Landcover Data Limits Our Understanding of Global Change

Lyndon Estes *, **Peng Chen** †, **Stephanie Debats** * , **Tom Evans** † , **Fanie Ferreira** ‡, **Gabrielle Ragazzo** * , **Justin Sheffield** * **Adam Wolf** * and **Kelly Caylor** *

*Princeton University, Princeton, NJ USA,†Indiana University, Bloomington, IN USA, and ‡GeoTerraImage, Pretoria, RSA

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Our knowledge of global change processes is built upon landcover data, but we lack the means to comprehensively assess landcover maps, particularly in rapidly developing regions, and thus do not fully comprehend how map quality impacts our understanding of global change. We used a unique, high quality, national-scale cropland map to assess the bias and accuracy in a) current generation landcover products, and b) analyses built on these. Maps have high bias and low accuracy, particularly at finer resolutions (1-10 km), which transfers into, and is often amplified, in downstream analyses of vegetative carbon stock, gridded crop production, and process-based food security models. Substantial map aggregation is needed to avoid potential misunderstandings of global change processes or misguided policy. Maps derived from coarse resolution sensors (such as MODIS and GlobCover) should be aggregated to at least 50-100 km resolution (and sometimes more). Map accuracy is lowest at intermediate levels of cover, thus coarser aggregation may be necessary in highly mixed landscapes. Higher resolution imaging systems (e.g. Landsat) produce maps that require less aggregation (1-25 km) to achieve high accuracy and low bias. For regional to global analyses, new fusion products, such as the GeoWiki cropland map, need just 5 km of aggregation to produce unbiased maps for most applications, but can require ≥25 km of aggregation to have sufficient accuracy. Widely used global datasets for carbon, crop area, and crop yield should be rebuilt on GeoWiki and similar maps, while new approaches methods that fuse human judgement, the latest computer vision-based algorithms, and growing high resolution image archives are needed to produce the next generation of landcover maps.

landcover | bias | remote sensing | agriculture | crop yield | harvested area | carbon | agent-based model | landscape

Abbreviations: GTI, GeoTerraImage; SSA, sub-Saharan Africa

The functioning of the Earth System is fundamentally connected to the characteristics of landcover (1), the physical constituents of the terrestrial surface. Human endeavors are strongly governed by and shape landcover, whether it be felling forests for timber or burning savannas to rejuvenate forage, and landcover is a primary factor in many climate and biogeochemical processes (2). Our increasing modification of the Earth's surface (2) means that socioeconomic and physical processes increasingly interact through landcover. To fully understand these processes and the nature of global change, it is essential to know the nature and distribution of landcover.

This importance is understood by a growing number of social, economic, and natural scientists, who are using landcover data to advance understanding of food security (3–5), carbon cycling (6, 7), biodiversity loss (8, 9), demographic shifts (10), and other important facets of global change.

The value of the insights resulting from such studies depends upon the veracity of their underlying landcover data, much as a house requires a solid foundation in order to remain standing. Unfortunately, the evidence to date indicates that much of our understanding of global change is built on shaky foundations. The reason for this is that landcover data can only practically be derived from satellite imagery, and in many regions the cover types of interest are smaller (e.g. smallholder's farms 11) than the sensor resolution, or spectrally indistinct from neighboring covers, which translates into substantial mapping errors (12–14). Landcover maps are therefore often inaccurate at finer scales and disagree widely between products, particularly in the world's most rapidly developing regions (15–17). These errors mean that we are still unable to obtain the granular, mechanistic understanding of global change processes that we need.

These problems with landcover products are known (12, 16–19), and there are a variety of map improvement efforts underway, particularly for agriculture (14, 20). What remains an open question is exactly how much the maps researchers typically use deviate from actual landcover, and how this in turn impacts our understanding of global change processes. Answering this question depends on having spatially comprehensive ground truth data, which are unavailable for most parts of the world, particularly over Africa and other developing regions (12). Our understanding of map accuracy is therefore built primarily on bottom-up tests made with a relatively small number of ground truth points (relative to the total mapped area), or from top-down "sanity checks" made in comparison to aggregated survey data. This allows us to quantify between-map discrepancies (e.g. 16, 21), or to understand map fidelity to country-level statistics (e.g. 16), but offers little direction for how to arrive at a true number.

Reserved for Publication Footnotes

Being unable to fully quantify the errors in landcover maps makes it difficult, if not impossible, to gauge their impact on downstream analyses. There has been some work examining how such error influences climate simulations (22), agricultural land use patterns (23), carbon flux measurements (24), and human population estimates (10), but these either use simulated landcover errors (22) or compare relevant differences in estimates between different satellite-derived landcover maps (10, 24). One exception is a Belgian study (23) that used ground-collected farm parcel data to assess how landcover errors bias measurements of agricultural land use patterns, but the study extent was fairly small and the validation data were discontiguous.

Fortunately, the recent, explosive growth in public and private initiatives to develop new Earth observing capabilities, which range from small drones[1] to new high resolution satellite arrays[2] and better mapping methods (15, 20, 25), are finally providing the means to comprehensively interrogate the accuracy and biases in the landcover products that have become commonplace in global change research–and which are often used to make policy decisions (26).

In this study, we take advantage of these recent advances to address the urgent need to more thoroughly and precisely quantify landcover map errors and how they might impact our understanding of change in the world's most dynamic regions. We use a unique, high accuracy landcover map of South African crop fields to conduct a spatially comprehensive, bottom-up quantification of error in several widely used landcover maps, and how these errors can propagate through "downstream" studies investigating into both the physical and socioeconomic components of global change. Our objective is to provide scientists and policy-makers who use landcover data with a better, more up-to-date, understanding of their appropriate uses and limitations.

Overview of study area and analyses

In the late 2000s, the South African government commissioned a cropland map that was made by manually interpreting and digitizing fields visible within high resolution satellite imagery (27). The resulting vectorized field boundaries provide highly accurate data on field sizes and distribution for the period 2009-2011. This dataset is particularly valuable because South Africa represents nearly 6% of sub-Saharan Africa's (SSA) area, which is a region that is poised to undergo rapid agricultural expansion (26), yet is notably lacking trustworthy maps of existing agriculture land (16). Moreover, South Africa's large agricultural sector represents the diversity of farming systems found throughout SSA, ranging from large commercial operations to smallholder farms (28, 29).

