Advanced InSAR Tropospheric Corrections from Global Atmospheric Models that Incorporate Spatial Stochastic Properties of the Troposphere

Yunmeng Cao¹, Sigurjón Jónsson¹, Zhiwei Li²

¹King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia

²School of Geoscience and Info-Physics, Central South University, Changsha, Hunan, China

Key Points:

- An improved InSAR troposphere correction method proposed, based on Global Atmospheric Models (GAM), that considers spatial stochastic properties of the troposphere at different altitude levels for better weighting the GAM samples.
- Comparisons with three other GAM corrections (PyAPS, d-LOS, and GACOS) show a significant reduction in average standard deviation of corrected interferograms and better time-series analysis results.
- The results demonstrate the importance of considering horizontal heterogeneities and spatial stochastic models of the troposphere when using GAM-corrections.

28 **Abstract**

29 Tropospheric delays are still the main error source of satellite-based Interferometric Synthetic
30 Aperture Radar (InSAR) mapping of Earth's surface movements. Recent studies have
31 demonstrated the potential of global atmospheric models (GAMs) in reducing InSAR
32 tropospheric delays. However, the importance of appropriate interpolation and weighting
33 strategies in GAM-corrections has largely been overlooked. Here we present a new GAM-
34 based tropospheric correction method that incorporates spatial stochastic models of the
35 troposphere into the weighting strategy of the correction. The method determines the
36 correlation between a pixel of interest and neighboring GAM grid locations (3D) according to
37 the spatial variability of the tropospheric random field, instead of subjectively using an
38 inverse distance method, a local spline function, or other standard interpolation scheme. Also,
39 our new method considers horizontal heterogeneities of the tropospheric field by estimating
40 the integral of the tropospheric delays along the satellite line-of-sight (LOS) direction, instead
41 of calculating projected zenith-delays. The method can be used with any GAM, but we here
42 implement it with the latest ECMWF (European Center for Medium-Range Weather
43 Forecasts) ERA5 reanalysis products. We validate the new method with hundreds of
44 Sentinel-1 images from 2015 to 2020 over the island of Hawaii, a location with variable
45 topography, surface conditions, local climate, and deformation, and explore the tropospheric
46 corrections for both interferograms and time-series analysis products (deformation velocities
47 and time-series solutions). Compared with other GAM-corrections (PyAPS, d-LOS, and
48 GACOS), our new method yields a larger reduction of the average standard deviation of the
49 corrected interferograms, i.e. from 2.55 cm to 1.91 cm, instead of 2.47 cm (PyAPS), 2.44 cm
50 (d-LOS), and 2.10 cm (GACOS). Also, a larger fraction of 87% of the interferograms (243
51 out of 280) is improved, compared with 52%, 53%, and 66% for the other GAM corrections,
52 respectively. These results demonstrate the importance of considering (1) tropospheric
53 stochastic models in GAM-corrections, (2) horizontal heterogeneities when estimating the
54 LOS delays, and (3) tropospheric delays when mapping long-wavelength or small-magnitude
55 deformations using InSAR.

56

1. Introduction

Satellite-based interferometric synthetic aperture radar (InSAR) observations are extensively used to map surface displacements associated with a wide range of geophysical processes, such as earthquakes (e.g., Massonnet et al., 1993; Simons et al., 2002; Jónsson et al., 2003; Jónsson, 2008; Feng et al., 2010; Bell et al., 2012), volcanic activity (e.g., Amelung et al., 2007; Sigmundsson et al., 2010; Plattner et al., 2013; Ruch et al., 2016; Rivera et al., 2017), slow tectonic movements (e.g., Wright et al., 2004; Elliott et al., 2008; Walters et al., 2013; Cavalié & Jónsson, 2014), and extraction of the subsurface fluids and mineral resources (e.g., Bawden et al., 2001; Liu et al., 2016; Yang et al., 2017). Although tremendous progress has been made on improving InSAR observations during the past three decades, radar path delays caused by the spatio-temporal variations of the troposphere still remain as the dominant source of error (Zebker et al., 1997; Hanssen, 2001; Jolivet et al., 2014; Bekaert et al., 2015b; Yu et al., 2018; Li et al., 2019; Zebker, 2020).

Tropospheric corrections of InSAR observations can be divided into three groups: i) empirical, ii) statistical, and iii) predictive corrections. Empirical corrections try to reduce the tropospheric effects by modeling the relationship between topographic height and InSAR phase values (Bekaert, et al., 2015a; Lin et al., 2010; Wicks, 2002). These methods can be quite successful but do not work well when atmospheric turbulence dominates the tropospheric effects (Liang et al., 2018) and can be troublesome when the deformation is correlated with topography (Delacourt et al., 1998). The second category of corrections aims to mitigate tropospheric delays based on time-series of SAR images or interferograms by using statistical, geo-statistical, or adjustment algorithms, such as stacking (Sandwell & Sichoix, 2000), a range of least-squares based methods with an empirical deformation model (Berardino et al., 2002; Cao et al., 2017; Li et al., 2019), or spatio-temporal filtering (Ferretti et al., 2001; Hooper, 2008; Ferretti et al., 2011; Cao et al., 2019). Unfortunately, statistical approaches are usually not very effective as they rely on averaging stochastic properties of the spatio-temporal tropospheric delays and normally require a large number of SAR images for obtaining satisfactory results (Cao et al., 2017; Siddique et al., 2018). The third group of methods, predictive corrections, use auxiliary atmospheric datasets to compute and correct

the InSAR tropospheric delays. Numerous algorithms have been presented that either use one type or multiple different external atmospheric information from local meteorological data (Li et al., 2004), GPS measurements (Webley et al., 2002; Onn and Zebker, 2006; Houlie et al., 2016), satellite-based multispectral observations from MERIS (MEdium Resolution Imaging Spectrometer onboard ENVISAT satellite) (Li et al., 2009; Li et al., 2012) and MODIS (MODerate resolution Imaging Spectrometer onboard Terra and Aqua satellites) (Li, 2005), and weather models (Wadge et al., 2002; Foster et al., 2006; Doin et al., 2009; Liu et al., 2009; Hobiger et al., 2010; Jolivet et al., 2011). The performance of the predictive corrections depends on the spatio-temporal resolutions and the intrinsic precision of the available external data or weather model outputs, but availability of sufficiently good external information is usually the major limitation of these methods.

With the development of global atmospheric models (GAMs), including both global and regional reanalysis and operational weather models (e.g., NARR (Mesinger et al., 2006), ERA-Interim (Dee et al., 2011), MERRA-2 (Gelaro et al., 2017), HRES (Haiden et al., 2018)), many studies have demonstrated the potential of using GAM-based predictive corrections for mitigating InSAR tropospheric delays (Jolivet et al., 2014; Parker et al., 2015; Yu et al., 2018; Shen et al., 2019; Hu and Mallorquí, 2019; Murray et al., 2019). Thanks to the global coverage, all-weather and all-time usability, spatial three dimensions (3D) of gridded estimates, and near real time availability of GAM outputs, these predictive approaches will most likely become routine for correcting InSAR tropospheric delays. However, as the spatial resolution of the GAMs (e.g., ~10s km) is still coarse compared to that of SAR/InSAR images (e.g., ~10s m), the challenge is to construct high-resolution InSAR tropospheric corrections based on the coarse GAM outputs.

Jolivet et al. (2014) used a spline interpolation in the vertical direction and a bilinear interpolation in the horizontal direction to correct for high-resolution InSAR delays. By incorporating both GPS and GAM observations, Yu et al. (2018) modeled the stratified delays first and then used an inverse distance weighting (IDW) method (within 150 km) to predict the turbulent components of each pixel. Hu and Mallorquí (2019) proposed to calculate the integral of the delays along the satellite line-of-sight (LOS) directly, and then

used a similar approach as Jolivet et al. (2014) to predict the high-resolution delay maps. However, the weighting strategies of these GAM-corrections were determined subjectively and did not quantitatively model the correlations between the GAM-based 3D tropospheric samples and the points of interest. Different weather conditions may cause different spatial variabilities in the tropospheric parameters, and therefore, a good predictor should adjust the weighting strategy according to the stochastic properties of the tropospheric random field.

In this paper, we present a new GAM-based troposphere correction method that incorporates spatial stochastic models (SSMs) of the troposphere at different altitude levels into the weighting strategy of the correction. The method flexibly determines the correlation between measurement pixels and GAM grid locations (3D) according to the spatial variabilities of the tropospheric random field, instead of subjectively using an inverse distance method or a local spline function. It also considers horizontal heterogeneities of the tropospheric field due to the effects of the atmospheric turbulence (Hobiger et al., 2010; Hu et al., 2019), by estimating the integral of the tropospheric delays along the satellite line-of-sight (LOS) direction directly, instead of calculating the projected zenith-delays. While the new method can easily be generalized to any GAMs, we here implement it with the latest ECMWF (European Center for Medium-Range Weather Forecasts) ERA5 reanalysis outputs, which has been improved from the ERA-Interim model and provides hourly estimates of the tropospheric parameters (temperature, partial pressure of water vapor, and geopotential height of pressure levels) on a global 0.25° grid (~ 25 km) at 37 pressure levels, from 1979 to within 5 days of present time.

