Satellite-detected gain in built-up area as a leading economic indicator

View the article online for updates and enhancements.

Environmental Research Letters

Satellite-detected gain in built-up area as a leading economic indicator

Qing Ying[1,5] , Matthew C Hansen[1] , Laixiang Sun[1,3,4] , Lei Wang[2] and Marc Steininger[1]

1.   Department of Geographical Sciences, University of Maryland, College Park, MD 20742, United States of America
2.   State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, 100101, People's Republic of China
3.   School of Finance and Management, SOAS University of London, London WC1H 0XG, United Kingdom
4.   International Institute for Applied Systems Analysis (IIASA), Laxenburg A-2361, Austria
5.   Author to whom any correspondence should be addressed.

E-mail: qying@umd.edu

Abstract

Leading indicators of future economic activity include measures such as new housing starts, managers purchasing index, money supply, and bond yields. Such macroeconomic and financial indicators hold predictive power in signaling recessionary periods. However, many indicators are constrained by the fact that data are often published with some delay and are subject to constant revision (Bandholz and Funke 2003, Huang *et al* 2018, Orphanides 2003). In this research, we propose a leading indicator derived from satellite imagery, the expansion of anthropogenic bare ground. Satellite-detected gain in built-up area, a major land cover and land use (LCLU) outcome of anthropogenic bare ground gain (ABGG), provides an inexpensive, consistent, and near-real-time indicator of global and regional macroeconomic change. Our panel data analysis across four major regions of the world from 2001 to 2012 shows that the logarithm of total ABGG, mostly owing to its major LCLU outcome, the expansion of built-up land in either year t, $t-1$ or $t-2$, significantly correlated with the year t logarithm of gross domestic product (GDP, de-trended by Hodrick–Prescott filter). Global ABGG between 2001 and 2012 averaged 7875 km^2 yr^{-1}, with a peak gain of 11 875 ($\pm$ 2014 km^2 at the 95% confidence interval) in 2006, prior to the 2007–2008 global financial crisis. The curve of global ABGG or its major LCLU outcome of built-up area in year $t-1$ accords well with that of the de-trended logarithm of the global GDP in year t. Given the 40 year archive of free satellite data, a growing satellite constellation, advances in machine learning, and scalable methods, this study suggests that analyses of ABGG as a whole or its LCLU outcomes can provide valuable information in near-real time for socioeconomic research, development planning, and economic forecasting.

Introduction

The 2007–2008 global financial crisis is considered the worst since the Great Depression of the 1930s and had dramatic impacts on global and regional economies and societies. Economists and policy makers seek indicators of overall economic health to help diagnose and forecast expected performance, with a goal to mitigate against volatility and to avoid shocks such as the crisis of 2007–2008. However, such efforts are limited as macroeconomic and financial variables are often reported with delays and constantly revised (Bandholz and Funke 2003, Orphanides 2003, Huang *et al* 2018). Recently, new data resources from satellite images (Jean *et al* 2016, Bennett and Smith 2017), cellphone metadata (Blumenstock *et al* 2015) and social media (Li *et al* 2013, Liu *et al* 2015) have emerged as indicators of economic activity. For example, proposed proxies for GDP, such as lit area (Elvidge *et al* 1997) and luminosity (Chen and Nordhaus 2011), have been derived from satellite data on nighttime light (Henderson *et al* 2012). However, the ability of

new big data sources to predict changes in economic outcomes at global or regional scale is a work in progress. Based on the presented research, we propose a leading indicator derived from freely available satellite imagery: the expansion of built-up area, a major LCLU outcome of anthropogenic bare ground gain (ABGG).

ABGG is a dynamic of land-cover and land-use change (LCLUC) that principally results from economic activities. Estimates of ABGG, based on the characterization of publicly available satellite imagery, can potentially serve as a low-cost, near-real time source for proxies of economic change from national to global scales. We define bare ground gain as a process of land-cover change featuring the removal and continued absence of vegetative cover for at least three years by either human or natural disturbances (Ying *et al* 2017). In our previous research (Ying *et al* 2017), globally and for each of seven regions over the 2001–2012 period, we partitioned bare ground gain into six components, defined by their LCLU outcomes: resource extraction; infrastructure development; commercial/residential built-up area; transitional bare ground gain, defined as new bare ground gain that had not yet been clearly put to some use; greenhouses; and one component for all natural gain (see Ying *et al* 2017 for detailed explanation). The five components of ABGG accounted for 95% of total bare ground gain over the study period. Examples of ABGG include expansion of urban areas, construction of new roads, mining, installation of oil wells, among other dynamics.

For the present study we estimated temporal trends of bare ground gain and its LCLU outcomes at the global scale (figure 1(a)) and for seven regions (figures 1(b)–(e), S3 is available online at stacks.iop. org/ERL/14/114015/mmedia) where time series of economic indicators are aggregated by the World Bank, including East Asia and the Pacific, North America, Europe and Central Asia, Latin America and the Caribbean, South Asia, Sub-Saharan Africa, and Middle East and North Africa. We conducted panel data analysis of the four regions that account for over 90% of ABGG and have relatively low uncertainties: East Asia and the Pacific, North America, Europe and Central Asia, Latin America and the Caribbean. This analysis shows that the logarithm of ABGG in either year t, $t-1$ or $t-2$ is significantly correlated to the de-trended (by Hodrick–Prescott (HP) filter) logarithm of GDP in year t, whereas such gains in year t or $t-1$ are significantly correlated to the de-trended logarithms of merchandise imports and exports, or energy consumption in year t. Globally, the annual ABGG between 2001 and 2012 was 7875 km^2 on average with a peak gain of 11 875 ($\pm$2014 km^2 for 95% CI) occurred in 2006 prior to the 2007–2008 global financial crisis. The curve of global ABGG in year $t-1$ accords well with that of the de-trended logarithm of the global GDP in year t. This predictive attribute of remotely sensed ABGG makes it an important LCLUC theme

