PNAS
www.pnas.org

Supplementary Information for

Near doubling of Brazil's intensive row crop area since 2000

Viviana Zalles[1], Matthew C. Hansen[1], Peter P. Potapov[1], Stephen V. Stehman[2], Alexandra Tyukavina[1], Amy Pickens[1], Xiao-Peng Song[1], Bernard Adusei[1], Chima Okpa[1], Ricardo Aguilar[1], Nicholas John[1], and Selena Chavez[1]

Author affiliations:
[1] Department of Geographical Sciences, University of Maryland, College Park, MD, 20742, United States
[2] Department of Forest and Natural Resources Management, SUNY, Syracuse, NY 13210, United States

Corresponding author:
Viviana Zalles (vzalles@umd.edu)

This PDF file includes:

Supplementary text

Figs. S1 to S11

Tables S1 to S3

References for SI reference citations

Other supplementary materials for this manuscript include the following:

Dataset S1

www.pnas.org/cgi/doi/10.1073/pnas.1810301115

Classification results and accuracy assessment:

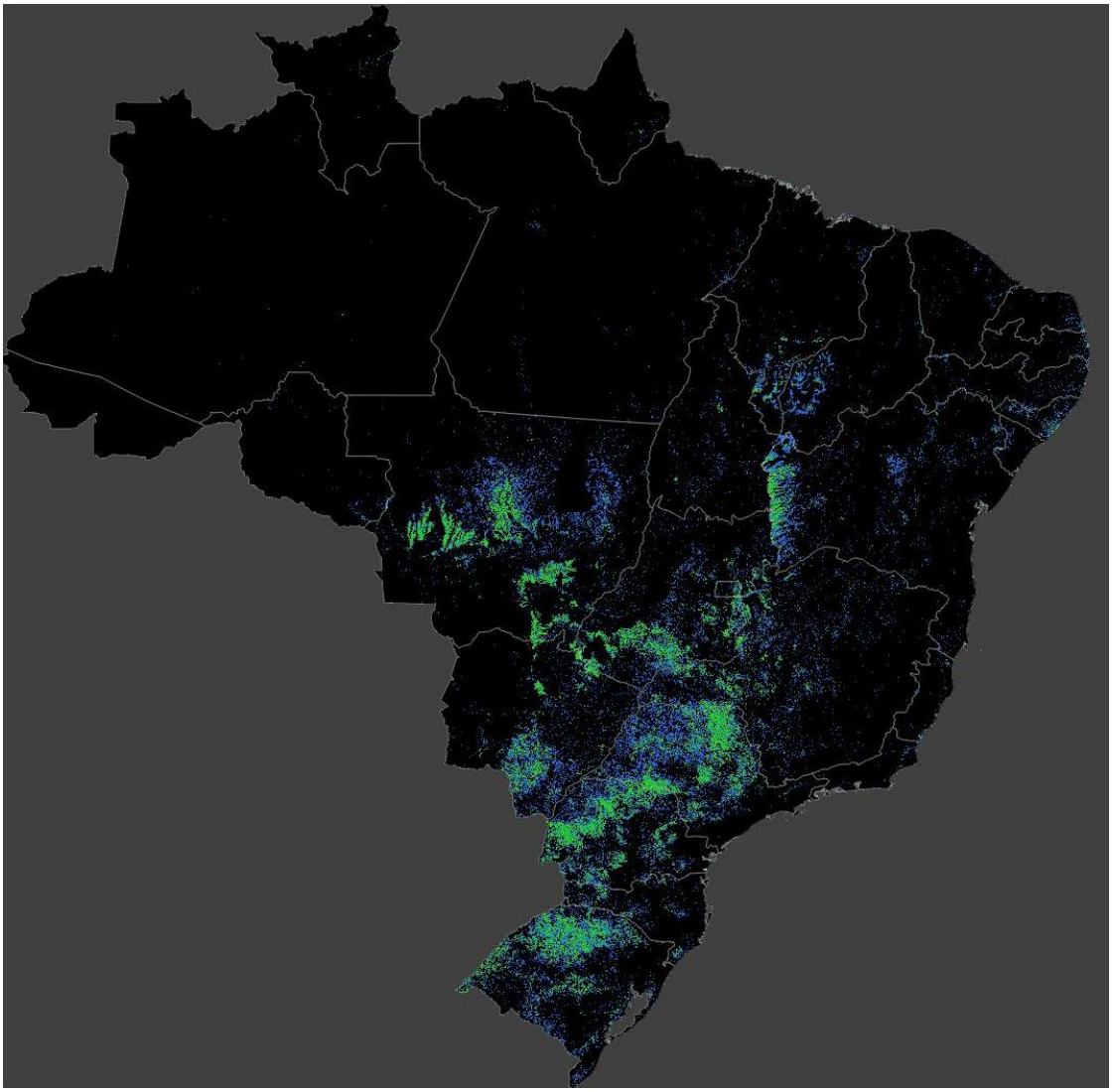
Classification results can be observed in Supplementary Fig. 1. Our sample-based accuracy assessment estimates overall map accuracy at 97.4% (SE=0.2%) (Supplementary Table 1). However, our map overestimates year 2000 cropland area, with user's accuracy of 89.9% (SE=3.5%), and producer's accuracy of 71.5% (SE=1.2%). User's and producer's accuracies of the cropland expansion class are more balanced – at 72.0% (SE=1.0%) and 71.2% (SE=5.3%) respectively. Map area of cropland expansion is 20.3 Mha, and sample-based area is 20.5 ± 1.6 Mha, demonstrating that the map does not under- or overestimate cropland expansion area.

Only 32 out of the 5000 sampled pixels were labeled as "cropland loss." All of these pixels had been mapped as cropland in the year 2000 and were counted as correctly classified pixels for the purpose of the confusion matrix.

