}|^2 + |u_{2,\mathbf{y}}|^2}{2} \tag{4} $$

and coherence

$$ \hat{\gamma}_{\mathbf{x}} = \frac{\left| \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} u_{1,\mathbf{y}} \bar{u}_{2,\mathbf{y}} \right|}{\sqrt{\sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} |u_{1,\mathbf{y}}|^2 \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} |u_{2,\mathbf{y}}|^2}} \tag{5} $$

can be obtained. One can think of the nonlocal filter as a selector for statistically homogeneous pixels for the averaging process.

When dealing with SAR images and InSAR images in particular, there are several additional factors to consider when applying the nonlocal filter paradigm. The next section highlights these pitfalls and describes how they are addressed specifically by the proposed method.

III. PROPOSED FILTER

In the following, we will refer to the proposed method as NL-SWAG, short for **N**on**L**ocal-**S**AR interferogram filter for **w**ell-performing **A**ltitude Map **G**eneration. Figure 2 shows a high-level flow graph of NL-SWAG. The following paragraphs describe in greater detail the individual operations and how they affect the
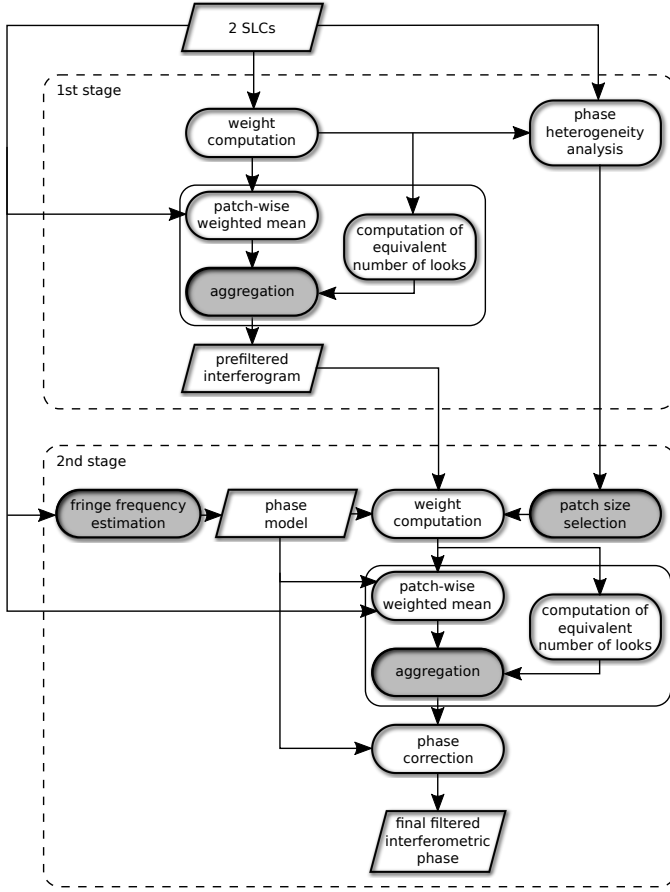
Figure 2. Flow graph of the proposed filter. Blocks that are highlighted in gray have their own respective subsections, which will also cover other related operations. The second stage uses the prefiltered output of the first stage for computing a new, more reliable set of weights.

filtering performance and outcome. We have highlighted in gray operations that are explicitly explained in the respectively named subsections, which will also cover other related blocks.

A. Aggregation

A common filtering artifact of nonlocal filters is the so-called *rare patch effect*, which occurs when only few similar patches are located within the search window, resulting in subpar filtering performance. The problem is especially prevalent near edges, as Fig. 1 illustrates, where for all patches that include the edge only few similar patches are found. Aggregating multiple estimates is one approach to counter this behavior [23].

Instead of the traditional pixel-wise nonlocal means filter as in Eq. (3), NL-SWAG computes the patch-wise weighted mean

$$\hat{\mathbf{z}}_{\mathbf{x}} = \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \mathbf{z}_{\mathbf{y}} \ . \tag{6}$$

The overlapping patch estimates $\hat{\mathbf{z}}$ are then aggregated into a single pixel estimate, weighted by their equivalent number of looks L

$$\hat{z}_{\mathbf{x}} = \frac{\sum_{\mathbf{y} \in \mathcal{P}_{\mathbf{x}}} L_{\mathbf{y}} \hat{\mathbf{z}}_{\mathbf{y},\mathbf{x}-\mathbf{y}}}{\sum_{\mathbf{y} \in \mathcal{P}_{\mathbf{x}}} L_{\mathbf{y}}} \ , \tag{7}$$

where $\mathcal{P}_{\mathbf{x}}$ denotes the set of all pixel indices within a patch centered at $\mathbf{x}$ and $\mathbf{x} - \mathbf{y}$ being the relative index inside the respective patch, i.e., $\mathbf{z}_{\mathbf{y},\mathbf{x}-\mathbf{y}} = z_{\mathbf{x}}$.

The weighting by L ensures that patch estimates with a higher number of looks, and therefore a smaller variance, have a larger impact on the final estimate. The effective number of looks, i.e., the variance reduction of the weighted mean, can directly be computed from the weight map [8]

$$L_{\mathbf{x}} = \frac{\left(\sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \right)^2}{\sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}}^2} \ . \tag{8}$$

Aggregation mitigates the rare patch effect as it also properly denoises pixels near features, such as edges, as long as they also belong to patches which do not contain said features.

B. Two Stage Filtering

SAR interferograms are affected by speckle and suffer from phase noise due to the innate coherence loss between two acquisitions, rendering the similarity estimates difficult and hereby degrading the denoising performance.

A solution that is often employed is a two-stage approach [11, 24, 25], where in the first step the so-called *guidance image* is generated by prefiltering the input image. In the second step, the guidance image is used to compute the patch similarities, which can now be more reliably estimated due to the reduced noise level.

The stages of NL-SWAG, which are also depicted in Fig. 2, employ the two similarity criteria derived in [8] for two single look complex images (SLC) in the first stage and a filtered interferogram in the second stage.

1) First stage: The similarity of two pixels in the first stage is the conditional likelihood of observing $u_{i,\mathbf{x}}$ and $u_{i,\mathbf{y}}$ ($i = 1, 2$), given that the true parameters, the coherence γ, the intensity I and the interferometric phase θ are identical [8]:

$$p(u_{1,\mathbf{x}}, u_{1,\mathbf{y}}, u_{2,\mathbf{x}}, u_{2,\mathbf{y}} | I_{\mathbf{x}} = I_{\mathbf{y}}, \theta_{\mathbf{x}} = \theta_{\mathbf{y}}, \gamma_{\mathbf{x}} = \gamma_{\mathbf{y}}) =$$
$$\delta_{\mathbf{x},\mathbf{y}}^1 = \sqrt{\frac{B}{C}}^3 \left(\frac{A+C}{A} \sqrt{\frac{C}{A-C}} - \arcsin \sqrt{\frac{C}{A}} \right), \tag{9}$$

where

$$A = \left(A_{1,\mathbf{x}}^2 + A_{2,\mathbf{x}}^2 + A_{1,\mathbf{y}}^2 + A_{2,\mathbf{y}}^2 \right)^2 \ ,$$
$$B = A_{1,\mathbf{x}} A_{2,\mathbf{x}} A_{1,\mathbf{y}} A_{2,\mathbf{y}} \text{ and}$$
$$C = 4 \left(A_{1,\mathbf{x}}^2 A_{2,\mathbf{x}}^2 + A_{1,\mathbf{y}}^2 A_{2,\mathbf{y}}^2 + 2B \cos\left(\varphi_{\mathbf{x}} - \varphi_{\mathbf{y}} \right) \right) \ .$$

The patch similarity in the first stage is computed as

$$\Delta_{\mathbf{x},\mathbf{y}}^1 = \sum_{\mathbf{o} \in \mathcal{O}} \log \delta_{\mathbf{x}+\mathbf{o},\mathbf{y}+\mathbf{o}}^1 \ , \tag{10}$$

where $\mathcal{O}$ denotes the set of all index offsets in the patch.

The dissimilarities are mapped into weights by an exponential kernel as in Eq. (1). As the purpose is only to reduce the noise level and remove outliers without introducing severe filtering artifacts before computing the similarities in the second step h is set to a comparatively small value. Except for the aggregation step, the first stage is identical to the non-iterative version of NL-InSAR and its guidelines for picking h can be used. The estimates of the phase, intensity and coherence are obtained via Eq. (3), Eq. (4) and Eq. (5) together with the aggregation in Eq. (6) and Eq. (7).

2) Second stage: The second stage computes the similarities as a function of the coherence $\hat{\gamma}$, intensity $\hat{I}$ and interferometric phase $\hat{\varphi}$ estimates produced by the first stage. The symmetric Kullback-Leibler divergence of two zero-mean complex circular Gaussian distributions, the underlying joint distribution of $\hat{\gamma}$, $\hat{I}$ and $\hat{\varphi}$, is given by [8]

$$\delta_{\mathbf{x},\mathbf{y}}^2 = \frac{4}{\pi} \left[\frac{\hat{I}_{\mathbf{x}}}{\hat{I}_{\mathbf{y}}} \frac{1 - \hat{\gamma}_{\mathbf{x}}\hat{\gamma}_{\mathbf{y}} \cos(\hat{\varphi}_{\mathbf{x}} - \hat{\varphi}_{\mathbf{y}})}{1 - \hat{\gamma}_{\mathbf{y}}^2} \right.$$
$$\left. + \frac{\hat{I}_{\mathbf{y}}}{\hat{I}_{\mathbf{x}}} \frac{1 - \hat{\gamma}_{\mathbf{y}}\hat{\gamma}_{\mathbf{x}} \cos(\hat{\varphi}_{\mathbf{y}} - \hat{\varphi}_{\mathbf{x}})}{1 - \hat{\gamma}_{\mathbf{x}}^2} - 2 \right] \quad (11)$$

and can be used as a similarity criterion. Instead of a fixed patch size, the second stage changes the patch size adaptively based on the local heterogeneity. The exact patch similarity and weight computation are covered in the following two sections since, as can be seen from Fig. 2, it is based on other operations.