We used this dataset as a reference layer for evaluating four landcover products representative of the type commonly used in Earth Systems research. The first was South Africa's own 30 m resolution 2009 National Landcover map (SA-LC)(30), which is typical of the higher-resolution, Landsat-based maps that are typically available only for individual countries (e.g. 31). The second and third were respectively the 300 m GlobCover 2009 (32) and 500 m resolution 2011 MODIS Landcover products, which are widely used global-scale products (e.g. 33, 34). The fourth dataset was the new 1 km GeoWiki hybrid-fusion cropland map for Africa (18), which incorporates the first three datasets and represents the current state-of-the-art in landcover mapping. For comparison, we converted all five datasets into 1 km gridded maps of cropland density, expressed as the pixel-wise percent cover.

In our first set of comparisons, we evaluated map quality by subtracting each of the four landcover product-derived percent cropland maps (hereafter we refer to these as the "the test maps") from the reference map, so that we could calculate each map's bias (the mean pixel-wise error) and accuracy (the mean absolute pixel-wise error, where a lower value indicates higher accuracy). We performed this analysis for the original 1 km resolution, and for maps that were further aggregated to 5, 10, 25, 50, and 100 km, in order to evaluate how error and bias changes with scale. We also assessed how landcover pattern impacts map performance by modeling the correlation between map accuracy and cropland density.

We then used these maps to conduct four further analyses typical of Earth System science research. The first two related to physical processes, namely the calculation of vegetative carbon stocks and the simulation of evapotranspiration by a hydrological model. The second two were socio-economic in nature: the estimation of agriculture yield and production, and an agent-based model assessment of household food security. The first and third analyses were relatively simple, in that the variable(s) of interest were mapped onto landcover using empirical relationships. The second and fourth relied on more complex numerical methods, where landcover was one of several variables needed to run each model. For the simpler analyses, we examined how results were influenced by map aggregation, while for the more complex cases, our assessments were confined to each numerical model's standard output resolution.

Map quality

Bias and accuracy. Our reference map indicated that crop fields covered 104,304 km^2, or nearly 10%, of the total mapped area in the 2009-2011 time period. The test maps derived from SA-LC and GeoWiki overestimated this area by 31 and 10%, respectively, while GlobCover and MODIS underestimated it by 18 and 23%. Subtracting each test map from the reference maps created pixel-wise residuals, where negative and positive values respectively represent overestimates and underestimates by the test map (Fig. 1A).

The most pronounced errors were in the MODIS and GlobCover maps, which showed large positive residuals in the center of the country where cropland is most concentrated (blue areas in Fig. 1A), and negative residuals (red areas) along the eastern and northern margins. These patterns translated into substantial map bias (Fig. 1.B), with GlobCover and MODIS mean error (weighted by the reference cropland density) exceeding 45% and 25% respectively at 1 km resolution, meaning that each map tends to underestimate cropland by that amount at that resolution. This bias declined with each level of map aggregation, being reduced to nearly 15% for GlobCover and 5% for MODIS at 100 km. The magnitude of (density weighted) mean absolute error (MAE) was somewhat higher in all cases. The GeoWiki map, in contrast, was the least biased overall, showing just a 7% bias at 1 km and near 0 for all other scales of aggregation, although its accuracy (23% MAE) was only half as good as SA-LC's at 1 km (11% MAE), which despite its uniform overestimation bias (Fig. 1A) was the most accurate map at aggregation scales < 10km. Above this, GeoWiki became slightly more accurate, having <5% MAE at 100 km resolution. The reason GeoWiki had relatively poor accuracy at 1 km resolution was due to the heterogeneity of residuals, which traded between positive and negative residuals over short distances, thereby inflating MAE at this scale.

Error and landcover density. Given that landcover characteristics can influence landcover map quality (12, 14, 25), we set

[1] e.g. 3DRobotics, DJIA

[2] Planet Labs, Skybox

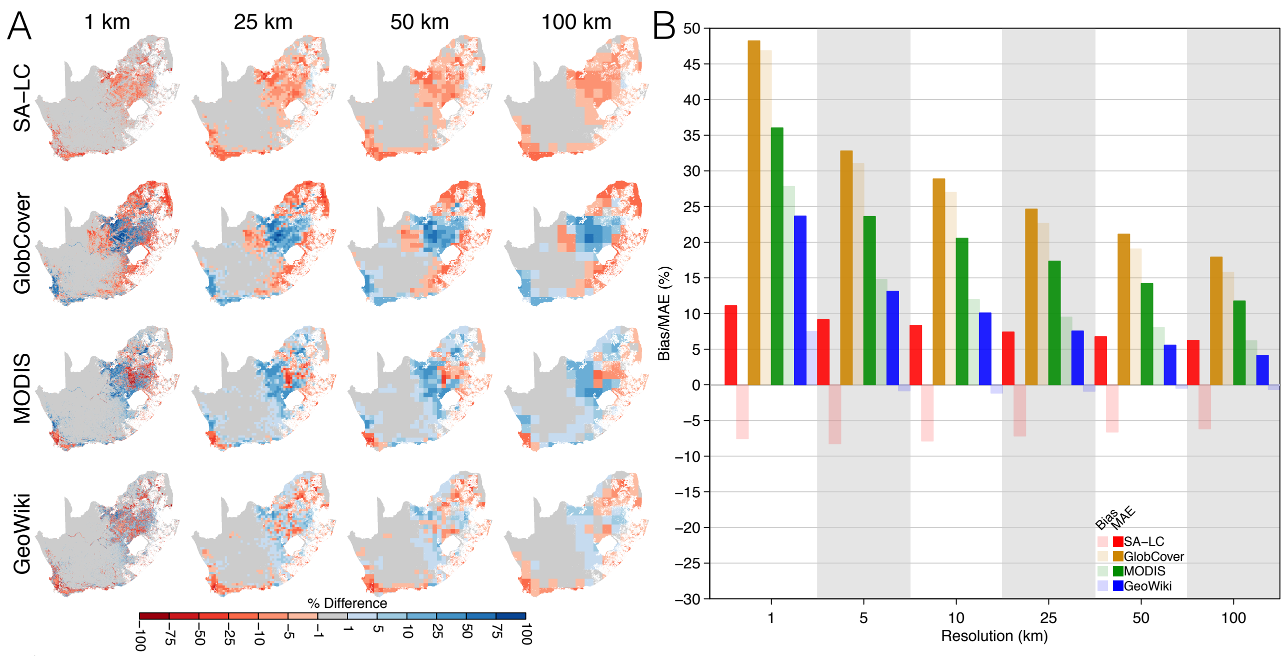

Fig. 1. (A) Errors in the percent cropland estimates resulting from each of the four test maps relative to the reference map at different scale of pixel aggregation. Rows indicate the test map being assessed (by subtraction from the reference map), while columns refer to resolution of aggregation. White indicates areas where areas under communal farmlands or permanent tree crops were removed from analysis. (B) The bias (mean error) and accuracy (mean absolute error [MAE]) of each test map at each scale of aggregation, weighted by the percentage cropland in each cell of the reference map. Bias estimates are indicated by the semi-transparent bars, accuracy (lower is more accurate) by the solid bars, with bar colors coded to specific cropland maps.

