

Supplementary Materials: Integrated Analyses of PALSAR and Landsat Imagery Reveal More Agroforests in A Typical Agricultural Production Region, North China Plain. *Remote Sensing* 2018, 8, Article No. remotesensing-319045

Zhiqi Yang^{1,2,3}, Jinwei Dong^{1,*}, Yuanwei Qin⁴, Wenjian Ni³, Guosong Zhao¹, Wei Chen³, Bangqian Chen⁵, Weili Kou⁶, Jie Wang⁴ and Xiangming Xiao^{4,7,*}

1. Justification of NDVI Threshold for Separating Vegetation and Non-Vegetation

While the forest and building pixels can be extracted by using PALSAR data, here we used a threshold-based approach to remove the disturbances of buildings. We calculated the $NDVI_{max}$ histograms for the training ROIs of forests and buildings and PALSAR-based forest. When the value of $NDVI_{max}$ more than 0.55 (the top ceiling value of the building NDVI), the misclassified buildings can be removed from PALSAR-based forest map.

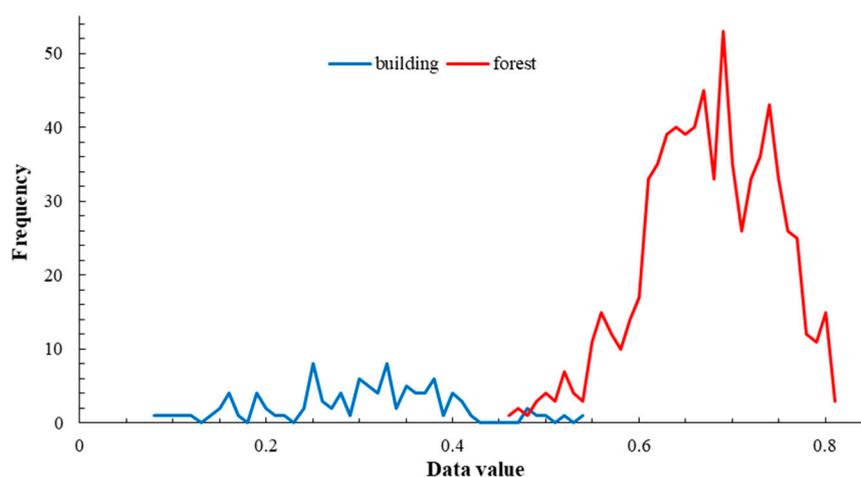


Figure S1. The $NDVI_{max}$ histogram of training ROIs of forests and buildings.

2. Mapping forest by fusing Landsat and PALSAR imagery

2.1. Image fusion

For PALSAR data, a radar image, the frost filtering and the terrain correction were conducted to reduce shadow of mountain and improve the problem of layover and foreshorten (Zhao *et al.* 2014). For Landsat data, atmospheric correction based on MOTRAN model and co-registration with PALSAR data were conducted for fusing different sensor datasets. After necessary preprocessing, we fused the mosaicked PALSAR data in 2010 (images from June to October mosaicked by JAXA) and the composited Landsat data in 2010 (median values of cloud-free observations during the peak growing season from June to October by using Google Earth Engine) based on an effective and classic fusion algorithm called High Pass Filter (HPF). Compared with other fusion algorithms for fusing PALSAR and Landsat datasets, this algorithm performs the best fused result, especially in the spatial information enhancement and spectral characteristics preservation (Zhao *et al.* 2014). The fused result was shown in Figure S2A.

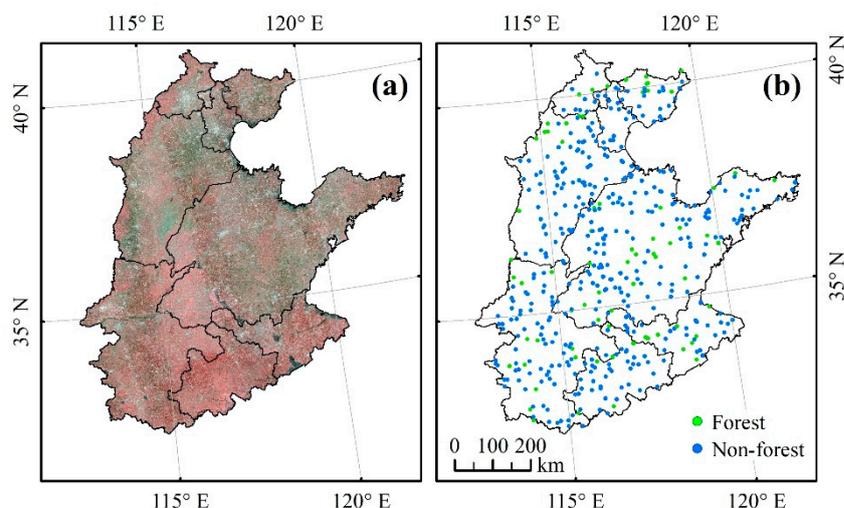


Figure S2. (a) RGB composite based on the fusion of Landsat and PALSAR imagery at data level in the North China Plain (NCP) in 2010 (R=Bnir, G=Bred, and B=Bgreen). (b) Spatial distribution of training ROIs in the North China Plain (NCP).

