STABILIZATION OF SENTINEL-1 SAR TIME-SERIES USING CLIMATE AND FOREST STRUCTURE DATA FOR EARLY TROPICAL DEFORESTATION DETECTION

J. Doblas^{1,*}, A. Carneiro¹, Y. Shimabukuro¹, S. Sant'Anna¹, L. Aragão¹, F.R.S. Pereira¹

¹ INPE, Remote Sensing Division, São José dos Campos, Brazil - (juan.doblas, arian.carneiro, yosio.shimabukuro, sidnei.santanna, luiz.aragao, francisca.pereira)@inpe.br

Commission III - ICWG III/7

KEY WORDS: Remote Sensing, Time-series Data, SAR, Modelling, Deforestation Detection, Change Detection

ABSTRACT:

In this study we analyse the factors of variability of Sentinel-1 C-band radar backscattering over tropical rainforests, and propose a method to reduce the effects of this variability on deforestation detection algorithms. To do so, we developed a random forest regression model that relates Sentinel-1 gamma nought values with local climatological data and forest structure information. The model was trained using long time-series of 26 relevant variables, sampled over 6 undisturbed tropical forests areas. The resulting model explained 71.64% and 73.28% of the SAR signal variability for VV and VH polarizations, respectively. Once the best model for every polarization was selected, it was used to stabilize extracted pixel-level data of forested and non-deforested areas, which resulted on a 10 to 14% reduction of time-series variability, in terms of standard deviation. Then a statistically robust deforestation detection algorithm was applied to the stabilized time-series. The results show that the proposed method reduced the rate of false positives on both polarizations, especially on VV (from 21% to 2%, α =0.01). Meanwhile, the omission errors increased on both polarizations (from 27% to 37% in VV and from 27% to 33% on VV, α =0.01). The proposed method yielded slightly better results when compared with an alternative state-of-the-art approach (spatial normalization).

1. INTRODUCTION

Weather and seasonal-related conditions of the surface can considerably affect SAR measurements modifying SAR timeseries characteristics (Benninga *et al.*, 2019). SAR variability can be either due to long-term factors, like seasonal droughts and rain scarcity and to short-term events, such as storm surges. Several authors have studied susceptibility of C-band measurements to dense rain cells (Atlas et al., 1993; Kasilingam *et al.*, 1997; Lin *et al.*, 1997), to intercepted precipitation water in the canopy (Dobson *et al.*, 1991; Henderson and Lewis, 1998; De Jong *et al.*, 2000; Cisneros Vaca and Van Der Tol, 2018), and to canopy humidity (see for example Quegan *et al.*, 2000).

Reportedly, C-band SAR backscattering can suffer attenuations between -2 and -2.4 dB when crossing dense storm cells (Moore *et al.*, 1997; Danklmayer *et al.*, 2009) and can increase 1 to 1.5 dB (Dobson *et al.*, 1991) due to intercepted rain. Benninga *et al.* (2019) found increases of 0.5 dB after medium to severe rain showers (1.8-4.5 mm/12 h). Seasonal changes on water content of the canopy can lead to oscillations of 2.5 dB (Quegan *et al.*, 2000, on ERS-C) to 1.5 dB (Benninga et al., 2019, on Sentinel-1) over mature temperate forests. Frolking *et al.* (2011) found a strong negative anomaly on SeaWinds active microwave Kuband backscatter data collected over the Amazon Basin during the 2005 drought and detected a striking correlation between water deficit measurements and Ku signal.

This reported instability can affect several kinds of techniques that make use of radar backscattering. In this work we will focus on deforestation detection on an Early Warning context. Early Warning Systems (EWS) are defined as a collection of algorithms and procedures able to identify tree loss or disturbance, on a periodic (monthly, weekly or daily) basis (Petersen *et al.*, 2018). EWS has been a crucial element to reinforce public policies that have led to significative deforestation rates decrease in Brazil (Soares-Filho et al., 2010; Assunção et al., 2013; Nepstad et al., 2014) and in Peru (Finer et al., 2018). A recent survey among users pointed to cloud cover as the most important effectiveness limiting factor to the actual EWS (Weisse et al., 2019). Hansen et al. (2016), reported 80% cloud cover during the wet season on Peru Landsat-7/8 data. In Brazil, frequent observations over Amazonian basin are seriously affected, as mean annual cloud cover on the Brazilian part of the biome is approximately 74%. This observational gap caused by cloud cover can be filled by orbital active microwave sensors, namely Synthetic Aperture Radar (SAR) satellites. However, the referred variability in backscattering time-series can mislead a potential automatic deforestation detection algorithm based on time-series analysis, decreasing its accuracy. Here we have reviewed some successful mitigation strategies: (1) Lin et al. (1997) were able to accurately model attenuation due to storm rain cells using meteorogical radar measurements and an attenuation model proposed by Atlas, Rosenfeld and Wolff (1993). (2) While searching for an optimal method to retrieve soil moisture using SAR data, Benninga et al. (2019) developed a complete set of rules and algorithms to deal with Sentinel-1 (S1) variability. Regarding seasonal trends, they proposed an 80-day window moving average computed over σ^0 values. Additionally, they propose a masking threshold based on the accumulated rain on the last 12 hours before SAR acquisition. (3) Reiche et al. (2017) proposed a different approach, using the regional backscatter P₉₅ percentile value to define a mean forest response for every recorded date and then normalizing all the image pixels by this value. This approach has an added advantage: it will deal with sensor oscillations as those reported by El Hajj et al. (2016). And (4) Reiche *et al.* (2018) fitted a harmonic function to S1 γ^{0}_{VH} time series to allow detrending of long-term variations.