out to elucidate the relationship between cropland configuration and map error. We first calculated the average 1 km reference cropland percentage within each of South Africa's 354 magisterial districts (the finest administrative unit, averaging 3,445 km^2; SI), which yielded a landscape-scaled metric of characteristic cropland density, and similarly calculated the district-wise MAE for each test map.

We then used a generalized additive model to evaluate the shape of the relationship between district MAE and cropland density. The model shows that map accuracy is typically lowest at intermediate levels of cropland density (50-60% cover) for all but the GlobCover map (where accuracy continues to decline with cropland cover), and is highest when the landscape is dominated either by cropland or by another type (Fig. 2). In other words, accuracy is generally lowest when cropland cover is mixed evenly with other cover types. GlobCover's accuracy continued to decrease with cropland density because the dominant agricultural cover class contributing to the test map was defined as 50-70% crops mingled with other vegetation, thus the maximum percentage was constrained by this mixture range.

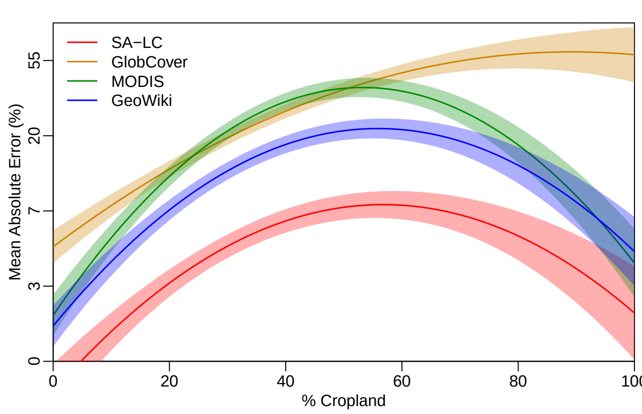

Fig. 2. The relationship between map accuracy (the mean absolute error) in test maps and the actual cropland cover within agricultural landscapes (reference map pixels having >0.5% cropland), here defined by the boundaries of magisterial districts (n = 345), as fit with a generalized additive model. Prediction curves are color-coded to the different test maps, with the solid line indicating predicted absolute bias, and the lighter shading the standard error of the coefficients.

The impact of map error on physical analyses

Estimating vegetative carbon stocks. To understand the carbon cycle and climate forcing due to land use change, it is important to have accurate, high resolution maps of vegetative carbon stocks (26). One widely used vegetative carbon dataset is that of (35), who mapped estimated carbon density values for different vegetation types to the classes of a global landcover product. The resulting data were intended to provide a baseline for climate policy by the Intergovernmental Panel on Climate Change (IPCC), as well as input to other land use and biogeochemical analyses (35).

We followed this method to create vegetative carbon maps for South Africa. Since our maps represented cropland percentage, we developed several variants by assigning the carbon densities of different vegetation types (forest, secondary forest, shrubland, grassland, and sparse vegetation (35)) to the non-cropland fraction of our maps. These hypothetical maps represented the range in potential carbon densities, and allowed us to investigate how carbon estimates can vary as a function of i) test map errors and ii) the properties of neighboring cover types. To assess carbon estimation error, we subtracted test map-derived carbon maps from those based on the reference map, and calculated bias and accuracy scores.

The spatial patterns of test map errors transmitted into carbon estimation errors, but the sign varied as a function of the density of carbon adjacent to croplands (SI). Where cropland was underestimated and the surrounding cover type was more carbon dense than cropland, carbon density was overestimated, but when the cover type was less dense than croplands (e.g. sparse vegetation), then carbon density was underestimated. The inverse was true where cropland was overestimated.

The magnitude of carbon errors varied as a function of the carbon density of surrounding cover, as demonstrated by the bias statistics (Fig. 3). Bias was near zero when grassland was the adjacent cover type (SI), as its carbon density is nearly the same as cropland. However, when forest was adjacent then bias was a three- to five-fold multiple of cropland map bias (Fig. 1B). At the most extreme, GlobCover's bias was -276% at 1 km, but even SA-LC and GeoWiki had biases of 22% and -50%, respectively. Bias could be substantial even for the least carbon dense vegetation type (sparse), as evidenced by the 15-25% bias at 1 km for MODIS and GlobCover under this class. The mean bias across the different potential adjacent vegetation classes ranged between -20 for GeoWiki and -123% for GlobCover at 1 km (with MODIS in between these), while SA-LC's average bias was 11%. Biases declined fairly rapidly with aggregation, with all datasets having an average (across cover types) bias magnitude of less than 10% by 25 km of aggregation, except for GlobCover, which was -12% at 100 km (SI). As with cropland percentages, GeoWiki produced the least biased carbon density estimates above above 1 km resolution.

In terms of accuracy, MAE values were essentially the same as bias magnitudes, except for GeoWiki's, which were twice as large. The average MAE across vegetation classes was 47% at 1 km, dropping to <10 only with 25 km of aggregation. In contrast, SA-LC's carbon estimates were twice as accurate at 1 km, and was slightly more accurate up to 25 km of aggregation were GeoWiki achieved parity.

Evapotranspiration estimates. Accurate estimation of hydrological fluxes is critical to understanding how land-atmosphere interactions impact the climate system and runoff (36). Land surface hydrological models, such as the Variable Infiltration Capacity (36), are used to simulate these processes, and depend on landcover maps to provide information on the characteristics of vegetation and other materials covering the surface, as these govern the rates of runoff, infiltration, and evapotranspiration. We used the VIC model to generate 25 km estimates of monthly evapotranspiration throughout South Africa, and examined how these were impacted by error in the test maps, which were used to determine landcover-specific values for leaf area index (LAI), plant rooting depth, aerodynamic roughness, and several other variables that VIC uses to partition water vapor fluxes into their evaporative and transpirative components.