We begin with a description of our new method, which includes five key steps. Then, we test its performance over the island of Hawaii, a location with complex climate and deformation, using hundreds of Sentinel-1 images spanning 5 years, and compare the results with corrections made by commonly-used GAM correction methods. Finally, we discuss the importance of considering (1) tropospheric spatial variabilities in GAM-corrections, (2) horizontal heterogeneities for estimating the LOS delays, and (3) tropospheric delays for mapping long-wavelength or small-magnitude deformation using InSAR. Unless otherwise stated, the interferograms used in this study have been processed from SLC data to

unwrapped interferograms using PyINT (PYthon based INterferometry Toolbox) (Cao, 2019) based on the GAMMA software following the standard two-pass differential procedure. The improved Goldstein filter was used to decrease decorrelation noise (Baran et al., 2003) and the minimum cost flow (MCF) method for unwrapping (Wegmüller et al., 2002). We used the POD Precise Orbit Ephemerides (AUX_POEORB) data to improve the orbit parameters and the 30 m Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) (Van Zyl, 2001) to remove topographic phases.

2. Methodology: Incorporating tropospheric stochastic models in GAM-corrections

2.1 Estimating line-of-sight tropospheric delay samples

The line-of-sight (LOS) tropospheric delay is the integral of the air refractivity along the LOS direction, from a point on the ground to the satellite, which can be estimated as (Baby et al., 1988):

$$\Delta L = 10^{-6} \int_{s_{LOS}} N ds = 10^{-6} \int_{s_{LOS}} \left(k_1 \frac{P_d}{T} + k_2 \frac{e}{T} + k_3 \frac{e}{T^2} \right) ds \quad (1)$$

where ΔL is the LOS tropospheric delay, N is the air refractivity, P_d and e are the partial pressures (in pascals) of dry air and water vapor, respectively, T is the temperature (in kelvin), and $k_1 = 0.776 K Pa^{-1}$, $k_2 = 0.716 K Pa^{-1}$, and $k_3 = 3.75e3 K^2 Pa^{-1}$ are empirical constants (Smith and Weintraub, 1953). Typically, according to its physical origin, the total delay is considered as a sum of the dry and the wet components, which can be described as (Bevis et al., 1992):

$$\Delta L_{dry} = 10^{-6} \frac{k_1 R_d}{g_m} P_{s_0} \quad (2)$$

$$\Delta L_{wet} = 10^{-6} \int_{s_{LOS}} \left[\left(k_2 - \frac{R_d}{R_v} k_1 \right) \frac{e}{T} + k_3 \frac{e}{T^2} \right] ds \quad (3)$$

where P_{s_0} is the total pressure at ground point s_0 , g_m is the local gravity at the center of the LOS path, and $R_d = 287.05 J kg^{-1} K^{-1}$, $R_v = 461.495 J kg^{-1} K^{-1}$ are the specific gas

constants for dry air and water vapor, respectively. It should be noted that, the above-mentioned dry delay (Eq. (2)) is partly related to the water vapor due to the nondipole component of water vapor refractivity (Bevis et al., 1992), although the largest contribution of ΔL_{dry} is that of the dry air.

GAMs provide spatial 3D-gridded tropospheric parameters several times a day, at different pressure levels; here, we consider the ECMWF ERA5 reanalysis products, which hourly provide 37 pressure levels of atmospheric estimates at a 0.25° spatial grid, and we download ERA5 outputs that extend 2 degrees (in all directions) beyond the coverage of the SAR images we plan to correct. Considering the largest time differences between the SAR acquisitions and the ERA5 outputs would be up to 30 min, we linearly interpolate the ERA5 outputs in time based on the two closest ERA5 outputs when the time difference is larger than 20 min, otherwise we use the closest ERA5 output directly. Using the gridded GAM products, we can predict the tropospheric parameters along the satellite LOS direction, and, once we obtain the LOS parameters, we can then calculate the LOS delay based on Eq. (1). To do this, we convert the ERA5 products into altitude levels, first using a linear interpolation along the vertical direction, and then, considering the atmospheric refractivity decreases significantly with increasing altitude in the lowermost 5 km and can be ignored above 40 km, we sample the altitude levels with increasing sampling steps from -200 m to 40 km: a sampling step of 200 m for altitudes between -200 m to 5 km, a 500 m step for 5 ~ 10 km altitudes, a 1 km step for 10 ~ 20 km, and a 5 km step for 20 ~ 40 km. We then calculate the locations of the tropospheric parameters along the LOS direction, at each altitude level, based on the SAR orbital state vectors and the ground point location of interest under the earth-centered Cartesian system (Hu and Mallorquí, 2019), and the start points of the LOS vectors are selected as the same locations of the GAM outputs at the lowest altitude level of -200 m (as Figure 1). After that, we estimate the LOS tropospheric parameters (temperature, pressure, and the partial pressure of water vapor), at each altitude level, using the Simple Kriging with local means (SKlm) algorithm (Kaymaz, 2005; Xu et al., 2011) and by considering spatial stochastic models (SSMs) of the tropospheric parameters. The predicted parameters at the k th altitude level can be described as:

$$z_{SKlm}^*(\varphi, L, h_k) = \sum_{i=1}^n u_i^{SK} [z(\varphi_i, L_i, h_k) - m_{SK}^*(\varphi_i, L_i, h_k)] + m_{SK}^*(\varphi, L, h_k) \quad (4)$$

where (φ, L, h_k) means the geographic coordinates of latitude, longitude, and the altitude of the k th level, $z_{SKlm}^*(\varphi, L, h_k)$ and $z(\varphi_i, L_i, h_k)$ represent the ‘to be predicted’ and the GAM-output parameters, respectively, $m_{SK}^*(\varphi, L, h_k)$ and $m_{SK}^*(\varphi_i, L_i, h_k)$ represent the local mean of the tropospheric parameters (e.g., P_d, T, e) at the ‘to be predicted’ location and the i th GAM sample location at the k th altitude level, respectively, and u_i^{SK} denotes the weight associated with the i th GAM output, and n is the total number of the available GAM outputs at the k th level. We model the local mean as a linear ramp for each altitude level by using all of the available grids within the level

$$m_{SK}^*(\varphi, L, h_k) = C_k + \alpha_k \varphi + \beta_k L \times \cos(\varphi) \quad k = 1, 2, \dots, M \quad (5)$$

where C_k , α_k , and β_k represent the ramp parameters of the k th altitude level, estimated using a least-squares algorithm from the gridded GAM outputs, $L \times \cos(\varphi)$ is used to get isometric coordinates, and M is the total number of the altitude levels. Here we consider a tropospheric ramp at each altitude level because 1) tropospheric parameters usually show a gradient in space, indicating organized large-scale air mass motion (e.g., Antonia et al., 1979), and 2) a ramp better presents the local variations in the mean value than a simple average. The weight u_i^{SK} is estimated based on a linear predictor by constraining 1) the theoretical variance of the prediction to be minimum, and 2) the sum of the weights equal to 1 for obtaining an unbiased prediction and the optimal kriging solution of the weights can be described as (Cressie, 1990):

$$\begin{bmatrix} \sigma_1^2 & \sigma_{1,2}^2 & \cdots & \sigma_{1,n}^2 & 1 \\ \sigma_{2,1}^2 & \sigma_2^2 & \cdots & \sigma_{2,n}^2 & 1 \\ \vdots & \vdots & \ddots & \vdots & 1 \\ \sigma_{n,1}^2 & \sigma_{n,2}^2 & \cdots & \sigma_n^2 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} u_1^{SK} \\ u_2^{SK} \\ \vdots \\ u_n^{SK} \\ \mu \end{bmatrix} = \begin{bmatrix} \sigma_{1,x}^2 \\ \sigma_{2,x}^2 \\ \vdots \\ \sigma_{3,x}^2 \\ 1 \end{bmatrix} \quad (6)$$

where σ_i^2 is the variance of the i th GAM output at the k th altitude level, $\sigma_{i,j}^2$ is the covariance between the i th and the j th GAM outputs, $\sigma_{i,x}^2$ is the covariance between the i th GAM output and the ‘to be predicted’ parameter, and μ is the Lagrange multiplier. The variance-

220 covariance components (σ_i^2 , $\sigma_{i,j}^2$, and $\sigma_{i,x}^2$) are estimated by modeling the variogram of the
221 tropospheric parameters (Hanssen, 2001; Kaymaz, 2005), which can be written as:

$$222 \quad \sigma_{i,q}^2 = \gamma(\vec{d}_{i,q}) \quad (7)$$

223 where $\gamma(\vec{d}_{i,q})$ means the variogram function, and $\vec{d}_{i,q}$ means the distance vector between the
224 two points (i.e., locations) i and q . Note that here the variance-covariance components do not
225 represent the uncertainty of the GAM outputs, instead they are determined by the spatial
226 variations of the troposphere. Here, we model the variogram function by using an empirical
227 isotropic spherical function, which includes three parameters (*nugget*, *sill*, *range*) and the
228 parameters are estimated using a least-squares algorithm (e.g., Cressie, 1990; Cao et al.,
229 2017). In order to make sure the spatial tropospheric parameters meet the second-order or
230 intrinsic stationary assumptions (Hanssen, 2001), which is the prerequisite of estimating a
231 variogram (i.e., structure function), we remove the ramp components (see Eq. (5)) from the
232 original observations before calculating the spatial variance samples. After estimating all the
233 variance-covariance components in Eq. (6), we can directly obtain the optimal solution for
234 the weights, and then calculate the SKlm-based prediction of $z_{SKlm}^*(\varphi, L, h_k)$. It should be
235 noted that the variogram (Eq. (7)) is estimated after removing the modeled ramp (Eq. (5)),
236 because a ramp may cause non-stationarity of the random field, which should at least fulfil
237 intrinsic stationarity when applying the Kriging algorithm (Cressie, 1990; Kaymaz, 2005).