Lin's approach, as any climatological modelling strategy dealing with rapid, small scale events such as storm surges should use climatological data with very high spatial and temporal

^{*} Corresponding author

resolution, such as those made available by meteorological radars. Unfortunately, this kind of data is far from being widely available over entire tropical basins. Regarding Benninga et al. approach, the suggested 80-day moving average is unfeasible on an EWS context, as we can't foresee future's SAR measurements. Reiche's spatial normalization seem very suited to EWS and would be used as a benchmark for our results. Reiche's harmonic fitting strategy can be very valuable (and, in fact, has inspired this work), but it can fail and produce significant errors as it doesn't account for inter-annual variability of rain and temperature regimes. As this, a harmonic fit trained on a humid year would probably overestimate backscatter values on the next, drier year. Climate change should potentialize this issue, as erratic rain patterns, altered dry seasons and extreme events gets more frequent (IPCC, 2014). Additionally, harmonic fit approach would not deal with short-term variations of backscattering due to heavy-rain episodes.

In this work we will investigate a novel approach to SAR variability mitigation. We will use precipitation, evapotranspiration, forest structure and drought indexes to model backscatter signal over intact forest tracks on the Amazonian basin, and then we will use this model to predict and stabilize SAR signal. The main advantage of a model built with such kind of data is that it won't depend on any optical measurement of land surface, and it will be sensible to seasonal changes and abrupt alterations of climate patterns.

2. MATERIALS AND METHODS

All data collection and pre-processing tasks were performed using the Google Earth Engine (GEE) platform (Gorelick *et al.*, 2017).

2.1 Study area and sampling space definition

The selected Area of Interest (AOI) covers a 67.000 km^2 tilted rectangle over the Brazilian Eastern Amazon (Figure 1). The area follows a ~500 km stretch of the Transamazonian highway (BR-230), between the cities of Altamira and Itaituba, on the Brazilian state of Pará.



Figure 1. Study area (red), located along the Transamazonian Highway (BR-230), between the cities of Altamira and Itaituba, on the Brazilian state of Pará. Background image: Google Inc.

The AOI belongs to a tropical region, characterized a wet season from November to May, followed by a drier (but not completely dry) season, from June to October. Total precipitation in Altamira varies around 1800 mm/yr. Regarding vegetation, dense ombrophylous forest covers all the intact vegetation areas. Flooded forests are not practically not present in the studied area. The main drive of deforestation in the AOI is land clearing for cattle-ranching in small to medium patches, corresponding to a low productivity, unsustainable production framed by land ownership concentration by largeholders (Brondizio *et al.*, 2009). Figure 2 shows a classic deforestation setup, with patches of forest being cleared on rectangular shapes next to previous deforested areas, along a secondary road.





Two different sampling sets were built for the purposes of this study (Figure 1): (1) An invariant forest data sample set, made of 6 hand-picked large patches of intact forest area, as assured by the PRODES/INPE records (INPE, 2019). The size of the intact forest patches ranges from 242 km² to 79 km², being the main patch size 125 km², large enough to diminish radiometric uncertainty and speckle effects when computing mean backscatter values (Benninga et al., 2019), and (2) a cleared forest data sample set, based on the DETER 2018 (INPE, 2019) deforestation data (August/17 to July/18 period). This last sampling set was built following a complex selection procedure: (i) The 2826 available polygons of the DETER/INPE 2018 dataset on the AOI were filtered to retain only clear-cutted areas with no previous degradation. (ii) The resulting 1203 polygon dataset was thinned, by removing all the areas which weren't observed, due to cloud cover, on the observation that predated the polygon creation date. This procedure, which retained 407 polygons, assured that the date of the polygon would be close to the date of the real deforestation; and (iii) the resulting polygons were intersected with the PRODES 2018 dataset. This final operation resulted on 198 fine-tuned, manually confirmed deforestation polygons, with a mean size of 0,92 km² and a total size of 182 km². These sets are illustrated in Figure 1 by yellow and black areas, respectively.