that can effectively support socioeconomic analysts and policy makers to develop financial plans and to allocate resources towards stable growth.
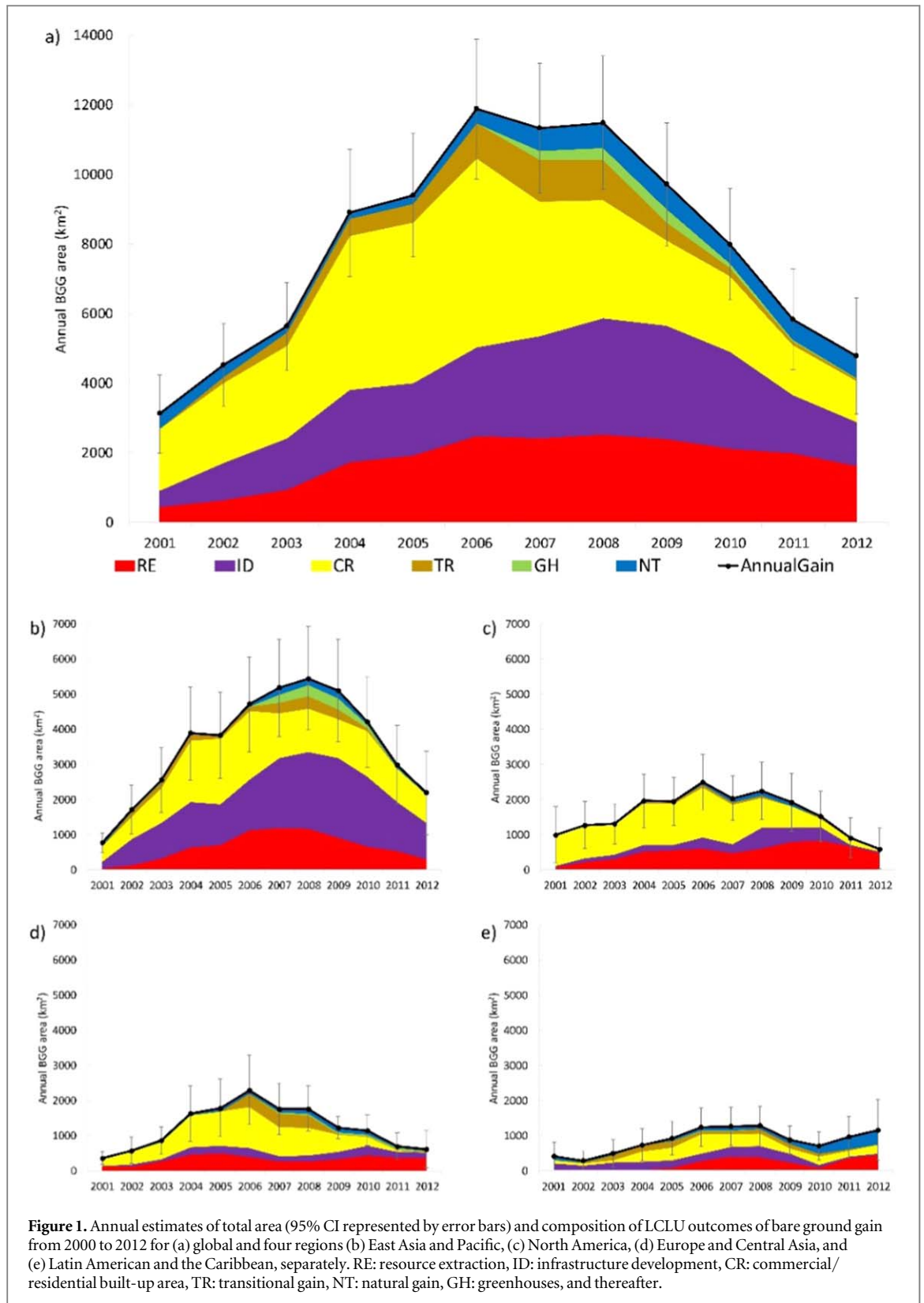
Method and data

Estimation of annual bare ground gain and attribution of LCLU outcomes

Unbiased estimates of areas of annual bare ground gain were produced from a set of probability-based samples (Ying *et al* 2017). Specifically, global land area was stratified by a set of global seamless bare ground gain layers that were produced through automatic classification methods using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) growing season composites between 2000 and 2012. Global annual composites were produced from 654 178 growing season Landsat scenes with per pixel detection of cloud, shadow, snow/ice, water or qualified observation (Hansen *et al* 2013). Ying *et al* (2017) calculated multi-temporal metrics out of the time series of annual composites that were then used to build tree models for classification of bare ground gain. Due to errors inherent with the produced layers of bare ground gain, we did not calculate the areas of bare ground gain from counting the pixels of labeled gain by classification models to avoid biased report of areas of bare ground gain. Instead, we employed those layers as a stratifier to help us efficiently distribute a set of sample for unbiased area estimation in bare ground gain. A total of 5750 sample pixels were selected globally in a stratified random design (25 strata) and then interpreted whether bare ground gain occurred or not (1635 gain pixels versus 4115 no-gain pixels) with time-series of Landsat images, 32 d normalized difference vegetation index (NDVI) time sequences from Google Earth Engine (Gorelick *et al* 2017) and high-resolution images on Google Earth (figure S1). For gain sample points, we decided the change year when NDVI dropped by over 50% and kept low for at least three years following our definition of bare ground gain. We attributed the LCLU outcomes of gain samples through combined information given by the characteristics of spectrum and configuration of Landsat and high-resolution images and even local photos from Google Earth. Recording the change year and attributing the LCLU outcomes of sample pixels that were labeled as bare ground gain enabled us to estimate annual change areas and disaggregate all changes to six components of direct bare ground gain drivers including resource extraction, infrastructure development, commercial/residential built-up area, transitional bare gain, greenhouses and natural bare ground gain (Ying *et al* 2017, Tyukavina *et al* 2018, Zalles *et al* 2019). We estimated the bare ground gain area attributable to the LCLU outcome i in year t in stratum h:

$$\hat{G}_{i,t,h} = A_h p_{i,t,h}, \qquad (1)$$

where A_h is the total area in stratum h, $p_{i,t,h} = n'_{i,t,h}/n_h$ is the sample proportion of pixels interpreted as bare

Figure 1. Annual estimates of total area (95% CI represented by error bars) and composition of LCLU outcomes of bare ground gain from 2000 to 2012 for (a) global and four regions (b) East Asia and Pacific, (c) North America, (d) Europe and Central Asia, and (e) Latin American and the Caribbean, separately. RE: resource extraction, ID: infrastructure development, CR: commercial/ residential built-up area, TR: transitional gain, NT: natural gain, GH: greenhouses, and thereafter.