238

Figure 1. Altitude levels of global atmospheric model (GAM) outputs (blue dots) and locations of ‘to be predicted’ parameters (orange dots) along line-of-sight (LOS) directions between ground points and an imaging satellite.

We estimate the LOS tropospheric parameters (temperature, pressure, and the partial pressure of water vapor) using Eq. (4) separately for each altitude level. Then, once we have obtained the LOS parameters for all the altitude levels, we compute the integral of the LOS delays (see Eq. (1)) to get the delay estimates at different altitude levels. It should be noted that the stochastic models of the tropospheric parameters at different altitude levels are estimated separately, for example, for 3 parameters at 51 altitude levels, we need to estimate 153 stochastic models.

2.2 Constructing high-resolution InSAR tropospheric delay maps

Because the preliminary densities of the estimated LOS delay samples at each altitude level are at a similar scale as that of the GAM outputs, i.e., at a coarse grid compared to the InSAR data, we need to interpolate the delay samples to a similar resolution as the InSAR measurements. To do this, we use the same SKlm method as the one used for predicting the

LOS tropospheric parameters in Section 2.1: 1) Using the LOS delay samples, we estimate the stochastic models of the delays, at different altitude levels, first, based on the empirical isotropic spherical function, we then estimate the densely gridded LOS delays using the SKlm method level by level. The prediction of the SKlm-based delay $S_k^*(x)$ at the k th altitude level and location x can be written as

$$S_k^*(x) = \sum_{i=1}^n \omega_i (S_{k,i} - R_{k,i}^*) + R_k^*(x) \quad (8)$$

where $S_{k,i}$ is the i th LOS delay sample of the k th level, $R_{k,i}^*$ is the modeled ramp component of the i th LOS delay sample with the ramp function (as in Eq. (5)) determined based on the delay samples at the k th altitude level, $R_k^*(x)$ is the expected ramp component at location x , and ω_i is the weight of the i th delay sample, which is determined by considering the stochastic model of the LOS tropospheric delays at the k th level (as in Eq. (6)).

In practice, considering the computational burden of estimating the SKlm weights pixel by pixel at a high-resolution, we do not interpolate the LOS delays to similar spatial resolution as the InSAR data. Instead, we only oversample 10 times of the original delay samples for each altitude level, leading to a spatial resolution of the SKlm interpolated LOS delays of approximately 2.5 km. In addition, we do not need to interpolate the delays for all of the altitude levels; we only select those levels that encompass the maximum and minimum elevations of the research area.

Lastly, using the densely 3D-gridded LOS delay samples, we can produce the high-resolution LOS delay maps based on the elevation and the geographic coordinate information of each SAR pixel. Here we use a trilinear interpolation to construct the high-resolution maps and the 30 m SRTM DEM. For producing the InSAR tropospheric delay maps, we use the two ERA-5 outputs closest to the two SAR-image acquisition times of each interferogram. We subsequently subtract the reference delay map from other one, which allows us to obtain the InSAR tropospheric delay map.

The workflow of the new method is presented as Figure 2 and is as follows:

- 1) Pressure levels of the GAM outputs are interpolated into altitude levels and the spatial variabilities of the tropospheric parameters modeled for different altitude levels.
- 2) The LOS locations, at different altitude levels, are calculated based on the orbit state vectors and the ground point geographic coordinates, and the tropospheric parameters at the LOS locations predicted using the SKlm method (see Eq. (4)).
- 3) The integral of the LOS tropospheric delays is then calculated using Eq. (1) to get LOS delay samples at different altitude levels.
- 4) The spatial stochastic models of the LOS tropospheric delays are estimated at different altitude levels and then the LOS delay samples interpolated by using a similar SKlm method as used in step 2 (see Eq. (8)).
- 5) Finally, the high-resolution InSAR tropospheric delay map is constructed using a 3D-trilinear predictor based on the 3D-gridded LOS delay samples by considering the SRTM DEM and the geographic coordinate information.

Figure 2. The general workflow of our new GAM-based InSAR tropospheric correction method.

3. Method validation and results

3.1 Study area, geographic setting, and experimental datasets

The island of Hawaii (or Hawai'i, Figure 3a) is an ocean island environment that includes the active shield volcanoes of Mauna Loa and Kīlauea, both located on the southern half of the island. Past InSAR-based studies have presented a range of deformation signals associated with the volcanic and tectonic activity on the island, often suggesting complex deformation sources and time-dependent processes (e.g., Lundgren et al., 2013; Farquharson et al., 2020). Meanwhile, strong tropospheric effects have also been commonly reported in InSAR data of the island (e.g., Foster et al., 2006; Shirzaei et al., 2013; Zebker, 2020). Hence, InSAR-derived deformation measurements are usually significantly contaminated by tropospheric delays. As an example of the strong tropospheric delays in Hawaii, Figure 3b shows a Sentinel-1 interferogram (SAR images from 1 August, 2016 and 13 August, 2016) with signals that are a complex mix of stratified and turbulent delays. This is due to the moist and heterogeneous tropical atmosphere in Hawaii, as well as the strong and highly variable interactions between winds around the large volcanoes on the island.

Due to its highly variable tropospheric field, both in time and space, and its active geological environment, the island of Hawaii is one of the best study areas to test our new method. Therefore, we used 197 Sentinel-1 SAR images (descending track 87) over the Hawaii area (Figure 3a, inset), spanning the years from 2015 to 2020, and generated more than 1000 interferograms for this purpose. Here we focus solely on the InSAR troposphere corrections, but do not attempt to interpret observed deformation due to the volcanic activity on the island during this period. We test the new tropospheric correction method for single interferograms, velocity estimations, as well as for the time-series solution estimations. In addition, we compare our results to those obtained by three other commonly-used GAM-corrections: PyAPS (Jolivet et al., 2014), d-LOS (Hu & Mallorquí, 2019), and GACOS (Yu et al., 2018). We implement PyAPS and d-LOS using the ERA-5 outputs in the comparison,

while we should note that, GACOS uses both the High Resolution ECMWF numerical weather model (HRES) and GPS data.

Figure 3. (a) Topographic map showing the five shield volcanoes on the island of Hawaii with red triangles representing the active volcanoes of Mauna Loa and Kīlauea. Top left inset shows the study location in the Hawaii archipelago and the coverage of the Sentinel-1 SAR images used (yellow rectangle). (b) An interferogram example (1-13 August 2016) of the island of Hawaii with strong tropospheric effects. (c) Unwrapped phase values (from the same interferogram as (b)) versus elevation, showing the complex relationship of the InSAR data with topography.

3.2 Generation of tropospheric delay maps for SAR acquisitions

The SAR imaging time of the Sentinel-1 images of Hawaii used here is UTC 16:16. Therefore, we downloaded ERA-5 raw outputs from UTC 16:00 for all of the 197 SAR acquisitions, with a coverage of the ERA-5 outputs extending by 2° from the original SAR image coverage, yielding 41×41 grid points for each horizontal layer. It should be noted that the extended area of the ERA5 data is determined subjectively in our case, and we suggest an extension of 1- to 2- degree from the raw SAR coverage.

Here, as an example, we show the generation of a LOS tropospheric delay map for a SAR image acquired on 30 October 2015 (Figure 4). The LOS tropospheric parameters

(Figure 4a-5c), which were estimated using the SKIm method and the original GAM outputs, show larger variabilities in partial pressure of water vapor (e) than in air temperature (T) and pressure (P), especially in the lower troposphere. The curve of the partial pressure of water vapor (located at latitude of 19.5° and longitude of -155°) versus the altitude (Figure 4j) shows that most of the water vapor activity is in the lower troposphere (i.e., below 5 km). This is consistent with the estimated spatial variograms at different altitude levels (Figure 4k), which show strong variabilities in partial pressure in the lower troposphere with variances at an altitude of 400 m reaching up to around $13 \times 10^4 Pa^2$, versus that of $6 \times 10^4 Pa^2$ at an altitude of 2000 m.