2.2 Input SAR data

The SAR data used on this study were delivered by ESA's Cband Sentinel-1A satellite. Sentinel-1 data are made available in the GEE platform as pre-processed, terrain-corrected σ^0 values, on VV and VH polarizations, and were converted to γ^0 values using the following expression (Woodhouse, 2006):

$$\gamma^0 = \frac{\sigma^0}{\cos \theta_i} \tag{1}$$

where θ_i represents the local incidence angle (LIA) and γ^0 , known as *gamma naught*, is the backscattering coefficient normalized by the incidence angle.

In order to reduce speckle noise, SAR data were filtered using a time-series filter (Quegan and Yu, 2001) and then a Frost 5x5 adaptative spatial filter (Frost *et al.*, 1982). The adopted spatial resolution of the Sentinel-1 pixel was 20 meters.

2.3 Climatologic and hydrological stress information

Daily precipitation data was obtained from the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) data set (Funk *et al.*, 2015). Hourly data, used to compute the 12 hours accumulated precipitation, was adquired from NASA's Integrated Multi-satellitE Retrievals for GPM (IMERG), on its version 06 (Huffman *et al.*, 2019). CHIRPS delivers global calibrated daily precipitation data on 0.05° (roughly 5x5 km) resolution. IMERG systems produce 0.1° (~10x10km) global precipitation data with a cadence of 30 minutes.

Water Deficit (WD) was computed as the accumulated negative difference between precipitation and evapotranspiration (ET), and represents a measure of the meteorologically-induced water stress (Aragão *et al.*, 2007). For every location and S1 acquisition date, WD was computed following the rule:

$$If WD_{i-1} + P_i - ET < 0$$

then $WD_i = WD_{i-1} + P_i - ET$
else $WD_i = 0$, $i \in \{1, ..., n\}$ (2)

where WD_i and P_i represent the n-sized 180-day daily time-series of WD and precipitation before the date of interest. For every location and acquisition date, WD will be considered as the WD_n value of the WD time-series computed following eq. 2. Daily ET was supposed to be constant and equal to the mean value of the PML-V2 (Zhang *et al.*, 2019) ET time series for every location.

Dry periods (DP) were computed by counting the number of days without a significative rain before the date of interest, for every location. Up to 5 thresholds were used to define what would be considered as a significative rain: 0, 1, 2, 5 and 10 mm·d⁻¹, generating 5 different indexes: DP0, DP1, DP2, DP5 and DP10.

2.4 Forest structure data

Above ground biomass density (AGBD) information was obtained from two different sources: (1) the 500 m. resolution WHRC Pantropical National Level Carbon Stock Dataset (Baccini et al., 2012), which was assembled from a combination of co-located field measurements, orbital LiDAR observations, and imagery recorded from the Moderate Resolution Imaging Spectroradiometer (MODIS), and (2) the 250 meter resolution EBA² dataset, computed using 836 LiDAR airborne transects randomly distributed across 3.5 million km² of the Amazon forests, and data from MODIS, SRTM, TRMM and ALOS-PALSAR datasets. It's worth noting that, although both datasets refer to the same variable, they can have significative differences, due to the different dates of reference (2010 for WHRC, 2016 for EBA), data, resolution and methods employed. Mean canopy height (CH) was extracted from the 2005 Global Forest Canopy Height data (Simard et al., 2011).

2.5 Backscatter modelling

To estimate backscattering on both VV and VH polarizations using the cited precipitation, drought and forest structures data we developed two different models. Firstly, we built a multivariate linear regression model. Due to the large number and strong collinearity of some of the predictors, an exploratory analysis was performed in order to reduce the number of variables and to evaluate eventual transformations to be performed on the variables before modelling. Non-normality of some of the selected variables lead us to build and use a second, non-parametric model using the Random Forest model (Breiman, 2001), which initially used the 26 computed variables (3 forest structure, 5 Dry Period, 16 precipitation indexes, plus Local Incidence Angle (LIA) and Water Deficit (WD). To avoid overfitting and computational oberburden we then reduced the number of predictor variables using the rfcv function of the *randomForest* R package. The training period ranged from November 2016 to November 2018. To measure the robustness of the computed models, we used them to predict backscatter values one year ahead the training period. We then measured the standard deviation and the range, in dB, associated to the original and predicted time-series, to check if the modelled backscattering values could contribute to decrease the variability of the S1 backscattering values.