ground gain of LCLU outcome i in year t in stratum h, $n'_{i,t,h}$ is the number of sample pixels interpreted as bare ground gain of LCLU outcome i in year t, and n_h is the number of sample pixels allocated to stratum h. Then the bare ground gain area of LCLU outcome i in year t is obtained by summing the area estimates over all strata:

$$\hat{G}_{i,t} = \sum_{h=1}^{25} \hat{G}_{i,t,h}. \qquad (2)$$

Area estimates and uncertainty quantification of bare ground gain were performed at global and regional level. To reduce the variation of annual bare ground

gain estimates from interpretation bias of change year due to noise and missing data in NDVI time sequences and historic high-resolution images, sample counts of the average of year t and 1 year neighbors ($t - 1$ and $t + 1$). Land cover change maps derived from remote sensed imagery are subject to omission and commission errors (Olofsson *et al* 2014). However, probability-base samples using maps as strata to target specific LCLU dynamics yield unbiased estimates of change areas (Stehman 2013). The approach of utilizing classified land cover change maps to target sampling of a rare class of interest compared to the overall land surface, in this case bare ground gain, raised the sampling efficiency by eleven times and greatly reduced the standard error in our change estimates (Ying *et al* 2017).

Producing highly accurate maps of bare ground gain is the key to increase sampling efficiency and reduce estimation uncertainty. For a historical data set, we employed the definition of three-year absence of vegetation in remote sensing classification of bare ground gain to eliminate commission errors from agricultural fallow being falsely classified as bare ground gain. Though the annual composite used for detection of bare ground gain was selected from growing season Landsat images, it is possible that for a consecutive three year, reflectance of pre-planting agricultural clear was selected for a pixel due to limited observations in some areas. To better serve as a leading indicator in near-real time, a number of enhancements are possible, including the use of monthly satellite observations of bare ground in place of the annual measure employed here, and land use maps to track ephemeral bare ground gain associated with established land uses such as agriculture and forestry.

Growing trends and business cycles of selected economic variables

Economic development is depicted by economic, trade and energy measures considered relating to ABGG. The World Bank is one of the leading groups for collecting and analyzing data of global economies at global, regional and national level. Indicators including GDP, merchandise imports and exports, energy use and production (table S3) were downloaded from the World Bank database at global and regional level from 2000 to 2012 (The World Bank 2019a). Annual statistics were recorded for 211 countries and regions, and then aggregated to seven regions grouped by geographic locations identified by the World Bank. We converted values of GDP, merchandise imports and exports, which were measured in current US dollars in original data, to constant US dollars in 2010 by adjusting for inflation.

The HP Filter (Hodrick and Prescott 1997) in the R statistical package (Balcilar 2015) was employed to perform economic growth-business cycle decomposition for time series of the five economic variables. The HP filter is arguably the most commonly used mathematical

tool in macroeconomics, especially in real business cycle theory, to separate the time-trend from cyclical component of a time series data (Hodrick and Prescott 1997, Williamson 2002). It is composed of two components. One controls the fitted trend close to the time series, which is measured by the residuals. The other one controls the smoothness of the trend that is measured by the second derivative of the trend. A parameter λ weighs the two components to control the trend between linear and the original time series. The set of λ value depends on data frequency. An optimal trend is the one that gives the minimum sum of the two components. We took natural logarithms of time series variables, removed the trend component, and derived the cyclical component using a λ of 100 for annual data (Backus and Kehoe 1992).

Panel data analysis of bare ground gain dynamics and economic fluctuations

Panel data analysis (Flanagan *et al* 2006) was used in a fixed effect mode to examine the correlation between the cyclical components of each economic variable and the natural logarithms of different combinations of LCLU outcomes of bare ground gain when detangling the unobservable time-invariant heterogeneity associated with each region. The fixed effect model, taking GDP as an example, is

$$\ln \mathrm{GDP}_{rt} = a_r + b * \ln \mathrm{BGG}_{r,\, t-\tau} + \varepsilon_{rt},$$
$$= \tau - 1, 0, 1, 2, \qquad (3)$$

where subscripts r and t denote region and year. In the model, a_r represents the characteristics such as resource endowments, laws and regulatory regimes, or culture that are unique to each region, not change much in a short period of about one decade, but correlated with the predictor variable. The fixed effect model controls for the effect of these characteristics to assess the net effect of changes in bare ground gain areas on the variation in economic variables. We used R 'plm' package for estimation (Yves and Giovanni 2008) and performed Hausman test to justify the selection of the fixed effect model.

We only used data in regions of East Asia and the Pacific, Europe and Central Asia, Latin America and the Caribbean and North America for their relatively low uncertainty of annual bare ground gain estimates. As our baseline year is 2000, the full-time frame of bare ground gain estimates is from 2001 to 2012.

Because our definition of bare ground gain requires an absence of vegetation for continuously three years, edge effects may have influenced in our estimates for the last two years, which could cause underestimation of bare ground gain areas in 2011 and 2012. We tested regression models for both full (2001–2012, table 1, S4) and cut-edge (2001–2010, table S5) time series. There was no significant difference in drawing the conclusion. Both results suggested the one-year leading characteristic of ABGG to economic variables. Nevertheless, the temporal trends

Table 1. Fixed effect regressions of economic variables on the 1 year-lagged sequences of different compositions of LCLU outcomes of bare ground gain (2001–2012).