Even though the two-step approach alleviates the problems caused by the high noise level in SAR images, we have to stress that a repeated application of any filter can potentially introduce staircase-like artifacts in the filtered output as we observed with NL-InSAR.

To elaborate a little further: Just like traditional neighborhood filters, nonlocal filters can also be seen as diffusion filters [26]. Diffusion filters have the interesting property that their repeated application steadily decreases the noise level and produces piecewise constant approximations of the original data [27]. While this can actually be a desired result for image segmentation or generating abstractions [28], for example, bilateral filters are often used to cartoonify photographs, in our case this phenomenon may lead to staircases in the generated DEM for iterative nonlocal algorithms, as errors of the phase estimate propagate and aggregate with every iteration.

C. Patch Size Selection

Patches contain information about the local texture and hence play a crucial role in distinguishing between suitable patches for averaging and patches that should be discarded. That raises the question: How to select the best patch size? In [29], the authors demonstrated that a global selection was suboptimal and that patch size should depend on the local neighborhood. The following paragraphs repeat their reasoning and puts it into the context of SAR interferogram denoising for DEM generation.

For the original nonlocal filter, patch similarity, just like Eq. (10), is essentially the sum of all contained pixel similarities. Naturally large patches reduce the variance and provide the most robust estimate of patch similarity. This is indeed the best strategy for plains, agricultural fields or other slowly varying terrain.

The situation is quite different for more complex terrain, for instance urban sites or mountain ridges. In these areas, a large patch size leads to the rare patch effect, since for every patch that contains some local structure only patches with similar features will have a significant impact on the averaging process. The likelihood of finding such patches decreases with increasing patch size.

NL-SWAG's solution is to adaptively select the patch size as a function of local scene heterogeneity. This way, a more robust patch similarity can be computed in flat regions or moderately hilly areas, due to the larger patch size, while at the same time the rare patch effect is alleviated in areas with many features and details. Yet we have to stress that small patches come at the cost of less reliable patch similarity estimates.

We would further like to draw attention to the fact that the argument for an adaptive patch size selection to avoid the rare patch

effect is identical to the one for aggregation. Both measures favor patches that exclude local structures by either shrinking the patch or including estimates where the patch is moved off-center with respect to the pixel that is to be denoised. This is somewhat contrary to the initial argument that patch-based methods perform so well because they take textures and details into account. Patches indeed provide an effective mean for discarding patches of different classes. But to maximize the number of patches that are classified as similar, both techniques also try to use the patch modification schemes we just mentioned.

To identify heterogeneous pixels and select the patch size accordingly we apply the local phase heterogeneity measure derived in [30]

$$\eta_{\mathbf{x}} = \frac{\text{Var}\,\{\varphi\}_{\mathbf{x}} - \sigma_{0,\mathbf{x}}^2}{\text{Var}\,\{\varphi\}_{\mathbf{x}}} \,, \quad (12)$$

which lies in the interval $[0, 1)$. $\text{Var}\,\{\varphi\}$ is the estimated variance of the phase in the search window and σ_0^2 the variance one would expected from the coherence [31]. For non-heterogeneous terrain, $\text{Var}\,\{\varphi\}$ is comparable in magnitude to σ_0^2 as only phase noise causes phase changes and Eq. (12) is close to 0. The situation changes when the search window contains structures, such as buildings. Their distinct phase profiles increase $\text{Var}\,\{\varphi\}$ resulting in larger phase heterogeneity values.

As the phase is wrapped, the filter first performs local unwrapping as in [30] to obtain the locally unwrapped phase $\tilde{\varphi}$ with respect to the average of the 5×5 pixels in the center. The phase variance is then estimated inside the search window, weighted by the respective weight map computed in the first stage

$$\text{Var}\,\{\varphi\}_{\mathbf{x}} = \text{E}\,\{\varphi_{\mathbf{x}}^2\} - \text{E}\,\{\varphi_{\mathbf{x}}\}^2$$
$$\approx \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \tilde{\varphi}_{\mathbf{y}}^2 - \left(\sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \tilde{\varphi}_{\mathbf{y}} \right)^2 . \quad (13)$$

As $\text{Var}\,\{\varphi\}$ is estimated in a local window due to insufficient sample size, Eq. (12) might be negative. In this case, the heterogeneity measure is set to zero.

To yield a more reliable estimate of σ_0 the coherence is estimated following the methodology in [32] as

$$\gamma = \frac{\text{E}\,\{|u_1|^2 \cdot |u_2|^2\}}{\sqrt{\text{E}\,\{|u_1|^4\}\,\text{E}\,\{|u_2|^4\}}} \,. \quad (14)$$

This way, the coherence is estimated from the speckle pattern and is not influenced by the topographic phase, which would yield an underestimation of the coherence if the common coherence estimator is used. Just like in Eq. (13), the expected value is replaced by the weighted mean over the respective quantities.

An example of the heterogeneity measure is depicted in Fig. 3. The urban area is clearly detected as being heterogeneous, the grassland is classified as the most homogeneous site and the forested areas are identified as moderately heterogeneous regions.

Instead of selecting a fixed patch size from a predefined set, depending on the local heterogeneity, NL-SWAG employs Gaussian windows of variable width. A possible mapping of the phase heterogeneity index into Gaussian window widths could be

$$\sigma_{\text{Gauss}} = 2 \cdot (1 - \eta) + 1 \,, \quad (15)$$

which gives strict lower and upper bounds for the window widths and is used in the remaining of the paper. Other mappings would
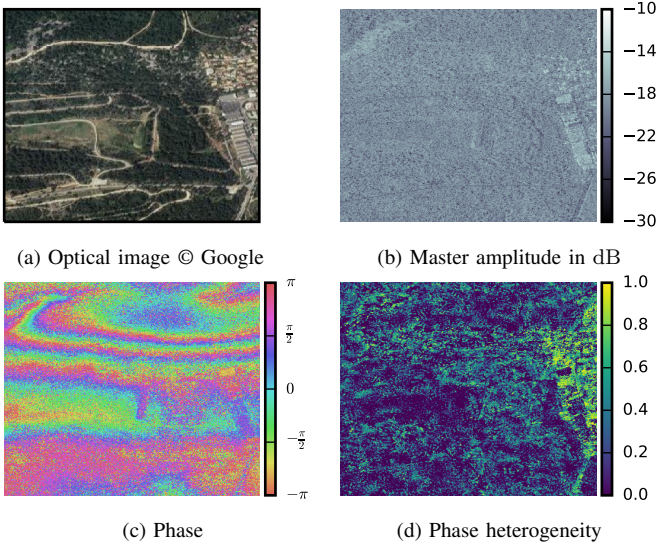
(a) Optical image © Google

(b) Master amplitude in dB

(c) Phase

(d) Phase heterogeneity

Figure 3. Phase heterogeneity computed in the first stage. Urban areas, forests and grassland show different levels of heterogeneity.

Figure 4. Relationship between the width of the Gaussian window σ_{Gauss} and the standard deviation of the resulting patch similarities σ_{Δ^2}. Due to the correlation of the pixel similarities there is no linear mapping.

also be possible as long as they result in wide Gaussian windows for homogeneous areas and the reverse for heterogeneous areas.

As an alternative approach for selecting the best effective patch size, the phase variance in Eq. (12) could be computed in Gaussian windows of successively increasing widths. This process is halted as soon as the heterogeneity level exceeds a predefined threshold, i.e., when significant phase changes, which most likely are the result of heterogeneous structures inside the patch, are detected. A similar approach was presented in [33] for adaptively selecting the search window size.

For Gaussian blurring, the reduction in variance is related to σ_{Gauss} by approximately $4\pi\sigma_{\text{Gauss}}^2$. So with Eq. (15), the variance of the patch similarity estimation is reduced by a factor ranging from 4π to 36π, roughly equivalent to 3×3 up to 11×11 patches.

Correspondingly to Eq. (10), the adaptive patch similarities are computed as the sum over the pixel similarities weighted by a Gaussian window $g_{\mathbf{x}}$

$$\Delta_{\mathbf{x},\mathbf{y}}^2 = \frac{\sum_{\mathbf{o}\in\mathcal{O}} g_{\mathbf{x},\mathbf{o}} \delta_{\mathbf{x}+\mathbf{o},\mathbf{y}+\mathbf{o}}^2}{\sum_{\mathbf{o}\in\mathcal{O}} g_{\mathbf{x},\mathbf{o}}} \ . \tag{16}$$

The patch dissimilarities still need to be mapped into weights, which in the second stage is also done by an exponential kernel. We now face the problem how to select the normalization factor h to compromise between bias and variance reduction.

The standard deviation of Δ^2 is reciprocally proportional to σ_{Gauss}, which effectively governs the patch size. Consequently, a fixed h for all heterogeneity levels will be insufficient and a method is needed that accounts for varying patch sizes. For this purpose, we selected a homogeneous training area and analyzed how the patch similarity's standard deviation σ_{Δ^2} changed with varying $\frac{1}{\sigma_{\text{Gauss}}}$. Fig. 4 shows the relationship for a fixed set of Gaussian window widths at a homogeneous test site without any topography. Clearly, the relationship is non-linear, due to the correlation between pixel similarities, but a second order polynomial, also depicted, is a good fit.