2.6 Pixel-wise time-series stabilization and deforestation detection

Once an optimal model was defined for VV and VH backscattering, we used a GEE script to extract filtered backscatter values and all the selected climatological and structural indexes on a set of 2000 points inside the AOI, being half of them intact forest and the other half deforested areas. The sampling areas for each of the forest/non-forest categories were built following the procedure described in **2.1**. The sampled data were then injected into the optimized random forest VV and VH models to build a predicted backscattering series ($\hat{\gamma}^0$). Finally, the predicted values and the observed values were combined to form a stabilized time-series. The expression used to compute the *k* element of a *n* sized time-series was:

$$\hat{\gamma}_{st_{k}}^{0} = \gamma_{f_{k}}^{0} - \hat{\gamma}_{N_{k}}^{0}; \ k \in \{1, \dots, n\}$$
(3)

where $\hat{\gamma}_{st_k}^0$ is the stabilized backscattering value, $\gamma_{f_k}^0$ is the original filtered backscatter and $\hat{\gamma}_{N_k}^0 = \hat{\gamma}_k^0 - \frac{1}{n} \sum_{i=1}^n \hat{\gamma}_i^0$, is the mean-normalized predicted value.

After stabilization we applied a change detection procedure to the data (detailed below) and compare the results with the results of the same methodology applied to the original backscatter series. The adopted deforestation pixel-wise detection procedure follows the steps detailed below:

- Fitting of a statistical distribution for every time-series using a designated training period (in our case, Nov/16 to Jul/17.
- Determination of a change threshold based on the fitted distribution and a significance level (for this study we tested 1%, 2% and 5% significance levels).
- Thresholding of every pixel's time-series during the detection period (in our case, Aug/17 to Jul/18).
- A 'direct alert' will be flagged if a backscatter value falls below the threshold.
- Two consecutives 'direct alerts' will trigger a 'confirmed alert'.

3. RESULTS

3.1 Predictors characterization

Regarding statistical characteristics of the predictor variables, none of the considered rain and drought indexes can be considered as normally distributed (Figure 3). This fact will affect modelling results.

All the time-dependent studied variables (backscattering, precipitation, WD and DP) show a periodic behaviour (see for example, VH in Figure 4). Sentinel-1 γ^0 values show a minimum-maximum range of 1.08-1.53 dB, with a mean of 1.27 dB for VV

Biome- MSA/Amazon Fund -Subproject 7- Estimating Biomass in the Amazon (EBA), to be available on Pangaea.

² The AGB map of Amazon Brazilian Forest was developed by Environmental Monitoring via Satellite in The Amazon

and 1.01-1.38 dB, with a mean of 1.17 dB for VH polarizations, slightly more stable than the temperate forests studied by Benninga *et al.* (2019), which showed mean ranges around 1.5 dB for both polarizations.

3.2 SAR backscatter modelling

Although precipitation indexes showed a strong linear correlation with backscattering values (120-day accumulated rain had $\rho{=}0.67$ correlation with $\gamma^0{}_{VV}$, for example), and $\gamma^0{}_{VH}$ values had also significative linear correlation with short-term rain indexes (12 hours had a $\rho{=}0.49$), non-linearity of the main predictor variables led to a biased multivariate linear regression model. The results of the linear regression modelling lead us to test a non-parametric approach, more suited to deal with the non-normality of the used predictors.



Figure 3. Probability density of accumulated mean daily rain indexes. Rain[i]d stands for "Accumulated [i] days rain/[i]". Rain_12h stands for the accumulated rain before12 before SAR measurement



Figure 4. Periodicity of the mean SAR VH backscattering signal, of the selected forests areas. Every colour represents a different forest area (see Figure 1 and Figure 4 for area colors).

In this sense, we tested a Random Forest (RF) approach (Breiman, 2001), following the methodology detailed in 2.5. The optimal *mtry* parameter (or number of variables available for splitting at each tree node1) was determined by iterative modelling using R^2 as a measure of model quality and was fixed

to half of the used predictors. The optimal number of predictors was fixed to 22 and 25 for VH and VV respectively. Figure 5 lists the selected predictors and their relative importance for each polarization. The values of explained variance of the final models were 71.64% and 73.28% for VV and VH respectively.



Figure 5. Relative importance of the predictors used on the final random forests models for VH (left) and VV (right) backscattering. Biomass: WHRC AGB datatet. Biomass_inpe: AGB EBA dataset. LIA: Local Incidence Angle.

3.3 Forest SAR measurements stabilization

Once defined the final random forest models, we use them to predict the Sentinel-1 measures corresponding to a whole year *after* the training period, as a validation process. Figures 6 and 7 detail the results of the validation. The predicted series showed a good overall accord with the actual values, being able to follow the seasonal trends and eventually to model peaks of backscattering due to short-term variations of precipitation. Extreme backscattering values were not so faithfully modelled as the moderate ones, probably due to the difficulties that random forest regression algorithms have to predict values out of the range of the training dataset.