We calculated the LOS delay components (Figure 4d-4e) directly based on Eqs. (2) and (3), using the LOS tropospheric parameters. A similar relationship between the altitude and wet delays (Figure 4l) can be seen as for partial pressure of water vapor (Figure 4j). Variograms of the wet delays (Figure 4m) show significantly different variabilities in the delays at different altitudes, although they show a similar maximum correlation distances (around 200 km). The wet-delay variance at an altitude of 400 m is over $15 cm^2$, whereas at an altitude of 2000 m, it has decreased to about $5 cm^2$. By using the SKIm method, we interpolated the coarse (~ 25 km) delay components to obtain finer delay estimations (Figure 4f-4g) at a grid with about 2.5 km spacing.

We consider a constant geoid height value (16.6 m) over the island of Hawaii and we added the geoid height to the SRTM DEM to get the same altitude reference as that of the gridded delay. Finally, we derived the SAR tropospheric delays (Figure 4h-4i) based on the densely 3D-gridded delays and the DEM by using a 3D-cubic interpolation method. The total LOS tropospheric delay map is presented in Figure 4n, showing that in this example they vary from about 1.6 m (at the peaks of Mauna Kea and Mauna Loa) to about 3.5 m (along the west coast). The LOS delays at the west coast are stronger than those at the east coast, mainly due to the differences in incidence angles, which is increases from about 32° at the east coast to about 42° at the west coast of the island. Similar changes between the two coastlines can be seen in the estimated dry delays (e.g., Figure 4d, 4f, and 4h), which dominate the total

371 delay magnitude, while the wet part of the delay does not exhibit this pattern. This indicates
 372 that the wet delays have stronger spatial variabilities than dry delays.

373

374 **Figure 4.** Generation of SAR LOS tropospheric delay map (30 October 2015) from GAM: (a)-(c) Estimated tropospheric
 375 parameters (P, T, and e) along LOS vectors using a SKlm method based on GAM outputs, (d)-(e) Calculated LOS delay
 376 components (dry and wet delays), (f)-(g) Interpolated densely gridded LOS delays using SKlm method, (h)-(i) High-
 377 resolution of SAR delay components, (j) Variation of partial pressure of water vapor with altitude (point location: latitude =
 378 19.5°, longitude = -155.0°); (k) Variograms of the water vapor partial pressure at different altitude levels (400 m, 1000 m,
 379 and 2000 m), (l) Variation of wet delays with altitude at the same location as for (j), (m) Variograms of wet delay
 380 components at different altitude levels, and (n) the total estimated SAR tropospheric delay along LOS.

381

3.3 Tropospheric corrections for single interferograms

We here present two example cases of correcting single interferograms with the method described above. Both interferograms have a short temporal baseline (12 days) and should thus be mostly without any deformation and mainly contain tropospheric signals (Figure 5).

The first interferogram is from the Winter of 2018 (29 Jan. - 10 Feb., 2018), well before the May 2018 eruption, and it exhibits (the first image in Figure 5a) strong tropospheric signals (~ 8 cm) along the northwestern and the northeastern coasts of the island. At first glance, the interferogram seems to imply substantial ground movements towards the satellite, which in reality is not the case, as these signals are purely of tropospheric origin. Also, some topography-correlated signals can be seen on Mauna Loa and Mauna Kea, which are mainly due to the effects of the stratified troposphere. Correcting the interferogram using the PyAPS and d-LOS methods mostly removes these coastal and topography-correlated signals, but an anomaly remains in the southeast (Figure 5). The GACOS-based result is less successful in this case leaving more tropospheric residuals than in the first two corrections, especially in the eastern and northern parts of the island. In comparison, our new method clearly shows the best performance, both visually as well as with lower standard deviation (STD) and mean value, which are 1.4 cm (from 3.0 cm) and 0.71 cm (from 1.47 cm), respectively. This indicates that over 50% of the tropospheric signals are corrected after using the new method. Comparison of the original data values with the four predictions (Figure 5b) show, as observed visually, that derived predictions from PyAPS, d-LOS, and the new method are closer to the original observations than GACOS. The best similarity is achieved by our new method, which is consistent with the lowest standard deviation obtained among these of four corrections.

The second interferogram case is from the Summer of 2018 (28 July - 9 Aug., 2018) and from within the May eruption period (from early May to early August). This interferogram is also strongly affected by the tropospheric delays (Figure 5j-5n), especially in the western and the active southeastern areas of the island. A local deformation signal is visible at Kīlauea (red rectangle and zoom at bottom right), which is clearly contaminated by tropospheric signals in the original interferogram. While the PyAPS and d-LOS corrections reduce the

411 tropospheric signals at Kīlauea, the overall improvements of the corrected interferograms
412 (Figure 5k and 5l) are limited with even a new anomaly appearing across the central part of
413 the island, which is not observed in the original data. The GACOS correction (Figure 5m)
414 shows a better performance than PyAPS and d-LOS in this case, particularly in the eastern
415 and southeastern areas of the island, but some anomalous signals occur in the west after the
416 GACOS correction. The new method yields substantially better results (Figure 5n) than the
417 other three corrections, especially in the southeastern areas, with the deforming pattern at
418 Kīlauea becoming more clear. The STD and mean values of the corrected results are 43.5%
419 (from 2.16 cm to 1.22 cm) and 87.7% (from 1.22 cm to 0.15 cm) lower, respectively, than in
420 the original interferogram and both values are lower than those for the other three corrections.
421 Comparisons of the original interferogram values with the predictions (Figure 5o-5r) show
422 that the PyAPS and d-LOS predictions are slightly overestimated for both positive and
423 negative values, while the GACOS prediction appears appropriate for the small values but
424 has a major bias in predicted large values. The new method presented here, on the other hand,
425 yields much better prediction for the entire range of values, than the other three methods, and
426 is without strong biases.

427

Figure 5. Two cases of tropospheric corrections for interferograms: (a-e) The first interferogram (29 January and 10 February 2018) before and after corrections; (f-i) Distribution of the raw observations of the first case along with the predictions from the four methods; (j-r) Same as (a-i), except for the second interferogram (28 July and 9 August 2018). The small black square in (a) shows the location of the reference point.

We statistically assess and compare the four correction methods on 280 small-baseline interferograms generated from the 197 SAR images, acquired from July 2015 to March 2020, using spatial- and temporal- baseline thresholds of ~ 120 m and ~ 12 days, respectively. Considering the strong deformation associated with the May 2018 eruption, we excluded six small-baseline interferograms containing the co- and early post-eruption deformation. Standard deviation values (Figure 6a) of the interferograms (before and after corrections) show that in most cases the new method (red circles) yields lower values than the other corrections, indicating better performance in mitigating the spatial variations of the tropospheric delays. The mean values of the residuals after correction (Figure 6b), which equal the average shift relative to a chosen reference pixel (always the same pixel), are

mainly caused by long wavelength residual signals. Here the results also show that the new method has the smallest deviation (around zero) among the corrections, while the mean values after the PyAPS and d-LOS corrections are often larger than that of the original uncorrected interferograms. This indicates that the new method can recover better those long wavelength tropospheric signals than the other three corrections. Quantitative comparison of the four correction methods (Table 1) confirms that, in general, all the corrections reduce spatial variations of the tropospheric signals in the data with the new method yielding the largest improvements. The STD values decrease on average from 2.54 cm to 1.91 cm and more than 85% (243 out of 280) of the input interferograms improved after the correction. The average absolute mean values show, on the other hand, that except for the new method the other three corrections yielded results are not improved (i.e., decreased); After the PyAPS and d-LOS corrections, the average absolute mean values are even larger than before corrections and fewer than 50% (about 37%) of the interferograms are improved. In comparison, around 58% (161 out of 280) of the interferograms show lower mean absolute values after correction when using the new method, with the average value decreasing from 1.38 cm to 1.15 cm.

Figure 6. Statistical assessment (before and after the troposphere corrections) of the 280 small-baseline interferograms, generated by data from June 2015 to March 2020. (a) standard deviations (STDs) and (b) absolute mean values. The

statistical values (STD and mean value) are calculated after masking ocean and low-coherence (coherence < 0.9) areas. Each bar represents an average of a group of 10 interferograms.

Table 1. Quantitative assessment and comparisons of the correction methods based on 280 small-baseline interferograms. Average standard deviation values (unit: cm) and average absolute mean values (unit: cm) are calculated based on the 280 standard deviation and mean values. The number in the ‘Imp. IFGs’ column represents the number of cases (of 280 in total) for which the correction yielded a smaller STD or mean values, compared to the original interferogram.

Corrections	Ave. STD Imp. IFGs		Ave. Mean Imp. IFGs	
Raw IFGs	2.54	\	1.38	\
PyAPS	2.47	148	2.18	100
d-LOS	2.44	149	2.08	104
GACOS	2.10	185	1.38	145
New-method	1.91	243	1.15	161

3.4 Tropospheric corrections for InSAR LOS velocity estimation

To evaluate the importance of tropospheric corrections for InSAR LOS velocity estimations, we calculated the pre-eruption velocities before and after applying the tropospheric corrections using the small baseline subset (SBAS) approach (Berardino et al., 2002; Zhang et al., 2019) for two different time periods or phases (phase 1: from June 2015 to June 2017; phase 2: from June 2017 to May 2018 eruption), determined according to the results of Farquharson and Amelung (2020). For the first phase we have 60 SAR images but only 27 SAR images for second phase due to its shorter time span. The derived LOS velocity maps and the related standard deviation maps are shown in Figure 7, with the standard deviations estimated using the misfits of the modeled linear time-series. Here, the positive velocity (in red) means ground movement towards satellite and the negative velocity (in blue) means movement away from the satellite.