The weights are computed as

$$w_{\mathbf{x},\mathbf{y}} = \exp\left\{-\frac{\Delta_{\mathbf{x},\mathbf{y}}^2}{h \cdot \xi\left(\sigma_{\text{Gauss},\mathbf{x}}^{-1}\right)}\right\} \ , \tag{17}$$
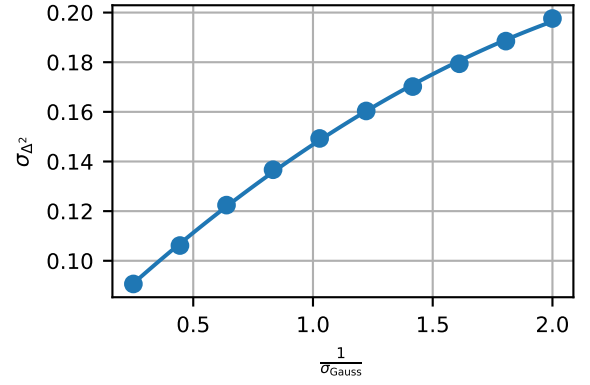
where ξ is the second order polynomial that accounts for the varying effective patch sizes and h provides a fixed compromise between detail preservation and noise reduction. In our experiments, we found that the interval $[1 \leq h \leq 2]$ provided the best trade-off.

To account for the fact that, due to the Gaussian window, not every pixel in the patch estimate contributed equally to the similarity computation in contrast to Eq. (7) the respective pixels are additionally weighted by their Gaussian weight in the final aggregation step

$$\hat{z}_{\mathbf{x}} = \frac{\sum_{\mathbf{y}\in\mathcal{P}_{\mathbf{x}}} L_{\mathbf{y}} g_{\mathbf{y},\mathbf{x}-\mathbf{y}} \hat{\mathbf{z}}_{\mathbf{y},\mathbf{x}-\mathbf{y}}}{\sum_{\mathbf{y}\in\mathcal{P}_{\mathbf{x}}} L_{\mathbf{y}} g_{\mathbf{y},\mathbf{x}-\mathbf{y}}} \ . \tag{18}$$

D. Fringe frequency estimation and compensation

Another obstacle hindering the use of nonlocal InSAR filters for DEM generation is the actual topography, which, together with the atmosphere, the deformation and noise, contributes to the measured interferometric phase. For the bistatic case, the acquisition mode of TanDEM-X interferograms for the generation of the global DEM, the deformation and the atmosphere components can be ignored, so that only the topography and noise components affect the similarity measure. Due to the topographic phase component it is considerably harder to detect statistically homogeneous pixels in regions with non-negligible height differences, that is pixels with identical noise distribution but different heights. Fig. 5 shows the symmetric Kullback-Leibler divergence from Eq. (11) with $\hat{I}_{\mathbf{x}} = \hat{I}_{\mathbf{y}}$ and $\hat{\gamma}_{\mathbf{x}} = \hat{\gamma}_{\mathbf{y}}$ as a function of the coherence and the phase difference $\Delta_{\hat{\varphi}} = \hat{\varphi}_{\mathbf{x}} - \hat{\varphi}_{\mathbf{y}}$, that is used in the second stage as the similarity criterion. Evidently the similarity quickly drops off with increasing $\Delta_{\hat{\varphi}}$ and the higher the coherence the more dramatic the decline. Consequently, the denoising performance suffers in hilly or mountainous terrain. This effect is quite pronounced for bistatic TanDEM-X data due to their generally high coherence.

This analysis is not exclusive to the Kullback-Leibler divergence. Similar arguments can be made for different similarity criteria, i.e., the one employed in [11] and Eq. (9).

To combat the reduced denoising performance for terrain with significant height changes, we incorporated a linear fringe model as in [34] that accounted for the deterministic, topographic phase component when computing the similarities and the weighted mean. Our approach is distantly related to [35], which employs affine transforms to find more similar patches.

For every pixel, the fringe compensation algorithm obtained an estimate of the fringe frequencies in azimuth and range $\mathbf{f} = [f_{\text{r}}, f_{\text{az}}]^T$
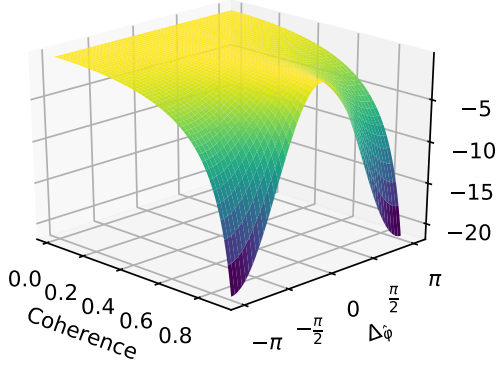
Figure 5. Symmetric Kullback-Leibler divergence from Equation (11) for two pixels with identical reflectivity and coherence, dependent on their phase difference.

using the 2D Fourier transform. To circumvent abrupt changes of the fringe frequency estimates, we smoothed $\mathbf{f}$ with a Gaussian kernel.

Without loss of generality we can consider Eq. (11) as a function of only the phase difference between two pixels

$$\delta^2_{\mathbf{x},\mathbf{y}}(\hat\varphi_{\mathbf{x}} - \hat\varphi_{\mathbf{y}}) . \qquad (19)$$

The fringe compensation takes the fringe frequencies at $\mathbf{x}$ into account by changing the pixel similarity function to

$$\delta^2_{\mathbf{x},\mathbf{y}}(\hat\varphi_{\mathbf{x}} - (\hat\varphi_{\mathbf{y}} - (\mathbf{x} - \mathbf{y})^T\mathbf{f}_{\mathbf{x}}) \bmod 2\pi), \qquad (20)$$

that is, we remove the phase component caused by the fringe frequency in azimuth and range.

The computation of the patch-wise weighted mean of the interferogram has to account for the phase model

$$\hat{\mathbf{z}}_{\mathbf{x}} = \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \mathbf{z}_{\mathbf{y}} \cdot e^{-\mathrm{j}(\mathbf{x}-\mathbf{y})^T\mathbf{F}_{\mathbf{x}}} . \qquad (21)$$

Here $\cdot$ denotes element-wise multiplication and $\mathbf{F} \in \mathbb{R}^{2 \times p \times p}$ is a three dimensional tensor that contains all fringe frequencies of the pixels inside the $p \times p$ patch centered at $\mathbf{x}$.

Figure 6 shows the effect that fringe frequency compensation has on the noise reduction. Denoising of a nonlinear phase ramp with constantly increasing frequency was performed using NL-SWAG with and without fringe frequency compensation. If the fringe frequency is not accounted for, the phase estimate's standard deviation increases steadily with increasing frequency. With fringe frequency compensation, the standard deviation is limited. Due to the discrete nature of the frequency estimation by fast Fourier transform in our implementation, the frequency was not perfectly estimated and the performance was not entirely frequency independent, which resulted in the wave-like pattern of the standard deviation. A more sophisticated frequency estimation algorithm would certainly alleviate this problem.

As a final note, we would like to point out the difference between the fringe frequency compensation and the local phase heterogeneity-based adaptive patch size selection. Both approaches address deterministic phase changes which can hamper the search for similar patches. But whereas the fringe frequency compensation strictly deals with large-scale phase changes due to topography by a linear compensation, the role of the phase heterogeneity is more to take care of arbitrary small-scale phase changes, which would not necessarily be captured by a simple linear approximation.

(a) without fringe compensation
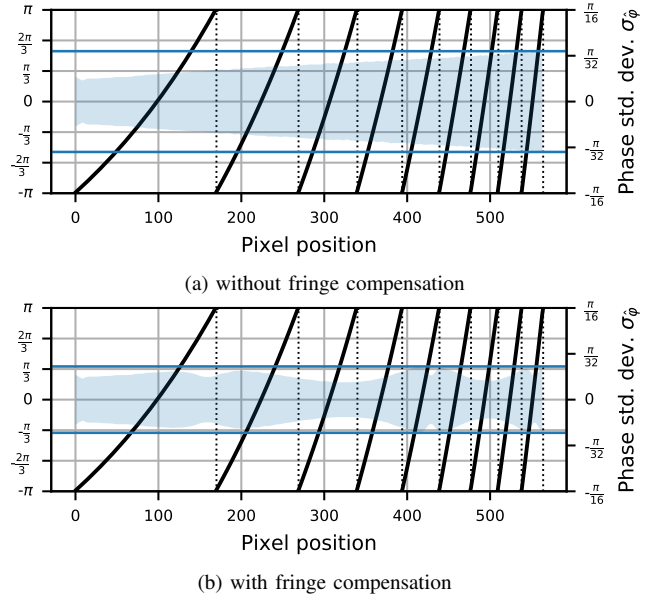

(b) with fringe compensation

Figure 6. Standard deviation (shaded blue area) of NL-SWAG's estimate of a nonlinear phase profile (in black) with and without compensating for the fringe frequency. The maximum value of the standard deviation are marked with a horizontal blue line. If the filter does not account for the deterministic phase change inside the search window the denoising performance decreases substantially with increasing frequency.

Table II
TANDEM-X STRIP MAP PARAMETERS OF THE TEST SITES

Parameter	Test site	Value
Range bandwidth	—	100 MHz
Ground range resolution	—	3 m
Azimuth resolution	—	3 m
Polarization	—	HH
Height of ambiguity	Marseille	30 m
	Munich	48 m
	Barcelona	48.5 m

IV. EXPERIMENTAL RESULTS

We compared NL-SWAG using simulations and real world data sets with existing nonlocal filters. We used TanDEM-X bistatic strip map interferograms of three different test sites: Marseille, Munich, and Barcelona for the evaluation. The most pertinent parameters are listed in Table II. The experiments also substantiated our claim of creating a DEM close in quality to the HDEM specifications in Table I.