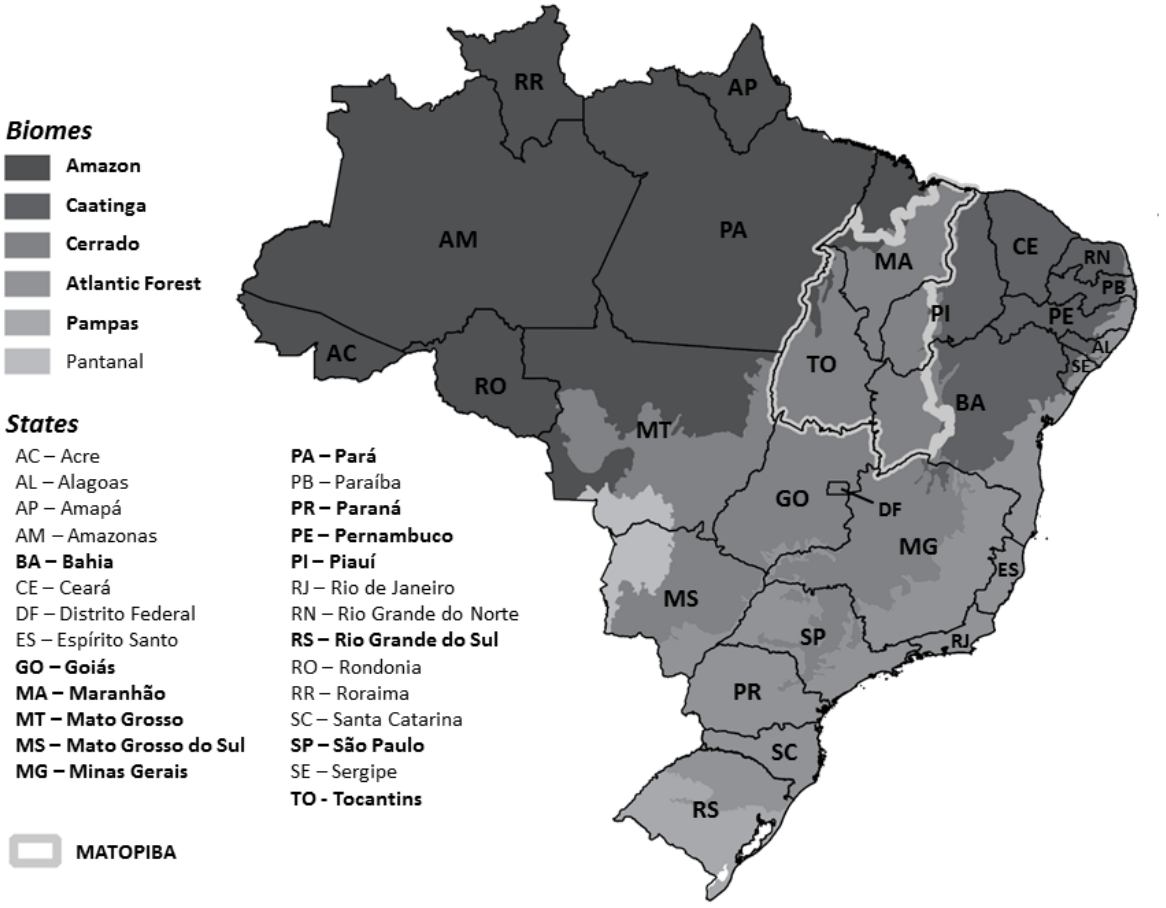
Despite the very large number of metrics used for each classification, the classification tree models decrease the majority of the training data's deviance based on a few select metrics. For the cropland 2000 classification, 74% of the overall deviance in the training data is decreased with only twenty of the >600 metrics used. The top metric, which corresponds to the average of the 75-90th percentile of the red band, an indicator of bare ground, provided 40% of the deviance decrease for this classification. For the cropland expansion classification, 78% of the deviance decrease is accounted by the top twenty metrics used by the classification tree, and 54% of the deviance decrease is explained by the top metric alone (the difference between the 2000 and the 2014 averages of the top 10% of SWIR2). In both cases, the classification tree captured cropland by targeting the spectral signatures of both high vegetation and bare ground, which is indicative of the agricultural cycle. Mapping of agricultural areas typically makes use of MODIS data because of its high temporal resolution, which provides NDVI time series needed for cropland characterization based on phenological responses. We get around the limitation of Landsat's lower temporal resolution by targeting the variation between the vegetated state and the unvegetated state typical of the agricultural cycle. -series metrics for cropland monitoring can be found for MODIS in (1–3), and for Landsat in (4–6).

Assessment of sample interpretations :

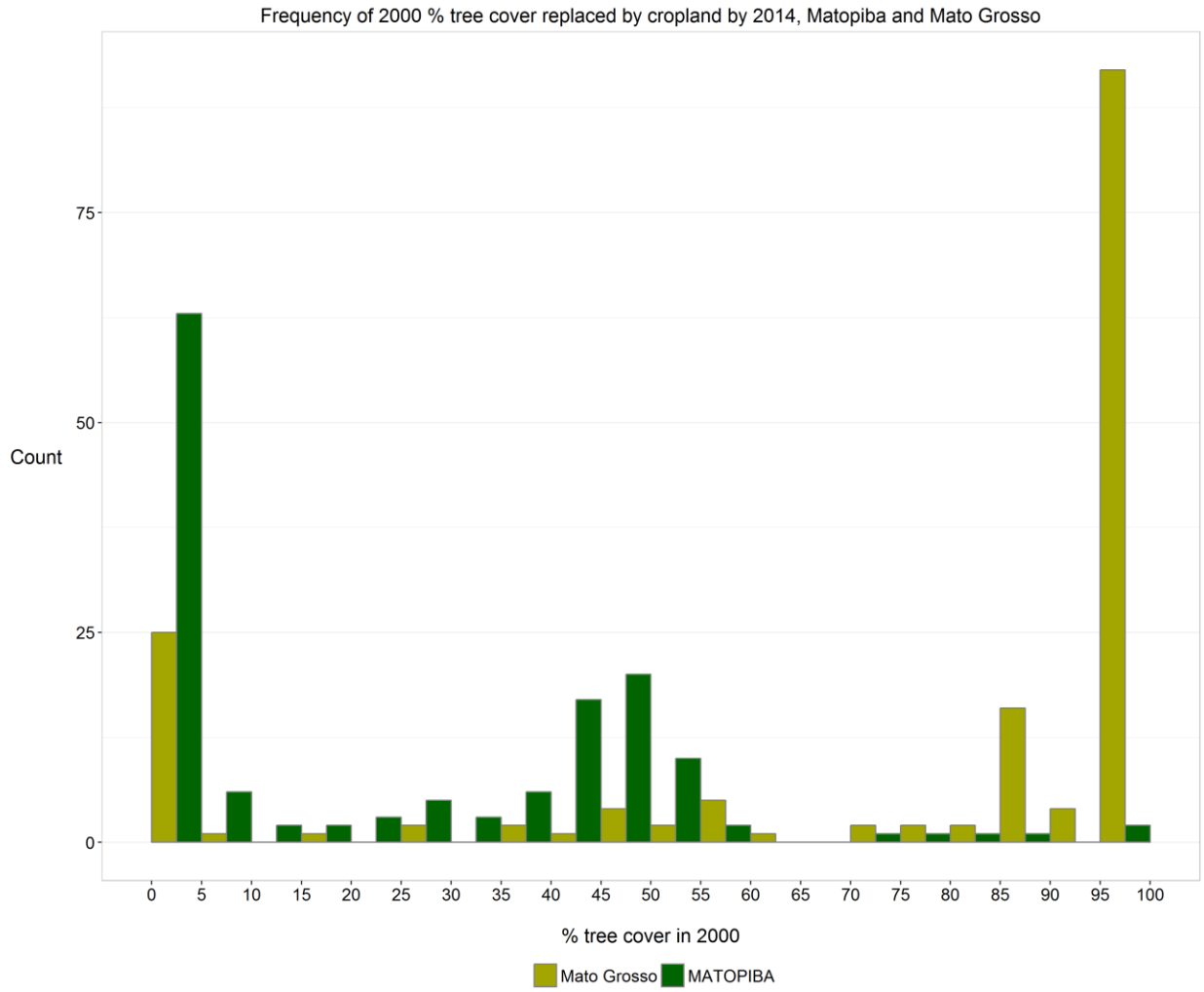
We were able to assess our sample interpretations against a dataset of field-verified samples. A stratified, two-stage cluster sampling design (consisting of 45 20x20 km blocks, each containing 20 randomly selected sample polygons corresponding to a Landsat pixel footprint) targeting soybean area was used to collect field data throughout Brazil in January 2017. We then used MODIS data to compare the spectral signatures of the field samples during the 2016/2017 growing season with those of the sample pixels we interpreted as cropland, pasture, and natural vegetation during the 2000/2001 growing season (SI Appendix, Fig. 11). Note that in addition to being from a different year, these data are not co-located. The plot illustrates how a competent interpreter can assign land cover and land use categories with spectral signatures consistent with those labeled in situ. The largest difference in distribution is in the cropland class. This is likely due to the fact that the 2016/2017 field-based random sample was stratified targeting soybean and not cropland, meaning most cropland field samples are likely to be soybean. However, interpreted samples from 2000/2001 are stratified on our cropland class which also includes sugarcane, cotton, rice and other crops that have spectral responses, particularly in the near-infrared, that are significantly lower than that of soybean.



