

An Analysis of Decadal Land Change in Latin America and the Caribbean Mapped from 250-m MODIS Data

M.L. Clark ^{a,*} and T.M. Aide ^b

^a Center for Interdisciplinary Geospatial Analysis, Sonoma State University, Rohnert Park, CA 94928 USA–
mateolclark@gmail.com

^b Dept. of Biology, University of Puerto Rico, Río Piedras, Puerto Rico – tmaide@yahoo.com

Abstract - We present a method to map annual land-use/land-cover (LULC) change using low-cost and consistent data from the MODIS satellite and Google Earth reference data. Our LULC maps cover Latin America and the Caribbean from years 2001 to 2009, have a 250-m pixel size, and include eight general classes. Maps were produced within 27 separate zonal biomes using the Random Forests classifier. We found that accuracy of the maps is sensitive to broad climatic and vegetation differences encapsulated by biomes, with an average 84.9% (s.d.±6.3%) overall accuracy for a five-class reclassification. Trends in woody vegetation area were analyzed using all years and pixels aggregated at ecoregion and municipality scales. The largest hotspots of deforestation were in South America, concentrated mostly in moist forests around the south to west Amazon basin and western edge of Atlantic forests, and dry forests in the Chaco ecoregions of northern Argentina. Driving factors included mostly pastures for cattle and industrial-scale agriculture for global export. We also detected increasing woody vegetation in northern Mexican deserts and in dry forests of southern Bolivia. Although not thoroughly evaluated, possible factors driving increased woody vegetation could be local increase in precipitation and the mitigating influence of protected areas.

Keywords: MODIS, land-cover change, Latin America, Random Forests

1. INTRODUCTION

In the 21st century, economic globalization and climate change have risen in importance as forces having broad-scale impacts on human and natural systems. Some aspects of economic globalization are positive for human systems, such as poverty reduction. However, globalization and climate change presage a host of negative effects for natural systems, such as land conversion for mechanized agriculture that supports global food supply and climate-induced shifts in species distributions, abundance and susceptibility to extinction.

The twin forces of globalization and climate change are producing shifts in Latin America and the Caribbean (LAC) land use that are not well understood, partially because there is an acute lack of land-use/land-cover (LULC) datasets across the region (Clark et al., 2010). To date, most analyses of land change have focused on loss of forest cover, (i.e., deforestation), ignoring reforestation or other land-use dynamics. At sub-national scales, the Brazilian Landsat-based PRODES program is a model deforestation monitoring program (<http://www.obt.inpe.br/prodes>), but this type of program is unique to LAC. Recently, Hansen and colleagues (2010) presented global forest cover loss assessment based on continuous percent forest cover estimated from 250-m MODIS data and calibrated with multi-temporal samples of

Landsat-derived forest-loss maps. The data from that study revealed that South America contains the largest global share of forest cover, but is also ranked third among continents in terms of deforestation rates. In particular, Brazil had the highest loss of forest cover of any nation, with deforestation in both the moist tropical forests (e.g., Amazon basin) and dry forests (e.g., Cerrado). Paraguay and Argentina also had clearing of both moist and dry forests, mainly for increased mechanized agriculture, such as soybean production, or cattle pastures.

We have developed a MODIS-based method to map recent land change at broad spatial scales with internally-consistent imagery and reference datasets. The method, explained originally in Clark et al. (2010) for the Dry Chaco ecoregion of Argentina, Paraguay and Bolivia, provides wall-to-wall LULC thematic maps on an annual basis starting in year 2001, with a 250-m pixel size. Here we present modifications of this method for continental-scale mapping of LULC across Latin America and the Caribbean. We assess the accuracy of the maps by regional biomes and then assess trends in woody vegetation (e.g., forest, woodlands, shrublands) change – both loss and gain – at municipality and ecoregion scales.

2. METHODS

2.1 Study area

The study area included all countries in Latin America and the Caribbean (LAC). The study area was divided into 11 “latitudinal” zones, and each zone was split by dominant biome (Olson et al., 2001) with boundaries modified to follow municipalities within countries (Table A). This created a total of 27 separate mapping zone-biome combinations.