In addition, the comparison included the result of a simple 5×5 Boxcar filter. Boxcar filters, the dimensions of which depend on range resolution, incidence angle and imaging mode, are employed in DLR's integrated processor (ITP) [14–16] for generating the global TanDEM-X DEM. For strip map data the dimensions of all employed Boxcar filters are close to 5×5 and their individual results will not be reported here.

We also analyzed NL-InSAR [8], the first nonlocal InSAR filter, where we set the search window size to 21×21, the patch size to 7×7 and used five iterations. We deviated from the suggested ten iterations in the original publication as in our experience the changes in estimation accuracy are negligible after about four to five iterations. Furthermore, the refinement provided by the iterations only resulted in improved detail preservation, which as we will show NL-InSAR already excels at, even with only five iterations. Also, more iterations aggrevate the aforementioned terrace-like artifacts.

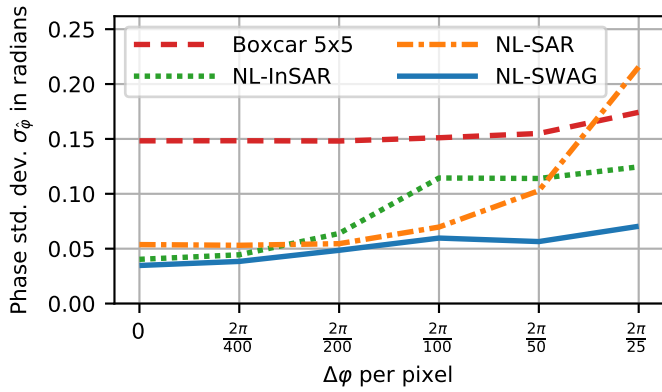The second nonlocal filter in the comparison was NL-SAR [11].

Figure 7. Standard deviation of the phase estimate as a function of a constant ramp's inclination. The steeper the incline the higher the standard deviation. The fringe frequency estimation of NL-SWAG alleviates this problem.

NL-SAR adaptively selects the best parameters from a predefined set, which includes the patch size, search window size and the strength of the initial prefiltering step. In our analysis, we used the same predefined set as in the original paper.

In all subsequent experiments concerning NL-SWAG, the search window size was set to 21×21, h to 4 in the first stage and to 2 in the second stage. The block size of the fringe estimation was 32×32 and the size of the discrete Fourier transform's was set to 64×64. This zero padding increases the accuracy of the fringe estimation.

A. Synthetic Data

Assuming fully developed speckle, the correlated complex normal distributed pixels of two SLCs have the covariance matrix [36]

$$\mathbf{C} = \begin{bmatrix} A^2 & A^2 \gamma e^{j\varphi} \\ A^2 \gamma e^{-j\varphi} & A^2 \end{bmatrix} \quad (22)$$

where A denotes the amplitude, φ the interferometric phase and γ the coherence.

Let $\mathbf{C} = \mathbf{L}\mathbf{L}^\dagger$ be the Cholesky decomposition of the covariance matrix $\mathbf{C}$, where $\dagger$ denotes conjugate transpose. A multiplication with $\mathbf{L}$ transforms two independent complex normal distributed samples r_1 and r_2 of zero mean and unit variance

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \mathbf{L} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = A \begin{bmatrix} 1 & 0 \\ \gamma e^{-j\varphi} & \sqrt{1-\gamma^2} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \quad (23)$$

to samples with the desired correlation properties, amplitude and phase defined by the covariance matrix.

An analysis for the slope-dependent noise suppression was carried out by denoising phase ramps of different inclinations. In the simulations, the intensity was constant for the whole slope and the coherence was set to 0.7. Figure 7 shows the standard deviation of the various filters' phase estimates for different inclinations, which is given as the phase change per pixel in radians.

The nonlocal filters are more sensitive to changes in inclination compared to the Boxcar filter as a result of their large search windows. NL-InSAR and NL-SAR in particular, since they do not compensate for the deterministic phase component. As mentioned earlier, the fringe estimation of NL-SWAG was not perfect due to the discrete nature of the fast Fourier transform used in the implementation and hence was still slope-dependent. Overall, we can see that NL-SWAG provided an improvement of roughly a factor of three compared to the Boxcar estimate over all frequencies.

Fig. 8 and Fig. 9 give an impression of the resolution preservation capabilities of the various filters. Both figures are the result of Monte-Carlo simulations with 10,000 repetitions when estimating a phase
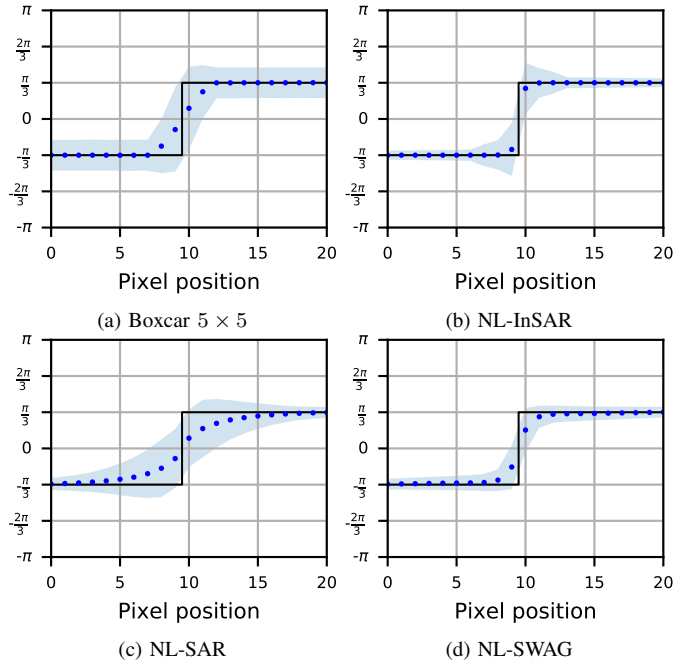

(a) Boxcar 5×5 (b) NL-InSAR

(c) NL-SAR   (d) NL-SWAG

Figure 8. Expected value of a step function's phase estimate, constant amplitude and coherence of 0.7. The shaded blue area delineates $\pm$ three times the estimate's standard deviation. We performed 10,000 simulations to obtain the statistics.

jump from $-\frac{\pi}{3}$ to $\frac{\pi}{3}$. The expected values are plotted as blue dots and their standard deviations as shaded blue areas. In Fig. 8, intensity and coherence are constant, with coherence having a value of 0.7, whereas in Fig. 9 coherence increases from 0.6 to 0.8 and the intensity difference is 6 dB.

Fig. 8 shows that the Boxcar filter's result exhibited the expected smoothing. Both NL-InSAR and NL-SWAG were unable to perfectly preserve the edge but fared much better than NL-SAR. The reason for NL-SAR's poor performance is that NL-SAR initially produces an intentionally oversmoothed result and then applies a bias-reduction step based on terrain heterogeneity. This heterogeneity test, however, only considers the intensity and therefore breaks down in this particular case, where only the phase changes.

The situation changed when the phase jump was accompanied by an intensity jump as in Fig. 9. The intensity change aids nonlocal filters in discriminating between similar pixels, resulting in sharper transitions. The benefit of setting the patch size adaptively is highlighted by NL-SAR and NL-SWAG, which do not exhibit a halo of high variance at the discontinuity. We could deduce that the rare patch effect was indeed the cause of this performance degradation, as the width of the halo for NL-InSAR was equal to the employed patch size minus one. All patches in this area included the edge and consequently suffered from the rare patch effect. NL-SWAG additionally benefited by the aggregation step, which further reduced the variance along the edge. Even with these measures in place, we could still see that the variance is increased near the edge.

To illustrate the propensity of the filters to produce the earlier introduced terrace-like features and other biasing artifacts, we simulated a noisy interferogram from a synthetic terrain created by the diamond-square algorithm [37]. Fig. 10 shows in the top row the simulated noisy interferogram with a constant coherence value of 0.7 and the filters' denoised results. The second row shows the true simulated phase and its difference compared to the filter output.

We also include a TanDEM-X interferogram whose phase resembles the simulation in our analysis to exemplify how these filtering
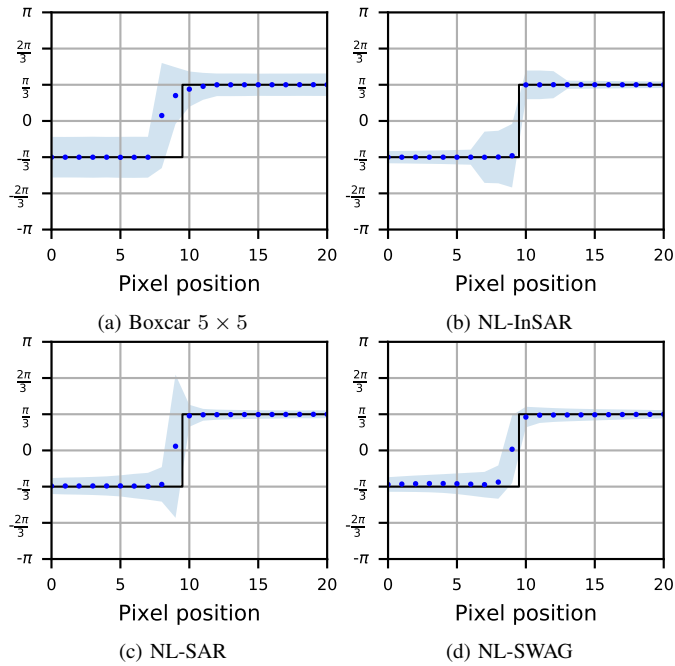
Figure 9. Estimated phase of a step function with a step in coherence from 0.6 to 0.8 and a intensity jump of 6dB. The additional change in intensity compared to Fig. 8 helped the nonlocal filters to preserve the edge.

characteristics affect real data, which is shown in the last row together with shaded reliefs of DEMs generated by the various filters.

