Potential and Limits of Non-local Means InSAR Filtering for TanDEM-X High-resolution DEM Generation

Xiao Xiang Zhu[a,b], Gerald Baier[a], Marie Lachaise[a], Yilei Shi[c], Fathalrahman Adam[a], Richard Bamler[a,c]

[a] *German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Oberpfaffenhofen, 82234 Wessling, Germany*
[b] *Technical University of Munich, Signal Processing in Earth Observation (SiPEO), Arcisstrasse 21, 80333 Munich, Germany*
[c] *Technical University of Munich, Chair of Remote Sensing Technology, Arcisstrasse 21, 80333 Munich, Germany*

Abstract

The primary objective of the German TanDEM-X mission is the generation of a globally available, highly accurate and detailed digital elevation model (DEM), with the final product having 12 m posting, 2 m relative and 10 m absolute vertical accuracy. The first version of this global DEM has been finalized by the German Aerospace Center (DLR), in September 2016. Our experience with the experimental application of non-local means filters to TanDEM-X data suggests that TanDEM-X has the potential of producing DEMs of even higher resolution and accuracy. The goal of this investigation is to explore the possibility of employing non-local InSAR filters to achieve an effective resolution of 6 m, with an equivalent posting, and a relative height error below 0.8 m, i.e. an increase of quality by a factor of 2×2 in resolution and a factor of $2\,\mathrm{m}/0.8\,\mathrm{m} = 2.5$ in height accuracy — all in all one order of magnitude.

Keywords: TanDEM-X, non-local filtering, interferometric SAR, phase noise, digital elevation models

1. Introduction

The primary objective of the German TanDEM-X mission Krieger et al. (2007) is the generation of a globally available, highly accurate and detailed

Table 1: Resolution and accuracy requirements of the standard global TanDEM-X DEM and the HDEM Hoffmann et al. (2016).

	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 $\leq$ 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)

digital elevation model (DEM), with the final product having 12 m posting, 2 m relative and 10 m absolute vertical accuracy Hoffmann et al. (2016). The first version of this global DEM has been finalized by the authors' institution, the German Aerospace Center (DLR), in September 2016. Due to sophisticated geometric calibration this global DEM exceeds the required absolute accuracy by an order of magnitude. Our first preliminary experiments with TanDEM-X data and processing with the earliest non-local InSAR filter Deledalle et al. (2011) suggest that TanDEM-X has the potential of producing DEMs of much higher resolution and accuracy (see Table 1). The goal of this investigation is to explore the possibility of achieving a resolution of 6 m and a relative error of 0.8 m, i.e. an increase of quality by a factor of 2×2 in resolution and a factor of $2\,\mathrm{m}/0.8\,\mathrm{m} = 2.5$ in accuracy — all in all one order of magnitude. The TanDEM-X mission scenario accounts for such requirements by acquiring so-called HDEM data with larger interferometric baselines for selected areas of the world. In this paper we investigate non-local (NL) InSAR filters as an alternative approach for increasing resolution while at the same time even better suppressing phase noise, compared to the default Boxcar filter.

In the TanDEM-X production workflow we distinguish between "Raw DEMs", i.e. DEMs generated from every individual bistatic interferometric TanDEM-X data pair, and the final DEM product for which Raw DEMs are calibrated, mosaicked and fused. The requirements on posting and accuracies of Table 1 refer to the final DEM. In this paper we work on individual scenes, i.e. on the Raw DEM level. Hence, any error assessment cannot directly refer to the requirements cited in Table 1, but is always *relative* to the standard TanDEM-X Raw DEMs.

Raw DEMs have an extent of about 30 km × 50 km. The whole Earth landmass has been mapped at least twice so that about 470,000 interferometric scenes have been acquired adding up to more than three petabytes of data. SAR and InSAR processing of all TanDEM-X data (i.e. bistatic SAR focusing, interferogram generation, phase unwrapping and geocoding) is performed by DLR's Integrated TanDEM-X Processor (ITP) Breit et al. (2010); Fritz et al. (2011); Lachaise et al. (2014). The large data volume of TanDEM-X necessitates employing computationally tractable processing algorithms in the ground segment. This way, the ITP is able to process 1,300 scenes or Raw DEMs per day when simple single-baseline phase unwrapping is used or more than 400 when the dual-baseline phase unwrapping correction Lachaise et al. (2014) is applied. In particular, a conventional 5×5 or 7×5 (depending on range resolution) boxcar filter denoises the interferometric phase. Whereas such a filter fulfills the TanDEM-X resolution and accuracy requirements (Table 1), by its very nature it also degrades the effective resolution of the stripmap acquisition from 3 m to about 12 m.

In this paper we investigate to which extent NL filters can improve TanDEM-X DEM quality. The NL filtering principle, first introduced by Buades et al. has become the foundation for most state-of-the-art denoising algorithms due its strong noise reduction and detail preservation. We call the resulting Raw DEM a "Prime Raw DEM" in allusion to its enhanced quality. An additional benefit of NL filters is the large number of pixels they include in their estimate, leading to a less biased and noisy coherence estimates Touzi et al. (1999).

A more accurate coherence estimate can then possibly aid subsequent phase unwrapping. The final TanDEM-X DEM mosaicked from several Raw Prime DEMs will be referred to as "Prime DEM". It should meet — or be close to — the HDEM requirements of Table 1 and will be the focus of another paper of the authors. This NL processing concept can also be applied to the HDEM acquisitions to further improve those data. As a last note, we found that NL filters to be computationally tractable even for large area processing, which is of relevance given the large data volume of the TanDEM-X mission.

In contrast to local neighborhood filters, NL filters use comparatively large areas for denoising a single pixel by a weighted average. The weights themselves are a function of a similarity measure, which helps to avoid smoothing over edges or other features. These two characteristics combined result in the aforementioned remarkable performance.

We use two versions of the NL filter, denoted "NL-InSAR", as described in Deledalle et al. (2011), and "NL-SAR"Deledalle et al. (2015). We show that NL-InSAR meets the requirements, but produces terrace-like artifacts on sloped terrain. NL-SAR avoids these artifacts but will be shown to be inferior to NL-InSAR in terms of resolution. We analyze these nonlocal filters using simulations, akin to the experiments in Deledalle et al. (2011), and by comparing the NL-filtered TanDEM-X Prime RAW DEMs at 6 m pixel spacing with the standard TanDEM-X 12 m Raw DEM. As sampling the standard TanDEM-X 12 m Raw DEM at 6 m provides no additional information, due to the boxcar filter's footprint, such a comparison is not included. The results suggest that the improved Raw DEMs fulfill or are at least close to the HDEM specifications — aside from the terrace effect for which we suggest an initial solution.

The remainder of the paper is organized as follows. In chapter II we review the concept of NL InSAR filtering, taking NL-InSAR Deledalle et al. (2011) as an example, and show some of the trade-offs NL filters have to make between noise reduction and detail preservation. Chapter III is devoted to the appeal of NL filters for DEM generation: the high achievable effective number of looks and the less biased coherence estimate. In chapter IV we showcase the difficulties

that NL filters face when processing areas with pronounced fringes. Chapter V analyzes two existing NL filters in terms of bias and noise reduction for their applicability to DEM generation on synthetic data. DEM examples from single TanDEM-X interferograms are presented in chapter VI.

The assessment of the real resolution of InSAR DEMs is a intricate topic. Due to the particular SAR imaging geometry there is no such thing as a DEM point response function. Overlay leads to a superposition of multiple scatterers and geocoding from the SAR coordinates range and azimuth to a DEM's longitude and latitude mixes the response of several pixels. Instead we use simulated step responses to assess the resolution of the filtered phase functions.

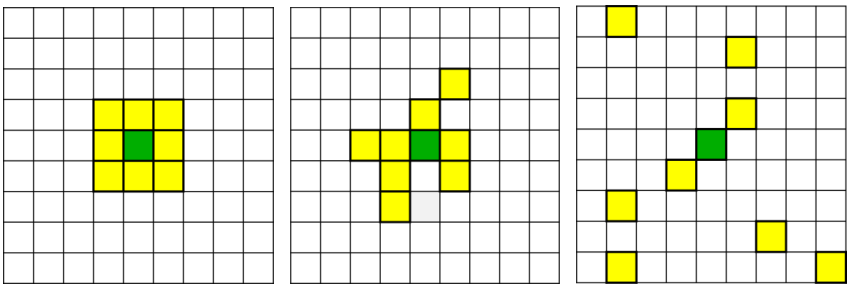
2. Non-local InSAR Filtering

2.1. The Non-local Filtering Concept

The NL-means concept proposed in Buades et al. (2005b, 2010) takes advantage of the inherent redundancy in natural images. In other words, natural images often have repetitive features such as edges, lines or points, which can jointly be used for denoising.