2.2 Reference data

Reference data for classifier training and accuracy assessment were collected with human interpretation of high-resolution imagery in Google Earth (GE, earth.google.com). Most of this imagery in our study area is from Digital Globe’s QuickBird (www.digitalglobe.com) and Geoeye’s IKONOS (www.geoeye.com) commercial satellites, and all samples included the image’s acquisition year. Sample centers were snapped to the closest MODIS pixel, described below, and no sample was closer than 1000 m to a neighboring sample. Interpretation criteria and error-checking protocols followed those in Clark et al., (2010), were conducted using an automated web-based tool called the Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT), and will be described in a forthcoming paper. Within VIEW-IT, samples were placed haphazardly across LAC in areas that had >80% cover seven class types: woody vegetation (*Woody*), herbaceous vegetation (*Herb*), agriculture (*Ag*), plantations (*Plant*), built-up areas (*Built*), bare areas (*Bare*), and water (*Water*). Areas with 20-80% *Woody*, with a *Bare*, *Herb* or *Ag* component,

* Corresponding author.

** Funded by the U.S. National Science Foundation (NSF #0709598).

were assigned to a mixed-woody vegetation class (*MixWoody*).

Table A. Zone and biomes (Olson et al., 2001) and total sample counts used in each separate Random Forest (n=27). These include only samples filtered by the Random Forests outlier removal process.

Zone-Biome	Geographic region	No. of Samples
<i>Tropical & Subtropical Moist Broadleaf Forests (TSMBF)</i>		
z01_b01	Mexico	1,354
z02_b01	Central America	2,082
z03_b01	Northern South America	2,140
z04_b01	Western South America	1,191
z06_b01	Amazon basin	5,035
z07_b01	Atlantic moist forest (Brazil)	1,030
z11_b01	Caribbean (all biomes)	1,486
<i>Tropical & Subtropical Dry Broadleaf Forests (TSDBF)</i>		
z01_b02	Mexico	1,577
z02_b02	Central America	1,062
z03_b02	Northern South America	1,809
z04_b02	Western South America	330
<i>Tropical & Subtropical Coniferous Forests (TSCF)</i>		
z01_b03	Mexico	1,532
z02_b03	Central America	1,422
<i>Temperate Broadleaf & Mixed Forests (TBMF)</i>		
z05_b04	Southern South America	1,935
<i>Trop. & Subtrop. Grasslands, Savannas, Shrubland (TSGSS)</i>		
z03_b07	Northern South America (Llanos)	1,130
z04_b07	Beni savanna (Bolivia)	485
z05_b07	Humid Chaco (Argentina, Paraguay)	443
z06_b07	Cerrado (Brazil)	1,460
z08_b07	Uruguayan savanna ecoregion	868
z10_b07	Dry Chaco (Arg., Bolivia, Paraguay)	3,396
<i>Temperate Grasslands, Savannas & Shrublands (TGSS)</i>		
z05_b08	Pampas and Patagonia (Argentina)	3,030
<i>Flooded Grasslands & Savannas (FGS)</i>		
z09_b09	Pantanal (Brazil, Bolivia)	447
<i>Montane Grasslands & Shrublands (MGS)</i>		
z04_b10	Andean puna (Arg., Chile, Bol., Peru)	802
<i>Mediterranean Forests, Woodlands, and Scrub (MFWS)</i>		
z05_b12	Chilean matorral	551
<i>Deserts & Xeric Shrublands (DXS)</i>		
z01_b13	Mexico	2,248
z04_b13	Western South America coast	753
z06_b13	Caatinga (Brazil)	724

2.3 Mapping process

We used the MODIS MOD13Q1 Vegetation Indices 250 m product (Collection 5) for LULC classification. The product is a 16-day composite of the highest-quality pixels from daily images and includes the Enhanced Vegetation Index (EVI), red, near infrared (NIR), and mid-infrared (MIR) reflectance and pixel reliability, with 23 samples per year (Huete et al., 2002). In this analysis, we used 39 MODIS tiles that cover LAC with all scenes from years 2001 to 2009. MODIS scenes were reprojected from their native Sinusoidal projection to the Interrupted Goode Homolosine projection (sphere radius of 6,378,137.0 m) using nearest-neighbor resampling. The original cell size of 231.7 meters was maintained in the reprojection.