We then measured the variability of the stabilized time-series, that is, the variability of the model residuals, as a measure of the performance of the stabilization process (Table 2).

Measure	Mean	std	Range
Observed VV	2.48E-01	1.90E-02	9.28E-02
Predicted VV	2.42E-01	1.20E-02	5.28E-02
Stabilized VV	6.13E-03	1.15E-02	8.04E-02
Observed VH	5.58E-02	4.23E-03	2.25E-02
Predicted VH	5.79E-02	3.04E-03	1.38E-02
Stabilized VH	-2.05E-03	2.60E-03	1.70E-02

Table 2. Results of the stabilization of the validation dataset (time-series from 10/2018 to 10/2019). Units are linear.

Overall, there is a reduction of 39% of the variability of both VV and VH polarizations, in terms of standard deviation.



Figure 7. Predicted backscattering time-series on each studied forest tract for VV (top) and VH (bottom) backscattering. The dashed vertical line specifies the start of the validation period.

3.4 Application to deforestation detection on stabilized timeseries

To test our method on operational conditions, we applied the stabilization procedure detailed on section 2.6 to time-series extracted on 2000 points of the AOI, half of them being recently deforested areas and the other half intact forests. For benchmark purposes, we applied the same methodology using the $\gamma^0 P_{95}$ value of the 2 km neighbourhood of the pixel as the stabilization value, instead of the climate predicted value.

This last technique, called spatial normalization, was successfully used by Reiche *et al.* (2017) in Bolivian tropical forest to improve deforestation detection. We will use "SpNorm"

to refer to this methodology, and "ClSt" (Climate Stabilization), for the methodology proposed here.

Figure 8 shows the initial result of the climate stabilization procedure, showing the original time γ^0 series, the filtered values and the predicted values of 8 randomly chosen samples.

Overall, the stabilization procedure lowered the γ^0_{VH} standard deviation values from $6.86\cdot10^{-2}$ to $6.16\cdot10^{-2}$, and the γ^0_{VV} standard deviation values from 0.029 to 0.025, which represents a 10% and 14% reduction on VH and VV instability, respectively.



Figure 8: Original (dashed), filtered (black) and predicted (blue) γ^0_{VH} values on 8 random locations. Upper 2 row time-series correspond to intact forest areas. Lower 2 row time-series come from deforested areas.

After stabilization, we followed the procedure described in 2.6 to determine the accuracy of deforestation detection on the stabilized time-series. The first step (statistical modelling) yielded the following results:

- None of the pixel-based backscatter time-series, in linear or logarithmic scale, can be considered as normal, having Shapiro-Wilk normality test p-values consistently above 0.4.
- Kolmogorov-Smirnoff tests against gamma and lognormal distributions are substantially more significant for the linear-scaled filtered and stabilized time-series, with p-values around 0.08.
- As this, for every pixel, thresholds were fixed based on a fitted lognormal distribution, with levels of significance of 1%, 2% and 5%, based on the predetection period (10/2016 to 07/2017) time-series.

Once the thresholds were fixed, a simple detection procedure was performed, being the alert triggering criteria the occurrence of two consecutive values under the threshold.

Figure 9 depicts the deforestation detection procedure results on 8 random points. Note how point 841 deforestation wasn't detected, causing an omission error. Remark how, in point 48, the proposed methodology triggered an SAR alert (red line) more than two months before the optical alert date (in blue).



Treatment -- Observed data - Filtered - Filtered+Stabilized

Figure 9. Results of the deforestation detection procedure on 8 random locations. Upper 2 row time-series correspond to intact forest areas. Lower 2 row time-series come from deforested areas. Horizontal brown line represents the detection threshold for the stabilized series. The vertical red line is a confirmed alert. Blue lines show when the DETER optical detection system flagged the location as deforested.

The results of the detection procedure using observed, filtered, spatially normalized and climatic stabilized values are detailed in the following table, in terms of commission error (CE), omission error (OE) and weighted overall error (WOE). WOE was computed as:

$$WOE = \frac{\sqrt{(3 \cdot CE)^2 + OE^2}}{2}$$
(4)

The rationale behind WOE computation is to penalise the occurrence of commission errors, which are to be absolutely avoided on an operational EWS. For brevity purposes, only the results for a significance level of 1% are shown in Table 3. Higher significance values showed inferior results.

Polarization	Processing	CE (%)	OE (%)	WOE (%)
VH	D	4	55	28
VH	F	10	27	20
VH	F+SpNorm	1	38	19
VH	F+ClSt	2	33	17
VV	D	4	62	32
VV	F	21	27	34
VV	F+SpNorm	2	38	19
VV	F+ClSt	2	37	19

Table 3. Results of the deforestation detection procedure for a 0.01 significance. D: Original data. F: Filtered data. SpNorm: Spatial normalization. ClSt: Climate Stabilization. CE: Commission errors. OE: Omission errors. WOE: Weighted overall error.