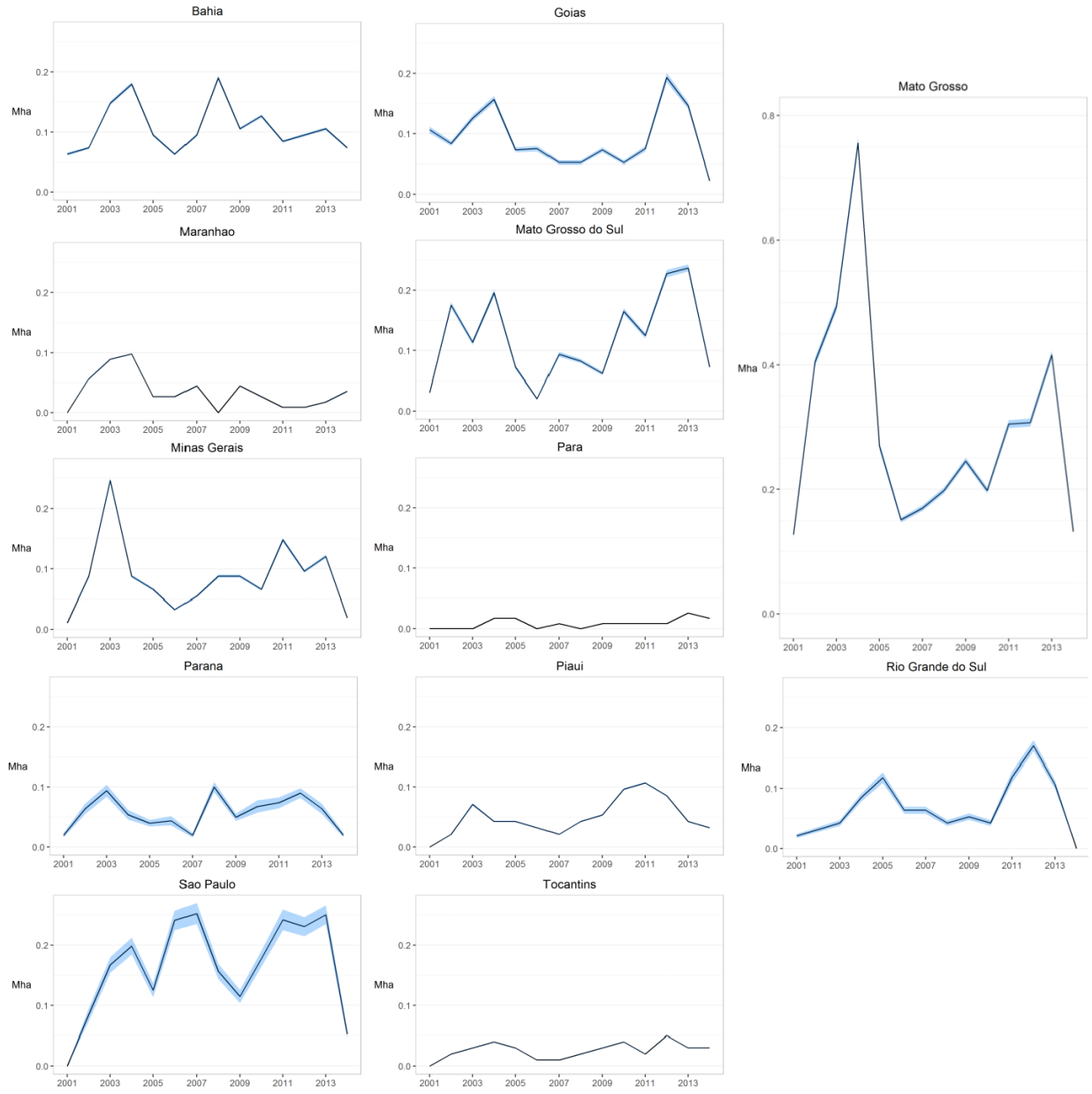
Supplementary Figure 1. Classification result. In green, cropland extent in the year 2000. In blue, cropland expansion through 2014.