Compared to the carbon analysis, the bias and accuracy in evapotranspiration (ET) calculated using the VIC model was negligible, averaging less than than +/-2%. However, there were several error hotspots in the resulting ET residual maps (Fig. 4). The most pronounced of these were the 5-15% overestimates in the center of the country caused when VIC was initialized with MODIS and GlobCover, while overestimates along the southern and western coasts reached 25%. These locations correspond primarily to the margins of major crop production regions–in the center is the westernmost boundary of the summer rainfall growing region, marked approximately by the 400 mm isohyet, where maize is the primary crop. The west coast hotspot falls at the western edge of the wheat-dominated winter rainfall region (28), where growing season rainfall is approximately 200 mm.

SA-LC and GeoWiki also resulted in ET errors estimates along the southern and western coasts, but here the tendency was to underestimate ET, while biases in the center of the

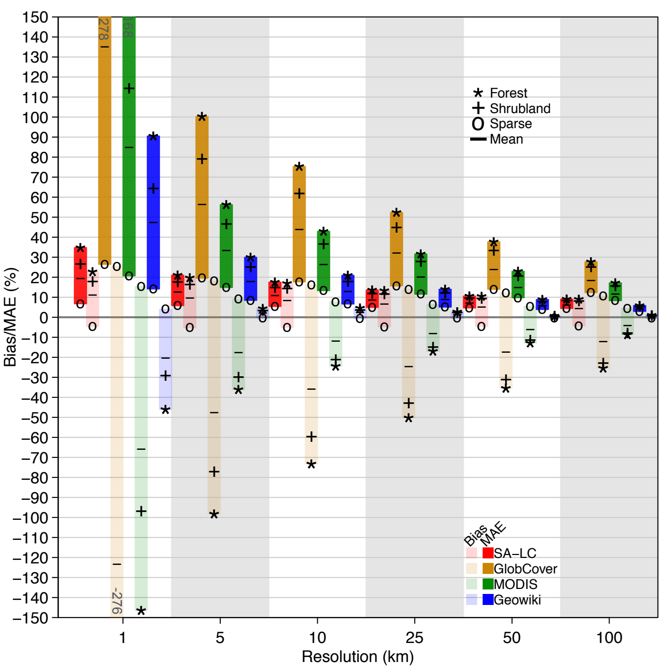

Fig. 3. Biases and accuracies (mean absolute errors) of carbon densities derived from cropland maps, calculated as percents relative to the reference map. Bias estimates (represented by symbols) fall within the semi-transparent floating bars, while accuracies are contained in the solid bars. Bar colors are coded to specific cropland map, symbols indicate which cover type was used to calculate cropland-adjacent carbon density. The bar represents the mean biases calculated across each of the 5 cover types. Shrubland and grassland bias values were near zero, while secondary forest values were close to forest values, and thus these are not shown for display clarity (but see Table S2). MODIS and GlobCover values at 1 km exceeding the plot's Y limits are provided near their truncated tops.

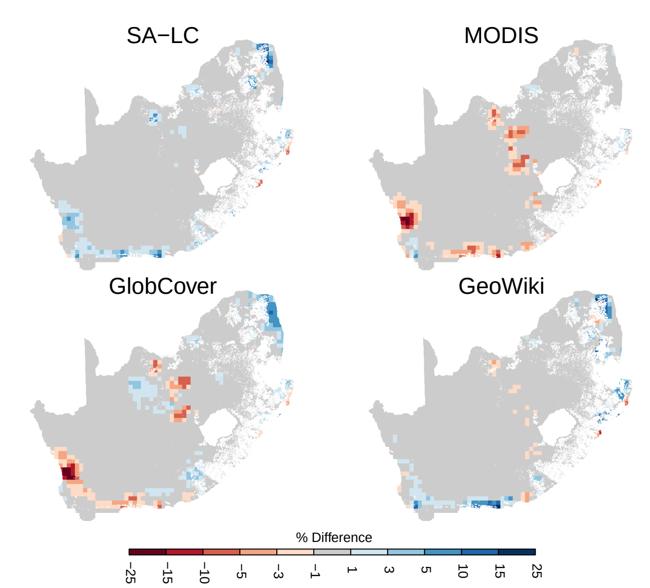

Fig. 4. Differences in annual mean evapotranspiration estimates from 29-year runs of the VIC land surface hydrology model when initialized with LAI response curves derived from the reference map, versus those from the four test maps.

country were either negligible to absent. All but MODIS underestimated ET by 5-15% in the northern tip of the country.

Socio-economic analyses

Gridded crop yield and production data. The spatial variability of crop productivity and production is critical for understanding a host of social, economic, and environmental issues, such as food security, trade, and the potential for agricultural expansion and intensification (5, 37). The most reliable source of such data are national to sub-national agricultural statistics, which are provided for relatively coarse-scaled administrative boundaries. To obtain higher spatial resolutions, efforts have been made to disaggregate these statistics using gridded landcover data (37, 38). These disaggregated datasets, which are constrained to match the agricultural statistics within the boundaries of the areas for which they are reported, have seen increasing use because they are considered to be more accurate than single source methods, and provide consistent data on which to base studies of global change (12, 38)

We used these same methods (37, 38) to disaggregate maize harvested area (South Africa’s largest crop (15)) on top of our reference and cropland maps, followed by yields, which were assigned to cells have harvested areas greater than zero. We then used these two layers to calculate maize production, and further aggregated the yield and production grids to 5, 10, 25, 50, and 100 km resolutions before quantifying the bias and accuracy of each test map’s yield and production values. In this case, we could not convert cell-wise errors into percentages of the reference map values (because many cells had zero values for one map but not the other), so calculated bias and accuracy from the map residuals and then normalized their values to the reference map means.

Maize yields disaggregated onto the test maps showed some marked differences relative to the reference map, but only at the margins of the major crop production areas where cropland is sparser (SI). These differences resulted when a yield value was mapped onto a grid cell where the reference map had no harvested area, and thus zero yield. In more densely cropped areas, such discrepancies were less frequent because both the reference and test maps were both likely to have some maize harvested area, and therefore a yield value. Yield biases were thus fairly low (and accuracy high), with the largest being 20% for MODIS at 1 km, following by GlobCover with 10% (Fig. 5). These dropped to below 10% with aggregation.