The results of the stabilized series are clearly superior to the nonstabilized ones, showing that the use of raw or filtered S1 backscattering for deforestation detection is non-advisable. The bigger improvement occurs on the commission errors rate, that had reductions of 9% and 19% in VH and VV respectively.

Finally, we computed the delay of detection compared to the optical DETER system. We wanted to determine if the more frequent but cloud-impacted optical observations could overcome our proposed method, in terms of rapidity of response. The results for the nearly 1000 sampled points in that the mean difference for stabilized alerts is -11 days, and the median is 11 days (positive values imply fastest SAR detection) (figure 10). The benchmark method, spatial normalization, showed similar values: a mean of -12 days and a median delay advantage of 11 days. Thus, we can say that most of the SAR detections came before the optical ones.



(SAR-Optical system), in days. Positive values imply faster SAR detection. D: Original data. F: Filtered data. ClSt: Climate Stabilization.

4. DISCUSSION

The performance of random-forest regression, with R² values superior to 0.7, proves that non-parametric modelling is well suited to deal with the proposed problem. The computed precipitation indexes have a strong bimodality, and WD and DP variables have a complex distribution. In this particular situation, non-parametric models, such as the one used here, should overpass parametric models like linear regression. Random trees optimization did not reduce significantly the number of used variables (from 26 to 23 in the case of VH and 25 for VV), showing the complexity of the modelled system. Also, the random forest model had difficulties predicting backscattering values lower or higher than the ones used for training. Other types of algorithms, like support vector machines (SVM), should be tested against this specific issue.

The most important information for the models was the above ground biomass (AGB). The optimal random forest models used two different AGB datasets (WHRC and EBA), bringing information of both datasets, indicating that they respond for different forest structure characteristics, as pointed out by Longo et al. (2016). Other than AGB, long term precipitation and water deficit were useful on VV modelling. The 3 to 4 months periodicity of VV-used rain variables is consistent with the seasonality detected by Benninga et al. (2019). VH was more susceptible to short-term precipitation (12 hours accumulated rain) and local angle of incidence (LIA), indicating that intercepted rain possibly enhances depolarization of the radar signal thus increasing volume scattering. The overall evaluation of the random forest variable importance (Figure 5) suggests that co-polarized backscattering is more sensible to forest long-term phenological changes, represented by the long-term precipitation indexes, while cross-polarized backscattering variability is mostly dominated by short-term precipitation. Drought indexes, namely WD (water deficit) and DP (dry period) were important as well on both polarizations, confirming the influence of phenology cycles on SAR backscattering variation.

Regarding detection after stabilization results, the most striking result of the stabilization procedure is the reduction of the false deforestation alerts, especially on the VV polarization, that goes from 21% to 2% for a 0.01 level of significance. This decrease should be ascribed to the seasonality reduction after stabilization. Meanwhile, this improvement comes with a price, as omissions grew 5% and 10% for VH and VV, respectively. These increases can be explained by the reduction on backscattering range following the stabilization, especially during the dry season (as precipitation diminish, backscattering signal decrease, increasing γ^0 values after stabilization, and making slightly more difficult to detect deforestation during this period). Luckily, during the dry season most optical sensors are active, and the consequences of the increased error can be controlled using combined EWS information.

It's worth noting that the thresholds of deforestation detection were based on pixel-wise statistical modelling of the original and stabilized series over a short period of time (10 months, from the start of regular acquisition of Sentinel-1A satellite over the AOI until the start of the detection period). Longer series should increase performance of the detection.

Regarding to the state-of-the-art alternative approach to the problem (spatial normalization), we consider that both approaches are equally efficient, in terms of precision and computational costs. Spatial normalization will spend more effort looking for reference forest areas, while climatic normalization will be costlier during the model development phase.

5. CONCLUDING REMARKS

SAR measurements instability can have different origins, among others: electronic noise, ionospheric activity, tropospheric attenuation, presence of water droplets on the canopy, changes on water content and structure of the vegetation and soil moisture. In this work we have tried to model and reduce the variability related to short-term and seasonal precipitation and the lack of it, making use of globally available climatological and forest structure data.

Modelling results were satisfying and contributed to reduce the variability of the S1 time-series, thus improving the deforestation detection accuracy by drastically reducing the commission errors. Additionally, the results of the deforestation detection procedures indicate that deforestation detection using stabilized SAR time-series will anticipate the alert trigging time by approximately 11 days, if compared to the stablished manual, optical-based approach. This value will be significantly higher for deforestation produced at the beginning of the rainy season, that will remain undiscovered until the dry season, several months later, if only optical sensors are used for deforestation detection.