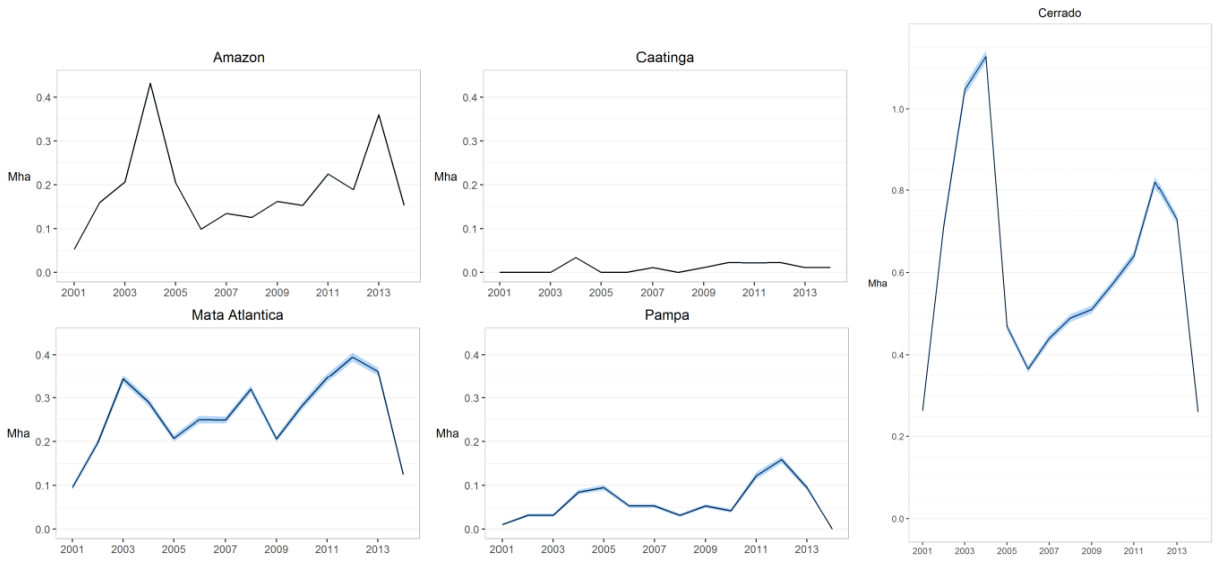
Supplementary Figure 2. **Brazilian states, biomes, and MATOPIBA.** States and biomes for which we are able to report sample-based area estimates are in bold.



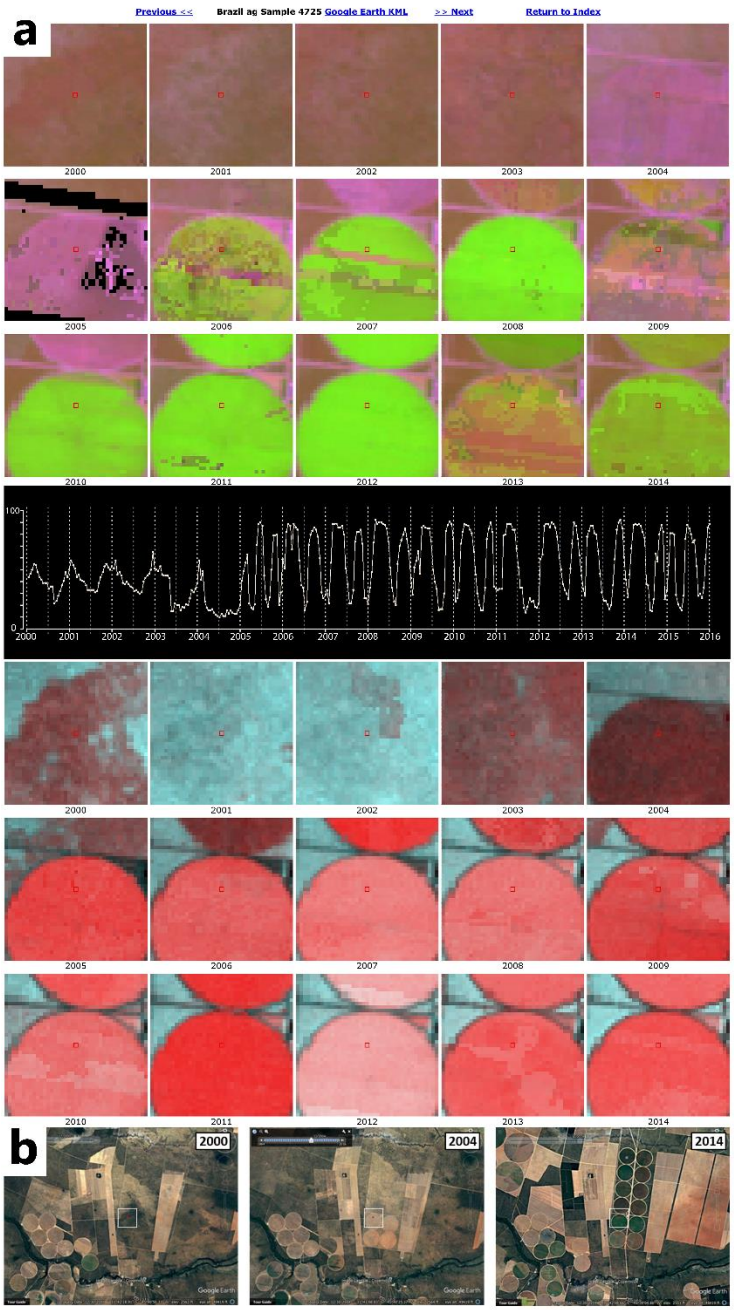
Supplementary Figure 3. **Antecedent tree cover for samples of naturally vegetated lands converted to cropland for Mato Grosso and Matopiba**. 2000 tree cover percent was obtained from the Global Forest Change maps(7).