The velocity maps of the first phase (Figure 7a) show that before any tropospheric corrections there are significant topography-correlated velocities (movement away from the satellite) on Mauna Kea and Mauna Loa, which are not real and mostly caused by the stratified tropospheric delays; The derived LOS velocity maps after the PyAPS and d-LOS corrections (the second and the third images of Figure 7a) are quite similar with both corrections successfully removing the topo-correlated velocities, yielding more realistic deformation patterns on the Island, especially at Mauna Loa. There are still some residual velocities in these maps that are not obviously associated with deformation, e.g., on the southwesternmost part of the island (~ 1.5 cm/year). The velocity map after the GACOS correction (the fourth image in Figure 7a) also shows that most of the topo-correlated signals are gone, but here the whole velocity map looks like slightly overestimated, i.e., most stable areas show slight positive velocities, as well as some local velocity anomalies (~ 2 cm/year), e.g., along the western coast. The velocity map from the new correction method (rightmost image in Figure 7a) appears similar to the PyAPS and d-LOS results, although the estimated velocities in stable areas are generally closer to zero (e.g., in the southwest), indicating that the new method performs better than the other three GAM-correction methods.

The standard deviation maps of the velocity derivations for phase 1 (Figure 7b) show that that for the uncorrected case (the leftmost image in Figure 7b) there are strong topography- correlated STD values on Mauna Loa and Mauna Kea with increasing STD with elevation (~ 1.0 cm), which is mainly due to increasing elevation differences with respect to the reference point (the black rectangular). The larger the elevation difference is, the stronger the tropospheric effects tend to be. Although these topography-correlated STD values disappear after the PyAPS and d-LOS corrections, most of the velocity STD values are not significantly reduced after these corrections compared with the uncorrected case, and the STD values on the eastern side of the island have even strongly increased. This implies that while PyAPS and d-LOS can mitigate the stratified tropospheric delays effectively, reductions of the turbulent signals are limited in this case. The standard deviation map after using the GACOS correction also shows obvious improvements regarding the topography-correlated STD values and the STD values to the southeast are also improved (i.e., decreased), especially along the coastline. This indicates that GACOS can mitigate temporal variations of

the tropospheric delays effectively, although the improvement in the velocity map is relatively small. The new method yields even lower STD values in the west, although the improvement is limited in the east. It is noteworthy that the velocities based on the new method in the east are quite similar to those from the PyAPS and d-LOS correction, whereas the velocity STD values in the same area are significantly smaller for our new method than those of PyAPS and d-LOS. This implies that the new method performs better in mitigating temporal deviation of the tropospheric delays than PyAPS and d-LOS and that it is important to consider a realistic spatial stochastic model of the troposphere in GAM corrections.

Figure 7. Estimated LOS velocity maps and the related standard deviations for two pre-eruption periods or phases: (a) Velocity maps of phase 1 (from June 2015 to June 2017) estimated based on 60 SAR images; (b) Standard deviations of the estimated velocities of phase 1; (c) Velocity maps of phase 2 (from June 2017 to the May 2018 eruption) estimated based on 27 SAR images; (d) Standard deviations of the estimated velocities of phase 2. Small black squares show the location of the reference point.

The velocity maps of phase 2 (Figure 7c) are more seriously affected by tropospheric delays than the maps of phase 1 (Figure 7a), due to the shorter period and the smaller dataset. The uncorrected velocity map shows again significant topography-correlated signals as well

as strong positive (movement towards satellite) signals along the coastline, which clearly is not real deformation. However, even after tropospheric corrections using the PyAPS, d-LOS, and GACOS methods, the velocity maps are not much improved and many residual tropospheric signals remain, especially along the southeastern coast. In addition, deformation velocities at Kīlauea appear overestimated. In comparison, the new method yields a velocity map (rightmost image in Figure 7c) with less tropospheric effects and the clearer deforming patterns at Mauna Loa and Kīlauea volcanoes, even though some local tropospheric signals remain, e.g., to the northwest and southwest.

The standard deviation maps for phase 2 (Figure 7d) have significantly larger magnitudes (~ 3.5 cm/year) than for phase 1 (Figure 7b), consistent stronger tropospheric effects on the velocities of phase 2. Again here, the STD values of the uncorrected velocity map (leftmost image in Figure 7d) show significant topography-correlated properties due to the effects of the stratified delays. The STD maps of corrected velocities all show large improvements, implying that all of the four corrections work well in reducing the temporal variations of the tropospheric delays during this phase, even though the improvements in the velocity maps are not that significant. When compared, the PyAPS and d-LOS STD maps are slightly better than that for the GACOS correction, especially in the East, but the STD map of the new method is clearly the best of the four corrections. This indicates the corrected velocities by the new method are more reliable than those of the other three corrections, with a more complete correction of the tropospheric delays.

In summary, the velocity and standard deviation maps of the two pre-eruption phases demonstrate that 1) stratified delays cause topography-correlated signals in the InSAR LOS velocity maps (see Figure 7a and 8c), even when 60 SAR images are used that span a time period of 2 years, 2) all four GAM corrections reduce the stratified delays effectively, whereas the method performance differs in mitigating the temporal variations of the troposphere, which is mainly related to the turbulent delays, 3) the new method performs the best among the four correction methods, yielding LOS velocity maps that have lower uncertainties, and 4) it is necessary to correct the tropospheric delays for estimating robust deformation velocities, especially when using small datasets of SAR images.

To quantitatively evaluate the accuracy of InSAR derived velocities, we compare them with GPS-derived velocities at permanent GPS stations on the island. Deformation velocities before the May 2018 eruption (June 2015 to April 2018) were calculated (Figure S1) at the 36 GPS stations having synchronous observations with InSAR and located in areas good InSAR coherence. By taking the GPS derived LOS velocities as true values, we estimate the absolute errors of the InSAR derived velocities (Figure 8). The mean absolute errors (MAE) of the InSAR velocities without corrections is 12.7 mm/year, which is mainly caused by the topography-correlated tropospheric delays as we analyzed before. The tropospheric corrections reduce the MAE to 4.9, 5.2, 5.7, and 4.5 mm/year, using the PyAPS, d-LOS, GACOS, and our new method, respectively. This shows that the four correction-based velocities yield similar velocity accuracies, with our new method leading to the smallest MEA. These results also indicate that topography correlated delays can cause large uncertainties in InSAR velocities, that after corrections the velocity uncertainties are mainly caused by tropospheric turbulence, and that turbulent effects on the velocity estimation is limited, especially when we have a large number of SAR acquisitions. We should also note that the differences between InSAR- and GPS- velocity could be partly due to inaccurate GPS station velocity.

Figure 8. Absolute errors of InSAR derived LOS velocities before the May 2018 eruption (June 2015 to April 2018): (a) no correction; (b) PyAPS correction; (c) d-LOS correction; (d) GACOS correction; (e) Our new method based correction.

3.5 Tropospheric corrections for InSAR time-series solutions

In addition to studying the improvement the tropospheric corrections have on single interferograms and InSAR velocity maps, we here assess the effect on InSAR time series solutions. We estimate five different time-series solutions (Figure S2-S6 in the supporting information) using all the 197 SAR acquisitions from June 2015 to March 2020 and 1134 interferograms, one without any tropospheric corrections and then four solutions using the four different tropospheric correction methods. Instead of analyzing the geological processes the time-series reveal, we focus here on evaluating the effect the tropospheric corrections have on the time-series solutions.

As examples, we present and evaluate the time-series results at two GPS stations (MKEA located at Mauna Kea and ELEP located at Mauna Loa) produced with different tropospheric corrections (Figure 9), and we fix the reference point at another GPS station PUH2 located at the western coastline (locations of the used GPS stations are presented in the leftmost subplot of Figure 7a). We take the GPS based displacements at the SAR acquisitions as real values, which are calculated using the daily GPS solutions from Nevada Geodetic Laboratory, by averaging six solutions that are the closest to the SAR acquisitions, and we calculate the differences that related to 197 SAR acquisitions between InSAR and GPS.

The results at the first GPS station MKEA located at Mauna Kea (Figure 9a) show that all of the time-series present similar secular trends (i.e., deformation velocity) of around zero, which is consistent with the stable geological environment at Mauna Kea volcano, but the 5 InSAR time-series have significantly different deviations (compared with GPS), which mainly is related to the temporal variations in the tropospheric delays. Histograms of the differences between InSAR and GPS (the right subplot in Figure 9a) show that, before tropospheric corrections, the standard deviation (STD) is 5.04 cm. In comparison, the STDs decrease to 3.52 cm (PyAPS), 3.41 cm (d-LOS), 3.27 cm (GACOS), and 2.21 cm (new

method), respectively, after the corrections. This indicates 1) all of the four methods can efficiently mitigate the tropospheric delays in the time-series at station MKEA, and 2) our new method performs the best among the four corrections, yielding the largest improvement of around 56%.