For NL-InSAR, a distinct pattern was visible in the difference plot which would manifest as terrace-like artifacts in a generated DEM. Indeed, the DEM produced by NL-InSAR from real data also exhibited similar patterns. Visually, we could asses that the overall noise level of all nonlocal filters was lower compared to the Boxcar filter, especially in regions where the fringe frequency was low. The difference plots also show that nonlocal filters suppressed the high-frequency component of the noise but created slowly varying undulations of spatially correlated noise.

To further shed light on some of the mechanisms of nonlocal filters, Fig. 11 shows the expected value and standard deviation of a Monte-Carlo simulation's phase estimate for the experiment with synthetic data in Fig. 10. All nonlocal filters biased the estimate along the ridge at the interferogram's diagonal. In general, nonlocal filters have a higher propensity to bias the estimate due to their comparatively large search windows. The standard deviation plots show the fringe-frequency dependent noise suppression of NL-InSAR and NL-SAR. NL-SWAG was much less affected by this aspect, although it was also not completely immune as noted earlier. Table III lists the mean standard deviations and the average equivalent number of looks, rounded to the nearest integer, over the whole image and all simulation runs. In accordance with our previous experiments, it was considerably lower for nonlocal filters. Contrasting Table III with Table I reveals that NL-SWAG would fulfill the noise reduction by a factor of 2.5, which is required for the production of a DEM according to the HDEM specifications.

B. Real Data

Experiments on TanDEM-X bistatic strip map interferograms were carried out for three test sites that were chosen to showcase the previously described qualities and phenomena when using nonlocal filters and NL-SWAG in particular for DEM generation. The interferograms from the test sites were processed with DLR's ITP, and

Table III
STANDARD DEVIATION IN RADIANS AND AVERAGE EQUIVALENT NUMBER OF LOOKS, ROUNDED TO THE NEAREST INTEGER, FOR THE MONTE-CARLO SIMULATION IN FIG. 11

	Boxcar 5×5	NL-InSAR	NL-SAR	NL-SWAG
$\sigma_{\hat{\varphi}}$ in rad	0.1482	0.0969	0.0768	0.0537
Number of looks	25	58	93	190

the aforementioned nonlocal filters were used in lieu of the default Boxcar filter.

The first test area was an industrial site near the French city of Marseille and it provided a visual impression of the performance increase that could be expected with nonlocal filters. Fig. 12 shows shaded reliefs of the generated DEMs, an optical image for better interpretation and a plot of the unfiltered phase. The resolution of the DEMs produced with the nonlocal filters was 6 m for longitude and latitude. The DEM generated using the 5×5 Boxcar filter had a resolution of 12 m, the default configuration for DLR's RawDEM. In the global TanDEM-X DEM processing chain, several RawDEMs are later combined to generate the final DEM product.

The higher level of details visible in the nonlocal DEMs is evident, as is the improved noise reduction for agricultural fields and the hill to the south. NL-InSAR produced clearly discernible terraces for the hill, a result of the staircasing effect. The road in the lower half of the image serves as an example for what kind of details can be preserved by the proposed filter.

Also noticeable are noisy artifacts near buildings for NL-InSAR at the industrial site, a consequence of the rare patch effect, which is avoided by NL-SAR and NL-SWAG. NL-SAR, however, tends to oversmooth some details, so that, for example, the road in the lower part of the test site is hardly distinguishable from its surrounding.

Fig. 13 sheds some more light on NL-SWAG's filtering characteristics. It shows the employed width of the Gaussian window used for computing the patch similarities and the final equivalent number of looks after the aggregation step. Both show that homogeneous areas benefit from wide Gaussian windows, resulting in accurate patch similarity estimates, and a large number of similar pixels within the search window, leading to low-noise estimates. The reverse is true for the industrial site, where narrow Gaussian windows were employed, due to the region's heterogeneity. This heterogeneity was also the cause of only a comparatively low number of looks. The impact that the fringe frequency estimation and compensation had on the estimate could be inferred from the equivalent number of looks for the hilly terrain to the south, which was virtually unaffected by the trend of the phase.

As a clearer example of detail preservation, Fig. 14 shows DEMs for an agricultural area near Munich, Germany. The resolution was the same as in the previous example: 6 m for the nonlocal DEMs and 12 m for the Boxcar filter. The data were acquired on August 19, 2011 when some of the fields had already been harvested so the outlines of different fields are clearly discernible, as electromagnetic waves in X-Band only marginally penetrate vegetation [38]. The shaded reliefs confirmed our simulation results in Fig. 8 and Fig. 9 that NL-InSAR provided the best result for this particular scenario, as it favors piecewise constant solutions and sharp edges. But this propensity was also the source of the highly unwelcomed staircasing for regions with a more interesting topographic profile.

We can also see the effect that a change of h has on the filtering result. A lower value of h produced a sharper transition at the edges of the field but reduced denoising in flat terrain.

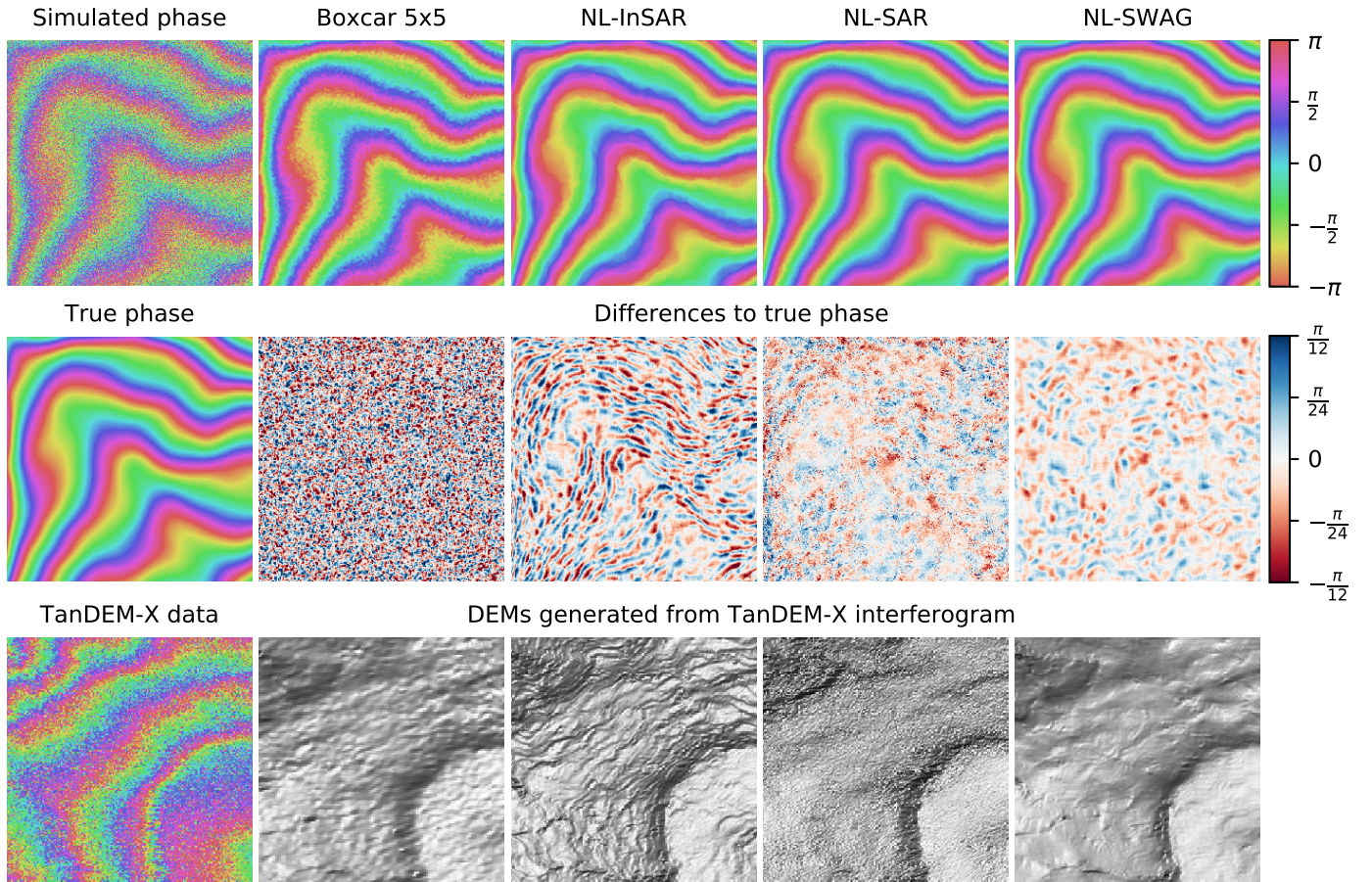As a last example we compared NL-SWAG to a high-resolution

Figure 10. Phase estimates of several filters for a synthetically-generated interferogram and their differences compared to the true phase are shown together with the noisy interferogram (the coherence was set to 0.7) and the true phase in the first two rows. The last row shows a comparable TanDEM-X strip map interferogram and the shaded relief of DEMs generated by the corresponding filter. The phase estimate of NL-InSAR shows a distinct staircase-like pattern, which is also clearly visible in the shaded relief plot. All nonlocal filters suppress the high-frequency component of the noise but produce low frequency undulations in the estimate.
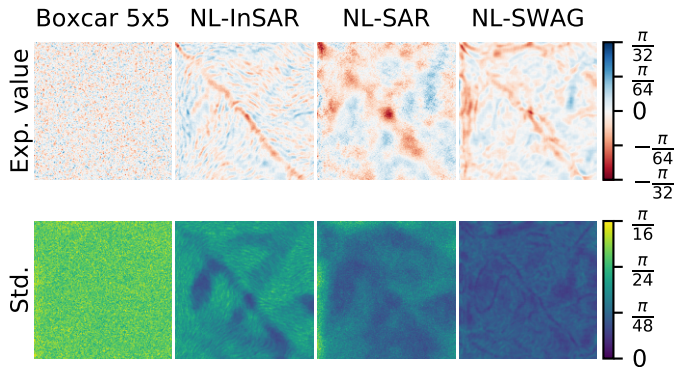

Figure 11. Expected values (top) and standard deviation (bottom) for a Monte-Carlo simulation of the simulated phase in Figure 10. Minor biases are present in the phase estimates. The slope dependent denoising performance of nonlocal filters is evident in the standard deviation plots.