Figure 1 contrasts the non-local filtering concept to convolutional (a) and adaptive filters (b), which both use a connected neighborhood of pixels for filtering. NL filters redefine this spatial neighborhood of a target pixel t (green in Figure 1) to a neighborhood in the patch space, where search pixels s, whose surrounding patches are more *similar* to the patch around t play a more significant role in the denoising process, regardless of their actual spatial distance. Close pixels in this generalized patch-based neighborhood, which can have arbitrary spatial positions (c), are then used to estimate the value at t.

Given a noisy image $\mathbf{v}$ on a discrete grid $\mathbf{S}$: $\mathbf{v} = \{v_s | s \in \mathbf{S}\}$, the estimated value $\hat{v}_{t,\mathrm{NL}}$ of a target pixel t is computed as a weighted average of all the pixels in the image



(a) Local (b) Adaptive (c) Non-local

Figure 1: Local vs. non-local concepts. Green: target pixel to be estimated, yellow: pixels considered to be similar to the target pixel. (a) rectangular window (local); (b) adaptive — but still local — window; (c) non-local set of pixels.

$$\hat{v}_{\mathrm{t,NL}} = \sum_{s \in \mathbf{S}} w(t,s) v_s \ , \tag{1}$$

where the weight $w(t,s)$ depends on the distance in the patch space of the pixels t and s and has to satisfy $0 \leq w(t,s) \leq 1$ and $\sum_{s \in \mathbf{S}} w(t,s) = 1$. In practice, due to computational constraints, the search for similar pixels is limited to a sufficiently large search window. In this case the symbol $\mathbf{S}$ in Equation (1) denotes the search window instead of the whole image. The distance in the patch space, i.e. the measure of patch similarity, is a function of the imaging process's noise characteristics. For example, if the underlying process is Gaussian and the weights are the normalized inverses of the variances of the pixels, Equation (1) is in fact the Maximum Likelihood Estimate. The problem though is how to estimate these weights. The original NL means algorithm Buades et al. (2005a) used the Euclidean distance between patches, weighted by a Gaussian kernel, to compute the similarity; this approach is not optimal for non-additive or non-Gaussian noise, as is the case for InSAR data. In Deledalle et al. (2011) Deledalle et al. introduced a method adapted to InSAR statistics, leading to the approach described in the following.

The algorithm is akin to an Expectation Maximization approach, resulting in estimates of reflectivity, coherence and phase for each pixel, which are iteratively

refined. In the following we distinguish between the *amplitude* of a pixel, i.e. the measured quantity including speckle, and its *reflectivity*, i.e. its expectation value.

2.2. Non-local InSAR Filtering (following Deledalle et al. (2011))

For determining a weight the two patches around the two concerned pixels are checked for similarity. The patches corresponding to the target pixel t and the search pixel s are denoted by O_{Δ_t} and O_{Δ_s}, respectively. The purpose of the similarity check is to find the likelihood that the examined patch and the target patch are both noisy versions of the same noiseless patch, i.e. they are two realizations of the same stochastic process.

Let $O_{t,k} = (A_{t,k}, A'_{t,k}, \phi_{t,k})$ and $O_{s,k} = (A_{s,k}, A'_{s,k}, \phi_{s,k})$ be the observations of the k^{th} pixel in O_{Δ_t} and O_{Δ_s}, respectively, where

- A is the amplitude of the first (master) image,
- A' is the amplitude of the second (slave) image, and
- ϕ is the interferometric phase.

$\Theta_{t,k} = (R_{t,k}, \beta_{t,k}, \gamma_{t,k})$ denotes the set of true values (expectations) of the three parameters at the k^{th} pixel in the patch O_{Δ_t} surrounding pixel t: the reflectivity $R_{t,k}$, the interferometric phase $\beta_{t,k}$ and the coherence $\gamma_{t,k}$. The expectations of the amplitudes of the master and the slave images are assumed to be identical.

The similarity is expressed as the conditional likelihood of observing $O_{t,k}$ and $O_{s,k}$ given that the true parameters $\Theta_{t,k}$ and $\Theta_{s,k}$ of the target and search pixel are identical. Assuming circular Gaussian scattering this leads to Deledalle et al. (2011):

$$p(O_{t,k}, O_{s,k} | \Theta_{t,k} = \Theta_{s,k}) = \sqrt{\frac{C}{B}}^3 \left(\frac{A+B}{A} \sqrt{\frac{B}{A-B}} - \arcsin \sqrt{\frac{B}{A}} \right), \qquad (2)$$

where

$$
\begin{aligned}
A &= \left(A_{t,k}^2 + A_{t,k}'^2 + A_{s,k}^2 + A_{s,k}'^2\right)^2 \,, \\
B &= 4\left(A_{t,k}^2 A_{t,k}'^2 + A_{s,k}^2 A_{s,k}'^2 + 2A_{t,k}A_{t,k}'A_{s,k}A_{s,k}'\cos\left(\phi_{t,k} - \phi_{s,k}\right)\right) \text{ and} \\
C &= A_{t,k}A_{t,k}'A_{s,k}A_{s,k}' \,.
\end{aligned} \tag{3}
$$

Instead of just using the similarity likelihood, Deledalle et al. (2011) followed a Bayesian approach and combined the similarity likelihood with a prior term to compute the *a posteriori* probability that two pixels are equal given a certain observation

$$
p(\Theta_{t,k} = \Theta_{s,k}|O) \propto p(O_{t,k}, O_{s,k}|\Theta_{t,k} = \Theta_{s,k}) \times p(\Theta_{t,k} = \Theta_{s,k}) \,, \tag{4}
$$

where the prior $p(\Theta_{t,k} = \Theta_{s,k})$ is iteratively estimated and is given by

$$
p(\Theta_{t,k} = \Theta_{s,k}) = \exp\left[-\frac{1}{T} SD_{KL}\left(\hat{\Theta}_{t,k}^{i-1}, \hat{\Theta}_{s,k}^{i-1}\right)\right] \,, \tag{5}
$$

with T being a smoothing parameter and $SD_{KL}\left(\hat{\Theta}_{t,k}^{i-1}, \hat{\Theta}_{s,k}^{i-1}\right)$ being the symmetrical Kullback-Leibler divergence, which depends on the estimate $\hat{\Theta}^{i-1}$ of the previous iteration.

For two zero-mean complex circular Gaussian distributions it is given by

$$
\begin{aligned}
SD_{KL}\left(\hat{\Theta}_{t,k}^{i-1}, \hat{\Theta}_{s,k}^{i-1}\right) = \frac{4}{\pi}\Bigg[&\frac{\hat{R}_{t,k}}{\hat{R}_{s,k}}\left(\frac{1 - \hat{\gamma}_{t,k}\hat{\gamma}_{s,k}\cos(\hat{\beta}_{t,k} - \hat{\beta}_{s,k})}{1 - \hat{\gamma}_{s,k}^2}\right) && (6) \\
&+ \frac{\hat{R}_{s,k}}{\hat{R}_{t,k}}\left(\frac{1 - \hat{\gamma}_{s,k}\hat{\gamma}_{t,k}\cos(\hat{\beta}_{s,k} - \hat{\beta}_{t,k})}{1 - \hat{\gamma}_{t,k}^2}\right) - 2\Bigg] \,. && (7)
\end{aligned}
$$

For the sake of brevity we omit the iteration index i for all quantities.

The weight of the patch centered on t is then given by the product over all pixel *a posteriori* probabilities, which for numerical stability reasons can be written the sum over the logarithms

$$
w(t,s) = \exp\sum_k \left[\log\frac{1}{h} p(O_{s,k}, O_{t,k}|\Theta_{t,k} = \Theta_{s,k}) - \frac{1}{T} SD_{KL}\left(\hat{\Theta}_{t,k}^{i-1}, \hat{\Theta}_{s,k}^{i-1}\right)\right] \,, \tag{8}
$$

with h being a second smoothing parameter.

With every iteration the weights are refined by the Kullback-Leibler divergence. Figure 2 shows the phase estimate after $n \in \{1, 2, 3, 5\}$ iterations for a sudden phase jump and a nonlinear smooth phase transition, while the reflectivity is constant and the coherence γ was set to 0.8. For both scenarios the initial estimate is an oversmoothed version of the original phase. With increasing iteration count the curvature of the estimate increases due to the more discriminant weights. For a distinct phase jump this is the desired result, whereas for the nonlinear smooth case more iterations tend to over-amplify curvature (see zoom-in in Figure 2).

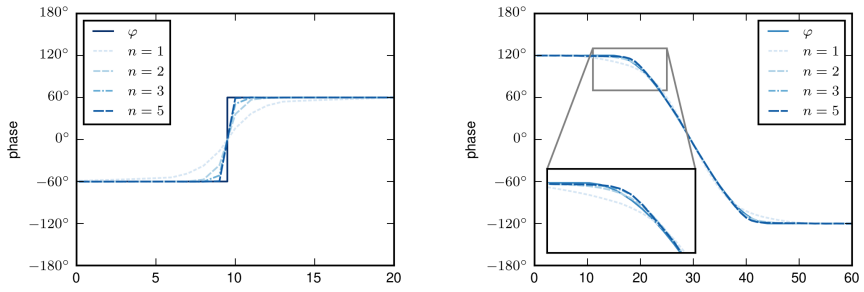


Figure 2: NL-InSAR phase estimates after $n = 1, 2, 3$, and 5 iterations for a jump in phase (left) and a nonlinear smooth transition (right). Parameters: $\gamma = 0.8, h = 12, T = 6$. With increasing iteration count transitions change from being oversmoothed to becoming more and more abrupt.