2.2. Regions of Interest (ROIs) for Algorithm Training

We randomly selected 650 sampling points for algorithm training on the fused image, each sampling point generated a buffer of 30 m radius circle as region of interest (ROI), and then the boundary of the ROI was overlapped on the fused image to obtain training samples. With reference to the very high resolution (VHR) remote sensing images of Google Earth in circa 2010, each ROI was labeled forest and non-forest. There were 240 forest sampling points with 810 pixels and 403 non-forest sampling points with 1335 pixels for training (Figure S2B).

2.3 Classification Method

Support Vector Machine (SVM), a non-parametric statistical learning technique as well as a large-margin classifier was used to classify forest and non-forest at the ENVI software. There are many advantages of SVM such as the stability of training parameters (Mountrakis *et al.* 2011) and the promising performances with limited number of samples (Foody and Mathur 2006).

2.4 Resultant Map and Its Validation

Through fusing the PALSAR and Landsat TM/ETM+ imagery at data level, we generated a new forest map (fusion data-based forest map) at the spatial resolution of 30-m in the NCP (Figure S3b). However, we found that the accuracy of that map is lower (85%) than the accuracy from this study (95%) (Table S1), by using the same validation samples described in the main text. That could be related to the time mismatch of the mosaicked PALSAR and Landsat images. A data fusion based on original PALSAR and Landsat images with closer dates could be more reasonable; however, it cannot be achieved due to the data unavailability of original PALSAR data and will be considered in our future study.

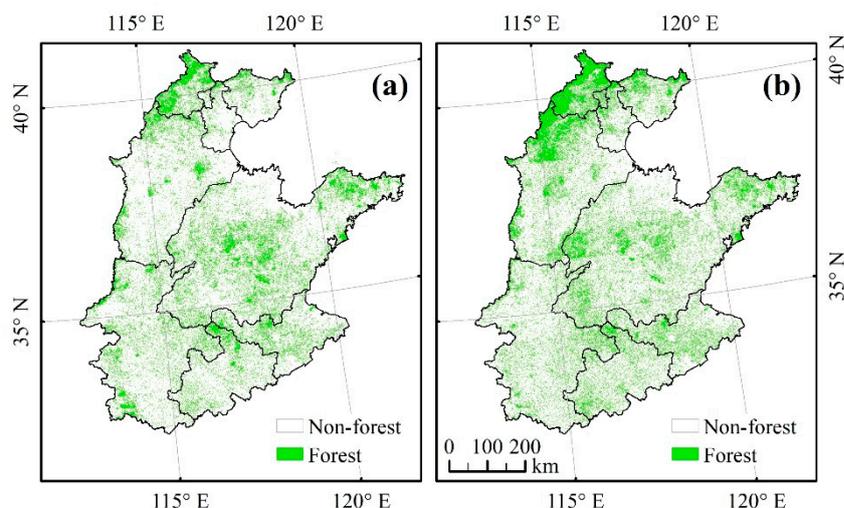


Figure S3. (a) Spatial distribution of the PL-based forest map, and (b) the fusion data-based forest map in the North China Plain (NCP) in 2010.

Table S1. Confusion matrix of accuracy assessments of the PL-based and the fusion data-based forest map forest maps.

Forest Products	Classes	GT Samples		UA	PA	OA	Kappa
		Forest	Non-Forest				
PL-based	forest	517	48	0.92	0.88	0.95	0.86
	non-forest	69	1536	0.96	0.97		
Fusion data-based forest map	forest	524	242	0.68	0.89	0.85	0.68
	non-forest	62	1342	0.96	0.85		

References

1. Foody, G.M.; Mathur, A. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sens. Environ.* **2006**, *103*, 179–189.
2. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259.
3. Zhao, P.; Liu, L.; Lu, D.; Du, H. A comparative analysis of data fusion techniques based on Landsat TM and ALOS PALSAR data. In Proceedings of the 2014 Third International Workshop on Earth Observation and Remote Sensing Applications (EORSA), Changsha, China, 11–14 June 2014; pp. 136–139.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).