Production biases were generally slightly higher, but still low, for most datasets, with the exception of GlobCover, which had an gigantic underestimation bias of >60% (relative to mean production) at 1 km, which remained above 10% even at 100 km of aggregation. MODIS production bias was above 20% at 1 km, but declined to below 10% at higher levels of aggregation.

The accuracy of production estimates was another story. Here all datasets but SA-LC had MAE values of 30% or higher below 25 km of aggregation (Fig. 5), reaching as high at 100, 65, and 45% at 1 km for GlobCover, MODIS, and GeoWiki, respectively. Only GeoWiki’s MAE fell below 10% with 100 km of aggregation. SA-LC estimated production was most accurate, having between 10-20% MAE between 1 and 10 km, and <10% at 25 km and higher. This low accuracy relative to the gridded yield measures relates to the disaggregation process for harvested area, which allocates a fractional value to each pixel, which is itself a fraction. The process of adjusting the gridded values so that their total match the statistics from which they are derived does not adjust map errors relative to the reference map, and the constraint in fact appears to shorten the distance between negative and positive residuals (SI), thereby increasing absolute errors.

Agent-based simulation of food security. Spatially-explicit agent-based model (ABMs) are frequently employed to understand land use decision-making and socio-ecological systems, often to facilitate improved policy, particularly in the arena of human development (39). To obtain valid insight into the functioning of such systems, it is important to calibrate an ABM to empirical data describing the characteristics of land and land users, so that the model realistic represents the social and biophysical features of the study region (39). In our example, we used an ABM of household food security that simulates the food production by individual farming households (the agents) in response to weather and management capacity, which varies as a function of household resources, including the area and physical characteristics of fields (40). The model is initialized such that each household is allocated its designated share of cropland, and then produces a crop over the course of one to many seasons. Like many spatial ABMs, the model is computationally intensive, and thus run over smaller geographic domains (e.g. districts, rather than an entire country) and at higher spatial resolutions (10s to 100s of meters) in order to represent the different land units of single farmers. To match these computational characteristics, we selected four contiguous magisterial districts (ranging from 1,040-1,343 km^2, Fig. S8) in the eastern part of the country with similar climate and 28-45% of their areas devoted to cropland. The initialization process iteratively assigns households to the landscape using a function that factors in neighbor and cropland proximity, to ensure that households are grouped into communities and that their fields are within a realistic proximity. The number of households and the cropland area per household was derived from surveys of farming communities where all cropland is used. The model was thus considered to

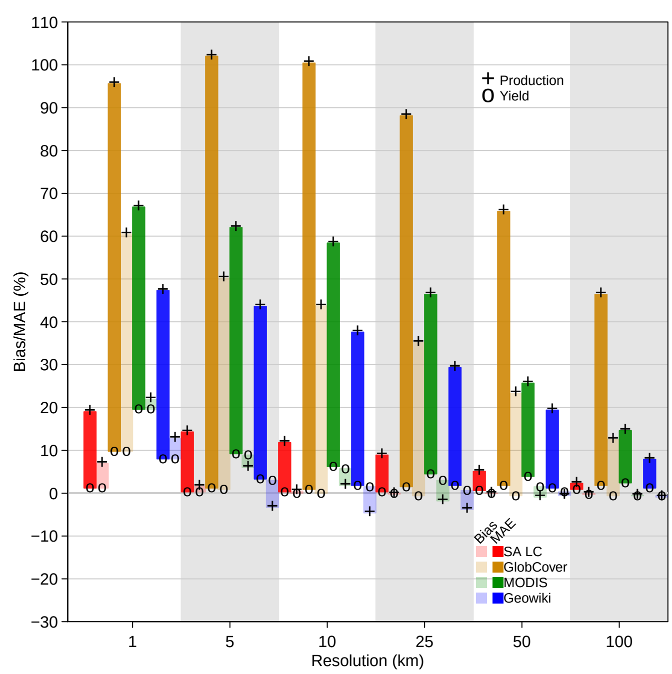

Fig. 5. Bias (mean error) and accuracy (mean absolute error [MAE]) in disaggregated maize yield and production estimates. Bias estimates (represented by symbols) fall within the semi-transparent bars, mean absolute errors in the solid bars, with bar colors coded to specific cropland maps. Symbols code the different variables (production and yield), normalized to their respective means.

well calibrated when all households were allocated their appropriate area of cropland, and all cropland was allocated.

We used the reference map and each cropland map to separately initialize the model, and examined how map errors impacted the land allocation process and household food production estimates. We examined three variables: the percent of unallocated cropland, land deficit, and food deficit. The percent of unallocated cropland measures how effective the model was in matching household agents to available cropland resources. Where error was negative (indicating a cropland overestimate by the test maps), the amount of land left unallocated had a straight one-to-one relationship with the percentage of overestimation (Fig. 6A). When cropland was underestimated, no land was left unallocated until the underestimation exceeded 50%. The MODIS-based simulation for districts 1 and 2 was most pronounced for this tendency, where 5-10% of cropland remaining unallocated despite the fact that cropland with such a high cropland underestimate (85%) the majority of households were not allocated cropland. This curious relationship occurred because croplands tend to cluster, and when underestimated clusters tend to be small, resulting in islands of cropland that fall outside of the search radius within which cropland is sought when agents are seeded onto the landscape.

The second measure, land deficit, describes the amount of land that the model should have assigned but couldn't due to the mismatch between the cropland map and the statistical estimate of cropland holdings. Land deficit increased exponentially in relation to cropland underestimation–reaching around 800% for MODIS in districts 1 and 2 (Fig. 6B)–and would become infinite in the case of a 100% underestimate. This contrasted with the third measure, food deficit, the shortfall in the average amount of food that should have been produced by each household but wasn't because of the land deficit; this particular deficit increased linearly with the percentage of cropland underestimate (Fig. 6C).

Discussion

This spatially comprehensive, bottom-up assessment of landcover map bias and accuracy provides new insight into their extent, causes, and consequences for understanding global change, made possible by a unique, high accuracy dataset that likely provides the truest measure of cropland area and distribution currently available for this region. This dataset is not perfect, being affected by the map-makers' occasional interpretation errors (mostly of omission), while some of the error we identified may be due to slight temporal mismatches between the reference and test datasets. However, our assessment (SI) suggests that these errors are small, and do not appreciably impact our findings, which is bolstered by previous work showing the substantial errors and inconsistencies between landcover maps (16, 33). These previous studies, which were conducted over broader extents than ours, also indicate that our findings should be generalizable beyond South Africa's borders, particularly to more rapidly developing, data poor regions.

