An Automated Approach to Map Winter Cropped Area of Smallholder Farms across Large Scales Using MODIS Imagery

Meha Jain, Pinki Mondal, Gillian L. Galford, Greg Fiske and Ruth S. DeFries

Table S1. Timing of inflection points associated with start of the season green up dates for each of the eleven regions for which we have high-resolution data. We selected 7,500 random pixels from each of the eleven sites for which we had high-resolution validation data. We examined whether the inflection point associated with the start of the season occurred between May through July (monsoon season), August through September (early winter season), or October through April (the time window considered in this study). This gives an indication of which crop phenologies may be missing from our cropped area dataset across each region. Percent of the 7,500 randomly selected pixels that fall within each of the three categories is reported for each state.

State	Monsoon Season	Early Winter	Time Window of Study
Andhra Pradesh	95.24	0.73	4.03
Bihar	43.73	1.73	54.53
Gujarat	74.89	0.027	25.08
Haryana	16.83	3.15	80.03
Karnataka	95.27	0.41	4.32
Maharashtra	95.79	0	4.21
Punjab	0.43	0.03	99.55
Rajasthan	21.35	0.03	78.63
Uttar Pradesh 1	28.8	0.05	71.15
Uttar Pradesh 2	35.75	1.09	63.16
West Bengal	95.64	0.19	4.17

Table S2. Confusion Matrix of cropped versus uncropped pixels in our cropped area product and high-resolution imagery. We provide results when we restrict analyses to pixels where crops are detected in our cropped area product (crop = 1, shaded as gray) as well as results when we consider pixels that were considered to be uncropped in our winter cropped area product. We are primarily concerned with the accuracy of pixels found to be cropped in the cropped area product as there are several reasons why cropped area may have been detected in the high-resolution data but not in our cropped area product (e.g. phenology mismatch, Table S1). The larger the number in the column "Total", the more likely there is a mismatch in phenology between our cropped area product and the crops that appear to be planted in the high-resolution dataset.

			High resolution Validation			
			Uncropped	Croppped	Total	Producer's Accuracy (%)
Cropped Area Product all pixels	Rajasthan	Uncropped	1	47	48	2.08
		Cropped	0	332	332	100
	Punjab	Uncropped	0	0	0	100
		Cropped	0	363	363	100
	Uttar Pradesh 1	Uncropped	0	18	18	0
		Cropped	1	270	271	99.63
	Uttar Pradesh 2	Uncropped	0	2	2	0
		Cropped	0	160	160	100
	Bihar	Uncropped	0	0	0	100
		Cropped	0	290	290	100
	Haryana	Uncropped	0	7	7	0
		Cropped	0	250	250	100
	Andhra Pradesh	Uncropped	0	186	186	0
		Cropped	0	62	62	100
	Karnataka	Uncropped	1	277	278	0.36
		Cropped	0	82	82	100
	Gujarat	Uncropped	0	160	160	0
		Cropped	0	283	283	100
	Maharashtra	Uncropped	0	124	124	0
		Cropped	0	17	17	100
	West Bengal	Uncropped	1	341	342	0.29
		Cropped	0	35	35	100

Table S3. R^2 and RMSE values for eleven high-resolution validation sites when implementing the MODIS Peak method without scaling developed in previous studies. The year and satellite for each high-resolution image, the number of points used for validation, and whether the phenology used in our peak-detecting algorithm matched the phenology of cropped pixels seen in the high-resolution imagery is also listed.

Location	Image Date	Satellite	MSA R^2	MSA RMSE	Phenology Match
Rajasthan	1-12-16	RapidEye	0.64	40.40	Yes
Punjab	1-31-11	WorldView-2	0.23	19.41	Yes
Uttar Pradesh 1	2-17-15	RapidEye	0.58	27.75	Yes
Uttar Pradesh 2	2-15-15	RapidEye	0.55	31.40	Yes
Bihar	2-11-16	SkySat	0.50	33.25	Yes
Haryana	1-18-15	RapidEye	0.21	34.44	Yes
Andhra Pradesh	2-25-15	RapidEye	0.52	26.63	No
Karnataka	2-18-15	RapidEye	0.53	27.17	No
Gujarat	1-18-10	WorldView-2	0.44	22.63	Yes
Maharashtra	2-17-15	RapidEye	0.25	24.54	No
West Bengal	1-12-08	Quickbird	0.10	23.80	No
All 11 Sites	-	-	0.44	30.08	-
Sites with Phenology Match	-	-	0.49	30.45	-