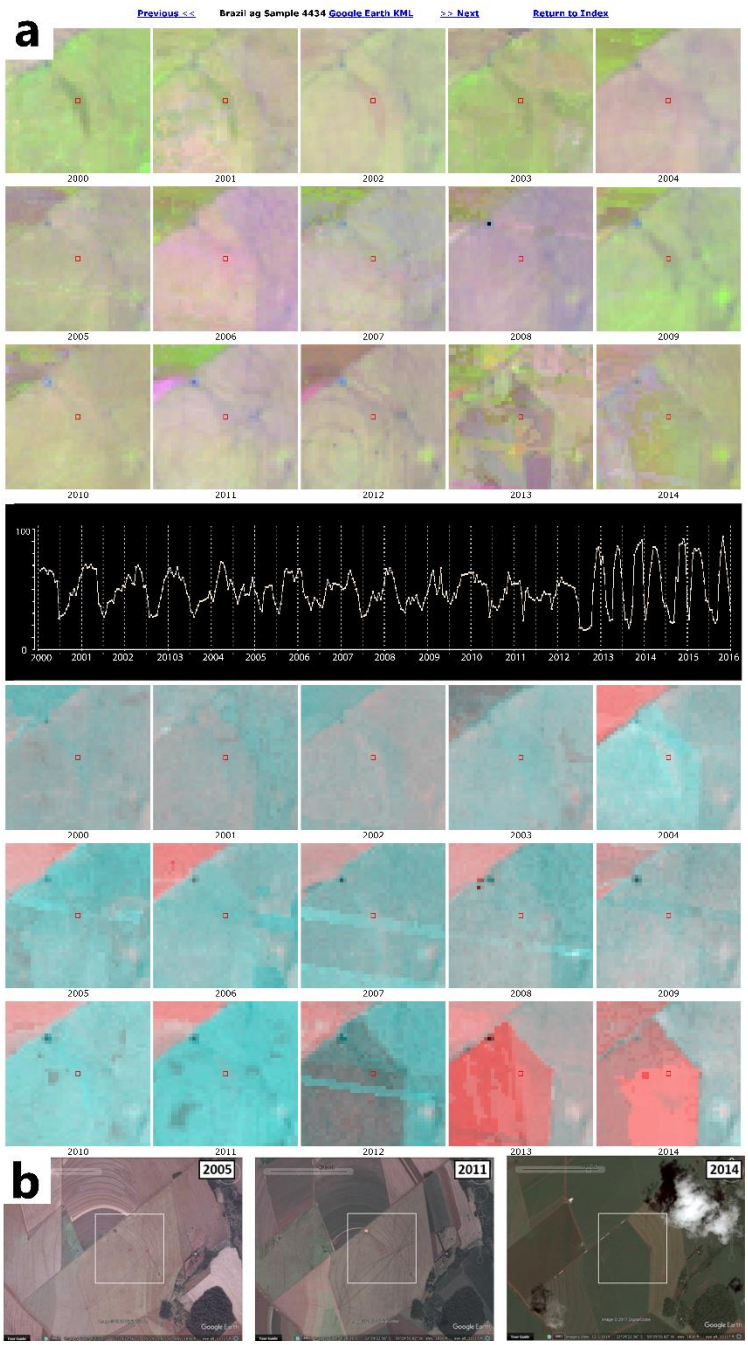
Supplementary Figure 4. **Trends of cropland expansion per state.** Sample-based cropland expansion area estimates for "cropland 2000" and "cropland expansion" strata (+/- one standard error) per state with >10 samples in "cropland expansion" strata. Year of expansion corresponds to year of planting (e.g. 2001 corresponds to the 2001/2002 growing season). Samples from the "no cropland" strata are not displayed. See Supplementary Table 3 for tabular data for all strata.



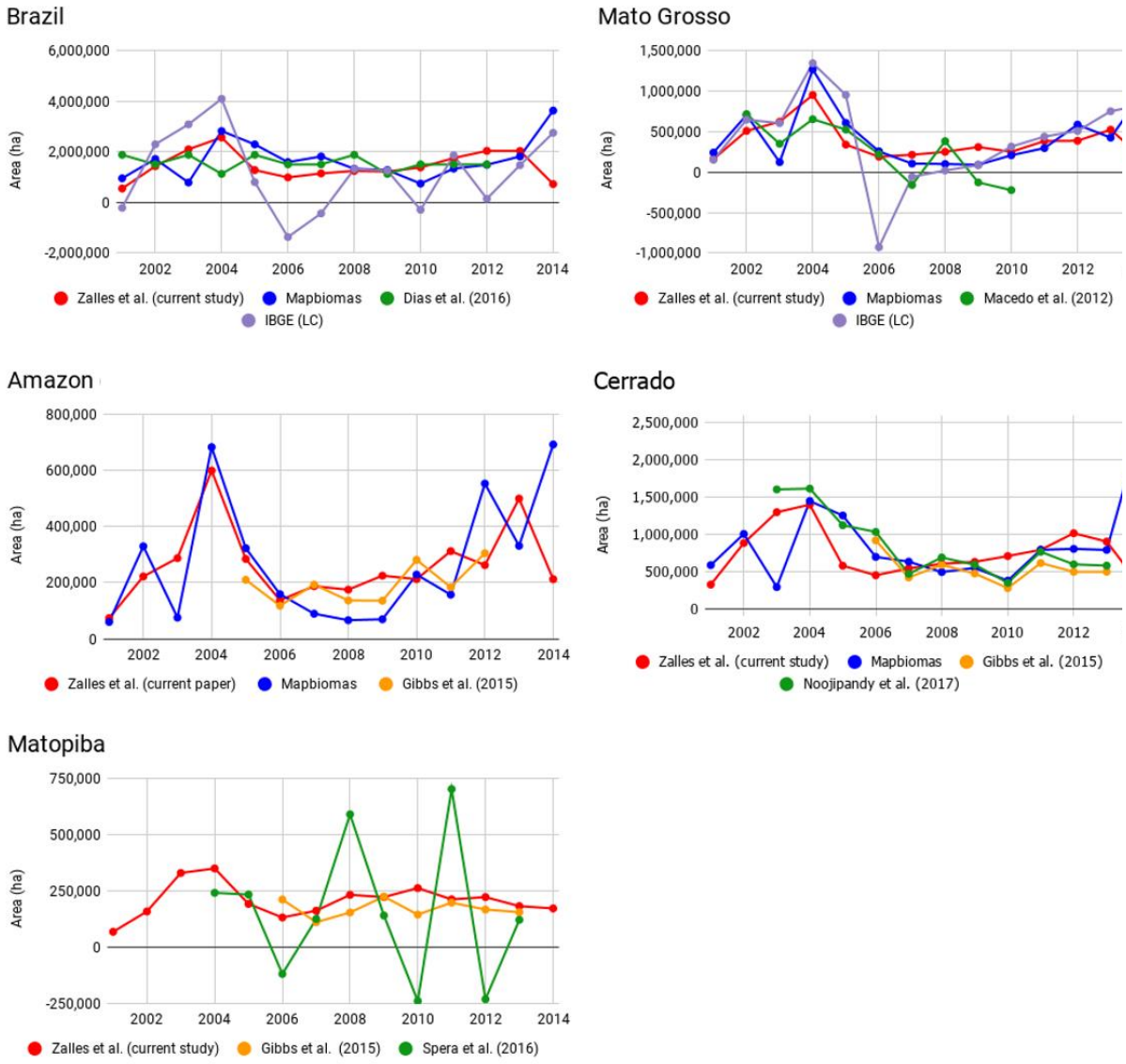
Supplementary Figure 5. **Trends of cropland expansion per biomes.** Sample-based cropland expansion area estimates for "cropland 2000" and "cropland expansion" strata (+/- one standard error) per biome with >10 samples in "cropland expansion" strata. Year of expansion corresponds to year of planting (e.g. 2001 corresponds to the 2001/2002 growing season). Samples from the "no cropland" strata are not displayed. See Supplementary Table 3 for tabular data for all strata.