The filtering parameters h and t are crucial as they define the trade-off between bias and variance. Figure 3 illustrates for different values of h the expected value of the phase estimate and its variance for a jump in phase with constant reflectivity and coherence ($\gamma = 0.8$) after the *first iteration* of NL-InSAR. For small values of h the phase estimates follow more closely the true phase. The price to pay is a weaker noise reduction, as shown by the standard deviation plot. Along the edge the noise reduction is less effective since fewer similar patches are available. Figure 4 further shows the impact h has on the phase standard deviation, which directly translates into height errors in the generated

DEM. We chose $h = 12$ with a patch size of 7×7 for our implementation, as in the original paper, and will show later that the resolution of this filter is sufficient (after 5 — 6 iterations). The reasoning for selecting T is similar and will therefore not be covered independently.

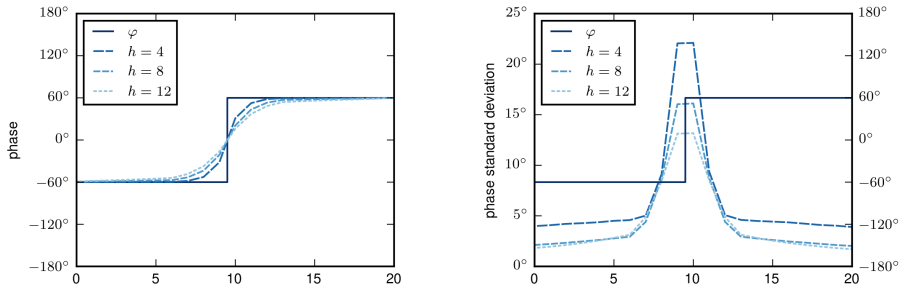


Figure 3: Bias-variance trade-off after the first iteration of NL-InSAR without any prior knowledge. The expected value (left) is closer to the true phase for smaller, whereas the standard deviation (right) increases.

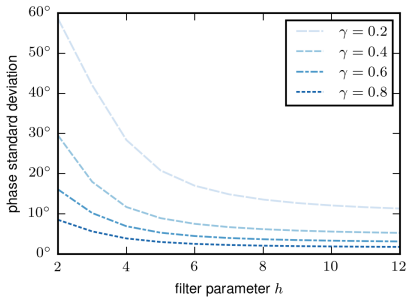


Figure 4: Phase standard deviation for a constant phase depending on h and different values of the coherence after the first iteration of NL-InSAR. For the remainder of the paper we use $h = 12$.

2.3. Equivalent Number of Looks of Non-local InSAR Filters

The equivalent number of looks achievable by a non-local InSAR filter — like every weighted average — is bounded by

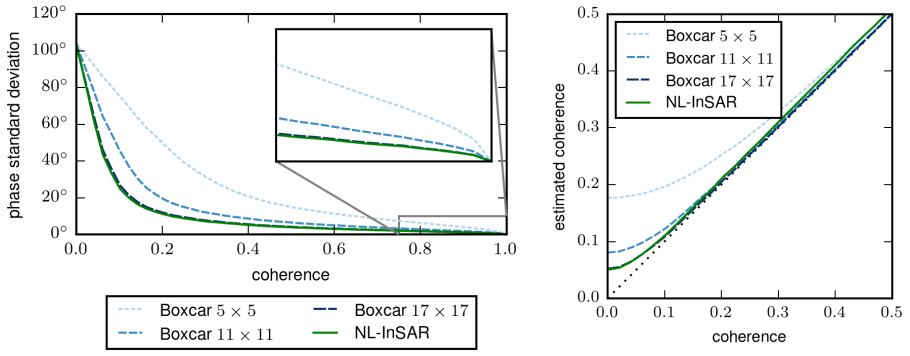
$$L(t) \leq \frac{\left(\sum\limits_{s\in\mathbf{S}} w(t,s)\right)^2}{\sum\limits_{s\in\mathbf{S}} w(t,s)^2} < |\mathbf{S}| \,, \tag{9}$$

where $|\mathbf{S}|$ stands for the number of pixels in the search window. The limit would only be reached, if all the search pixels were realizations of the same process. The right hand side limit $|\mathbf{S}|$ cannot be achieved at all, because it would require that all weights are identical $1/|\mathbf{S}|$. Since the weights are estimated from noisy data, they are noisy themselves and never equal.

In this remainder of this section we analyze the noise reduction power of the original NL-InSAR filter in terms of effective number of looks and its impact on the coherence estimate. We fixed the patch size to 7×7, the search window to 21×21 pixels, set $h = 12$ and $T = 6$ and used five iterations.

As mentioned in the introduction, our goal is an improvement of noise reduction by a factor of at least 2.5 compared to the standard processing by a 5×5 boxcar filter (as mostly used in standard TanDEM-X processing). Hence, our target is to achieve an effective number of looks of $L = 5 \times 5 \times (2.5)^2 \approx 156$, approximately equivalent to a boxcar filter of $13 \times 13 = 169$.

To gain an understanding of the level of improvement that can be achieved by NL-InSAR, simulations with constant interferometric phase but different coherence levels γ were conducted. This is the best case scenario for non-local filters as they can take full advantage of their large search windows and can serve as an upper bound on what level of improvement can be achieved. Figure 5(a) contrasts the noise standard deviation (STD) of the NL-InSAR filter output as a function of coherence γ to the results of boxcar filters of different sizes. The NL curve follows approximately the one of a 17×17 boxcar filter. The ratio of phase STDs of a 5×5 boxcar and the NL filter at different coherence levels is constant and equals approximately $17/5 = 3.4$, i.e. we exceed the requirement of noise reduction of 2.5 as mentioned in the Introduction chapter.



(a) (b)

Figure 5: Phase standard deviation (a) and coherence estimates (b) as a function of coherence for different number of looks and the original NL-InSAR filter. NL-InSAR with our parameter settings is comparable to a 17×17 Boxcar filter.

Coherence estimates obtained with small windows suffer from an inherent bias Touzi et al. (1999), which increases with decreasing coherence. Such biased coherence estimates pose a problem to the Minimum Cost Flow phase unwrapping algorithm, as used in the ITP, which relies on coherence-based weights for guiding the branch-cuts in low coherence areas. Figure 5(b) compares the coherence estimates of the NL-InSAR filter with boxcar estimates and shows the advantage of using a higher number of pixels in the estimate. The simulation confirms that the NL-InSAR filter with the aforementioned parameter setting achieves an effective number of looks of about $17 \times 17 = 289$.

Figure 6 compares the coherence estimate of a 5×5 boxcar filter to NL-InSAR and serves as real world example for the reduced bias of NL-InSAR in low coherent areas, such as water bodies and forested areas.

3. Shortcomings of the NL-InSAR filter when processing areas with Pronounced fringes

The high noise reduction capability of non-local InSAR filter derived in the preceding chapter was also substantiated by our initial experiments Zhu et al.

Figure 6: Coherence estimates for the test site St Lawrence: Optical image © Google (left), coherence estimate of a 5×5 boxcar filter (middle) and estimate by NL-InSAR (right). The grayscale from black to white indicates a coherence value of 0 to 1.

(2014a,b).

The results of one of them is shown in Figure 7, which shows shaded reliefs of the TanDEM-X raw DEM (middle) and the improved NL TanDEM-X raw DEM (right) for a rural-agricultural site near the city of Jülich, Germany. The NL TanDEM-X DEM shows a significant higher number of details and remarkably less noise as evidenced in the flat areas.

This preliminary but promising result suggests that one might apply such a filter straight away. However, with regard to global DEM generation, it is important that results of the same quality can be achieved for all terrain types. Through extensive simulations and visual inspection of various test sites with different height profiles we have found several shortcomings of the NL-InSAR filter for global DEM generation. Figure 8 shows as an example a zoom-in of three TanDEM-X Raw DEMs, one produced from data filtered by the standard boxcar kernel and the other two from the data filtered by two non-local InSAR filters.