The results at the second station ELEP located at Mauna Loa (Figure 9b) show that the deforming velocity from June 2015 to March 2020 was almost constant at the station and that the velocities derived from all of the time-series are very similar at approx. 3.0 cm/year, moving up towards the satellite. However, as at the first station, the InSAR derived time-series (before and after corrections) have different deviations relative to that of the GPS, and again, our new method shows the best performance in reducing the temporal tropospheric variations with the largest reduction of the STD among the four GAM corrections, i.e., a reduction of about 49%, from 4.39 cm in the uncorrected data to 2.23 cm.

In the raw InSAR time-series (without corrections) at both locations, there are significant seasonal variation signals, which are mainly related to seasonal changes in the topography-correlated (or stratified) delays. These seasonal signals are substantially reduced in all the four corrected time-series at the two sites, implying that all four GAM-methods can efficiently mitigate the stratified delays. The residuals in the corrected time-series are likely mainly caused by tropospheric turbulence, and thus, the smallest STDs of the time-series derived from our new correction method also indicates that it performs better in mitigating the tropospheric turbulence than the other three GAM-methods.

Figure 9. Time-series analysis of the ground displacements (related to the 197 SAR acquisitions from June 2015 to March 2020) at two GPS stations (MKEA and ELEM) referred to the station PUH2 (located at the western coastline): (a) MKEA located at Mauna Kea (stable) and (b) ELEM located at Mauna Loa (actively deforming). The black dots represent the original InSAR time-series displacements without troposphere correction, whereas the purple, green, blue, and red dots represent the displacement time-series after troposphere corrections from PyAPS, d-LOS, GACOS, and our new method, respectively. The gray dots represent the GPS daily solutions of the displacements projected into the radar LOS direction, and the red triangles represent the estimated GPS LOS displacements at the SAR acquisitions, by averaging 3 daily solutions before and after each SAR acquisition. The histograms and the standard deviations (σ) are calculated based on the differences between InSAR and GPS LOS displacements. The first point of each InSAR time series is vertically offset and made to start at 0.3, 0.6, 0.9, 1.2, and 1.5 m, respectively.

4. Discussion

4.1 Importance of considering horizontal heterogeneities of the tropospheric field

Estimating InSAR LOS delays using projected zenith-delays, e.g., $d_{LOS} = d_{zenith} / \cos(\theta_{inc})$, ignores horizontal heterogeneities in the tropospheric field and assumes that the troposphere is purely stratified or a plane-parallel refractivity medium (Williams et al., 1998). However, atmospheric turbulence, mainly in the lower troposphere (i.e., below ~5 km), causes horizontal tropospheric heterogeneities at a range of spatial scales. For a Sentinel-1 SAR image from 30 October 2015 (Fig. 4), for example, the variance of the water vapor

pressure between two points in the lower troposphere at an altitude of 400 m (Figure 4k) is 10^5 Pa^2 for distances of ~ 100 km, leading to a LOS tropospheric delay variance of about 10 cm^2 (Figure 4m). This implies that the standard deviation of the tropospheric delay difference between two points with distance of 10 km can be as large as 1 cm. Therefore, it is important to consider horizontal tropospheric heterogeneities for accurately estimating InSAR LOS delays.

Distances between the LOS and zenith directions at different altitudes are determined by the incidence angle. The greater the incidence angle is, the larger the distance and stronger the tropospheric heterogeneities between the two locations tend to be. In the above, the incidence angles of the Sentinel-1 data (Figure 10a) range from 32.6° in the east to 41.3° at the west coast of the island of Hawaii (Figure 10a). This means that for an effective tropospheric thickness of 5 km, the maximum distance between the LOS and zenith directions is 3.1-4.3 km (Figure 10b). These distances are larger than many small scales of the atmospheric turbulence, indicating that tropospheric parameters (e.g., partial pressure of water vapor) at the LOS location may be significantly different than those at the zenith location. For SAR data with larger incidence angles, the tropospheric heterogeneities along the LOS and zenith directions would be even larger. For example, the imaging incidence angle of ALOS-2 data can be as large as 70° , which means the maximum distance between the LOS and zenith directions would be up to 13.7 km for an effective tropospheric thickness of 5 km (see Figure 10b).

Figure 10. (a) The range of incidence angles for a Sentinel-1 image (Descending track 87) of the island of Hawaii. (b) Maximum distances between LOS and zenith locations on the ground for different incidence angles when assuming an effective tropospheric thickness of 5 km.

In our experiments, the d-LOS derived results turned out to be slightly better than those of PyAPS (see Figure 6, Table 1, and Figure 9). The d-LOS method calculates the integral of the InSAR delays along the LOS direction (Hu & Mallorquí, 2019) while PyAPS estimates them based on the projected zenith delays (Jolivet et al., 2014), but otherwise these two methods use the same interpolation strategy. Despite the different delay integral paths of the two methods, we found the d-LOS results only marginally better than the PyAPS results. This is probably due to the large spacing between output locations of the present ERA-5 model (~ 25 km) that is much larger than the distances between the LOS and zenith directions in lower troposphere (~ 5 km), meaning the zenith and LOS projections map pretty much the same tropospheric heterogeneities in the ERA5 output. However, as the spatial resolution of GAMs continues to improve, considering the tropospheric horizontal heterogeneities will become more and more important, because higher spatial resolution GAMs include more small-scale tropospheric variations.

4.2 Importance of incorporating spatial stochastic models for the interpolations

There are two steps in our new method that consider spatial stochastic models of the troposphere. The first step is when the LOS tropospheric parameters are derived from the raw GAM outputs, and the second step when dense LOS tropospheric delays are estimated from the coarse LOS delay samples of the first step. Based on the structure functions of the GAM outputs (e.g., partial pressure of water vapor, Figure 4k) and the tropospheric delays (see Figure 4m) at different altitude levels, we find that both are highly variable, especially in the lower troposphere (e.g., ~ 1 km), and that the variations at different altitude levels are also different. Therefore, interpolating GAM outputs or tropospheric delays using always the same weighting strategy, i.e., at all altitude levels and for the whole atmospheric environment, is far from ideal; it is clearly of advantage to consider spatial stochastic models for flexible determination of the weights in these estimations.

Both the d-LOS method and our new method calculate the tropospheric delays along the LOS direction, but the major difference between them is that the new method includes spatial stochastic models of the troposphere when interpolating the coarse GAM products. This has a dramatic effect, as our new method performs much better than the d-LOS method, both when correcting single interferograms and for time-series products. The average standard deviations of corrected interferograms derived from the d-LOS and our new method are 2.44 cm and 1.91 cm, respectively, which indicates an improvement of 21.7%. In addition, the number of improved interferograms derived using these two methods are 149 (d-LOS) and 243 (new method), out of a total of 280 interferograms, i.e., 87% of the interferograms were improved after using our new method, while only 53.2% of interferograms got better after the d-LOS correction.

In addition to the improvements of our new method outlined above, we also include a deterministic tropospheric-ramps in the interpolations, both when interpolating the atmospheric parameters (see Eq. (4)) and the tropospheric delays (see Eq. (8)). This is because tropospheric parameters and delays at different altitude levels usually exhibit a trend, which can easily be modeled using a polynomial (e.g., Eq. (5)) determined from the GAM outputs. When the tropospheric random field is dominated by a ramp, it is important to separate the estimation of the deterministic tropospheric-ramp and the interpolation of the turbulent component (or residual tropospheric component after removing the ramp), since a model-based estimation is usually better than an interpolation-based estimation. In addition, a tropospheric ramp can cause non-stationarity of the tropospheric random field as the local mean of tropospheric parameters (or delays) might differ significantly between two locations.

4.3 Limitations of correcting the smaller scales of the tropospheric turbulence

Compared with the general ground-based geodetic measurements of the troposphere, like GPS, one advantage of the GAM-based outputs is that we can obtain tropospheric predictions at different elevations, e.g., we can get tropospheric predictions at 37 different altitude levels from the raw ERA-5 data for a given point. Therefore, it is not difficult to predict (or model) topography-correlated (i.e., stratified) tropospheric delays based on the 3D GAM outputs, even with limited horizontal samples. In contrast, it is challenging to predict

the turbulent delays (or tropospheric turbulence), which are highly nonlinear and irregular in space. The interpolation precision of the tropospheric turbulence depends strongly on the spatial resolution of the GAM outputs, particular in the horizontal direction. According to the Nyquist-Shannon sampling theorem, the maximum recoverable signal frequency is half of the sampling frequency. Thus, the shortest wavelength of the ERA-5 recovered (or interpolated) signals from ERA-5 would be around 50 km for a given horizontal layer. Therefore, the present GAM-based correction methods are still limited for correcting the tropospheric turbulence, especially the smaller spatial scales of the turbulence.