	GDP lag (1)	Imports lag (1)	Exports lag (1)	Energy use lag (1)	Energy produce lag (1)
Anthropogenic BGG	0.066[a]	0.083[a]	0.082[a]	0.013[b]	0.010
	(0.016)	(0.029)	(0.027)	(0.006)	(0.006)
N	44	44	44	40	40
R-sq	0.31	0.21	0.19	0.13	0.07
F statistic	17.13	10.36	9.35	5.19	2.81
ID + CR + TR	0.055[a]	0.071[b]	0.062[b]	0.014[b]	0.010
	(0.015)	(0.027)	(0.025)	(0.006)	(0.007)
N	44	44	44	40	40
R-sq	0.26	0.15	0.14	0.13	0.068
F statistic	13.96	6.96	6.24	5.21	2.55
ID + CR	0.056[a]	0.072[b]	0.062[b]	0.016[b]	0.010
	(0.016)	(0.028)	(0.026)	(0.006)	(0.007)
N	44	44	44	40	40
R-sq	0.25	0.14	0.13	0.15	0.06
F statistic	12.80	6.58	5.65	6.39	2.13
CR	0.051[a]	0.066[a]	0.055[a]	0.016[a]	0.009
	(0.011)	(0.021)	(0.020)	(0.005)	(0.006)
N	44	44	44	40	40
R-sq	0.33	0.20	0.16	0.22	0.07
F statistic	19.64	9.56	7.50	9.94	2.71
ID	0.014	0.019	0.021	0.003	0.003
	(0.012)	(0.021)	(0.019)	(0.004)	(0.004)
N	44	44	44	40	40
R-sq	0.03	0.02	0.03	0.01	0.01
F statistic	1.37	0.82	1.20	0.46	0.44
RE	0.020[c]	0.036[c]	0.027	0.004	−0.002
	(0.010)	(0.019)	(0.018)	(0.004)	(0.004)
N	42	42	42	38	38
R-sq	0.09	0.085	0.06	0.04	0.01
F statistic	3.64	3.45	2.26	1.37	0.17

[a] Significant at the $p < 0.01$ level.
[b] Significant at the $p < 0.05$ level.
[c] Significant at the $p < 0.1$ level.

were not hindered by the edge effect as the major changes occurred between 2003 and 2010.

Results

Global distributions, trends and compositions of bare ground gain

A probability-based, stratified-random sample (figure S2) was selected from mapped bare ground strata from 2000 to 2012. Each sample was visually interpreted using reference imagery, specifically Landsat Enhanced Thematic Mapper Plus (ETM+) imagery with a moderate spatial resolution (~30 m) and commercial imagery with very high spatial resolution freely viewable on Google Earth. Each sample location of bare ground gain was assigned to a type of LCLU outcome and to a year of initial vegetation removal (figure S1).

From 2000 to 2012, global bare ground gain averaged 7881 km^2 yr^{-1}. That is, every year a land area close to the size of the Yellowstone national park semi/permanently lost its vegetation cover. About half of this average rate, 3550 km^2 yr^{-1}, was in the East Asia and the Pacific region, and the smallest portion, 154 km^2 yr^{-1}, was in the Middle East and North Africa. The temporal trend of total bare ground gain globally is unimodal (figure 1(a)), as are those for all regions (figures 1(b)–(e), figure S3), with peaks from 2006 to 2008, except Sub-Saharan Africa which is bimodal. Global year on year bare ground gain increased from 3118 ± 1132 km^2 in 2001 to a peak of 11 878 ± 2014 km^2 in 2006 with a growth rate of 31% per year on average and fell by 60% of the peak to 4772 ± 1673 km^2 in 2012. Total bare ground gain in East Asia and the Pacific increased fivefold from 2001 to 2004, and then slowed between 2005 and 2008. Latin America and the Caribbean was unique in its recovery of a positive trend in bare ground gain after 2010.

The greatest LCLU outcome of bare ground gain in most regions was commercial/residential built-up area (49% in North America, 44% in East Asia and the Pacific, 29% in Latin America and the Caribbean), followed by resource extraction (North America 32%, Europe and Central Asia 26%, Latin America and the Caribbean 23%). East Asia and the Pacific differed, where infrastructure was the largest (39%), followed

by commercial/residential built-up area (34%). Different LCLU outcomes of bare ground gain, however, showed different patterns of peak time and change rate (figure 1, figure S4). The trends in the expansion of commercial/residential built-up area varied among the four regions in the years following their 2006 peaks (figure S4). For example, those in North America and Europe and Central Asia gradually declined, that in East Asia and the Pacific temporarily stabilized in 2007 through 2010 and then resumed its decline, and that in Latin America and the Caribbean appeared to begin to recover beginning in 2010 (figure S4). New infrastructure development generally peaked in 2008–2010. A shorter cycle of transitional bare ground gain appeared in each region following the decline of commercial/residential built-up area. Growth in resource extraction resumed about two years after the peak in each region except East Asia and the Pacific and was the source of the recovery in overall bare ground gain in Latin America and the Caribbean.

Gain in built-up area foreshadowed the Great Recession

The cyclic patterns of global ABGG foreshadowed the decade's macroeconomic fluctuations dominated by the 2007–2008 global financial crisis. The rise in ABGG, driven mainly by commercial/residential built-up area, transitioned to a decline prior to the 2008 crash (figure 3(a)). Furthermore, inter-annual ABGG and commercial/residential built-up area were both significantly correlated to the de-trended global GDP (figures 3(b), (c)).

To explore how these patterns were related to the economic activities during this study period, we carried out panel data regressions of fluctuations on individual economic variables versus different combinations of LCLU outcomes of ABGG, and their leading or lagging terms. GDP was used in the analysis because among the four components of expenditure, investment is most related with infrastructure development and commercial/residential built-up area, whereas consumption and net exports are more associated with all LCLU outcomes of ABGG. Merchandise exports and imports were included to further account for domestic land use and displacement (Lambin and Meyfroidt 2011; Yu *et al* 2013). Energy use and production partly accounted for fossil fuel extraction, one major component of bare ground gain from resource extraction. Merchandise imports and exports, and energy use are significantly correlated with GDP (figure S7).

Panel analysis indicates that trends in regional, ABGG are significant leading indicators, by one year, of the tested economic variables, especially GDP (table 1, S4, figure 4). Alternating the lag length of bare ground gains, panel regressions show that commercial/residential built-up area leads GDP by one year or two years at the highly significant levels (figure 4).