The higher resolution and lower noise level of the latter are obvious. Yet there are two effects of the NL-InSAR filter applied to interferometric data pairs in the presence of pronounced fringes, a minor obvious one and less intuitive, yet annoying, one. The obvious problem is that in case of a phase ramp the

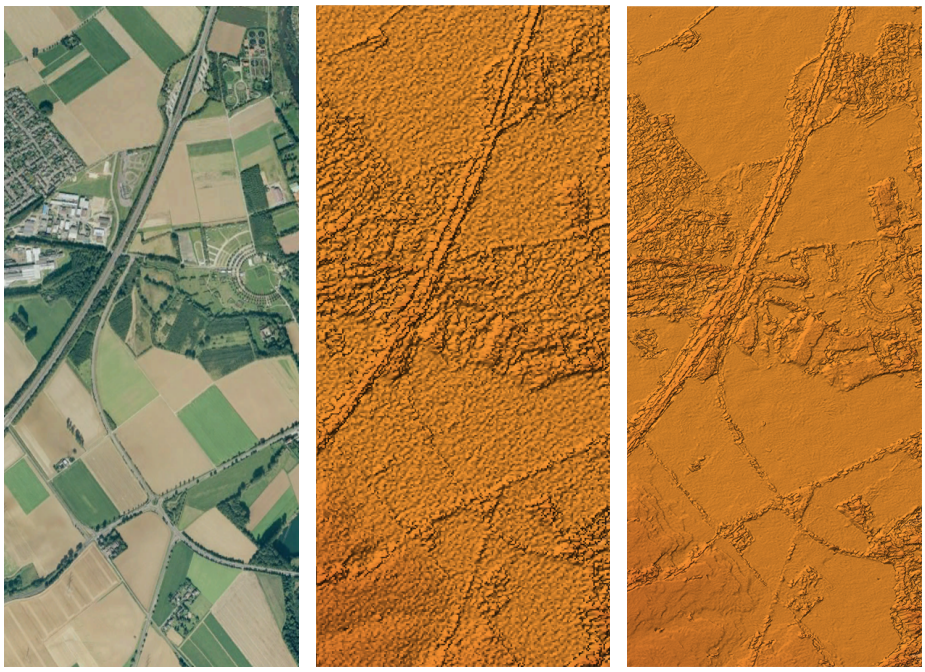
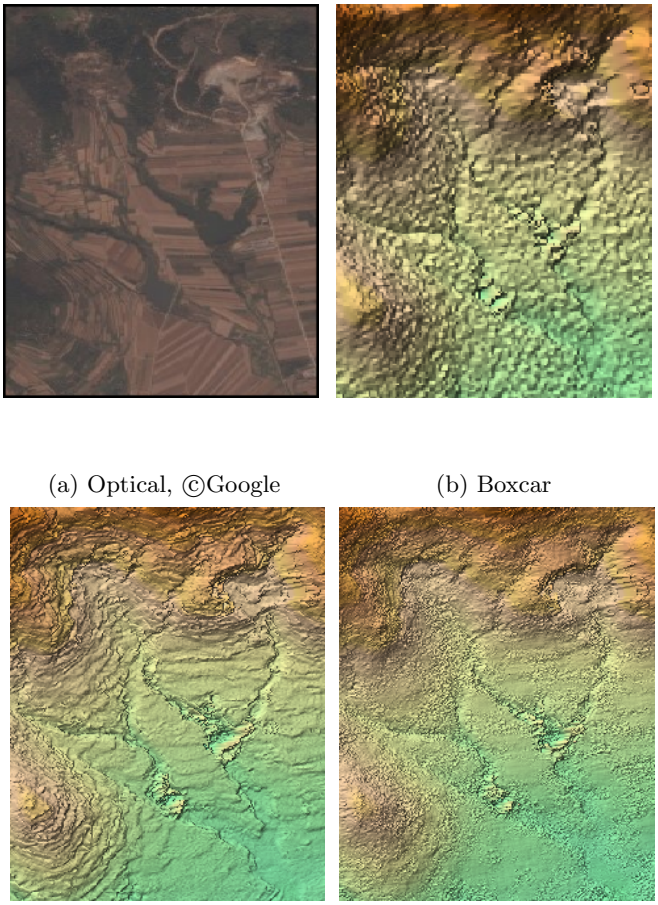


Figure 7: Mixed rural-agricultural test site: Optical image © Google (left), TanDEM-X raw DEM (middle) and improved non-local TanDEM-X raw DEM (right).

similarity of patches drops quickly when the search patch lies "above" or "below" the target patch on the ramp. This is a consequence of the particular similarity measure which renders two patches different if they have a mutual constant phase offset, even if they are otherwise identical. The weight map narrows in areas of high phase gradient (see fig. 9, right, pixels 25 to 35) leading to a reduced number of effective looks. Note, however, that convolutional filters suffer from a similar effect; a phase ramp across a convolution kernel reduces the accuracy of the estimate, following a sinc-function in the case of a boxcar filter; in the extreme case, when a full 2π phase cycle extends across the averaging window, the filter output is only noise.

A second unwanted effect is caused by the large search window of NL filters compared to the kernels of traditional filters. Large windows come with a larger



(a) Optical, ©Google (b) Boxcar

(c) NL-InSAR (d) NL-SAR

Figure 8: Zoom-in of a standard TanDEM-X Raw DEM and DEMs produced from NL-InSAR and NL-SAR filtered data. Note (i) the strong noise reduction, (ii) the higher resolution (iii) the staircase effect of the NL-InSAR DEM and (iiii) an increase in noise for NL-SAR along edges between homogeneous areas.

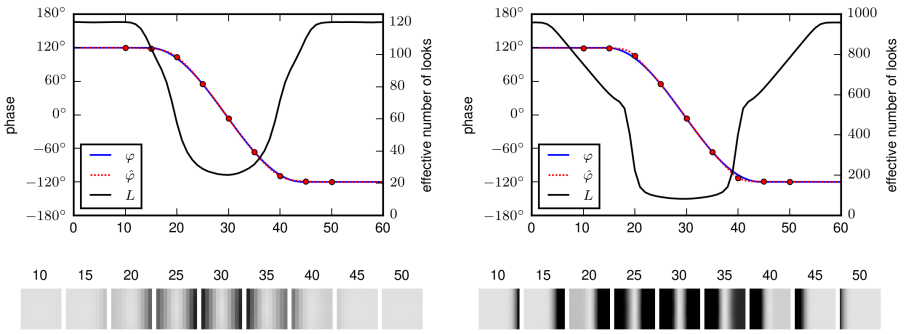


Figure 9: Monte-Carlo simulations of NL-InSAR filter results of a nonlinear transition between a flat phase area (pixels 0 — 15) and a linear phase ramp. Search window sizes are 11×11 (left) and 31×31 (right). The bottom row shows the weight maps for different pixels (pixel numbers 10 — 50 shown above weight maps), for convenient display rescaled to the same size. Two effects can be observed: (i) In the flat phase area the weight maps are quite homogeneous leading to a high effective number of looks L, while on the slope the weight map shrinks according to the phase gradient. (ii) For large search windows the weight maps become asymmetrically skewed to the smaller gradient in the nonlinear transition zone (pixels 15 — 25). This leads to a tendency to weigh the flat (or low slope) areas higher and down weigh the steeper areas. The effect is a kind of overshoot of the filtered phase — part of the reason for the staircase effect.

propensity to bias the estimate. This is especially evident for nonlinear changes of the phase (see Figure 9 which compares the results of two Monte-Carlo simulations for search window sizes of 11×11 and 31×31 and a nonlinearly changing phase).

From Figure 9, right, it becomes obvious that the NL filter favors low gradient phase functions, because the weight map gets asymmetrically skewed toward the area of the lower gradient, exaggerating nonlinear phase changes. Convolutional filters of a comparable spatial extent of the kernel show a similar behavior, but in the NL-InSAR filter this effect is amplified, as at every iteration the effect increases producing a significantly biased final phase estimate. In the Raw DEM generated from these NL filtered data this leads to terrace-like artifacts as seen in Figure 8. In image processing this artifact of iterative signal-dependent filters

is also known as "staircasing", which is often employed to create cartoon-like images from real photographs by repeatedly applying a bilateral filter. NL-SAR is unaffected by these filtering artifacts but suffers from an increased variance at the border between two different homogeneous areas as a result of its bias reduction step. As we will show in later experiments it also tends to oversmooth the resulting phase and can thus preserve less of the original resolution.

Since the terrace-effect only shows up in sloped terrain, a remedy would be to first demodulate the interferogram by a low-pass version of the phase. This low-pass phase can either be tapped off the filter after the first iteration or could be estimated in a separate pre-processing step. We will treat this problem in a follow-up paper. Here we focus on noise reduction and resolution. The terrace-effect, though, is the reason why we include the alternative NL-SAR filter from Deledalle et al. (2015) in our investigation. This filter avoids the terrace-effect, but has other disadvantages, as will be shown in the next chapter.

4. Analysis of Non-local Filters

We analyze two existing non-local filters, namely NL-InSAR Deledalle et al. (2011) and NL-SAR Deledalle et al. (2015), for their suitability to produce highly accurate DEMs and compare them with a conventional 5×5 boxcar filter. As mentioned in the introduction we use simulated phase step responses to assess the resolution of the filters. Since a phase discontinuity often comes along with some image features in reflectivity and coherence we simulated several scenarios to gauge the influence of reflectivity and coherence changes on the phase estimate. Note that in Figure 10 to Figure 14 the true values are delineated by solid lines and the estimation error is indicated as shaded areas, representing the 95%-quantile of the estimates.