While we have demonstrated the advantages of our new method over the island of Hawaii, where the tropospheric delays include both stratified and turbulent components, less improvements can be expected in areas where the tropospheric delays are dominated by the turbulent components, e.g., over flat terrains. Future GAMs with higher resolution outputs will help addressing this issue and lead to better InSAR results in general, also in areas without significant topographic relief.

5. Conclusions

We have here proposed a new and an advanced method to correct InSAR tropospheric delays that significantly improves on earlier GAM correction methods. The key advancements of the new method are that it 1) incorporates spatial stochastic models of the troposphere at different altitude levels for better weighting of GAM tropospheric samples, and 2) calculates the integral of the tropospheric delays directly along the satellite LOS direction, instead of using projected zenith delays. Comparisons between our new method and the other three previous GAM correction methods (PyAPS, d-LOS, and GACOS) show significant improvements for both single interferograms and time-series analysis products (deformation velocities and time-series solutions). The results demonstrate that: 1) it is important to consider stochastic tropospheric models in constructing high-resolution InSAR tropospheric delay maps from coarse GAM outputs, 2) GAM corrections have the potential of becoming a routine method for InSAR tropospheric corrections, and 3) it is essential to correct the InSAR tropospheric delays for mapping small magnitude or long-wavelength ground displacements.

752

753

754

755

756

757 Acknowledgements

758 This research was supported by King Abdullah University of Science and Technology
759 (KAUST), under award number BAS/1/1353-01-01, and the National Science Fund for
760 Distinguished Young Scholars (grant 41925016). The Sentinel-1 images are from the
761 European Space Agency (ESA) and downloaded from the Alaska Satellite Facility (ASF)
762 (<https://asf.alaska.edu/data-sets/sar-data-sets/sentinel-1/>), and the EAR5 data are provided by
763 ECMWF(<https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>). The GACOS
764 data were obtained from Newcastle University (<http://www.gacos.net/index.html>) and the
765 GPS time series from the Nevada Geodetic Laboratory
766 (http://geodesy.unr.edu/gps_timeseries).
767

768 References

- 769 Amelung, F., Yun, S.-H., Walter, T. R., Segall, P., & Kim, S.-W. (2007). Stress Control of
770 Deep Rift Intrusion at Mauna Loa Volcano, Hawaii. *Science*, 316, 1026-1030.
- 771 Antonia, R. A., Chambers, A. J., Friehe, C. A., & Van Atta, C. W. (1979). Temperature
772 ramps in the atmospheric surface layer. *Journal of the Atmospheric Sciences*, 36(1),
773 99-108.
- 774 Baby, H., Gole, P., & J.A., L. (1988). Model for tropospheric excess path length of radio
775 waves from surface meteorological measurements. *Radio Science*, 22, 1023-1038.
- 776 Baran, I., Stewart, M. P., Kampes, B. M., Perski, Z., & Lilly, P. (2003). A modification to the
777 Goldstein radar interferogram filter. *IEEE TRANSACTIONS ON GEOSCIENCE AND*
778 *REMOTE SENSING*, 41(9), 2114-2118.
- 779 Bawden, G. W., Thatcher, W., Stein, R. S., Hudnut, K. W., & Peltzer, G. (2001). Tectonic
780 contraction across Los Angeles after removal of groundwater pumping effects. *Nature*,
781 412, 812-815.
- 782 Bekaert, D. P. S., Hooper, A., & Wright, T. J. (2015). A spatially variable power law
783 tropospheric correction technique for InSAR data. *Journal of Geophysical Research:*
784 *Solid Earth*, 120(2), 1345-1356.
- 785 Bekaert, D. P. S., Walters, R. J., Wright, T. J., Hooper, A. J., & Parker, D. J. (2015).
786 Statistical comparison of InSAR tropospheric correction techniques. *Remote Sensing*
787 *of Environment*, 170, 40-47.

- Bell, J. W., Amelung, F., & Henry, C. D. (2012). InSAR analysis of the 2008 Reno-Mogul earthquake swarm: Evidence for westward migration of Walker Lane style dextral faulting. *Geophysical Research Letters*, 39(18).
- Berardino, P., Fornaro, G., Lanari, R., & Sansosti, E. (2002). A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 40(11), 2375-2383.
- Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., & Ware, R. H. (1992). GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. *Journal of Geophysical Research: Atmospheres*, 97(D14), 15787-15801.
- Cao, Y. (2019). PyINT: Python&GAMMA based interferometry toolbox. *Remote Sensing Code Library*.
- Cao, Y., Li, Z., & Amelung, F. (2019). Mapping ground displacement by a multiple phase difference-based InSAR approach: with stochastic model estimation and turbulent troposphere mitigation. *Journal of Geodesy*.
- Cao, Y., Li, Z., Wei, J., Hu, J., Duan, M., & Feng, G. (2017). Stochastic modeling for time series InSAR: with emphasis on atmospheric effects. *Journal of Geodesy*, 92(2), 185-204.
- Cavalié, O., & Jónsson, S. (2014). Block-like plate movements in eastern Anatolia observed by InSAR. *Geophysical Research Letters*, 41(1), 26-31.
- Cressie, N. (1990). The origins of kriging. *Mathematical Geology*, 22(3), 239-252.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553-597.
- Delacourt, C., Briole, P., & Achache, J. (1998). Tropospheric corrections of SAR interferograms with strong topography: Application to Etna. *Geophysical Research Letters*, 25(15), 2849-2852.
- Doin, M. P., Lasserre, C., Peltzer, G., Cavalié, O., & Doubre, C. (2009). Corrections of stratified tropospheric delays in SAR interferometry: Validation with global atmospheric models. *Journal of Applied Geophysics*, 69(1), 35-50.
- Elliott, J. R., Biggs, J., Parsons, B., & Wright, T. J. (2008). InSAR slip rate determination on the Altyn Tagh Fault, northern Tibet, in the presence of topographically correlated atmospheric delays. *Geophysical Research Letters*, 35(12), n/a-n/a.
- Farquharson, J. I., & Amelung, F. (2020). Extreme rainfall triggered the 2018 rift eruption at Kilauea Volcano. *Nature*, 580(7804), 491-495.
- Feng, G., Hetland, E. A., Ding, X., Li, Z., & Zhang, L. (2010). Coseismic fault slip of the 2008 Mw7.9 Wenchuan earthquake estimated from InSAR and GPS measurements. *Geophysical Research Letters*, 37(1), n/a-n/a.
- Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., & Rucci, A. (2011). A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 49(9), 3460-3470.
- Ferretti, A., Prati, C., & Rocca, F. (2001). Permanent Scatterers in SAR Interferometry. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 39(1), 8-20.
- Foster, J., Brooks, B., Cherubini, T., Shacat, C., Businger, S., & Werner, C. L. (2006). Mitigating atmospheric noise for InSAR using a high resolution weather model. *Geophysical Research Letters*, 33(16).

- Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J Clim*, Volume 30(Iss 13), 5419-5454.
- Haiden, T., Janousek, M., Bidlot, J., Buizza, R., Ferranti, L., Prates, F., & Vitart, F. (2018). Evaluation of ECMWF forecasts, including the 2018 upgrade.
- Hanssen, R. F. (2001). RADAR INTERFEROMETRY: Data Interpretation and Error Analysis.
- Hooper, A. (2008). A multi-temporal InSAR method incorporating both persistent scatterer and small baseline approaches. *Geophysical Research Letters*, 35(16).
- Hobiger, T., Kinoshita, Y., Shimizu, S., Ichikawa, R., Furuya, M., Kondo, T., & Koyama, Y. (2010). On the importance of accurately ray-traced troposphere corrections for Interferometric SAR data. *Journal of Geodesy*, 84(9), 537-546
- Houlié, N., Funning, G. J., & Burgmann, R. (2016). Use of a GPS-Derived Troposphere Model to Improve InSAR Deformation Estimates in the San Gabriel Valley, California. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 54(9), 5365-5374.
- Hu, Z. B., & Mallorquí, J. J. (2019). An Accurate Method to Correct Atmospheric Phase Delay for InSAR with the ERA5 Global Atmospheric Model. *Remote Sensing*, 11(17).
- Jolivet, R., Agram, P. S., Lin, N. Y., Simons, M., Doin, M.-P., Peltzer, G., & Li, Z. (2014). Improving InSAR geodesy using Global Atmospheric Models. *Journal of Geophysical Research: Solid Earth*, 119(3), 18.
- Jolivet, R., Grandin, R., Lasserre, C., Doin, M. P., & Peltzer, G. (2011). Systematic InSAR tropospheric phase delay corrections from global meteorological reanalysis data. *Geophysical Research Letters*, 38(17), n/a-n/a.
- Jónsson, S. (2008). Importance of post-seismic viscous relaxation in southern Iceland. *Nature Geoscience*, 1(2), 136-139.
- Jónsson, S., Segall, P., Pedersen, R., & Björnsson, G. (2003). Post-earthquake ground movements correlated to pore-pressure transients. *Nature*, 424, 179-183.
- Kaymaz, I. (2005). Application of kriging method to structural reliability problems. *Structural Safety*, 27(2), 133-151.
- Li, Z. (2005). Interferometric synthetic aperture radar (InSAR) atmospheric correction: GPS, Moderate Resolution Imaging Spectroradiometer (MODIS), and InSAR integration. *Journal of Geophysical Research*, 110(B3).
- Li, Z., Cao, Y., Wei, J., Duan, M., Wu, L., Hou, J., & Zhu, J. (2019). Time-series InSAR ground deformation monitoring: Atmospheric delay modeling and estimating. *Earth-Science Reviews*, 192, 258-284.
- Li, Z., Fielding, E. J., & Cross, P. (2009). Integration of InSAR Time-Series Analysis and Water-Vapor Correction for Mapping Postseismic Motion After the 2003 Bam (Iran) Earthquake. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 47(9), 3220-3230.
- Li, Z. W., Ding, X. L., & Liu, G. X. (2004). Modeling atmospheric effects on InSAR with meteorological and continuous GPS observations: algorithms and some test results. *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(11), 907-917.
- Li, Z. W., Xu, W. B., Feng, G. C., Hu, J., Wang, C. C., Ding, X. L., & Zhu, J. J. (2012). Correcting atmospheric effects on InSAR with MERIS water vapour data and elevation-dependent interpolation model. *Geophysical Journal International*, 189(2), 13.
- Liang, H., Zhang, L., Ding, X., Lu, Z., & Li, X. (2018). Toward Mitigating Stratified Tropospheric Delays in Multitemporal InSAR: A Quadtree Aided Joint Model. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 1-13.