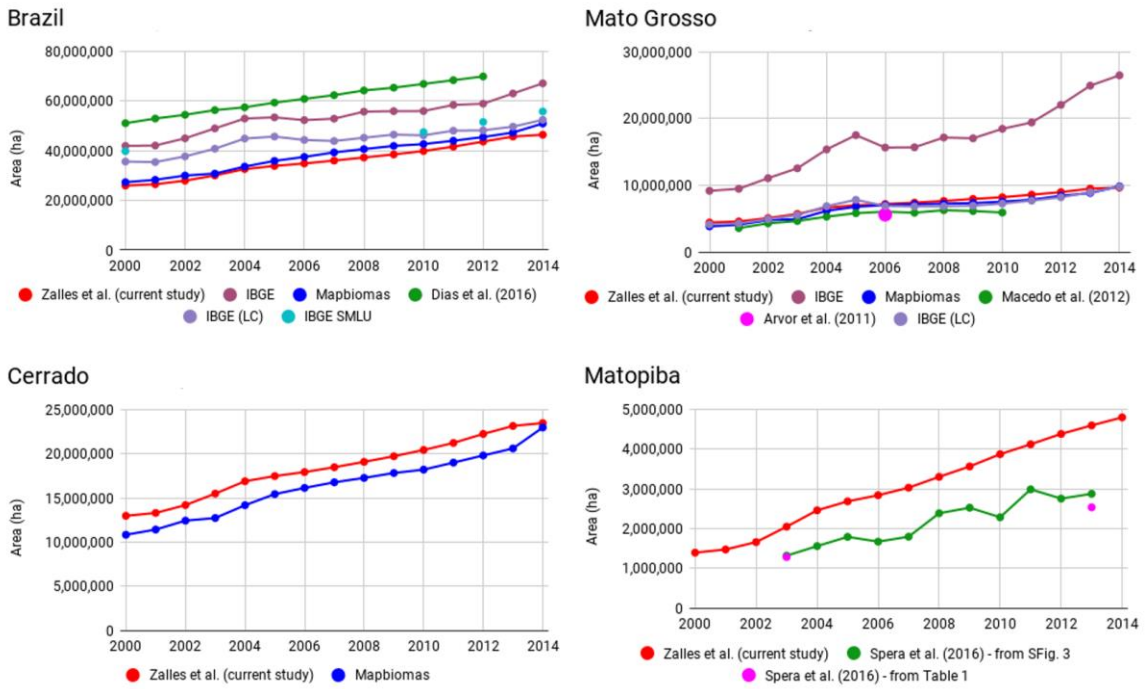
Supplementary Figure 6. Sources of remote sensing data used for sample interpretation shown through the example of a sample pixel. (**a**) shows web interface for an example sample pixel. Composites on top are annual cloud-free Landsat composites with SWIR 1- NIR - red loaded in RGB from 2000 to 2014. Bottom composites are annual cloud-free Landsat composites where yearly maximum NDVI is loaded in the red band, and yearly minimum NDVI is loaded in the green and blue bands from 2000 to 2014. Time series graph shows MODIS 16-day NDVI time series. The link on the top right downloads a .kml file which allows the interpreter to visualize the sampled pixel on Google Earth. This example shows conversion from natural vegetation to cropland. (**b**) shows Google Earth imagery corresponding to this sample pixel for the beginning of the time period (2000), year of change from natural vegetation to cropland (2004), and end of the time period (2014). White boxes on (**b**) are 1.1 x 1.1 km and correspond to the size of a Landsat subset on the web interface.



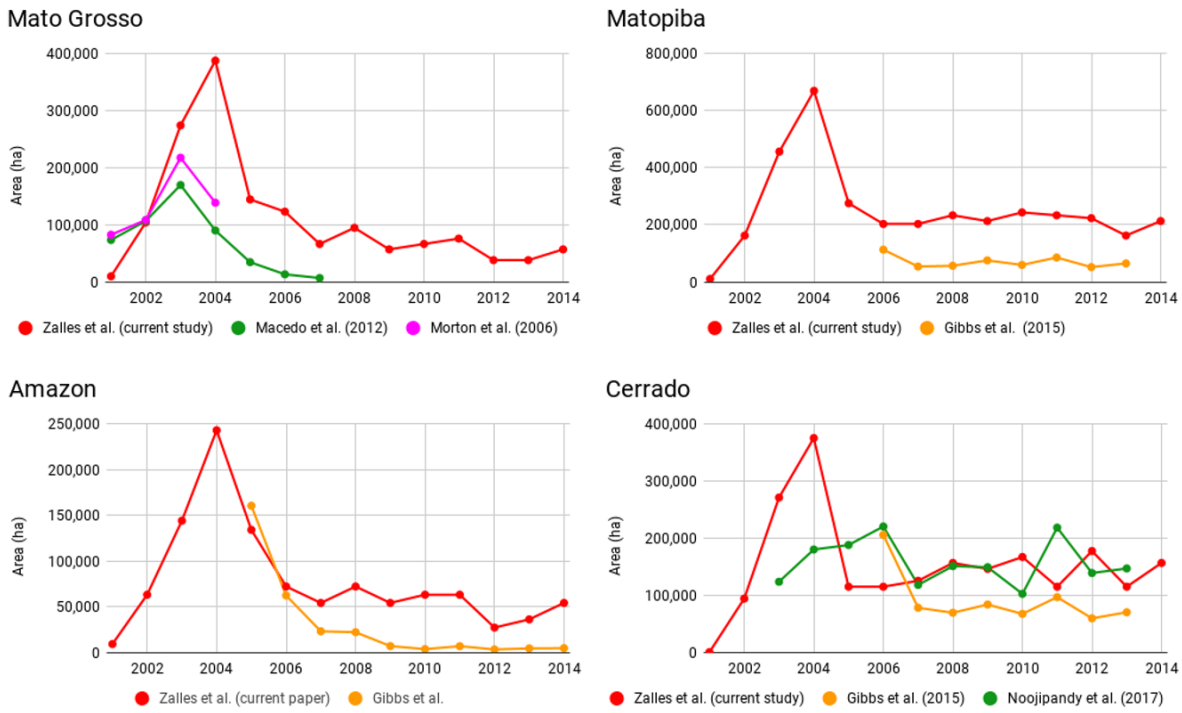
Supplementary Figure 7. Sources of remote sensing data used for sample interpretation shown through the example of a sample pixel. (**a**) shows web interface for an example sample pixel. Composites on top are annual cloud-free Landsat composites with SWIR 1- NIR - red loaded in RGB from 2000 to 2014. Bottom composites are annual cloud-free Landsat composites where yearly maximum NDVI is loaded in the red band, and yearly minimum NDVI is loaded in the green and blue bands from 2000 to 2014. Time series graph shows MODIS 16-day NDVI time series. The link on the top right downloads a .kml file which allows the interpreter to visualize the sampled pixel on Google Earth. This example shows conversion from pasture to cropland. (**b**) shows Google Earth imagery corresponding to the sample pixel for the earliest available time (2005), year before change (2011), and end of the time period (2014). White boxes on (**b**) are 1.1 x 1.1 km and correspond to the size of a Landsat subset on the web interface.