4.1. constant reflectivity and coherence (Figure 10)

The boxcar filter shows the expected blur by the extent of its averaging window. NL-InSAR gives a better resolution of about two sample intervals.

NL-SAR blurs the edge even more than the boxcar filter. Also visible is the superior noise reduction capability of NL-InSAR as well as its better coherence estimate.

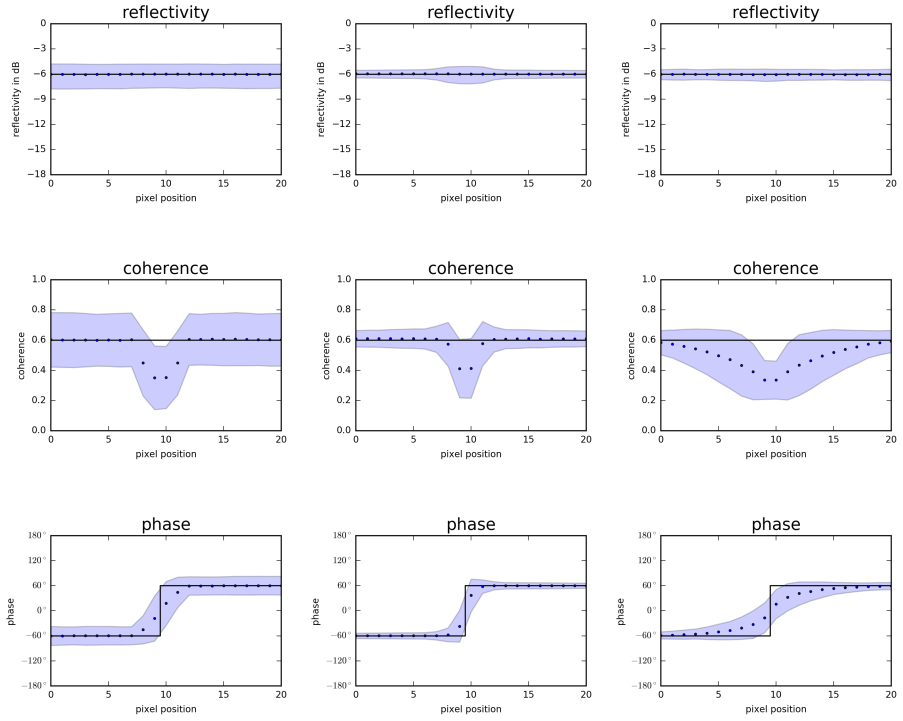


Figure 10: Scenario with constant reflectivity and coherence. From left to right the results produced by a 5×5 Boxcar filter, NL-SAR and NL-InSAR. The true values are delineated by the solid lines, the areas in light blue show 95% percentile of the estimates. The improved noise reduction that non-local filters provide is evident. However the result of NL-SAR shows unacceptable smoothing of the edge.

4.2. constant reflectivity, step in coherence (Figure 11)

The change in coherence skews the phase transition towards the higher coherence values for all filters, as these pixels add up more coherently when computing the weighted mean. For NL-InSAR the width of the transition is shortened compared to the scenario in Figure 10 whereas it remains largely the same for the

other filter and only changes its position.

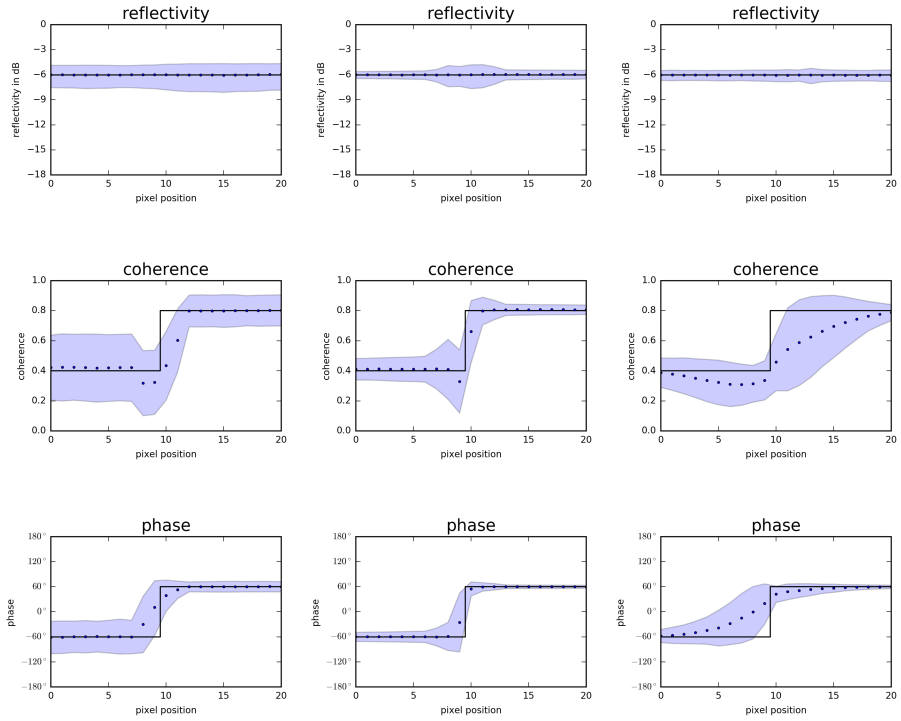


Figure 11: Constant reflectivity but step in coherence. The change in coherence skews the phase transition towards the higher coherence areas for all filters, as these pixels add up more coherently when averaging. For NL-InSAR the width of the transition is shortened compared to the scenario in Figure 10, whereas it remains largely the same for the other filter and only changes its position.

4.3. step in both reflectivity and coherence (Figure 12)

A change in reflectivity additionally refines the weight maps of the non-local filters along the edge, leading to sharper transitions. The boxcar filter shows no improvement, due to its indiscriminate selection of pixels, and its performance actually worsens compared to Figure 11 as pixels with high coherence have also high reflectivity further biasing the averaging in favor of pixels on the right side of the edge.

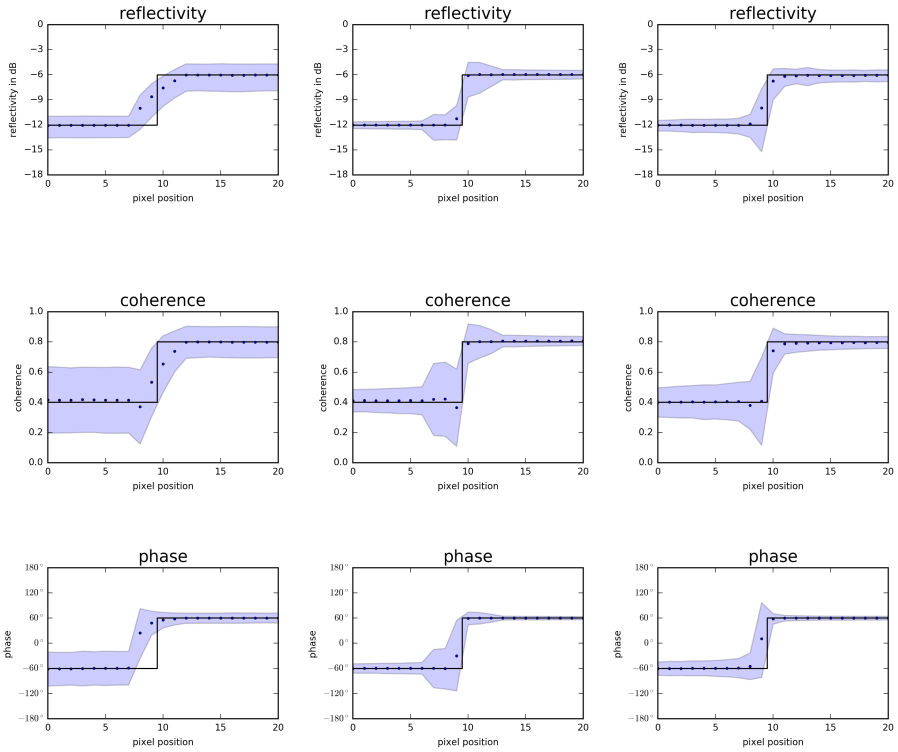


Figure 12: Reflectivity, coherence, and phase (from top to bottom), all with a concurrent jump, filtered by boxcar (left), NL-SAR (middle) and NL-InSAR (right). The change in reflectivity additionally refines the weight maps of the non-local filters along the edge, leading to sharper transitions. The boxcar filter shows no improvement, due to its indiscriminate selection of pixels, and its performance actually worsens compared to Figure 11 as pixels with high coherence have also high reflectivity further biasing the averaging in favor of pixels on the right side of the edge.