Comparing different compositions of LCLU outcomes of ABGG, panel regressions show that commercial/residential built-up area leading GDP yield the highest r^2 values (figure 4). A 10% increase in ABGG in an antecedent year is associated with a growth in the following year of 0.6% ($\pm$0.2%) for GDP, 1% ($\pm$0.3%) for merchandise imports, and 0.9% ($\pm$0.3%) for merchandise exports (table 1). For one-year lagged terms ($t-1$), total ABGG, as well as gains in commercial/residential built-up area, alone or combined with those in infrastructure development and transitional land, are all significantly correlated with GDP, merchandise imports and exports, and energy use (table 1), showing a predictive capability in economic changes. Compared to resource extraction and infrastructure development, one-year lagged term of commercial/residential built-up area outperforms the coincident term with an increase in r^2 of 0.21 (table 1, S4).
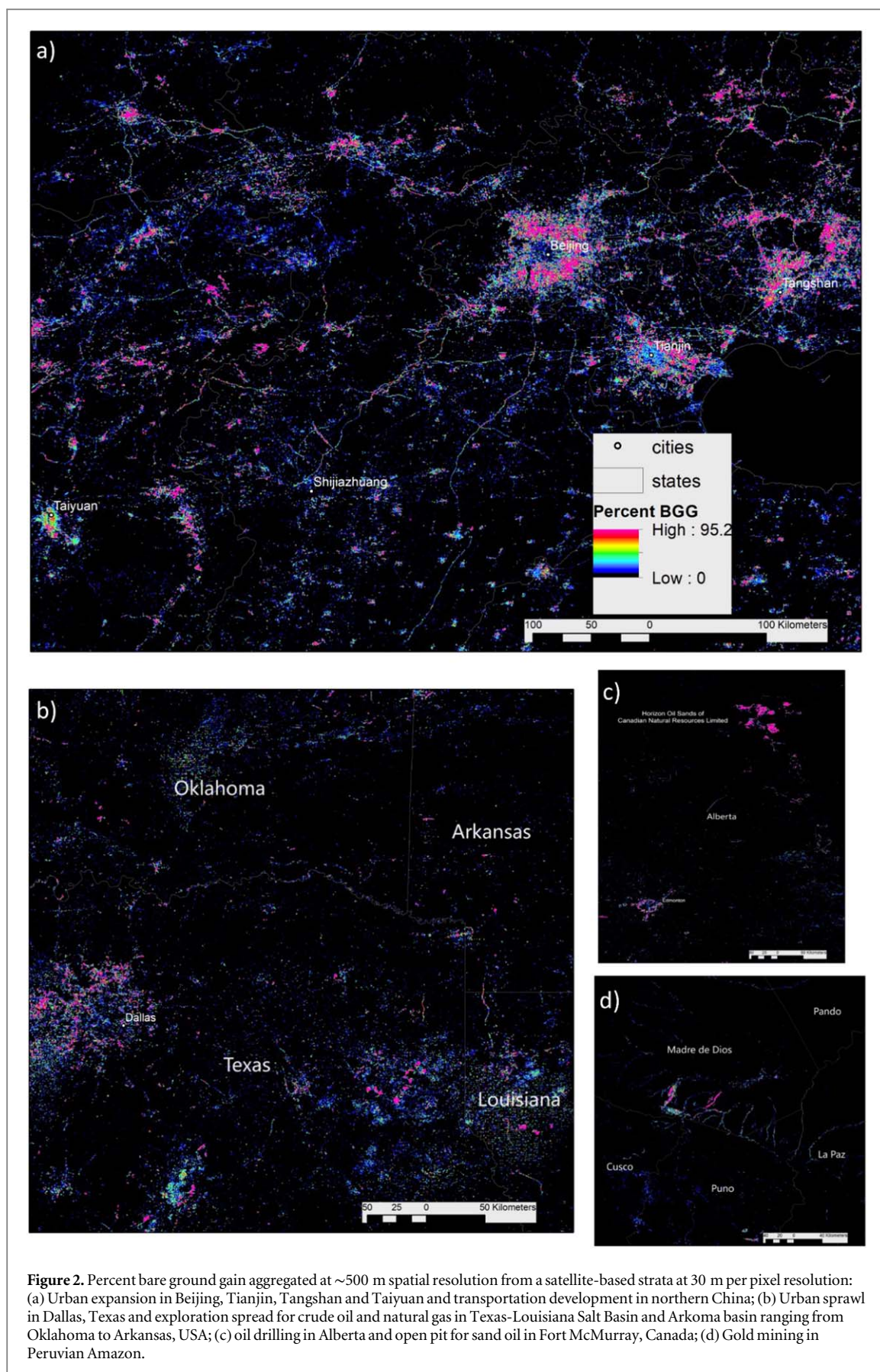
Leading indicators performed differently among different regions. For example, commercial/residential built-up area peaked in 2006 in Europe and Central Asia and Latin America and the Caribbean (figures 1(d)–(e), figure S4), two years earlier than the GDP peak in these regions (figure S6). Also, the magnitude of changes in ABGG was modest compared to the magnitude of GDP changes. For example, the decrease of newly commercial/residential built-up area in 2008 was weak compared to the plunge in the GDP of Europe and Central Asia in 2009 (figure S8).