- Lin, Y.-n. N., Simons, M., Hetland, E. A., Muse, P., & DiCaprio, C. (2010). A multiscale approach to estimating topographically correlated propagation delays in radar interferograms. *Geochemistry Geophysics Geosystems*, 11(9).
- Liu, P., Li, Q., Li, Z., Hoey, T., Liu, G., Wang, C., et al. (2016). Anatomy of Subsidence in Tianjin from Time Series InSAR. *Remote Sensing*, 8(3), 266.
- Liu, S., Hanssen, R., & Mika, Á. (2009). On the value of high-resolution weather models for atmospheric mitigation in SAR interferometry. *Geoscience and Remote Sensing Symposium, 2009 IEEE International, IGARSS 2009*, 2.
- Lundgren, P., Poland, M., Miklius, A., Orr, T., Yun, S. H., Fielding, E., et al. (2013). Evolution of dike opening during the March 2011 Kamoamo fissure eruption, Kīlauea Volcano, Hawaii. *Journal of Geophysical Research: Solid Earth*, 118(3), 897-914.
- Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K. L., & Rabaute, T. (1993). The displacement field of the Landers earthquake mapped by Radar interferometry. *Nature*, 364(6433), 138-142.
- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American Regional Reanalysis. *Bulletin of the American Meteorological Society*, 87(3), 343-360.
- Murray, K. D., Bekaert, D. P. S., & Lohman, R. B. (2019). Tropospheric corrections for InSAR: Statistical assessments and applications to the Central United States and Mexico. *Remote Sensing of Environment*, 232.
- Onn, F., & Zebker, H. A. (2006). Correction for interferometric synthetic aperture radar atmospheric phase artifacts using time series of zenith wet delay observations from a GPS network. *Journal of Geophysical Research*, 111(B9).
- Parker, A. L., Biggs, J., Walters, R. J., Ebmeier, S. K., Wright, T. J., Teanby, N. A., & Lu, Z. (2015). Systematic assessment of atmospheric uncertainties for InSAR data at volcanic arcs using large-scale atmospheric models: Application to the Cascade volcanoes, United States. *Remote Sensing of Environment*, 170, 102-114.
- Plattner, C., Amelung, F., Baker, S., Govers, R., & Poland, M. (2013). The role of viscous magma mush spreading in volcanic flank motion at Kīlauea Volcano, Hawai'i. *Journal of Geophysical Research: Solid Earth*, 118(5), 2474-2487.
- Rivera, A. M. M., Amelung, F., Mothes, P., Hong, S.-H., Nocquet, J.-M., & Jarrin, P. (2017). Ground deformation before the 2015 eruption of Cotopaxi volcano detected by InSAR. *Geophysical Research Letters*, 44, 6607-6615.
- Ruch, J., Wang, T., Xu, W., Hensch, M., & Jónsson, S. (2016). Oblique rift opening revealed by reoccurring magma injection in central Iceland. *Nature Communications*, 7(1).
- Sandwell, D. T., & Sichoix, L. (2000). Topographic phase recovery from stacked ERS interferometry and a low-resolution digital elevation model. *Journal of Geophysical Research: Solid Earth*, 105(B12), 28211-28222.
- Shen, L., Hooper, A., & Elliott, J. (2019). A Spatially Varying Scaling Method for InSAR Tropospheric Corrections Using a High-Resolution Weather Model. *Journal of Geophysical Research: Solid Earth*, 124(4), 4051-4068.
- Shirzaei, M., Bürgmann, R., Foster, J., Walter, T. R., & Brooks, B. A. (2013). Aseismic deformation across the Hilina fault system, Hawaii, revealed by wavelet analysis of InSAR and GPS time series. *Earth and Planetary Science Letters*, 376, 12-19.
- Siddique, M. A., Strozzi, T., Hajnsek, I., & Frey, O. (2018). A case study on the correction of atmospheric phases for SAR tomography in mountainous regions. *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), 416-431.

- Sigmundsson, F., Hreinsdottir, S., Hooper, A., Arnadottir, T., Pedersen, R., Roberts, M. J., et al. (2010). Intrusion triggering of the 2010 Eyjafjallajökull explosive eruption. *Nature*, 468(7322), 426-430.
- Simons, M., Fialko, Y., & Rivera, L. (2002). Coseismic deformation from the 1999 M w 7.1 Hector Mine, California, earthquake as inferred from InSAR and GPS observations. *Bulletin of the Seismological Society of America*, 92(4), 1390-1402.
- Van Zyl, J. J. (2001). The Shuttle Radar Topography Mission (SRTM): a breakthrough in remote sensing of topography. *Acta Astronautica*, 48(5-12), 559-565.
- Wadge, G., Webley, P. W., James, I. N., Bingley, R., Dodson, A., Waugh, S., et al. (2002). Atmospheric models, GPS and InSAR measurements of the tropospheric water vapour field over Mount Etna. *Geophysical Research Letters*, 29(19), 11-11-11-14.
- Walters, R. J., Elliott, J. R., Li, Z., & Parsons, B. (2013). Rapid strain accumulation on the Ashkabad fault (Turkmenistan) from atmosphere-corrected InSAR. *Journal of Geophysical Research: Solid Earth*, 118(7), 3674-3690.
- Webley, P. W., Bingley, R. M., Dodson, A. H., Wadge, G., Waugh, S. J., & James, I. N. (2002). Atmospheric water vapour correction to InSAR surface motion measurements on mountains: results from a dense GPS network on Mount Etna. *Physics and Chemistry of the Earth, Parts A/B/C*, 27(4-5), 363-370.
- Wegmüller, U., Werner, C. L., Strozzi, T., & Wiesmann, A. (2002). Processing strategies for phase unwrapping for INSAR applications. In *Proceedings of the European Conference on Synthetic Aperture Radar (EUSAR 2002)*, 1, 353-356.
- Wicks, C. W. (2002). Magmatic activity beneath the quiescent Three Sisters volcanic center, central Oregon Cascade Range, USA. *Geophysical Research Letters*, 29(7).
- Williams, S., Bock, Y. and Fang, P., (1998). Integrated satellite interferometry: Tropospheric noise, GPS estimates and implications for interferometric synthetic aperture radar products. *Journal of Geophysical Research: Solid Earth*, 103(B11), pp.27051-27067.
- Wright, T. J., Parsons, B., England, P. C., & Fielding, E. J. (2004). InSAR Observations of Low Slip Rates on the Major Faults of Western Tibet. *Science*, 305, 236-239.
- Xu, W. B., Li, Z. W., Ding, X. L., & Zhu, J. J. (2011). Interpolating atmospheric water vapor delay by incorporating terrain elevation information. *Journal of Geodesy*, 85(9), 555-564.
- Yang, Z., Li, Z., Zhu, J., Yi, H., Hu, J., & Feng, G. (2017). Deriving Dynamic Subsidence of Coal Mining Areas Using InSAR and Logistic Model. *Remote Sensing*, 9(2), 125.
- Yu, C., Li, Z. H., Penna, N. T., & Crippa, P. (2018). Generic Atmospheric Correction Model for Interferometric Synthetic Aperture Radar Observations. *Journal of Geophysical Research: Solid Earth*.
- Yunjun, Z., Fattahi, H., & Amelung, F. (2019). Small baseline InSAR time series analysis: Unwrapping error correction and noise reduction. *Computers & Geosciences*, 133, 104331.
- Zebker, H. A., Rosen, P. A., & Hensley, S. (1997). Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps. *Journal of Geophysical Research: Solid Earth*, 102(B4), 7547-7563.
- Zebker, H. Accuracy of a Model-Free Algorithm for Temporal InSAR Tropospheric Correction. Preprints 2020, 2020120526