4.4. step in reflectivity and coherence but in opposite direction (Figure 13)

As both reflectivity and coherence have not only an influence on the weight map, but also in the weighted means of non-local filters their influence can counterbalance given appropriate profiles. If they change in opposite ways as in Figure 12 pixels on the left of the edge sum up more coherently, biasing the average. This is offset by pixels on the right side due to their larger reflectivity in the weighted means, leading to a sharp and symmetric transition.

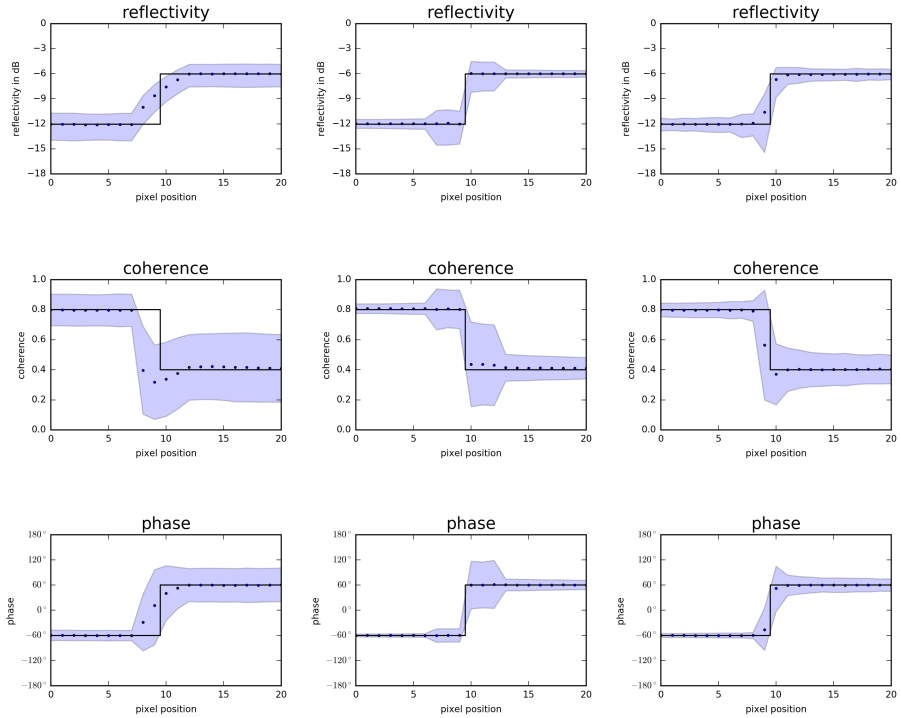


Figure 13: For a jump in reflectivity and coherence, but in opposite directions, their influences counterbalance leading to a sharp and symmetric transition for the non-local filters. In detail the high reflectivity on one side of the edge negates the effect of the high coherence on the other side when computing the weighted means, so that this scenario does not exhibit the skewed transition as in Figure 12

4.5. Reflectivity: rect-function, coherence: rect-function (Figure 14)

Compared to the somewhat artificial examples in the previous figures, Figure 14 shows a scenario that resembles more realistic reflectivity, coherence and phase profiles. We assume that lay-over areas of excibih higher reflectivity, due to multiple reflections, and lower coherence. To investiage the resolution and distortion at edges the phase profile is kept as a step function, even though in a real overlay scenario it depends on the reflectivity of the various reflections. Evidently both the boxcar filter and NL-SAR are incapable of retaining the resolution of the phase profile, whereas NL-InSAR produces a highly accurate and unbiased estimate.

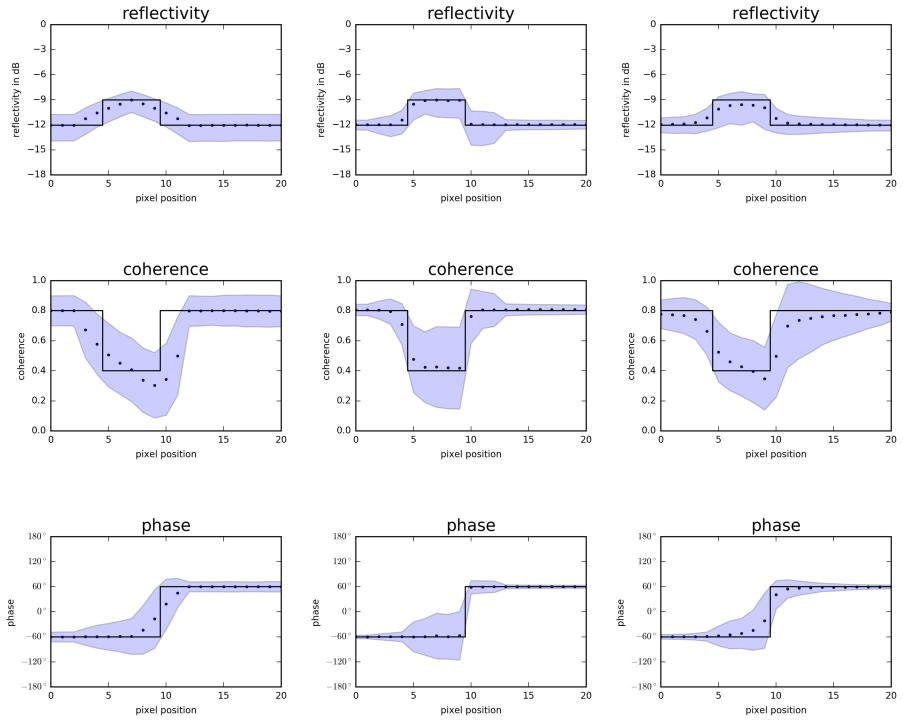


Figure 14: Scenario which resembles a lay-over area with increased reflectivity and lower coherence. The phase is conveniently chosen to be a step function to investigate resolution and distortion at edges.

From the previous simulations it seems as if NL-InSAR fulfilled all the needs for generating a high resolution DEM. In particular, the filter results of a phase step suggest that NL-InSAR maintains the inherent resolution, i.e. the phase changes within one sample interval in most of the simulation scenarios. Since the original TanDEM-X resolution is in the order of 3m, the 6m resolution goal can be easily achieved. Yet when applied to real data we observed terrace-like artifacts in the generated DEM as in Figure 8. To show that these are indeed filtering artifacts and not features of the terrain we created a simple synthetic terrain using the diamond-square algorithm. Reflectivity and coherence ($\gamma = 0.8$) were set to constant values for the whole image. Figure 15 shows the true and noisy phase, the estimates produced by the denoising filters and the difference of their results to the true phase. Clearly visible are artifacts for NL-InSAR along iso-height lines, which after phase-unwrapping would manifest as terraces in the DEM.

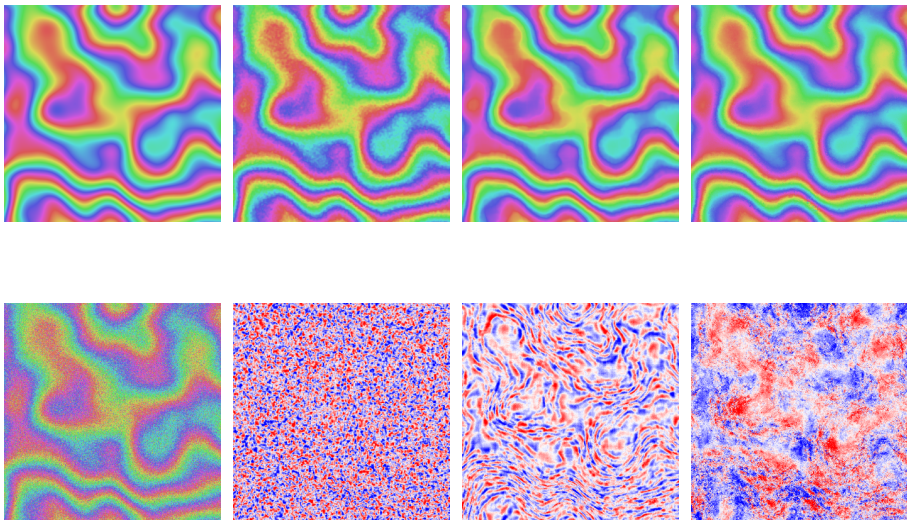


Figure 15: Synthetic terrain; top row: true phase and filtering results obtained by 5×5 Boxcar, NL-InSAR, and NL-SAR (from left to right); bottom row: noisy phase and phase differences of the respective filters to the true phase.

Table 2: Test sites acquisition dates and locations

Name	Acquisition date	Latitude and longitude
Salar de Uyuni	2011-11-03	-20.043, -67.65
Munich	2011-08-19	48.33, 11.64
Marseille	2012-05-07	43.24, 5.50
Hambach	2012-10-31	50.86, 6.36

Table 3: Phase standard deviation in degrees for a flat and homogeneous area of the salt flat Salar de Uyuni. NL-InSAR provides roughly a factor of three compared to the standard boxcar filter.

unfiltered phase	5×5 Boxcar	NL-InSAR	NL-SAR
56.10	12.51	4.64	3.76

5. 6m TanDEM-X DEMs

In addition to the previous synthetic experiments we showcase in this section, that the aforementioned qualities and peculiarities of non-local filters also show up for real data by a careful selection of test sites, which Table 2 lists.