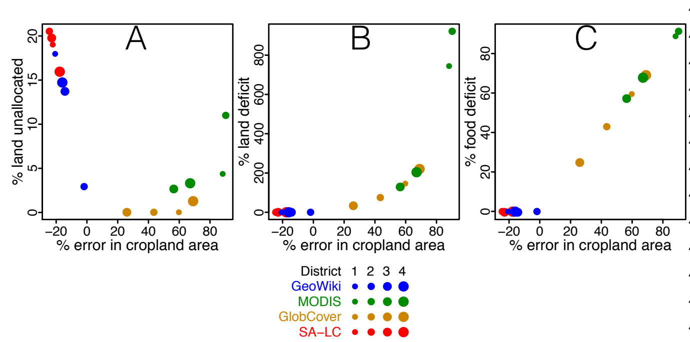

Fig. 6. Biases in agent-based model results relative to the district-wise errors (as a percent) in total cropland area, in terms of A) the percent of cropland in each district that was not allocated to any household, B) the land deficit, or the percent of cropland that should have been allocated to households but wasn't, and C) the food deficit, or the average food production across households resulting from inadequate cropland allocation. Dot sizes correspond to district numbers, colors represent the landcover map.

General guidelines for selecting and using landcover data. Our results suggest several guidelines for selecting and using landcover data in global change research. First, landcover products derived from coarse resolution sensors, such as MODIS and GlobCover, need substantial aggregation to reduce the large biases and inaccuracies that manifest at finer resolutions. Researchers might need to use these datasets for a variety of reasons, for instance, because they provide consistent, multiple landcover classes at a global scale, or allow changes to be tracked across large scales over time (9). The appropriate scale of aggregation then depends on which error property the user wants to minimize. For instance, if the objective is to minimize systematic errors in order to calculate an unbiased mean (e.g. average carbon density), then at least 50-100 km of aggregation is needed to reduce map bias to within +/-10% (Fig. 1B, 3). If the user wants to maximize accuracy (accuracy is important for comparing differences between locations on a single map, or between dates at fixed locations), then even coarser aggregations are needed to reduce mean absolute error below 10% (Fig. 1, 3, 5).

The characteristics of cover at the landscape scale may also require greater levels of aggregation to achieve sufficient accuracy. We found that accuracy is lowest where mixing of cover types is greatest (Fig. 2), a result echoed in an Africa-wide change detection study made with MODIS (33). Outside of South Africa, where smallholder-driven agriculture predominates (41) such mixed landscapes are more common, thus greater levels of aggregation may be needed over these areas. Coarser aggregations are also necessary if the cover type of interest belongs primarily to a mixed class (as with GlobCover). This leads to underestimation bias (Fig. 2) and high inaccuracy, which will persist until the aggregated pixel exceeds the characteristic size of the geographical area within which the cover type dominates. In South Africa's farming regions, this can exceed 1000 km^2.

Maps derived from higher resolution sensors, such as the SA-LC dataset, do not have this mixture problem because the pixels are fine enough to be assigned to a single cover type. This characteristic leads to higher accuracy, which is what we observed here: the SA-LC dataset needed the least aggregation–just 5-10 km–to achieve <10% MAE in most of the example applications we tested (note, however, another guideline in this result: even with high resolution maps some aggregation is needed to achieve higher accuracy).

Unfortunately, Landsat-scaled maps are typically developed for specific countries, using varying methods, and can be hard to obtain (an exception to this is the newly released GLC30, which has a reported classification accuracy of 80% (42)). Our results also suggest that, despite their higher accuracy, they are not necessarily the least biased maps. Except for a few results (notably the 1 km carbon and maize production estimates), GeoWiki-based maps had the least amount of bias across most scales. GeoWiki's lower bias results from its process of calibrating the cropland percentages to reported agricultural statistics (17, 18). The procedure is similar to the one we used in disaggregating maize harvested areas (37, 38), which led to fairly unbiased estimates of maize production and yield (Fig. 5) above 1 km (except for GlobCover's production estimates). This result, together with GeoWiki's unbiased cropland estimates (Fig 1.B), indicates the value of fusing inventory data with remote sensing to reduce map bias.

Of course, for the constraint to be effective, the inventory data themselves must be faithful to reality, but in many places, particularly Africa, agricultural statistics are suspect (43, 44), which is a vulnerability of the fusion approach (12). Beyond this limitation, this method does not improve map accuracy, which is evident in GeoWiki's relatively high MAE for cropland percentage at 1 km (23%, Fig. 1B), and by all maps' (except SA-LC's) highly inaccurate maize production estimates (Fig. 5). The reason that accuracy is not improved is because the statistical constraint cannot correct the omission and commission errors within the landcover maps.

These last observations suggest two further guidelines: 1) for high accuracy, users should select well-validated datasets from higher resolution sensors–this option should be increasingly available as computational power grows and high resolution image archives expand (42, 45); 2) where low bias is needed over large areas, the best option is to use newer generation, statistically constrained maps such as GeoWiki (and possibly the GLC-Share [3] datasets for other cover types). But in either case some aggregation–5-25 km, depending on application and whether low bias or high accuracy is the objective–will still likely be necessary to achieve sufficiently low error. Additionally, these higher accuracy, lower bias product are made infrequently, thus applications requiring high frequency landcover data will depend on MODIS and confreres, as previously mentioned, and therefore much coarser aggregations.

Implications for global change analyses. Besides the preceding landcover data selection and usage guidelines, our findings also have broader implications for global change analyses and associated policy decisions. The simpler, most direct uses of landcover data have substantial to skew our understanding of global change processes. For example, the global carbon carbon density map of (35), of which we created facsimiles here, has been widely used, but our results suggest that any analyses of the spatial variability of carbon stocks, or of mean carbon densities, can be highly misleading if the map is not first sufficiently aggregated. For example, a study of deforestation risks in the Democratic Republic of the Congo estimated forest rents as a function of carbon densities based on this map. Although the map was aggregated to $\sim$50 km, we found that mean absolute errors can still reach 40% at this scale (Fig. 3). The same carbon map provided country-specific mean forest carbon densities for an evaluation of climate mitigation policies (46), which our results suggest could be underestimated by 70% (Fig. 3). In contrast to these, an assessment of the congruence between carbon and biodiversity (47) aggregated the carbon data to $\sim$110 km resolution, where accuracy is much higher.