For each pixel, we calculated the statistics mean, standard deviation, minimum, maximum and range for EVI, and red, NIR and MIR reflectance values from calendar years 2001 to

2009. These statistics were calculated for all 12 months (annual), 2 six-month periods (bi-annual), and 3 four-month periods (tri-annual). The MOD13Q1 pixel reliability layer was used to remove all unreliable samples (value = 3) prior to calculating statistics.

We mapped LULC with the Random Forests (RF) tree-based classifier (Breiman, 2001), following methods in Clark et al., (2010) with the following modifications. The RF classifier was implemented using R (v. 2.11.1; R Development Core Team, 2010), the “randomForest” package (v. 4.5-34; Liaw and Wiener, 2002), and 1999 trees. Predictor variables were MODIS-based full-year, bi-annual and tri-annual statistics for EVI, red, NIR and MIR, and were extracted for the year corresponding to the high-resolution image year (2001 to 2009) for each GE reference sample. Samples were assigned to a zone-biome (Table A). For each zone-biome mapping area, an initial RF was generated with the reference sample class and MODIS predictor variables. The outlier function in randomForest was used to eliminate samples with an outlier metric greater than 10, and a final RF was generated from the remaining samples (Table B). Accuracy statistics presented here are based on the RF “out-of-bag” statistics, which is a form of cross-validation that has been shown to have minimal bias (Breiman, 2001, Clark et al., 2010).

Table B. Class sample counts from all zones-biome Random Forests and after the outlier removal process.

Class	No. of Samples	Class	No. of Samples
<i>Ag</i>	5,140	<i>MixWoody</i>	5,033
<i>Bare</i>	1,541	<i>Plant</i>	2,436
<i>Built</i>	2,639	<i>Water</i>	4,159
<i>Herb</i>	7,931	<i>Woody</i>	11,443
Total	40,322		

We used R and the RGDAL library to apply the final RF object to every pixel in the zone-biome for each year, 2001 to 2009. The final maps for each zone-biome were then mosaicked together for continental-scale analysis. For our accuracy and trend analyses, we reclassified maps to have five classes: *Woody*, *Plant/MixWoody*, *Ag/Herb*, *Bare/Built*, and *Water*.

2.4 Trend analysis

We focused on trends in the *Woody* class, as it represents change associated with natural vegetation, such as deforestation or reforestation, which has important implications for species habitat and carbon dynamics. Country municipality (n = 16,052) and ecoregions (n = 183) (Olson et al., 2001) polygons were overlaid on the mosaicked LULC maps. For each polygon (municipality or ecoregion), a linear regression of *Woody* area (dependent variable) against time (independent variable, 9 years) was conducted. Reported trends are the Pearson correlation coefficients per polygon, which is a standardized metric that allows comparison among variably-sized polygons. We present only those polygons with a significant slope ($p \leq 0.05$). Absolute areas of *Woody* reported for 2001 and 2009 are estimates from the linear regression model developed for each polygon.

3. RESULTS AND DISCUSSION

3.1. Accuracy of zonal-biome maps

Based on RF sample OOB statistics, the average overall accuracy of 27 separate RFs with five classes was 84.9% (s.d. \pm 6.3%, Fig. 1). Overall accuracy was lowest in z01_b13,

which includes Mexico's deserts (e.g., Chihuahua, Sonora) and z11_b01, which was all of the Caribbean islands. The highest overall accuracy was in z04_b07 and z08_b07, which are the Beni and Uruguayan savanna ecoregions, respectively (Fig. 2). Accuracy for *Woody*, which is the class that we focus on in subsequent trend analyses, had an average producer's accuracy of 84.3% (s.d. \pm 22.1%) and an average user's accuracy of 82.7% (s.d. \pm 12.4%). *Woody* user's accuracy was generally $>80\%$ for regions dominated by trees (e.g. moist forests, dry forests) while its accuracy was less for more sparsely vegetated, drier areas with more shrubs (e.g., deserts and Andean puna). *Water* had consistently high and less variable accuracy across zone-biome regions. In contrast, the *MixWoody/Plant* combined class had relatively low accuracy, as these classes had a lot of sub-pixel mixing and were generally confused with *Woody*, *Herb* and *Ag*, depending on the zone-biome region.