The phase noise reduction of all filters is analyzed by an interferogram of Salar de Uyuni, a salt flat in Bolivia, almost perfectly flat and homogeneous. Table 3 shows the phase standard deviation in degrees for the unfiltered phase and the aforementioned filters for a selected area of $4,000 \times 4,000$ pixels with no discernible elevation and reflectivity change. Compared to the unfiltered phase the 5×5 Boxcar filter reduces the phase standard deviation by a factor close to the theoretical value of five. NL-InSAR and NL-SAR further improve on this by a factor roughly equal to three and four, respectively, showing that the targeted noise reduction by a factor of 2.5 is achievable.

To better grasp the improvements in DEM quality that NL-InSAR and NL-SAR provide we used both filters to generate 6m Raw DEMs. The 6m Raw DEM generation is identical to the standard raw DEM generation and uses the

aforementioned ITP of DLR Breit et al. (2010); Fritz et al. (2011), where the filters under analysis replace the default boxcar phase filter. In addition the DEM is geocoded to a finer 6 m grid to adhere to the HDEM specifications. For comparison we further provide the default 12m Raw DEM output of ITP which relies on a Boxcar filter for phase denoising.

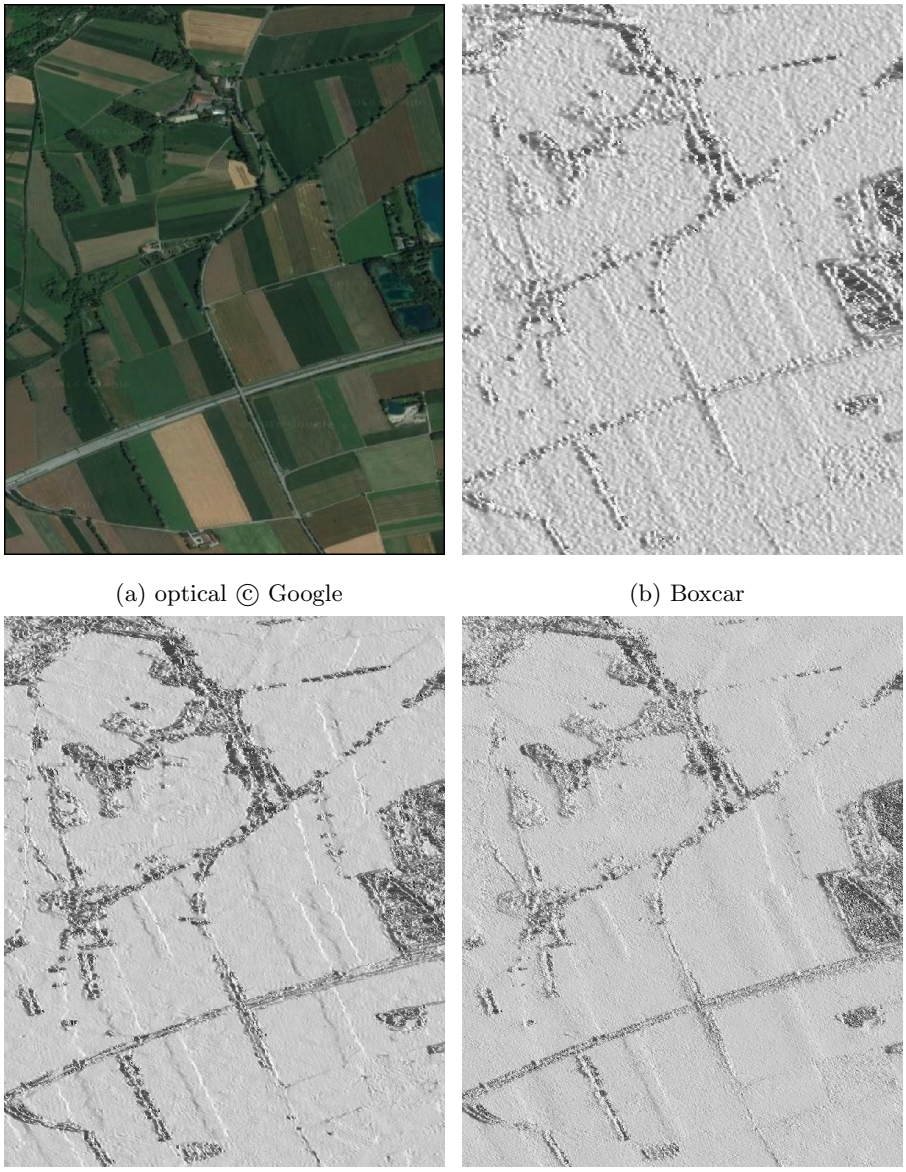
Figure 16 shows an optical image of a rural, agricultural area in Southern Germany and the respective DEMs produced by the filters, which show the improved noise reduction and detail preservation of non-local filters in comparison to the currently employed boxcar filter in the ITP chain. Since different varieties of crops are grown in the area, height changes are observable between the agricultural fields. As with the simulations of Figure 10 to Figure 14 the estimate of NL-InSAR exhibits much sharper edges than NL-SAR.

For demonstrating the terrace-effect of NL-InSAR we selected a mountainous area near Marseille, France. Figure 17 depicts an optical image and the DEMs of the filters under analysis. Again the superior noise reduction and detail preservation of both non-local filters is evident, by looking at the buildings in the upper half. For hilly terrain though NL-InSAR produces a DEM with distinct terraces, which are visible as gray-level fluctuations due to the applied shading.

One last example of the benefits of non-local filters over the Boxcar filter is highlighted by Figure 18, which shows DEMs of an open-pit mine in Western Germany. Many more details, such as the conveyor belts in the center, are visible in the DEMs generated by the non-local filters for two reasons: They drastically reduce the noise floor revealing structures that might have remained hidden otherwise and they don't smooth small features and details as the Boxcar filter does.

6. Conclusion

We have shown that the quality of TanDEM-X Raw DEMs can be improved by about a factor of ten by applying the NL-InSAR filter with appropriately