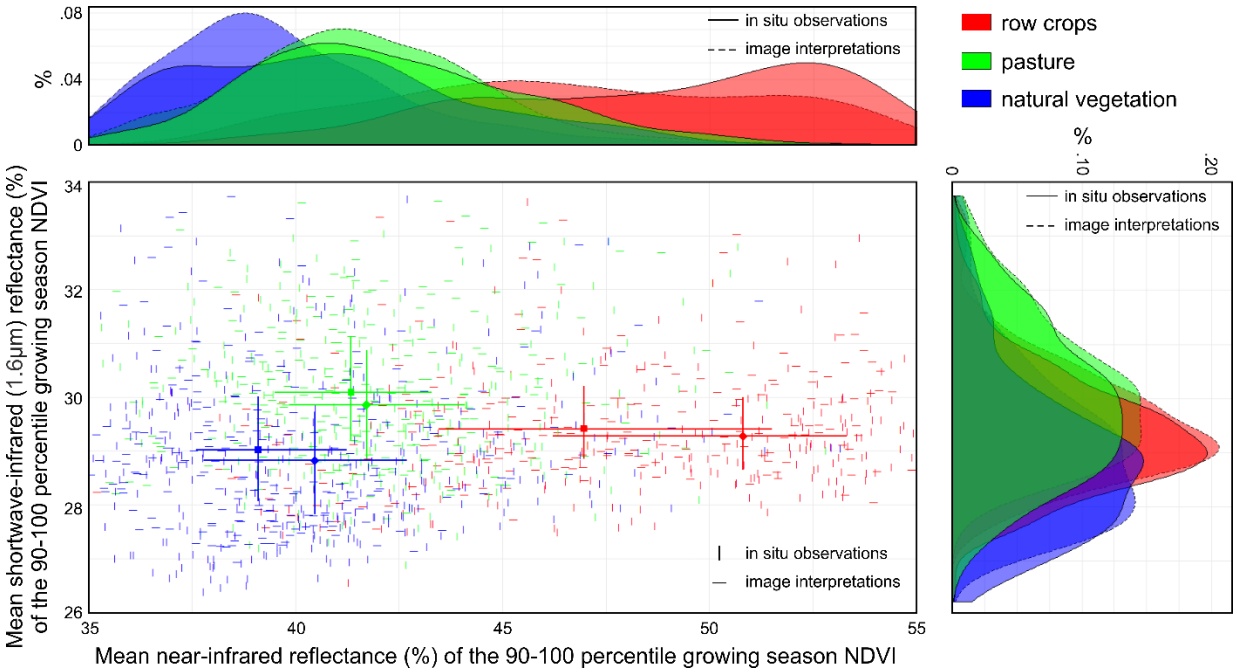
Supplementary Figure 8. Comparison of cropland expansion area between current study and others (8–13). “IBGE LC” (IBGE cropland Land Cover) area corresponds to 1st season corn, 1st season bean, soy, cotton, sugarcane, and rice areas from the IBGE SIDRA database. For comparison purposes, area of the cropland expansion coming from the "no cropland" stratum was distributed across all years proportionally to the area of expansion from the "cropland expansion" and "cropland 2000" strata combined.



Supplementary Figure 9. Comparison of total cropland area between current study and others (8, 11, 12, 14–16). “IBGE” data corresponds to area of soy, corn, sugarcane, cotton, rice, manioc, beans, and rice from the IBGE SIDRA database. “IBGE LC” (IBGE cropland Land Cover) area corresponds to 1st season corn, 1st season bean, soy, cotton, sugarcane, and rice areas from the IBGE SIDRA database. IBGE SMLU corresponds to IBGE’s Systematic Monitoring of Land Use project. Spera et al. (2016) report different areas in Table 1 and in Supplementary Figure 3 of their study. Both are included. For comparison purposes, area of the cropland expansion coming from the "no cropland" stratum was distributed across all years proportionally to the area of expansion from the "cropland expansion" and "cropland 2000" strata combined.



Supplementary Figure 10. Comparison of natural vegetation conversion to cropland area between current study and others (9, 10, 16, 17). For comparison purposes, area of the cropland expansion coming from the "no cropland" stratum was distributed across all years proportionally to the area of expansion from the "cropland expansion" and "cropland 2000" strata combined.



Supplementary Figure 11. Comparison of separate probability-based samples of in situ observations and interpretations using multi-source earth observation data using a common reference of growing season MODIS data. For in situ samples, MODIS data are from the 2016/2017 growing season (November-April) as part of a field campaign estimating soybean cultivated area. For multi-source interpreted samples (this study), MODIS data are from the 2000/2001 growing season and consist of a subset of samples from the 2000 cropland and 2000-2014 cropland gain strata. The x-axis corresponds to the mean MODIS near-infrared reflectance (%) of the 90-100 percentile growing season NDVI. The y-axis corresponds to the mean MODIS shortwave-infrared (1.6µm) reflectance (%) of the 90-100 percentile growing season NDVI. These time-series metrics represent the near-infrared and shortwave-infrared reflectance of each sample at the time of peak vegetation greenness. Bars on the scatterplot correspond to the 25th and 75th percentiles of the respective distributions.

		Reference				
Map		Cropland 2000	Cropland Expansion	Not Cropland	Total	User's accuracy (SE)
	Cropland 2000	2.81	0.15	0.98	3.93	71.5 (1.2)
	Cropland expansion	0.07	1.75	0.62	2.43	72.0 (1.0)
	Not cropland	0.25	0.56	92.83	93.64	99.1 (0.2)
	Total	3.12	2.46	94.42	Overall accuracy: 97.4 (0.2)	
	Producer's accuracy (SE)	89.9 (3.5)	71.2 (5.3)	98.3 (0.1)		

Supplementary Table 1. Confusion matrix for cropland 2000 and cropland expansion validation. Values shown are % of the study area.

	Cropland extent, 2000		Crop expansion (2001-2014)	
	Area	**SE**	**Area**	**SE**
Amazon	459,489	67,719	3,691,086	735,555
Bahia	1,049,878	103,800	1,500,385	89,041
Brazil	26,007,115	1,109,094	20,482,178	1,577,001
Caatinga	172,698	57,448	148,056	41,306
Cerrado	12,978,620	776,110	10,513,213	1,044,223
Goias	2,504,490	126,930	2,315,471	725,427
Maranhao	259,189	52,952	484,523	49,791
Mata Atlantica	10,027,647	217,241	5,356,662	978,253
Mato Grosso	4,428,442	556,760	5,255,843	770,853
Mato Grosso Do Sul	1,407,947	106,198	1,672,412	75,507
Matopiba	1,395,333	95,010	3,401,240	524,577
Minas Gerais	1,435,500	123,200	1,210,611	81,690
Pampa	2,601,462	729,403	862,150	57,678
Para	14,792	14,792	121,981	22,586
Parana	4,127,574	113,881	1,854,300	734,782
Piaui	177,227	51,524	690,899	59,235
Rio Grande Do Sul	5,381,663	820,042	959,638	61,673
Sao Paulo	4,658,543	479,890	2,753,393	466,900
Tocantins	103,483	39,113	838,209	478,874
Pernambuco	53,371	35,302	33,439	13,218

Supplementary Table 2. Area and standard error estimates. Estimates of area and SE of cropland extent in 2000 and cropland expansion through 2012 for all regions with >10 samples in the expansion strata. All area and standard error estimates are in hectares.