The results of our work suggest as well that filtering and stabilization of backscattering time-series are effective and should be a premise of any SAR-based deforestation detection method that seeks operational reliability.

Regarding future developments, longer SAR time-series, new forest structure measurements coming from the Global

Ecosystem Dynamics Investigation (GEDI) and refined algorithms should improve future models built upon the proposed methodology. More research is needed to determine the performance of the proposed climatological stabilization method in other regions of the Amazon basin having different rain regimes. Analogous modelling of non-forest land covers (grass, crops, deserts) can be an interesting way to gain insight on SAR backscattering variability.

ACKNOWLEDGEMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The first author would like to thank Dr. Guido Lemoine for his support on GEE coding.

REFERENCES

Aragão, L.E.O.C., Malhi, Y., Roman-Cuesta, R.M., Saatchi, S., Anderson, L.O., Shimabukuro, Y.E., 2007. Spatial patterns and fire response of recent Amazonian droughts. *Geophys. Res. Lett.* 34, 1–5. doi: 10.1029/2006GL028946.

Assunção, J., Gandour, C. and Rocha, R., 2013. DETERring Deforestation in the Brazilian Amazon. Climate Policy Initiative Report, pp. 1–36.

Atlas, D., Rosenfeld, D. and Wolff, D. B., 1993. C-band attenuation by tropical rainfall in Darwin, Australia, using climatologically tuned Ze-R relations. *Journal of Applied Meteorology*, 32(2), pp. 426–430. doi: 10.1175/1520-0450(1993)032.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S., Houghton, R.A., 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Chang.* 2, 182–185. doi: 10.1038/nclimate1354.

Benninga, H.-J. F., van der Velde, R. and Su, Z., 2019. Impacts of Radiometric Uncertainty and Weather-Related Surface Conditions on Soil Moisture Retrievals with Sentinel-1. *Remote Sensing*, 11(17), pp. 2025–2053. doi: 10.3390/rs11172025.

Breiman, L., 2001. Random forests. *Machine learning*, 45(1), pp. 5–32. doi: 10.1007/978-3-662-56776-0_10.

Brondizio, E.S., Cak, A., Caldas, M., Mena, C., Bilsborrow, R., Futemma, C.T., Moran, E.F., Batistella, M., Ludewigs, T., 2009. Small Farmers and Deforestation in Amazônia. *Geophys. Monogr.* Ser. 186 186, 117–143. doi: 10.1029/2008GM000716.

Cisneros Vaca, C., Van Der Tol, C., 2018. Sensitivity of Sentinel-1 to rain stored in temperate forest, in: International Geoscience and Remote Sensing Symposium (IGARSS). IGARSS, Valencia, Spain, pp. 5330–5333. doi: 10.1109/IGARSS.2018.8517859.

Danklmayer, A., Doring, B.R.J., Schwerdt, M., Chandra, M., 2009. Assessment of atmospheric propagation effects in SAR images. *IEEE Trans. Geosci. Remote Sens.* 47, 3507–3518. doi: 10.1109/TGRS.2009.2022271.

De Jong, J., Klaassen, W. and Ballast, A., 2000. Rain storage in forests detected with ERS tandem mission SAR. *Remote Sensing of Environment*, 72(2), pp. 170–180. doi: 10.1016/S0034-4257(99)00100-5.

Dobson, M.C., Pierce, L., McDonald, K., Sharik, T., 1991. Seasonal change in radar backscatter from mixed conifer and hardwood forests in northern Michigan. *Dig. - Int. Geosci. Remote Sens. Symp.* 3, 1121–1124. doi: 10.1109/igarss.1991.579268.

El Hajj, M., Baghdadi, N., Zribi, M., Angelliaume, S., 2016. Analysis of Sentinel-1 radiometric stability and quality for land surface applications. *Remote Sens.* 8. doi: 10.3390/rs8050406.

Finer, B.M., Novoa, S., Weisse, M.J., Petersen, R., Mascaro, J., Souto, T., Stearns, F., Martinez, R.G., 2018. Combating deforestation: From satellite to intervention. *Science* (80-.). 360, 1303–1305. doi: 10.1126/science.aat1203.

Frolking, S., Milliman, T., Palace, M., Wisser, D., Lammers, R., Fahnestock, M., 2011. Tropical forest backscatter anomaly evident in SeaWinds scatterometer morning overpass data during 2005 drought in Amazonia. *Remote Sens. Environ.* 115, 897–907. doi: 10.1016/j.rse.2010.11.017.

Frost, V.S., Stiles, J.A., Shanmugan, K.S., Holtzman, J.C., 1982. A Model for Radar Images and Its Application to Adaptive Digital Filtering of Multiplicative Noise. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-4, 157–166. doi: 10.1109/TPAMI.1982.4767223.