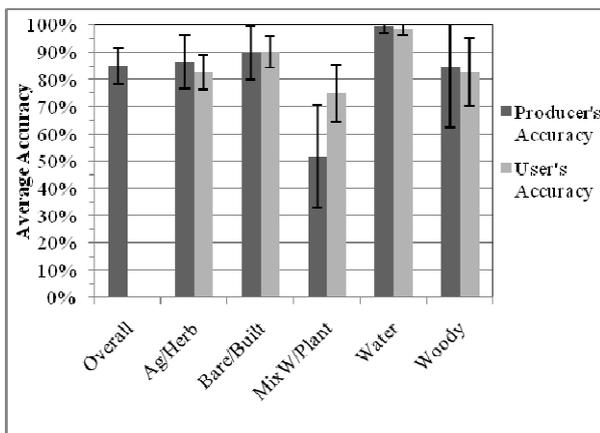


Figure 1. Random Forests out-of-bag average and standard deviation (bars) accuracy for the 5-class maps across all 27 zone-biomes regions.

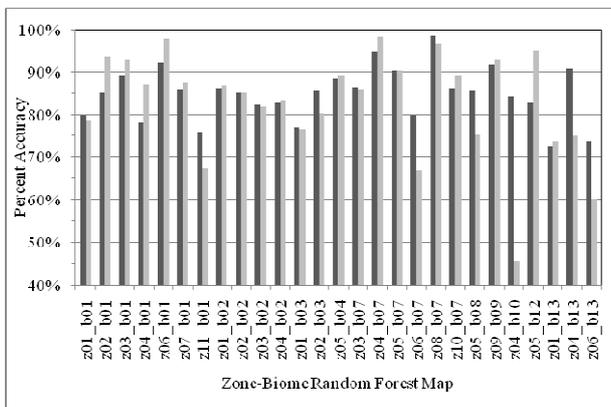


Figure 2. Overall accuracy (dark gray) and *Woody* user's accuracy (light gray) for the 27 zone-biome mapping regions (see Table A for definitions).

3.2. Trends in woody vegetation over the last decade

From 2001 to 2009, *Woody* showed steep declines in the Madeira-Tapajós and adjoining Purus-Madeira moist forests ecoregions (TSMBF biome), as well as the Dry Chaco ecoregion (TSGSS biome), all with correlation coefficients $r < -0.90$ (Fig. 3). The municipality-level trend analysis revealed finer spatial patterns (Fig. 4), with clusters of municipalities with high deforestation rates in TSMBF around the Amazon basin (e.g. "arc of deforestation") and the Ucayali-Huanuco-Pasco-Junin region of lowland Peru, both areas with extensive

cattle pastures. In this area, the municipalities Porto Velho (Rondônia), Novo Progresso (Para), Aripuanã (Mato Grosso, Fig. 5A) in Brazil had both steep downward trends ($r < -0.90$) and the highest ($> 3,950 \text{ km}^2$) area of *Woody* loss from 2001 to 2009.

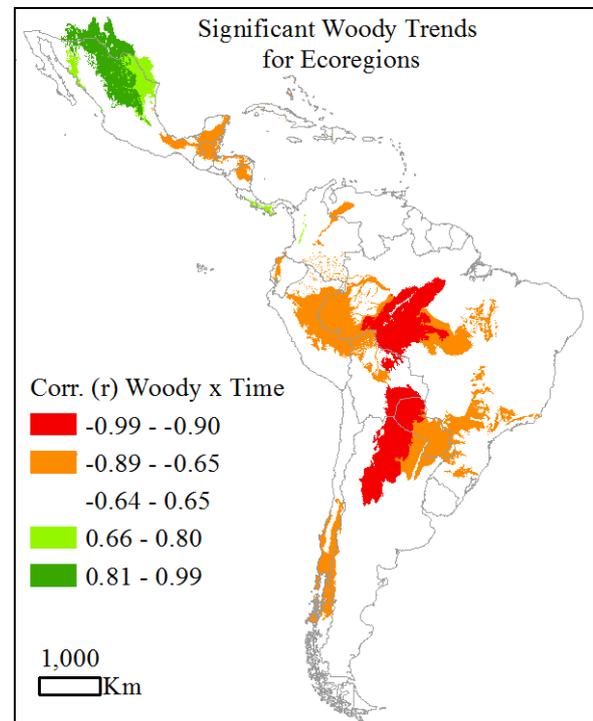


Figure 3. Linear correlation (r) in *Woody* vs. year for ecoregions (Olson et al., 2001) with significant slopes ($p \leq 0.05$).