LiDAR DEM, which served as a gold standard for our analysis. The test site was the town Terrassa close to Barcelona in Spain. The top row in Fig. 15 shows an optical image from Google maps and the LiDAR DEM with $5\,\mathrm{m}$ spacing plus DEMs generated from a single TanDEM-X interferogram by a 5×5 Boxcar filter and our proposed method. The DEMs were resampled to the grid of the LiDAR DEM. As LiDAR and SAR have fundamentally different

imaging geometries and properties, we tried to remove areas with systematic errors, such as urban areas suffering from layover and shadowing or vegetation, where the LiDAR's last returns differed from the scattered wave's phase center at X-Band. In order to do so, we compared the LiDAR DEM to the global TanDEM-X DEM and excluded points with a height difference larger than $2\,\mathrm{m}$. The result is depicted in the bottom row of Fig. 15 and a cleaned mask, using morphological operations, right next to it. The height differences are the remaining two pictures annotated with the standard deviation of the height difference computed over the masked area.

This experiment had several noteworthy results. As expected, the SAR DEMs differed substantially from the LiDAR DEM for buildings, and the height values were unusable. However the SAR DEMs could still be used to detect buildings as the test site near Marseille (Fig. 12) showed as well. On the masked-out, moderately hilly, homogeneous terrain, NL-SWAG improved the noise level roughly by a factor of $1.3420\,\mathrm{m}/0.7980\,\mathrm{m} \approx 1.6817$ almost equivalent to a filter with three times as many looks, which is, however, insufficient for completely fulfilling the requirements in Table I.

At first glance, this improvement in noise reduction contradicted our findings reported in Table III. We could exclude systematic height differences due to the different physical properties of LiDAR and SAR as the penetration depth of electromagnetic waves at X-Band is negligible as an error source [39]. Coregistration errors of the LiDAR and SAR DEMs might also be a contributing factor for height

(a) Optical, © Google

(b) Unfiltered phase

(c) Boxcar 5×5

(d) NL-InSAR
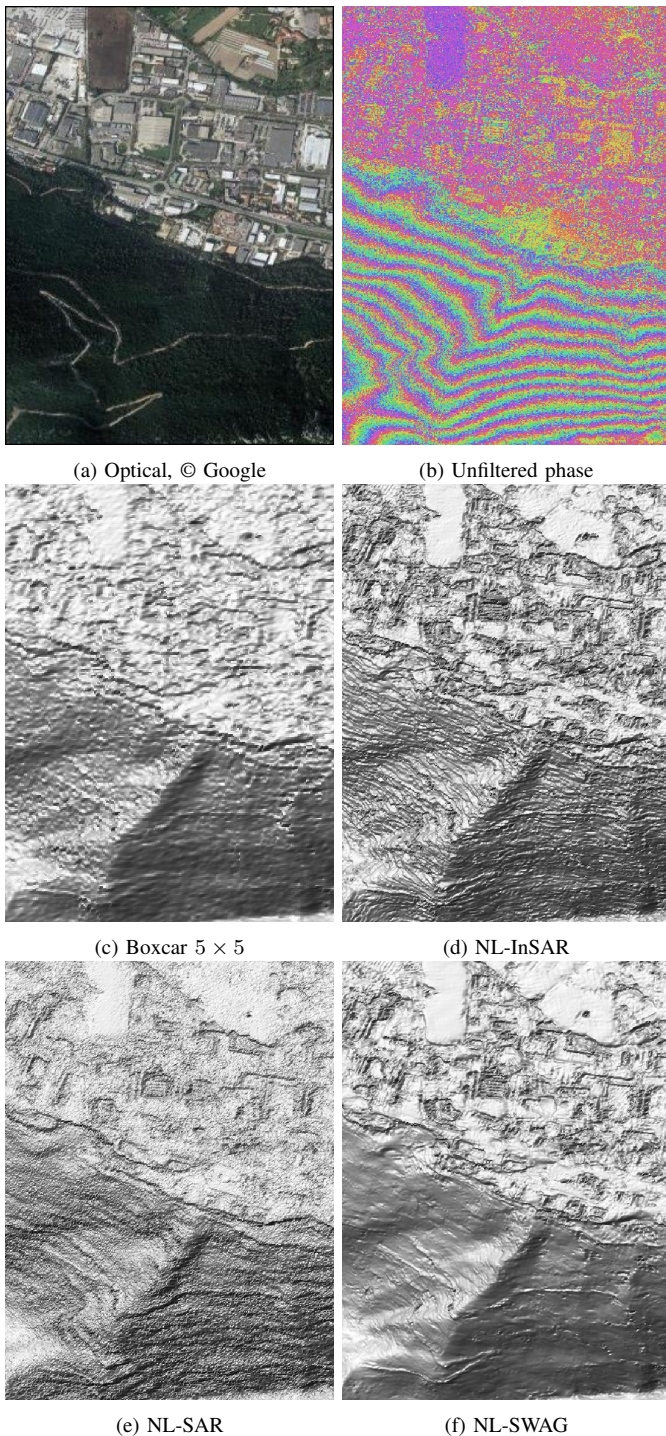
(e) NL-SAR

(f) NL-SWAG

Figure 12. Shaded reliefs of DEMs generated with the various filters. The nonlocal filters improved the resolution and noise level compared to the Boxcar estimate. NL-InSAR suffered from the rare patch effect near structures due to its fixed patch size.

differences but for the moderately hilly terrain they would only play a minor role. Such error sources would equally increase the difference compared to the LiDAR DEM, leading to a misrepresented noise level reduction. The true reason for this discrepancy is the resampling from approximately $3\,\mathrm{m}$ pixel spacing in range and azimuth to the $5\,\mathrm{m}$ LiDAR pixel spacing, which essentially increased the footprint of the Boxcar filter. For NL-SWAG, this effect was imperceptible due to its comparatively large search window.

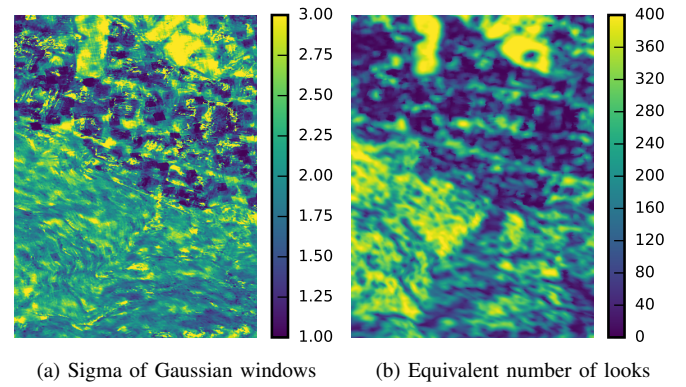(a) Sigma of Gaussian windows

(b) Equivalent number of looks

Figure 13. Width of the Gaussian windows used for computing the patch similarities and the equivalent number of looks for the test site from Fig. 12.

V. DISCUSSION

The initial goal of our investigation was to ascertain whether nonlocal filters were suitable for generating a DEM close to the HDEM standard (see Table I) from the globally available TanDEM-X data. In the following paragraphs, we will detail how the proposed filter held up to these challenges.

All conducted experiments confirmed that nonlocal filters were able to deliver a vastly improved noise reduction over the exemplary local Boxcar filter. The reason is that, due to their large search windows, nonlocal filters found a multitude of pixels for the averaging process, even for comparatively heterogeneous terrain. To further quantify this improvement: For the experiments on synthetic data (Fig. 7 and Table III) the standard deviation was lower by a factor of three and for the real data set of Fig. 15 on moderately complex terrain it was still reduced by a factor of approximately 1.7. Relating this to the level of noise reduction we aimed for in Table I, our filter fell short of reaching the target of 2.5 roughly by a factor of $\sqrt{2}$. Depending on the type of terrain, this might still be sufficient to obtain a DEM that fulfills the requirements of the HDEM, as already the globally available TanDEM-X DEM often overfulfills its accuracy requirements. In any case, having twice as many acquisitions available would also satisfy the specification.

Our proposed filter implemented several techniques to reach this level of noise reduction. It reduced the detrimental effect of topography by its fringe frequency compensation accounting for the deterministic topographic phase component, as evidenced by Fig. 7 and Fig. 11. Furthermore, even on flat, homogeneous terrain, the high inherent noise level of InSAR hampered denoising, which was countered by the two-step approach.

Fig. 11 shows that nonlocal filters bias the estimate for nonlinear phase profiles. The bias is limited by approximately $\pm\frac{\pi}{100}$. With a height of ambiguity of $40\,\mathrm{m}$, which is a typical value for TanDEM-X interferograms, this translates to deterministic height errors of $\pm$ $20\,\mathrm{cm}$, well within the HDEM specifications.