A related example can be seen in studies that create high resolution maps (10 km) showing the potential food production benefit of closing yield gaps (e.g. Figure 3 in 48). Such maps give an inflated sense of the precision with which such locations can be identified. Although the study in this example is careful to report aggregated statistics, in which we can have greater confidence, it is not inconceivable that the maps themselves are used to inform policy. For this reason, it is advisable to degrade maps until their accuracy is of an acceptable standard (i.e. $\geq$100 km).

Aggregating until acceptable accuracy is achieved will be particularly important in growing efforts to find tradeoffs between agricultural development and ecological protection (e.g. 26, 49). Finding win-wins fundamentally requires accurate, fine-scale data in order to identify "win-win" solutions, which are comprised of areas where there is some deviation from an expected relationship; for example, an area of unusually low carbon density but high crop fertility (26). Without sufficient aggregation, it is difficult to know whether such efficiencies are due to real biophysical properties, or simply landcover error.

What then, of the impacts of landcover data on more complex analysis? Here the results appear to be mixed. Map error appears to have had little impact on the overall bias and accuracy of VIC's evapotranspiration (ET) estimates, which contrasts with a related study that examined how landcover maps impact rainfall simulations (22). There are two reasons for this. First, map accuracy increased nearly two-fold by aggregating to 25 km, the resolution of VIC. Second, VIC's landcover-derived variables (e.g. LAI curves, effective rooting depth) could dampen the effect of map errors on ET calculations when the values of these variables in the adjacent landcovers are similar to that of cropland, which is the case in much of VIC's landcover scheme for South Africa. In other regions, where the matrix landcover's LAI values depart substantially from cropland's (e.g. in forest), the ET errors would likely be much larger. The carbon example illustrates how cover properties can mute or amplify map error.

Despite the low values of the overall error metrics, the hotspots of ET errors (Fig. 4) are important, because they occurred at the climatic margins of the major crop growing regions, where a substantial share of farms use irrigation. Since VIC does not simulate irrigation, the magnitude of these errors is likely to have been underestimated, because irrigation substantially increases latent heat flux (50). This also suggests the potential for map error to magnify in a coupled land-atmosphere model (which could explain the difference between this result and that of Ge et al. (22).), and thereby give misleading insight into the climatic effects of land use.

Relative to the ET example, map errors had a much larger impact on the agent-based model's results. For the model's most important metric of food security–household-level crop production–the result was fairly straightforward for a class of models where the interactions of multiple agents often produce analytically intractable outcomes (51). In this case, the model was only sensitive to cropland underestimates, which had the perfectly predictable effect of lowering average household food production, which would in turn overstate the degree of food *in*security. Overestimates did not matter in this case, because the number of households remained fixed, and their cropland holdings were constrained to match the survey statistics (but this would matter for models that allow these parameters to vary). Less predictable was the finding that the spatial distribution of cover can impact model parameterization, with land being left unallocated even when needed, thus such models also require spatial, not just statistical, accuracy in their base landcover maps.

The errors in the cropland data we tested were large, and it is unlikely that current ABM-based studies, mostly of small regions, are based on such inaccurate landcover data. However, as computing power increases, so, too, does the feasible application extent of ABMs, which then become more reliant on larger-scaled, more erroneous, landcover data, with the attendant risk that insights into socio-ecological systems arising from these models could be highly misleading.

Beyond these examples, there are other applications where landcover map errors could alter understanding and policy.

[3] GLC-Share. www.glcn.org

One of these lies in assessments of land availability for agricultural expansion, which could be particularly erroneous if they use a "residual approach" (41), where existing cropland and other occupied cover types are used to filter a map of potential suitability. Our results suggest that the pronounced underestimation biases (Fig. 1) we found could inflate land availability estimates, and therefore increase the risk of unjust land allocation policies (52). Another example is the growing efforts to minimize the environmental and climate impacts on new agricultural expansion frontiers, particularly Africa's savannas, which requires identifying areas that provide the highest agricultural benefit for the lowest environmental cost (26). In order to be effective, such analyses must be undertaken at the fine scales–hectares, rather than square kilometers–at which land use decisions are made, thus their value for informing land use policy is contingent on the accuracy of their underlying landcover/land use data (26).

The way forward. Our analysis demonstrates that the landcover data that increasingly inform much global change research and policy can be substantially misleading, depending on the dataset selected, its application case, the scale to which it is aggregated, and whether the insight being drawn from the map depends on how accuracy or low bias. We provide some basic guidelines for selecting and using landcover data that can help increase confidence in the conclusions drawn from landcover-based studies.

Beyond these guidelines, we recommend updating several important, global-scale datasets using the latest generation landcover maps. For example, rebuilding the widely used cropland distribution and yield maps (37, 38) using the GeoWiki map could substantially improve our knowledge about current global agricultural distribution and productivity. Similarly, reconstructing the carbon map of (35) using the GLC-Share datasets (in which GeoWiki provides the cropland layer) would improve our confidence in regional to global assessments of vegetative carbon stocks. Alternatively, the new high-resolution, Landsat-based GLC30 (42), aggregated and converted to fractional cover classes at 1 km–and fused with high quality national landcover maps, where available (similar to 17, 18)–may provide a preferable base map for rebuilding these datasets.

Fusing human judgement with expanding high resolution satellite data archives, improved classification algorithms, and expanding computing power appears to offer the greatest promise for improving the quality of landcover maps. The GeoWiki project and the new high resolution global forest change map (45) are the most advanced examples of these newer methods. Another approach is that of the Mapping Africa project[4], which, inspired by quality of the reference map used in this study, enlists internet-based workers to manually digitize cropland boundaries (14). Merging this crowd-sourced approach to vector map creation with newer computer vision-based mapping algorithms (25) is a near-term goal of this project, as the former can interactively provide high quality training and test data (by using humans' superior capacity to delineate discrete cover types in noisy images (14)) for the latter, which has the advantage of speed and the ability to handle high-dimensional datasets. Approaches such as these will be needed to provide the next generation of landcover maps, so that we can gain a clearer, unbiased understanding of global change.