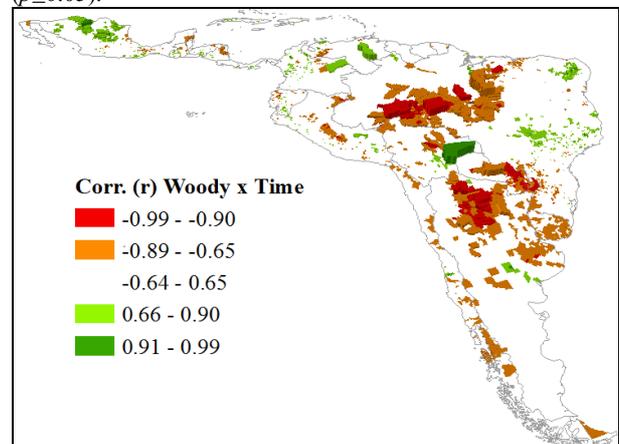


Figure 4. Linear correlation (r) in *Woody* vs. year for municipalities with significant slopes ($p \leq 0.05$). Height of municipalities is estimated area of *Woody* change, 2001 to 2009 (exaggerated $\text{km}^2 \times 25$).

There were also several municipalities with rapid deforestation in the Alto Paraná Atlantic forests ecoregion (TSMBF biome), concentrated in the provinces of Misiones, Argentina and Caaguazú and Canindeyú, Paraguay. These are areas with mechanized agriculture and plantations. In the TSGSS biome, hot spots of deforestation were in the northern Argentina part of the Dry Chaco ecoregion, with several municipalities in Santiago del Estero and Chaco provinces, and in the Humid Chaco ecoregion to the east, including municipalities in Santa Fe and Formosa provinces. In particular, the municipalities Alberdi, Moreno and Juan Felipe

Ibarra (Santiago del Estero) and Anta (Salta) had both steep deforestation trends ($r < -0.95$) and relatively high estimated *Woody* loss ($> 3,162 \text{ km}^2$). This TSGSS region has extensive dry forests that are mostly converted via initial degradation to mechanized agriculture, such as soy beans, although some livestock pastures are also in the landscape.

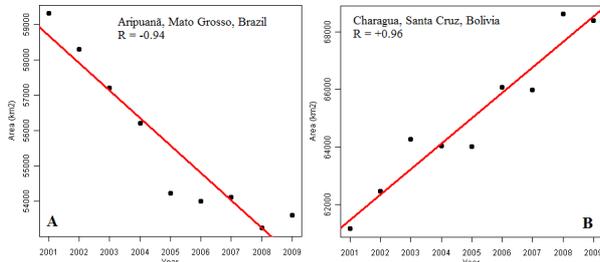


Figure 5. Woody vegetation area over time for the municipality A) Aripuanã, Brazil and B) Charagua, Bolivia.

Over the past decade, there was a strong *Woody* increase (e.g., woody regrowth or encroachment) in the DXS biome of northern Mexico, particularly in the Chihuahuan desert and Meseta Central matorral ecoregions (Fig. 3). Municipalities with strong reforestation were in the states of Chihuahua, Durango, Nuevo Leon, Sonoma and Zacatecas (Fig. 4). For example, the municipality of Carmargo in Chihuahua had an $r = +0.94$ and $4,037 \text{ km}^2$ of *Woody* increase. These increases in woody vegetation could be attributed to more precipitation; between 2001 and 2009, the Chihuahua and Nuevo Leon experienced an average increase in precipitation of 38% and 21%, respectively, relative to average precipitation during the previous 60 years. There were multiple municipalities in the Cerrado, Caatinga and Alto Paraná Atlantic forest ecoregions of Brazil that showed significant trends of increasing *Woody* (Fig. 4). Most of these municipalities are very small and so do not represent large increases in *Woody* area relative to other LAC municipalities, but trends may be from shifts in agriculture and local precipitation (Redo et al., *in review*).