Discussion

Link between gain in built-up area and regional economic activities

As with the global trend, ABGG and gains from commercial/residential built-up area in North America and Europe and Central Asia, where the economies were first hit by the subprime mortgage crisis in 2007 and substantial European debt crisis, all peaked in 2006, earlier than these financial crises. According to the Standard & Poor/Case-Shiller Composite Home Price Index, a measure of the aggregate market for single family homes in 10 U.S. major cities, the real estate market entered a price boom in late 1990s, and abruptly turned down after mid-2006 (Shiller 2008). Commercial/residential built-up area is closely related to the housing market, thus the trends in this component of ABGG mirrored the market, immediately suggesting a downturn in investment in real estate, which is further reflected in macroeconomic accounts.
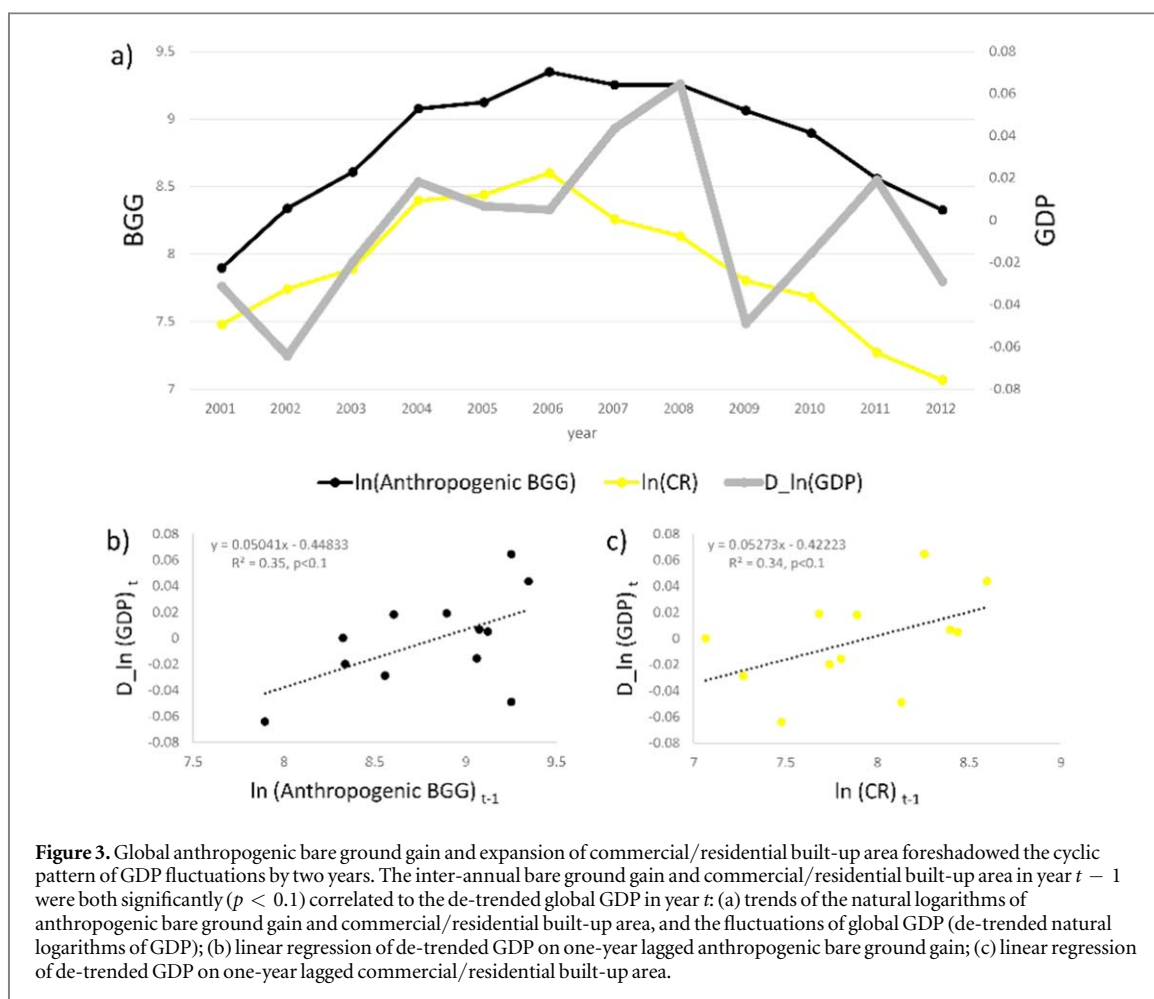
The temporal dynamics of commercial/residential built-up area, infrastructure development and transitional bare ground gain were inherently coupled over our study period. For example, the peak in infrastructure gain lagged that of commercial/residential built-up area by two years. This suggests that flexible housing markets are more sensitive to economic

Figure 2. Percent bare ground gain aggregated at ∼500 m spatial resolution from a satellite-based strata at 30 m per pixel resolution: (a) Urban expansion in Beijing, Tianjin, Tangshan and Taiyuan and transportation development in northern China; (b) Urban sprawl in Dallas, Texas and exploration spread for crude oil and natural gas in Texas-Louisiana Salt Basin and Arkoma basin ranging from Oklahoma to Arkansas, USA; (c) oil drilling in Alberta and open pit for sand oil in Fort McMurray, Canada; (d) Gold mining in Peruvian Amazon.

changes than infrastructure projects are, the latter typically requiring more planning and equipment (Mok *et al* 2015). Gain in transitional land peaked between 2006 and 2008, likely a reflection of ceased construction and other projects during the onset of the financial crisis. Decreases in every LCLU outcome of

Figure 3. Global anthropogenic bare ground gain and expansion of commercial/residential built-up area foreshadowed the cyclic pattern of GDP fluctuations by two years. The inter-annual bare ground gain and commercial/residential built-up area in year $t - 1$ were both significantly ($p < 0.1$) correlated to the de-trended global GDP in year t: (a) trends of the natural logarithms of anthropogenic bare ground gain and commercial/residential built-up area, and the fluctuations of global GDP (de-trended natural logarithms of GDP); (b) linear regression of de-trended GDP on one-year lagged anthropogenic bare ground gain; (c) linear regression of de-trended GDP on one-year lagged commercial/residential built-up area.
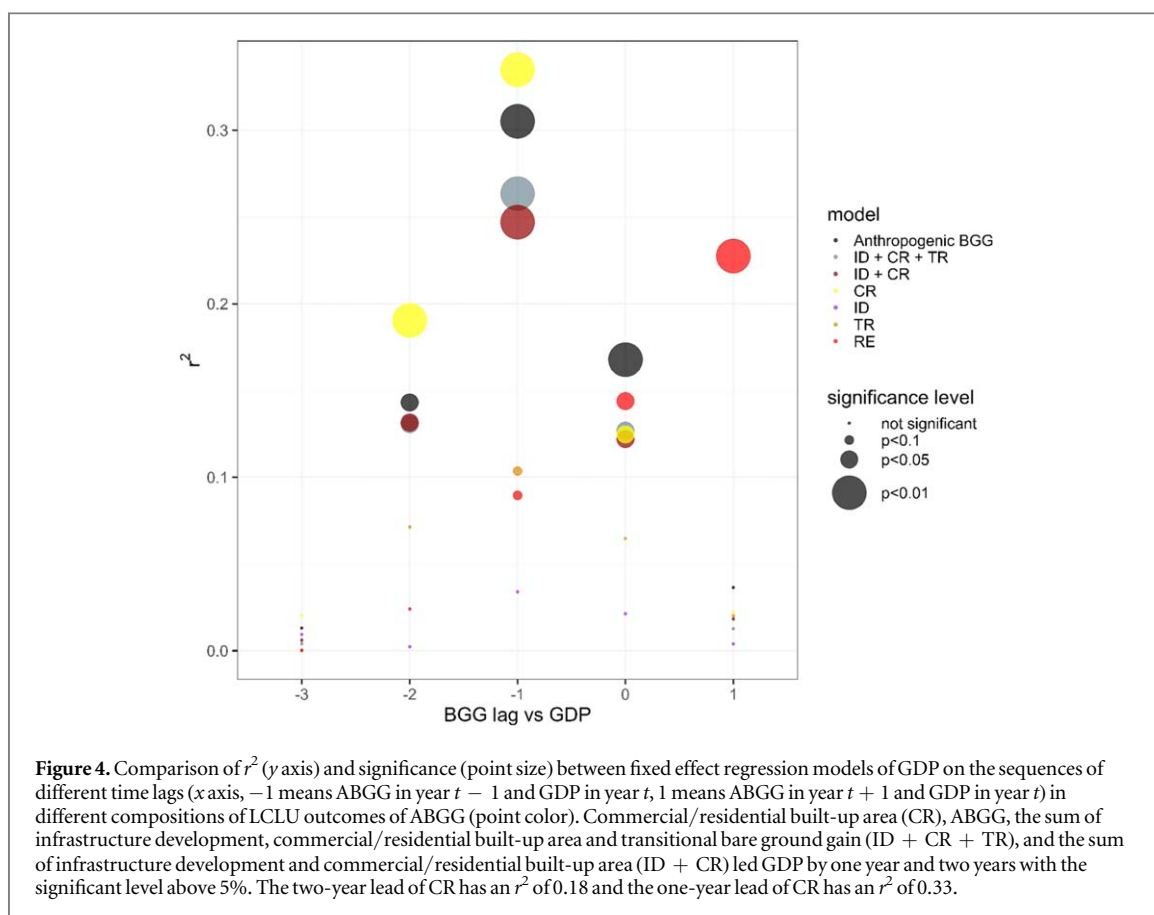
ABGG occurred after 2008 when the financial crisis triggered wide impacts on all major sectors of economies.