Figure S1. Available Landsat imagery from 2000–2001 to 2015–2016. Number of years for which at least one cloud-free pixel is available during the main period of the winter growing season, January 1 to March 31. Pixels range from those where no imagery were available (blue) across the 16 years considered in our study (2000-2001 to 2015-2016) to pixels where imagery were available for all 16 years (red). A large proportion of the country has at least one missing year across time, highlighting the difficulty in relying on Landsat imagery to produce annual cropped area maps for much of the country.

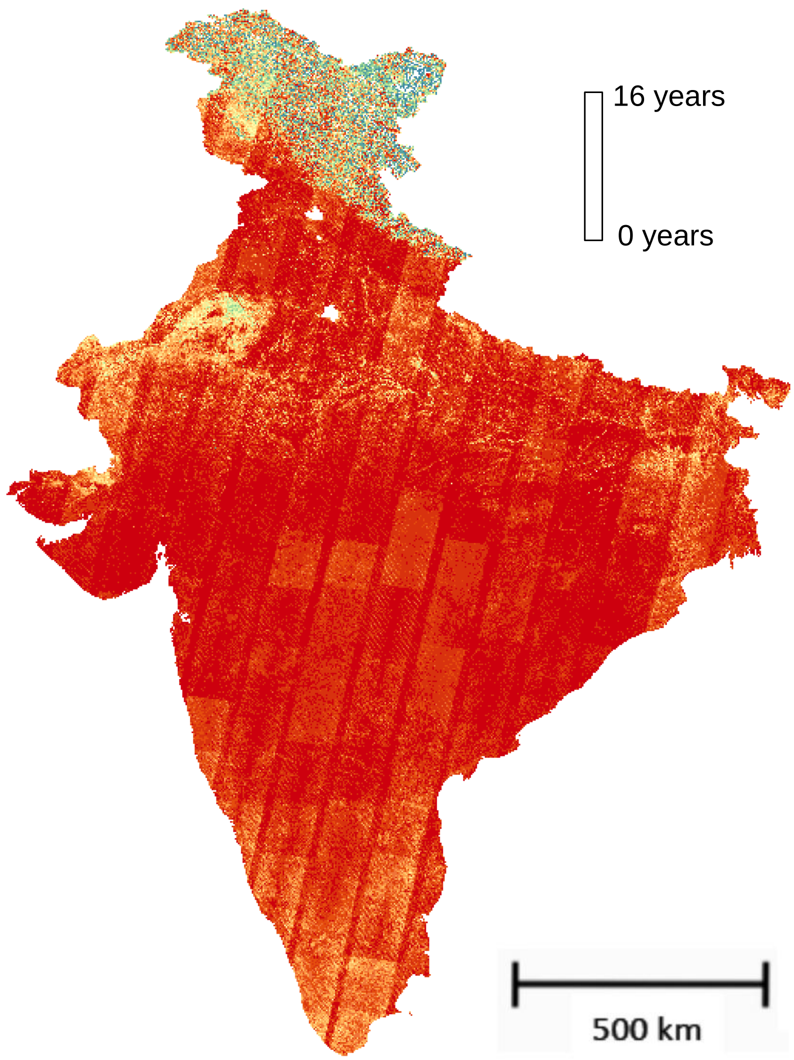

Figure S2. Map of eleven high-resolution validation sites. Footprints of each of the eleven high-resolution images used for validation are highlighted in black.

Figure S3. Scatterplots of our MSA-derived estimates of cropped area versus high-resolution estimates of cropped area at a 1 × 1 km² across eleven validation sites. R^2 for each region is located within each plot and the one-to-one line is represented as a dashed line.