Although not affecting the overall ecoregion change, *Woody* cover also increased significantly in the Sucre municipality, Bolivar state, Venezuela ($r = +0.73$, $4,786 \text{ km}^2$ *Woody* gain) and, to a lesser extent, the Cumaribo municipality, Vichada state, Colombia ($r = +0.78$, $1,047 \text{ km}^2$ *Woody* gain, Fig. 4). It is not clear which factors contribute to these trends, as these are remote areas with low population density and large protected areas. However, both of these municipalities include the sharp southern ecotone of the Llanos savanna ecoregion (grasslands) with moist forests, which may be susceptible to over-mapping forest relative to grasslands with 250-m pixels.

The Charagua municipality in Santa Cruz, Bolivia had a strong increasing trend in *Woody* ($r = +0.96$), with a regression-estimated increase of $7,030 \text{ km}^2$ from 2001 to 2009 (Fig. 4, Fig. 5B). Although the Santa Cruz state includes heavy deforestation from mechanized agriculture, the Charagua municipality includes the Kaa-Iya del Gran Chaco National Park and Integrated Management Natural Area, established in 1995 and the largest protected area in Bolivia, which prevents broad-scale deforestation in this remote area. Some of the reforestation may be due to forest recovery from past fire disturbance or from an increase in precipitation. This municipality (and its protected area) is an anomaly in the Dry Chaco ecoregion, which has otherwise seen extensive deforestation over the last decade (Figs. 3 & 4).

4. CONCLUSION

Our method of mapping land change in Latin America uses low-cost and free data sources for both base imagery and reference data, making it useful for mapping at sub-national to global scales. Accuracy of the maps is sensitive to broad climatic and vegetation differences encapsulated by biomes, with the best accuracy for moist forests and savannas, and lower accuracy for deserts. In our cursory analysis of *Woody* trends, targeting the most salient ecoregions and finer-scale municipalities, we corroborated work by Hansen et al. (2010) using MODIS data from 2000-2005 which were aggregated to 18.5-km grids: the largest hotspots of deforestation are in South America, concentrated mostly in moist forests around the south to west Amazon basin and western edge of Atlantic forests, and dry forests in the Chaco ecoregions of northern Argentina. The causes for these changes are well established – mostly industrial-scale agriculture for global food markets and cattle production. However, our data also revealed an interesting signal of increasing woody vegetation that may be related to local increase in precipitation and the influence of protected areas. More research is needed, with finer-scale maps from sensors such as Landsat and field data, to corroborate our observed trends.

REFERENCES

- Breiman, L. "Random Forests," *Machine Learning*, vol. 45, pp. 5-32, 2001.
- Clark, M.L., Aide, T.M, Grau, H.R., and Riner, G. "A scalable approach to mapping annual land-cover at 250 m using MODIS time-series data: A case study in the Dry Chaco ecoregion of South America," *Remote Sens. of Env.*, vol. 114, pp. 2816-2832, 2010.
- Hansen, M.C., Stehman, S.V., and Potapov, P.V. "Quantification of global gross forest cover loss," *PNAS*, vol. 107, 2010.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., and Ferreira, L.G. "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sensing of Environment*, vol. 83, pp. 195-213, 2002.
- Liaw, A. and M. Wiener. "Classification and Regression by randomForest," *R News*, 2(3), pp. 18-22, 2002.
- Olson, D.M, Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. and Kassem, K.R. "Terrestrial ecoregions of the world: A new map of life on earth," *Bioscience*, vol. 51, pp. 933-938, 2001.
- Redo, D., Aide, T.M., and Clark, M.L. "Vegetation Change in Brazil's Dryland Ecoregions and the Relationship to Crop Production and Environmental Factors: Cerrado, Caatinga and Mato Grosso, 2001-2009," *Agr. Ecosyst. Environ.* (in review).

ACKNOWLEDGEMENTS

This research was funded by a Dynamics of Coupled Natural and Human Systems grant from the U.S. National Science Foundation (NSF #0709598). We thank George Riner at Sonoma State University for assistance with data processing.