East Asia and the Pacific accounted for 45% of global bare ground gain, 78% of which was in China. China has experienced an excessive rural-urban migration since the economic reform. The urbanization rate increased from 36% in 2000 to 52% in 2012, while the average urban household income grew fourfold (The World Bank 2019b). To accommodate massive urban in-migration, China carried out urban housing reform, pushing the provision of urban housing from a welfare to market-oriented system (Chen *et al* 2011). The demographic, economic and institutional changes resulted in an average gain of 976 km^2 yr^{-1} in commercial/residential built-up area (figure 2(a) and yellow samples in figure S2). Infrastructure development expanded even greater at 1233 km^2 yr^{-1}. China overtook the US in 2007 and Germany in 2009 to become the world's largest exporter since its accession to the World Trade Organization (WTO) in 2001. The enormous boom in manufacture plants in tandem with fast growing infrastructure investment, e.g. in transportation (figure 2(a)), has facilitated the relocation of populations from inland rural villages to coastal cities engaged in the global marketplace (Wang *et al* 2012, Song *et al* 2018).

The increase in bare ground gain from resource extraction in North America post 2008 coincides with an expansion of shale-gas projects and new crude-oil exploration (red samples in figure S2 and figures 2(b), (c)). Led by new technologies of hydraulic fracturing and horizontal drilling, shale gas extraction has developed quickly in the US and spread to Canada and other continents. The number of horizontal wells in US alone doubled between 2008 and 2012 (US Energy Information Administration (EIA) 2018). Local resource extraction driven by demand in distance and trade in global markets affected domestic and global investment and outcome in resources sector (Sonter *et al* 2014). For example, Australia, where minerals are the largest export, had bare ground gain in resource extraction accounting for 60% of its total bare ground gain. Gold mining in Madre de Dios of Peru (figure 2(d)), driven by a constant rate (~18%) of increasing gold prices (Swenson *et al* 2011, Asner *et al* 2013), was a source of the upward trend post-financial crisis in Latin America and the Caribbean.

Implications for near-real time monitoring of global and regional economic health

In the long run, the accumulated change of bare ground from anthropogenic demands is driven by population growth and economic development

Figure 4. Comparison of r^2 (y axis) and significance (point size) between fixed effect regression models of GDP on the sequences of different time lags (x axis, -1 means ABGG in year $t - 1$ and GDP in year t, 1 means ABGG in year $t + 1$ and GDP in year t) in different compositions of LCLU outcomes of ABGG (point color). Commercial/residential built-up area (CR), ABGG, the sum of infrastructure development, commercial/residential built-up area and transitional bare ground gain (ID + CR + TR), and the sum of infrastructure development and commercial/residential built-up area (ID + CR) led GDP by one year and two years with the significant level above 5%. The two-year lead of CR has an r^2 of 0.18 and the one-year lead of CR has an r^2 of 0.33.

(Seto *et al* 2012, Verburg *et al* 2004). Our results also confirmed the coincident positive correlation between ABGG and changes in GDP at annual intervals (table S4). Nevertheless, short-term factors that cause fluctuations in GDP, such as market anticipation, monetary system, technology innovation and policy decision, also affect annual changes in the quantity, attributes and spatial allocation of new bare ground. More importantly, the ability of the satellite-detected built-up area changes to signal economic recession with a one-year lead can effectively help policy makers to initiate counter-recession measures in a much more timely manner compared with the current policy-making practice. This ability can also help financial institutions and specialists to make much more informative investment decisions, and help the public to better prepare for economic recessions. Thus, it is important to implement near-real time monitoring of bare ground gain at global and regional scales.

The ongoing earth observation programs of Landsat and Sentinel 2 satellites enable near-real time monitoring of land change at large scale as exemplified by an alert system of forest disturbance in operation on a weekly basis (Hansen *et al* 2016). The surge of CubeSat (Hand 2015) technology also provides high resolution images that are especially important for the validation and land-use attribution of bare ground gain. The latency of confirming bare ground gain and LCLU change attribution may be facilitated by high resolution observations and contextual inference. Our approach is scalable and bridges the relationship between socio-economics and land-use change at national or local scale, with a potential for operational implementation.

Data availability

Any data that support the findings of this study are included within the article. Sample reference images and plots are openly available at https://glad.geog.umd.edu/bare-ground-gain-as-leading-economic-indicator.