Study	Map data source	Study region	Study period	Provides some form of accuracy assessment	Follows "good practice" recommendations[+]			
					Probability sample	SE/CI of accuracy metrics reported	Uses reference sample data for area estimation	Estimates SE/CI for area estimates
Morton et al. (2006)	MODIS	Deforested areas in Mato Grosso	2001 - 2004	✓	×	×	×	×
Macedo et al. (2012)	MODIS	Mato Grosso	2001 - 2010	✓	×	×	×	×
Arvor et al. (2011)	MODIS	Mato Grosso	2006/2007	✓	×	×	×	×
Gibbs et al. (2015)	MODIS	Amazon, Matopiba, Cerrado	2001 - 2014	✓	×	×	×	×
Spera et al. (2016)	MODIS	Matopiba	2003 - 2013	✓	×	×	×	×
Noojipady et al. (2017)	MODIS	Cerrado	2003 - 2013	✓	×	×	×	×
Graesser and Ramankutty (2017)	Landsat	South America	2000/2001	✓	×*	×	×	×
IBGE SMLU	MODIS	Brazil	2000, 2010, 2012, 2014	×	×	×	×	×
TerraClass Amazon	Landsat	Brazilian Legal Amazon	2004, 2008, 2012, 2014	✓	×	×	×	×
TerraClass Cerrado	Landsat	Cerrado	2013	✓	✓	×	×	×
Mapbiomas (3.0)	Landsat	National level and every biome, state and municipality	1985-2017	×	×	×	×	×
Zalles et al. (current study)	Landsat	National level, 5 biomes, 13 states, and Matopiba	2000 - 2014	✓	✓	✓	✓	✓

Supplementary Table 3. Comparison of studies on cropland area in Brazil (8, 10, 11, 13, 15, 16, 18–21) to the current study. "SE/CI" refers to standard error/confidence interval. *Graesser and Ramankutty provide two accuracy assessments. The information on this chart reflects the accuracy assessment of their thematic map. [+] Per (22–26).

References

1. Chang J, Hansen MC, Pittman K, Carroll M, DiMiceli C (2007) Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets. *Agron J* 99(6):1654–1664.

2. Pittman K, Hansen MC, Becker-Reshef I, Potapov PV, Justice CO (2010) Estimating global cropland extent with multi-year MODIS data. *Remote Sens* 2(7):1844–1863.

3. Dempewolf J, et al. (2014) Wheat Yield Forecasting for Punjab Province from Vegetation Index Time Series and Historic Crop Statistics. *Remote Sens* 6(10):9653–9675.

4. Song X-P, et al. (2017) National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. *Remote Sens Environ* 190:383–395.

5. King L, et al. (2017) A multi-resolution approach to national-scale cultivated area estimation of soybean. *Remote Sens Environ* 195:13–29.

6. Khan A, Hansen MC, Potapov P, Stehman SV, Chatta AA (2016) Landsat-based wheat mapping in the heterogeneous cropping system of Punjab, Pakistan. *Int J Remote Sens* 37(6):1391–1410.

7. Hansen MC, et al. (2013) High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342(6160):850–853.

8. MapBiomas Available at: http://mapbiomas.org/ [Accessed July 6, 2017].

9. Gibbs HK, et al. (2015) Brazil's Soy Moratorium. *Science* 347(6220):377–378.

10. Noojipady P, et al. (2017) Forest carbon emissions from cropland expansion in the Brazilian Cerrado biome. *Environ Res Lett* 12(2):025004.

11. Spera SA, Galford GL, Coe MT, Macedo MN, Mustard JF Land-use change affects water recycling in Brazil's last agricultural frontier. *Glob Change Biol* 22(10):3405–3413.

12. IBGE SIDRA - Produção Agrícola Municipal Available at: https://sidra.ibge.gov.br/pesquisa/pam/tabelas.

13. Mapeamento Sistemático do Uso da Terra *Inst Bras Geogr E Estat*. Available at: https://ww2.ibge.gov.br/home/geociencias/recursosnaturais/usodaterra/default.shtm [Accessed September 27, 2018].

14. Dias LCP, Pimenta FM, Santos AB, Costa MH, Ladle RJ (2016) Patterns of land use, extensification, and intensification of Brazilian agriculture. *Glob Change Biol* 22(8):2887–2903.

15. Arvor D, Jonathan M, Meirelles MSP, Dubreuil V, Durieux L (2011) Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. *Int J Remote Sens* 32(22):7847–7871.

16. Macedo MN, et al. (2012) Decoupling of deforestation and soy production in the southern Amazon during the late 2000s. *Proc Natl Acad Sci* 109(4):1341–1346.

17. Morton DC (2016) Forest carbon fluxes: A satellite perspective. *Nat Clim Change* 6(4):346–348.

18. Morton DC, et al. (2006) Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *Proc Natl Acad Sci* 103(39):14637–14641.

19. Graesser J, Ramankutty N (2017) Detection of cropland field parcels from Landsat imagery. *Remote Sens Environ* 201(Supplement C):165–180.

20. Projeto TerraClass Available at: http://www.inpe.br/cra/projetos_pesquisas/dados_terraclass.php [Accessed July 6, 2017].

21. Projeto TerraClass Cerrado Available at: http://www.dpi.inpe.br/tccerrado/index.php?mais=1 [Accessed July 6, 2017].

22. Stehman SV (2013) Estimating area from an accuracy assessment error matrix. *Remote Sens Environ* 132:202–211.

23. Olofsson P, Foody GM, Stehman SV, Woodcock CE (2013) Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens Environ* 129:122–131.

24. Penman J, et al. (2016) Integration of remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: methods and guidance from the Global Forest Observations Initiative.

25. Map Accuracy Assessment and Area Estimation: A Practical Guide (2016) *Natl For Monit Assess Work Pap FAO* (No.46/E). Available at: http://www.fao.org/3/a-i5601e.pdf.

26. 2006 IPCC Guidelines for National Greenhouse Gas Inventories (National Greenhouse Gas Inventories Programme, IGES, Japan).