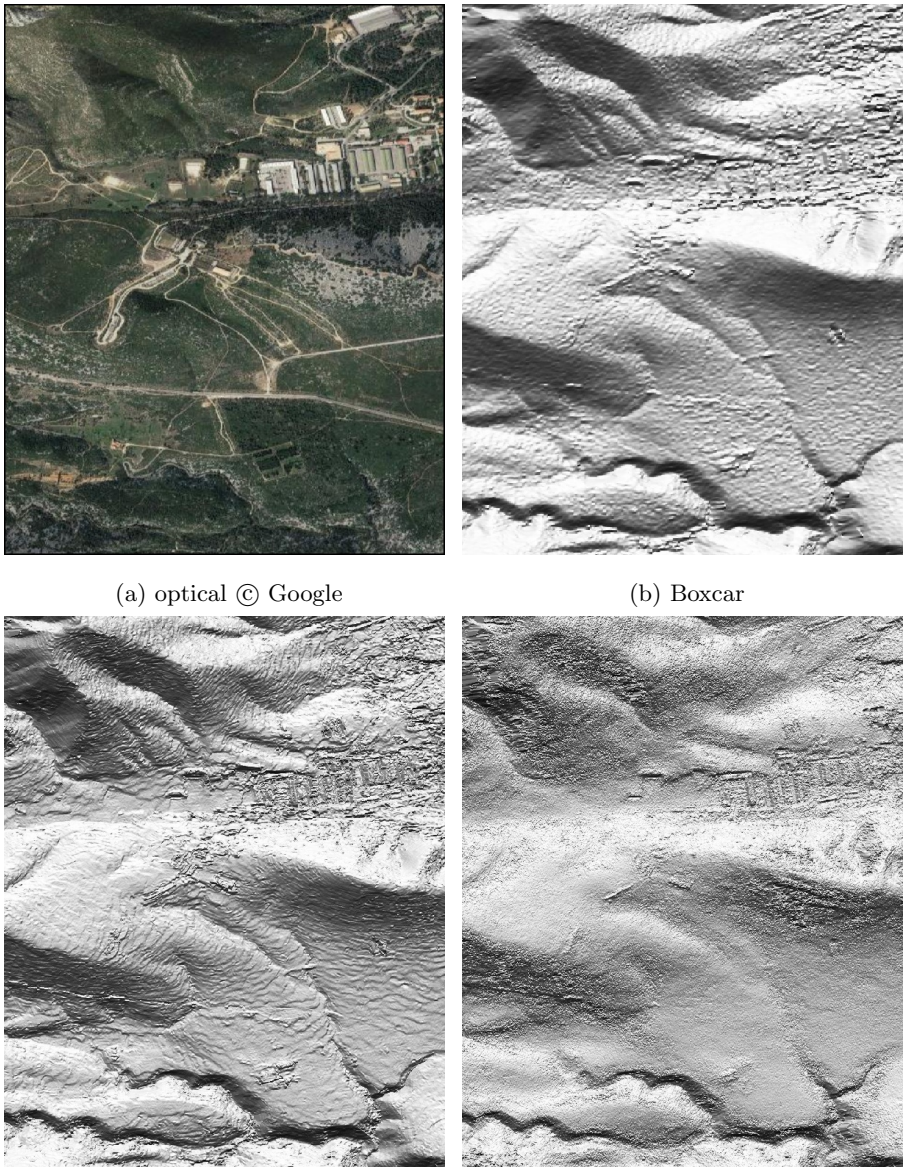
(a) optical © Google

(b) Boxcar

(c) NL-InSAR

(d) NL-SAR

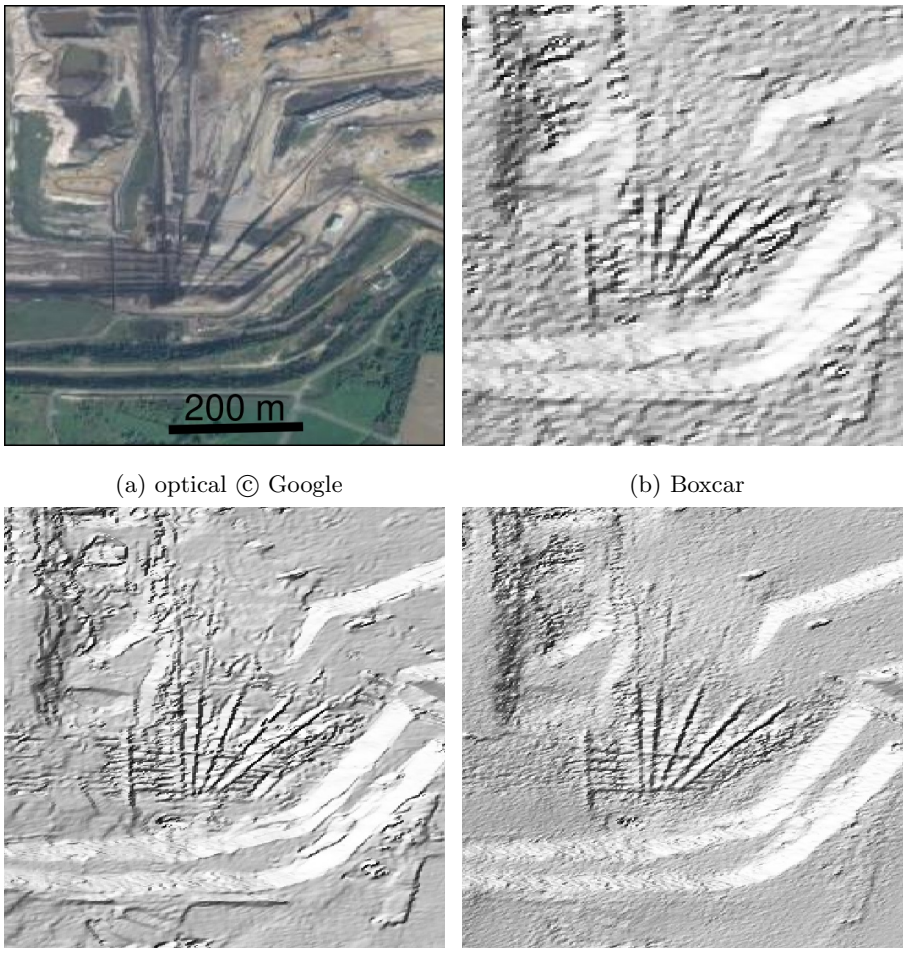
Figure 16: Agricultural area with fields. Optical image ©Google and DEMs with shading produced by a 5×5 Boxcar filter, NL-InSAR and NL-SAR. Clearly visible is the improved noise reduction of non-local filters, which makes it possible to discern fields of different height. From the resolution point of view the NL-InSAR filter is the superior one.



(a) optical © Google (b) Boxcar

(c) NL-InSAR (d) NL-SAR

Figure 17: Mountainous area. Optical mage © Google and DEMs with shading produced by a 5×5 Boxcar filter, NL-InSAR and NL-SAR. Again the non-local filters provide better noise reduction, yet NL-InSAR produces an estimate with distinct terraces.



(a) optical © Google (b) Boxcar

(c) NL-InSAR (d) NL-SAR

Figure 18: DEMs of an open-pit mine in western Germany produced by a 5×5 Boxcar filter, NL-InSAR and NL-SAR. Again the non-local filters exhibit a greater number of details and less noise, yet NL-SAR smoothes some details in comparison to NL-InSAR.

chosen parameters on the interferometric complex data. The NL-InSAR filter also produces a significantly less biased coherence estimates, which can ease the crucial and error-prone phase unwrapping step. We have observed and explained an unwanted terrace-like artifact produced by the original NL-InSAR filter. In a follow-up paper we will investigate possibilities to avoid this effect, e.g. by a special defringing pre-processing step. The effect of other filtering parameters, namely the weighting kernel smoothing h and the number of iterations, were also explained and shown in experiments, which are also generalizable to other non-local filers.

Acknowledgement

This work is supported by the Helmholtz Association under the framework of the Young Investigators Group "SiPEO" (VH-NG-1018).

References

Breit, H., Fritz, T., Balss, U., Niedermeier, A., Eineder, M., Yague-Martinez, N., & Rossi, C. (2010). Processing of bistatic TanDEM-X data. In *2010 IEEE International Geoscience and Remote Sensing Symposium* (pp. 2640–2643). doi:10.1109/IGARSS.2010.5653602.

Buades, A., Coll, B., & Morel, J. M. (2005a). A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (pp. 60–65 vol. 2). volume 2. doi:10.1109/CVPR.2005.38.

Buades, A., Coll, B., & Morel, J. M. (2005b). A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, *4*, 490–530. doi:10.1137/040616024.

Buades, A., Coll, B., & Morel, J. M. (2010). Image denoising methods. a new nonlocal principle. *SIAM Review*, *52*, 113–147. doi:10.1137/090773908.

Deledalle, C., Denis, L., & Tupin, F. (2011). NL-InSAR: Non-Local Interferogram Estimation. *IEEE Transaction on Geoscience and Remote Sensing*, *49*, 1441–1452. doi:10.1109/TGRS.2010.2076376.

Deledalle, C. A., Denis, L., Tupin, F., Reigber, A., & Jäger, M. (2015). NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In)SAR denoising. *IEEE Transactions on Geoscience and Remote Sensing*, *53*, 2021–2038. doi:10.1109/TGRS.2014.2352555.

Fritz, T., Rossi, C., Yague-Martinez, N., Rodriguez-Gonzalez, F., Lachaise, M., & Breit, H. (2011). Interferometric processing of TanDEM-X data. In *2011 IEEE International Geoscience and Remote Sensing Symposium* (pp. 2428–2431). doi:10.1109/IGARSS.2011.6049701.

Hoffmann, J., Huber, M., Marschalk, U., Wendleder, A., Wessel, B., Bachmann, M., Bräutigam, B., Busche, T., González, J. H., Krieger, G., Rizzoli, P., Eineder, M., & Fritz, T. (2016). *TanDEM-X Ground Segment DEM Products Specification Document*. Technical Report German Aerospace Center. URL: https://tandemx-science.dlr.de/pdfs/TD-GS-PS-0021_DEM-Product-Specification_v3.1.pdf.

Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M., & Zink, M. (2007). TanDEM-X: A satellite formation for high-resolution SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *45*, 3317–3341. doi:10.1109/TGRS.2007.900693.

Lachaise, M., Fritz, T., & Breit, H. (2014). InSAR processing and dual-baseline phase unwrapping for global TanDEM-X DEM generation. In *2014 IEEE Geoscience and Remote Sensing Symposium* (pp. 2229–2232). doi:10.1109/IGARSS.2014.6946912.

Touzi, R., Lopes, A., Bruniquel, J., & Vachon, P. W. (1999). Coherence estimation for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *37*, 135–149. doi:10.1109/36.739146.

Zhu, X. X., Bamler, R., Lachaise, M., Adam, F., Shi, Y., & Eineder, M. (2014a). Improving TanDEM-X DEMs by Non-local InSAR Filtering. In *EUSAR 2014; 10th European Conference on Synthetic Aperture Radar; Proceedings of* (pp. 1–4). URL: http://elib.dlr.de/93782/1/06856998.pdf.

Zhu, X. X., Lachaise, M., Adam, F., Shi, Y., Eineder, M., & Bamler, R. (2014b). Beyond the 12m TanDEM-X DEM. In *2014 IEEE Geoscience and Remote Sensing Symposium* (pp. 390–393). doi:10.1109/IGARSS.2014.6946440.