We also highlighted that for nonlocal filters it is far easier to denoise homogeneous terrain than heterogeneous targets, as more similar pixels are found. Nonetheless, nonlocal filters were well suited for preserving heterogeneous targets as shown by the simulation results in Fig. 8 and Fig. 9, where the adaptive patch size and the aggregation step played a significant role to avoid the rare patch effect near the edge.

For filtering SAR interferograms of urban areas or terrain with man-made structures, nonlocal filters were especially appealing as such heterogeneous targets exhibit a very high radar cross-section compared to their surroundings. These high intensity variations aid nonlocal filters to preserve details as their weight maps are more

(a) Optical, © Google

(b) Boxcar 5×5

(c) NL-InSAR

(d) NL-SAR
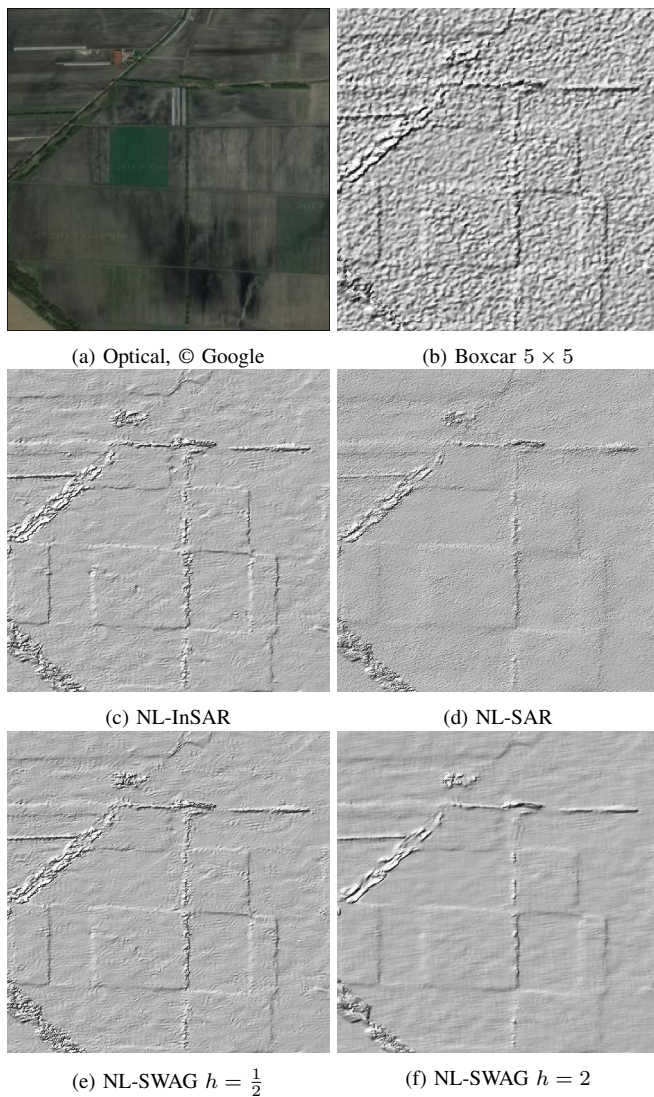
(e) NL-SWAG $h = \frac{1}{2}$

(f) NL-SWAG $h = 2$

Figure 14. Shaded reliefs of DEMs of an agricultural site. Clearly visible are height changes between fields. The bottom row shows the effect that changing h in the second stage has on detail preservation and noise reduction.

discriminant. The gain in resolution, compared to simple boxcar averaging, is evident for real data in Fig. 12 and Fig. 14.

It wold also be rather straightforward to extend existing nonlocal filters with the proposed modifications. The fringe compensation requires only a minor adaption of the similarity criterion. Changing the patch size adaptively is an isolated modification, which could also be performed based on the intensity heterogeneity criterion derived in [40], for example, in a nonlocal SAR despeckling filter. The aggregation step is an extension of the pixelwise weighted mean and can be treated separately from all other adjustments.

VI. CONCLUSION

We showed that applying existing nonlocal filters led to artifacts when generating DEMs. Our analysis highlighted the mechanisms behind the encountered phenomena, like the topographic phase component and the myriad types of terrain and settings, from agricultural fields to city landscapes, in which InSAR filters have to operate. The proposed filter addressed these issues by accounting for the deterministic fringe frequency and setting its filtering parameters adaptively. We demonstrated the effectiveness of these measures which resulted in a comparable noise reduction and detail preservation compared to

other nonlocal InSAR filters without any of the undesired properties. The derived DEMs also far surpassed the RawDEMs produced with the existing global TanDEM-X processing chain, which relies on conventional boxcar multilooking.

We will further evaluate the proposed method on a wider array of real data, which will also highlight some of the characteristics of SAR compared to LiDAR for generating DEMs. Such an extensive evaluation is essential for considering nonlocal filters as a total replacement for the boxcar filter in the TanDEM-X processing chain.

Promising paths for future research include exploiting spatial redundancies within a patch as is the case with SAR-BM3D [6] and also taking into account the slope-dependent reflectivity when computing similarities. The robustness of the filter could be increased by also setting the search window dimensions adaptively depending on the local scene heterogeneity, which could also be achieved by designing a weighting kernel with thresholding. Furthermore, the proposed filter only relies on the interferometric phase to classify scenes as heterogeneous, also taking the intensity into account might provide more accurate estimates, especially in urban areas.

REFERENCES

[1] J. S. Lee. "A simple speckle smoothing algorithm for synthetic aperture radar images". In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-13.1 (Jan. 1983), pp. 85–89. ISSN: 0018-9472. DOI: 10.1109/TSMC.1983.6313036.

[2] J. S. Lee et al. "Improved Sigma Filter for Speckle Filtering of SAR Imagery". In: *IEEE Transactions on Geoscience and Remote Sensing* 47.1 (Jan. 2009), pp. 202–213. ISSN: 0196-2892. DOI: 10.1109/TGRS.2008.2002881.

[3] J. S. Lee et al. "Polarimetric SAR Speckle Filtering and the Extended Sigma Filter". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.3 (Mar. 2015), pp. 1150–1160. ISSN: 0196-2892. DOI: 10.1109/TGRS.2014.2335114.

[4] A. Buades, B. Coll, and J.-M. Morel. "A non-local algorithm for image denoising". In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005*. Vol. 2. June 2005, pp. 60–65. DOI: 10.1109/CVPR.2005.38.

[5] C.A. Deledalle, L. Denis, and F. Tupin. "Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights". In: *IEEE Transactions on Image Processing* 18.12 (2009), pp. 2661–2672. DOI: 10.1109/TIP.2009.2029593.

[6] S. Parrilli et al. "A Nonlocal SAR Image Denoising Algorithm Based on LLMMSE Wavelet Shrinkage". In: *IEEE Transactions on Geoscience and Remote Sensing* 50.2 (Feb. 2012), pp. 606–616. ISSN: 0196-2892. DOI: 10.1109/TGRS.2011.2161586.

[7] G. Di Martino et al. "Scattering-Based Nonlocal Means SAR Despeckling". In: *IEEE Transactions on Geoscience and Remote Sensing* 54.6 (June 2016), pp. 3574–3588. ISSN: 0196-2892. DOI: 10.1109/TGRS.2016.2520309.

[8] C.-A. Deledalle, L. Denis, and F. Tupin. "NL-InSAR: Nonlocal Interferogram Estimation". In: *IEEE Transactions on Geoscience and Remote Sensing* 49.4 (Apr. 2011), pp. 1441–1452. ISSN: 0196-2892. DOI: 10.1109/TGRS.2010.2076376.
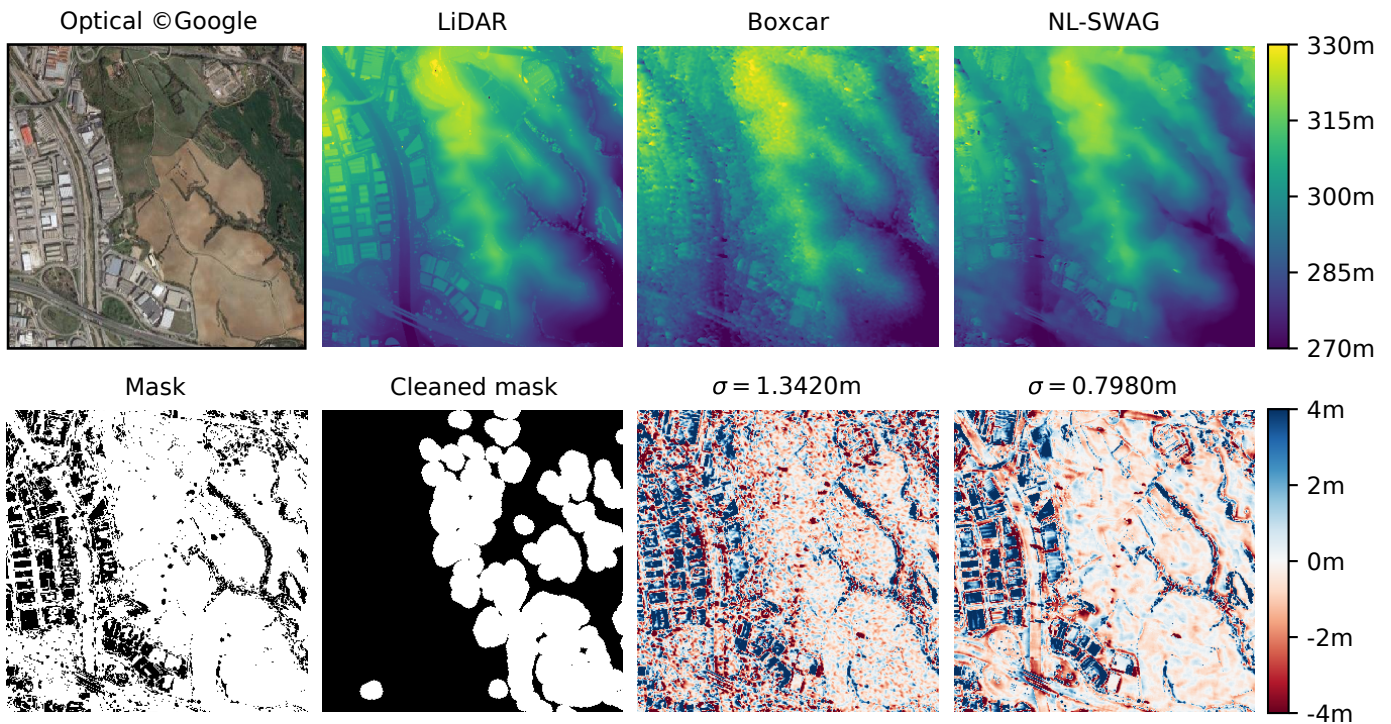
Figure 15. DEMs generated by NL-SWAG and a 5×5 Boxcar filter from a TanDEM-X interferogram are compared to a LiDAR DEM. The bottom row shows the height differences compared to the LiDAR DEM. For the masked-out area, standard deviations were computed for the height differences.