Funk, C. et al. (2015) 'The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes', *Scientific Data*, 2, pp. 1–21. doi: 10.1038/sdata.2015.66.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. doi: 10.1016/j.rse.2017.06.031.

Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P. V, Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., Moore, R., 2016. Humid tropical forest disturbance alerts using Landsat data. *Environ. Res. Lett.* 11, 034008. doi: 10.1088/1748-9326/11/3/034008.

Henderson, F. M. and Lewis, A. J., 1998. Principles and applications of imaging radar. Manual of remote sensing: Third edition, Volume 2. John Wiley and Sons, Inc., Somerset, NJ.

Huffman, G.J., Stocker, E.F., Bolvin, D.T., Nelkin, E.J., Tan, J., 2019. GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06 (GPM_3IMERGHH) at GES DISC. doi: 10.5067/GPM/IMERG/3B-MONTH/06.

INPE, 2019. Programa de monitoramento da Amazônia e demais biomas. Desmatamento – Amazônia Legal – Available in: http://terrabrasilis.dpi.inpe.br/downloads/.

IPCC, 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Kasilingam, D., Lin, I.I., Lim, H., Khoo, V., Alpers, W., Lim, T.K., 1997. Investigation of tropical rain cells with ERS SAR imagery and ground-based weather radar, in: Third ERS Symposium on Space at the service of our environment, vols. II & III, ESA special publications. pp. 1603–1608.

Lin, I.I., Kasilingam, D., Alpers, W., Lim, T.K., Lim, H., Khoo, V., 1997. A quantitative study of tropical rain cells from ERS SAR imagery, in: International Geoscience And Remote Sensing Symposium, Proceedings Vols I-IV: Remote Sensing - A Scientific Vision For Sustainable Development. pp. 1527–1529.

Moore, R.K., Mogili, A., Fang, Y., Beh, B., Ahamad, A., 1997. Rain measurement with SIR-C/X-SAR. *Remote Sens. Environ.* 59, 280–293. doi: 10.1016/S0034-4257(96)00147-2.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Da Motta, R.S., Armijo, E., Castello, L., Brando, P., Hansen, M.C., McGrath-Horn, M., Carvalho, O., Hess, L., 2014. Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science* (80-.). 344, 1118–1123. https://doi.org/10.1126/science.1248525

Petersen, R., Renschler, K. and Weisse, M., 2018. Early Warning Systems for Deforestation: An explainer. Available at: https://blog.globalforestwatch.org/data-and-research/early-warning-systems-for-deforestation-an-explainer.

Quegan, S., Toan, L., T., Y., J, J., Ribbes, F., Floury, N., 2000. Multitemporal ERS SAR analysis applied to forest mapping. *Remote Sens. Environ.* 38, 741–753. doi: 10.1109/36.842003.

Quegan, S., Yu, J.J., 2001. Filtering of multichannel SAR images. IEEE Trans. Geosci. Remote Sens. 39, 2373–2379. doi: 10.1109/36.964973.

Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., Herold, M., 2017. Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. *Remote Sens. Environ.* 204, 147–161. doi: 10.1016/j.rse.2017.10.034.

Reiche, J., Verhoeven, R., Verbesselt, J., Hamunyela, E., Wielaard, N., Herold, M., 2018. Characterizing tropical forest cover loss using dense Sentinel-1 data and active fire alerts. *Remote Sens.* 10, 1–18. doi: 10.3390/rs10050777.

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. *J. Geophys. Res. Biogeosciences* 116, 1–12. doi: 10.1029/2011JG001708.

Soares-Filho, B., Moutinho, P., Nepstad, D., Anderson, A., Rodrigues, H., Garcia, R., Dietzsch, L., Merry, F., Bowman, M., Hissa, L., Silvestrini, R., Maretti, C., 2010. Role of Brazilian Amazon protected areas in climate change mitigation. *Proc. Natl. Acad. Sci.* 107, 10821–10826. doi: 10.1073/pnas.0913048107.

Weisse, M.J., Noguerón, R., Eduardo, R., Vicencio, V., Castillo Soto, D.A., 2019. Use of near-real-time deforestation alerts: a case study from Peru. World Resources Institute. Washington, DC.

Woodhouse, I. H., 2006. Introduction to Microwave Remote Sensing. Boca Raton, Florida, USA: CRC Press.

Zhang, Y., Kong, D., Gan, R., Chiew, F.H.S., McVicar, T.R., Zhang, Q., Yang, Y., 2019. Coupled estimation of 500 m and 8day resolution global evapotranspiration and gross primary production in 2002–2017. Remote Sens. Environ. 222, 165–182. doi: 10.1016/j.rse.2018.12.031.