ORCID iDs

Qing Ying ⓘ https://orcid.org/0000-0002-9752-8973
Matthew C Hansen ⓘ https://orcid.org/0000-0003-0042-2767
Laixiang Sun ⓘ https://orcid.org/0000-0002-7784-7942

References

Asner G P, Llactayo W, Tupayachi R and Luna E R 2013 Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring *Proc. Natl Acad. Sci.* **110** 18454–9

Backus D K and Kehoe P J 1992 International evidence on the historical properties of business cycles *Am. Econ. Rev.* **82** 864–88

Balcilar M 2015 Miscellaneous Time Series Filters (https://cran.r-project.org/package=mFilter)

Bandholz H and Funke M 2003 In search of leading indicators of economic activity in Germany *J. Forecast.* **22** 277–97

Bennett M M and Smith L C 2017 Advances in using multitemporal night-time lights satellite imagery to detect, estimate, and monitor socioeconomic dynamics *Remote Sens. Environ.* **192** 176–97

Blumenstock J, Cadamuro G and On R 2015 Predicting poverty and wealth from mobile phone metadata *Science* **350** 1073–6

Chen J, Guo F and Wu Y 2011 One decade of urban housing reform in China: Urban housing price dynamics and the role of migration and urbanization, 1995–2005 *Habitat Int.* **35** 1–8

Chen X and Nordhaus W D 2011 Using luminosity data as a proxy for economic statistics *Proc. Natl Acad. Sci.* **108** 8589–94

Elvidge C D, Baugh K E, Kihn E A, Kroehl H W, Davis E R and Davis C W 1997 Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption *Int. J. Remote Sens.* **18** 1373–9

Flanagan W, McIntosh C N, Le Petit C and Berthelot J-M 2006 Deriving utility scores for co-morbid conditions: a test of the multiplicative model for combining individual condition scores *Population Health Metr.* **4** 13

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D and Moore R 2017 Google earth engine: planetary-scale geospatial analysis for everyone *Remote Sens. Environ.* **202** 18–27

Hand E 2015 Startup liftoff *Science* **348** 172–7

Hansen M C, Krylov A, Tyukavina A, Potapov P V, Turubanova S, Zutta B, Ifo S, Margono B, Stolle F and Moore R 2016 Humid tropical forest disturbance alerts using Landsat data *Environ. Res. Lett.* **11** 34008

Hansen M C *et al* 2013 High-resolution global maps of 21st-century forest cover change *Science* **342** 850–3

Henderson J V, Storeygard A and Weil D N 2012 Measuring economic growth from outer space *Am. Econ. Rev.* **102** 994–1028

Hodrick R J and Prescott E C 1997 Postwar US business cycles : an empirical investigation *J. Money, Credit Bank* **29** 1–16

Huang M Y, Rojas R R and Convery P D 2018 News Sentiment as Leading Indicators for Recessions (arXiv:1805.04160)

Jean N, Burke M, Xie M, Davis W M, Lobell D B and Ermon S 2016 Combining satellite imagery and machine learning to predict poverty *Science* **353** 790–4

Lambin E F and Meyfroidt P 2011 Global land use change, economic globalization, and the looming land scarcity *Proc. Natl Acad. Sci.* **108** 3465–72

Li L, Goodchild M F and Xu B 2013 Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr *Cartogr. Geogr. Inf. Sci.* **40** 61–77

Liu Y, Liu X, Gao S, Gong L, Kang C, Zhi Y, Chi G and Shi L 2015 Social sensing: a new approach to understanding our socioeconomic environments *Ann. Assoc. Am. Geogr.* **105** 512–30

Mok K Y, Shen G Q and Yang J 2015 Stakeholder management studies in mega construction projects: a review and future directions *Int. J. Proj. Manage.* **33** 446–57

Olofsson P, Foody G M, Herold M, Stehman S V, Woodcock C E and Wulder M A 2014 Good practices for estimating area and assessing accuracy of land change *Remote Sens. Environ.* **148** 42–57

Orphanides A 2003 Monetary policy evaluation with noisy information *J. Monet. Econ.* **50** 605–31

Seto K C, Guneralp B and Hutyra L R 2012 Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools *Proc. Natl Acad. Sci.* **109** 16083–8

Shiller R J 2008 Historic turning points in real estate *East. Econ. J.* **34** 1–13

Song X-P, Hansen M C, Stehman S V, Potapov P V, Tyukavina A, Vermote E F and Townshend J R 2018 Global land change from 1982 to 2016 *Nature* **560** 639–43

Sonter L J, Moran C J, Barrett D J and Soares-Filho B S 2014 Processes of land use change in mining regions *J. Clean. Prod.* **84** 494–501

Stehman S V 2013 Estimating area from an accuracy assessment error matrix *Remote Sens. Environ.* **132** 202–11

Swenson J J, Carter C E, Domec J-C and Delgado C I 2011 Gold mining in the Peruvian Amazon: global prices, deforestation, and mercury imports *PLoS One* **6** e18875

The World Bank 2019a World Bank Open Data | Data (WWW Document) (https://data.worldbank.org/) (Accessed: 10 September 2019)

The World Bank 2019b Urban population (% of total) | Data (WWW Document) (https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS) (Accessed: 25 January 2019)

Tyukavina A, Hansen M C, Potapov P, Parker D, Okpa C, Stehman S V, Kommareddy I and Turubanova S 2018 Congo Basin forest loss dominated by increasing smallholder clearing *Sci. Adv.* **4** eaat2993

US Energy Information Administration (EIA) 2018 (WWW Document) (https://eia.gov/petroleum/wells/) (Accessed: 29 August 2018)

Verburg P H, Schot P P, Dijst M J and Veldkamp A 2004 Land use change modelling: current practice and research priorities *GeoJournal* **61** 309–24

Wang L *et al* 2012 China's urban expansion from 1990 to 2010 determined with satellite remote sensing *Chinese Sci. Bull.* **57** 2802–12

Williamson S D 2002 *Macroeconomics* (London: Pearson)

Ying Q, Hansen M C, Potapov P V, Tyukavina A, Wang L, Stehman S V, Moore R and Hancher M 2017 Global bare ground gain from 2000 to 2012 using Landsat imagery *Remote Sens. Environ.* **194** 161–76

Yu Y, Feng K and Hubacek K 2013 Tele-connecting local consumption to global land use *Glob. Environ. Change* **23** 1178–86

Yves C and Giovanni M 2008 Panel data econometrics in R: the plm package *J. Stat. Softw.* **27** 1–43

Zalles V *et al* 2019 Near doubling of Brazil's intensive row crop area since 2000 *Proc. Natl Acad. Sci. USA* **116** 428–35