[9]   X. Lin et al. "Nonlocal SAR Interferometric Phase Filtering Through Higher Order Singular Value Decomposition". In: *IEEE Geoscience and Remote Sensing Letters* 12.4 (Apr. 2015), pp. 806–810. ISSN: 1545-598X. DOI: 10.1109/LGRS. 2014.2362952.

[10] J. Chen et al. "Nonlocal Filtering for Polarimetric SAR Data: A Pretest Approach". In: *IEEE Transactions on Geoscience and Remote Sensing* 49.5 (May 2011), pp. 1744–1754. ISSN: 0196-2892. DOI: 10.1109/TGRS.2010.2087763.

[11] C.-A. Deledalle et al. "NL-SAR: A Unified Nonlocal Framework for Resolution-Preserving (Pol)(In)SAR Denoising". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.4 (Apr. 2015), pp. 2021–2038. ISSN: 0196-2892. DOI: 10.1109/ TGRS.2014.2352555.

[12] F. Sica et al. "Nonlocal Adaptive Multilooking in SAR Multipass Differential Interferometry". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8.4 (Apr. 2015), pp. 1727–1742. ISSN: 1939-1404. DOI: 10.1109/ JSTARS.2015.2421554.

[13] O. D'Hondt et al. "Nonlocal Filtering Applied to 3-D Reconstruction of Tomographic SAR Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 56.1 (Jan. 2018), pp. 272–285. ISSN: 0196-2892. DOI: 10.1109/TGRS.2017.2746420.

[14] H. Breit et al. "Processing of bistatic TanDEM-X data". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*. July 2010, pp. 2640–2643. DOI: 10.1109/ IGARSS.2010.5653602.

[15] T. Fritz et al. "Interferometric processing of TanDEM-X data". In: *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. July 2011, pp. 2428–2431. DOI: 10. 1109/IGARSS.2011.6049701.

[16] Cristian Rossi et al. "TanDEM-X calibrated raw DEM generation". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 73 (2012), pp. 12–20. DOI: http://dx.doi.org/10.1016/ j.isprsjprs.2012.05.014.

[17] G. Krieger et al. "TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry". In: *IEEE Transactions on Geoscience and Remote Sensing* 45.11 (Nov. 2007), pp. 3317–3341. ISSN: 0196-2892. DOI: 10.1109/TGRS.2007.900693.

[18] M. Lachaise and T. Fritz. "Update of the Interferometric Processing Algorithms for the Tan-DEM-X High Resolution DEMs". In: *EUSAR 2016; Proc. of*. June 2016, pp. 1–4.

[19] Xiao Xiang Zhu et al. "Improving TanDEM-X DEMs by Nonlocal InSAR Filtering". In: *EUSAR 2014; Proc. of*. June 2014.

[20] Xiao Xiang Zhu et al. "Potential and Limits of Non-local InSAR Filtering for TanDEM-X High-resolution DEM Generation". Submitted.

[21] J. Hoffmann et al. *TanDEM-X Ground Segment DEM Products Specification Document*. Tech. rep. German Aerospace Center, 2016.

[22] G. Baier and X. X. Zhu. "GPU-based nonlocal filtering for large scale SAR processing". In: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. July 2016, pp. 7608–7611. DOI: 10.1109/IGARSS.2016.7730984.

[23] M. Lebrun et al. "Secrets of image denoising cuisine". In: *Acta Numerica* 21 (2012), pp. 475–576. DOI: 10 . 1017 / S0962492912000062.

[24] K. Dabov et al. "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering". In: *IEEE Transactions on Image Processing* 16.8 (Aug. 2007), pp. 2080–2095. ISSN: 1057-7149. DOI: 10.1109/TIP.2007.901238.

[25] J. Salmon, R. Willett, and E. Arias-Castro. "A two-stage denoising filter: The preprocessed Yaroslavsky filter". In: *2012 IEEE Statistical Signal Processing Workshop (SSP)*. Aug. 2012, pp. 464–467. DOI: 10.1109/SSP.2012.6319733.

[26] Danny Barash and Dorin Comaniciu. "A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering

and mean shift". In: *Image and Vision Computing* 22.1 (2004), pp. 73–81. ISSN: 0262-8856. DOI: http://dx.doi.org/10.1016/j.imavis.2003.08.005.

[27] Joachim Weickert. "A review of nonlinear diffusion filtering". In: *Scale-Space Theory in Computer Vision: First International Conference, Scale-Space'97 Utrecht, The Netherlands, July 2–4, 1997 Proceedings*. Ed. by Bart ter Haar Romeny et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 1–28. ISBN: 978-3-540-69196-9. DOI: 10.1007/3-540-63167-4_37.

[28] Holger Winnemöller, Sven C. Olsen, and Bruce Gooch. "Real-Time Video Abstraction". In: *ACM Transactions on Graphics (Proc. of SIGGRAPH'06)* 25.3 (July 2006), pp. 1221–1226. DOI: http://doi.acm.org/10.1145/1141911.1142018.

[29] Vincent Duval, Jean-François Aujol, and Yann Gousseau. "A bias-variance approach for the Non-Local Means". In: *SIAM Journal on Imaging Sciences* 4.2 (Jan. 2011), pp. 760–788. DOI: 10.1137/100790902.

[30] J. S. Lee et al. "A new technique for noise filtering of SAR interferometric phase images". In: *IEEE Transactions on Geoscience and Remote Sensing* 36.5 (Sept. 1998), pp. 1456–1465. ISSN: 0196-2892. DOI: 10.1109/36.718849.

[31] R. Bamler and P. Hartl. "Synthetic aperture radar interferometry". In: *Inverse Problems* 14 (1998), R1–54.

[32] A. M. Guarnieri and C. Prati. "SAR interferometry: a "Quick and dirty" coherence estimator for data browsing". In: *IEEE Transactions on Geoscience and Remote Sensing* 35.3 (May 1997), pp. 660–669. ISSN: 0196-2892. DOI: 10.1109/36.581984.

[33] Charles Kervrann and Jérôme Boulanger. "Local Adaptivity to Variable Smoothness for Exemplar-Based Image Regularization and Representation". In: *International Journal of Computer Vision* 79.1 (Aug. 2008), pp. 45–69. ISSN: 1573-1405. DOI: 10.1007/s11263-007-0096-2.

[34] Z. Suo, Z. Li, and Z. Bao. "A New Strategy to Estimate Local Fringe Frequencies for InSAR Phase Noise Reduction". In: *IEEE Geoscience and Remote Sensing Letters* 7.4 (Oct. 2010), pp. 771–775. ISSN: 1545-598X. DOI: 10.1109/LGRS.2010.2047935.

[35] V. Fedorov and C. Ballester. "Affine Non-Local Means Image Denoising". In: *IEEE Transactions on Image Processing* 26.5 (May 2017), pp. 2137–2148. ISSN: 1057-7149. DOI: 10.1109/TIP.2017.2681421.

[36] N. R. Goodman. "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)". In: *Ann. Math. Statist.* 34.1 (Mar. 1963), pp. 152–177. DOI: 10.1214/aoms/1177704250.

[37] Alain Fournier, Don Fussell, and Loren Carpenter. "Computer Rendering of Stochastic Models". In: *Commun. ACM* 25.6 (June 1982), pp. 371–384. ISSN: 0001-0782. DOI: 10.1145/358523.358553.

[38] C. Rossi and E. Erten. "Paddy-Rice Monitoring Using TanDEM-X". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.2 (Feb. 2015), pp. 900–910. ISSN: 0196-2892. DOI: 10.1109/TGRS.2014.2330377.

[39] M. Nolan and D. R. Fatland. "Penetration depth as a DInSAR observable and proxy for soil moisture". In: *IEEE Transactions on Geoscience and Remote Sensing* 41.3 (Mar. 2003), pp. 532–537. ISSN: 0196-2892. DOI: 10.1109/TGRS.2003.809931.

[40] Jong-Sen Lee. "Speckle analysis and smoothing of synthetic aperture radar images". In: *Computer Graphics and Image Processing* 17.1 (1981), pp. 24–32. ISSN: 0146-664X. DOI: http://dx.doi.org/10.1016/S0146-664X(81)80005-6.