Materials and Methods

Methods. To establish the accuracy of the cropland boundary reference map, we randomly selected 609 cell from a 1 km^2 grid covering South Africa, weighted by the probability of cropland occurrence (14). We sub-divided each cells into 200 m sub-cells, and visually interpreted whether crop fields presence/absence within these against Google Earth imagery, and compared these results to the digitized field polygons to determine the reference map's sensitivity (0.89) and specificity (0.99), overall map accuracy (97%), and its true skill statistic (0.88; 53). We intersected the field vectors with a 1 km grid, and calculated the percent of each cell occupied by fields to create the gridded cropland reference map. In the resulting grid, we masked out areas classified as communal farmland, because only their outer perimeters were digitized (27), which posed a risk of overestimating cropland extent. We also excluded the permanent tree crop class, to avoid confusion with landcover maps that include these in their cropland classes. For the same reason, we used other high resolution landcover maps for selected provinces of South Africa to mask out commercially afforested areas and sugarcane plantations, which were also not included in the reference map. The resulting masked reference grid covered 90% of South Africa (1,081,000 km^2), of which 104,304 km^2 was cropland. We extracted the cropland classes from SA-LC, MODIS, and GlobCover, and converted these into percent cropland estimates at 1 km resolution (GeoWiki is already a cropland percentage map). Both MODIS and GlobCover had mixed pixel classes, thus we used the upper, mean, and lower cropland estimates from these classes to produce three versions of the gridded percentages. We used the mean map for the main analysis, but estimated error variability using all three versions (SI).

To uncover the relationship between map error and cover density, we used magisterial district boundaries (n=354, mean area=3,445 km^2, see SI) to provide a landscape-scaled unit for calculating characteristic cover density. We filtered out pixels with $<$0.05% (0.5 ha) cropland, to prevent the much larger areas of non-agricultural districts from drowning out any relationship. We extracted the absolute values of test map errors and corresponding reference cropland percentages, calculated their district-wide means, and then modeled the relationship between mean absolute error (response) and cropland density (predictor) as a polynomial function within a generalized additive model. We accounted for potential spatial autocorrelation by fitting a two-dimensional smoothing spline to each district's centroid coordinates, and weighted the model by the number of agricultural pixels per district (**?**).

To develop the carbon maps, we followed the methods used by Ruesch and Gibbs (35) to create their vegetative carbon map. We took the average of the biome specific carbon density values they provided for African forests, secondary forests, shrublands, grasslands, sparse habitats, and croplands. We multiplied cropland densities by cropland fractions and added these to each of the five other densities multiplied by the residual non-cropland fractions to create five different carbon density maps. We aggregated each carbon map up to the five coarser resolutions for scaling comparisons.

We used the Variable Infiltration Capacity (36) land surface hydrology model run with the Africa Flood and Drought Monitor's meteorological data (54) to produce monthly gridded evapotranspiration estimates for South Africa for the years 1979-2010. We adjusted VIC's native landcover scheme so that its cropland fractions matched those of the 25 km reference and test maps (each reprojected and resampled to VIC's 0.25° resolution). We ran one instance of VIC for each of the five landcover permutations, and compared the mean annual ET produced by the reference map variant with those from the test maps.

To create the gridded maize yield and production maps, we followed (38) to first calibrate the test cropland maps against cropland areas reported by administrative-level agricultural censuses. In place of census reports, we used the reference map to calculate total cropland areas for South Africa's nine provinces, then adjusted the pixel-wise cropland percentages in the four test maps so that their province-wise sums matched these totals (SI). We then followed Monfreda et al's procedure (37) to disaggregate maize (South Africa's largest crop 55) planted area and yields onto the reference and adjusted test maps. Per (37), the gridded planted area and yield estimates were obtained from magisterial district-level agricultural censuses, in this case for the most recent census of 2007 (56). We then aggregated gridded harvested areas and yields (weighted by the corresponding harvested area) to each coarser resolution, and multiplied yield by harvested area to calculate production.

The agent-based model of food security used here calculates average household-level food production over multiple seasons based on the biophysical environment (soil properties, weather), management inputs (planting decisions, labor availability), and farm assets (total field area) (40). The frequency distributions of household cropland holdings were derived from household survey data obtained from Zambia (no equivalent data were available for South Africa). We used these statistics to calculate the "true" number of households per district by dividing reference cropland areas by the mean cropland area (2.2 ha), and preserved the cropland area distributions by multiplying the total number of households by the frequencies. To create cropland surfaces for the, we disaggregated all five cropland maps to binary cropland/non-cropland cover types with 100 m resolution, in order to match the survey-recorded average field size of 1 ha. In initializing, the model takes a weighted (by cropland area frequency) random

[4] http://mappingafrica.princeton.edu

draw of 100 households and places these within a district, assigning each household its required number of "fields" (cropland pixels), which must be within 1.5 km of the household's location and not already assigned to another household. This process is iterated until all households are assigned cropland, or all available cropland is allocated. The model then runs separately for each district, with the household agents planting and managing a maize crop throughout the growing season, with possible final maize yields drawn from a look-up table that relates yield to particular climatic outcomes and management decisions. We ran the model for a single season, once per district and when initialized by each of the cropland maps (20 simulations total).

In calculating map error statistics, and with the exception of the agent-based model results, we use the reference cropland percentages to calculate the weighted mean map error (bias) and mean absolute map error (accuracy). For the comparisons between aggregated map values, we applied a further weight–the number of pixels contributing to each aggregated pixel, to prevent pixels close to national boundaries or where non-target cover types were masked out from having outsize influence on the statistics.

ACKNOWLEDGMENTS. This work was supported by funds from the Princeton Environmental Institute Grand Challenges program, the NASA New Investigator Program (NNX15AC64G), and the National Science Foundation (SES-1360463 and BCS-1026776).

References

1. Lambin EF (1997) Modelling and monitoring land-cover change processes in tropical regions. *Progress in Physical Geography* 21(3):375–393.
2. Lambin EF, Geist HJ, Lepers E (2003) Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources* 28(1):205–241.
3. Lark TJ, Salmon JM, Gibbs HK (2015) Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environmental Research Letters* 10(4):044003.
4. Wright CK, Wimberly MC (2013) Recent land use change in the Western Corn Belt threatens grasslands and wetlands. *Proceedings of the National Academy of Sciences* 110(10):4134–4139.
5. Licker R et al. (2010) Mind the gap: how do climate and agricultural management explain the yield gap of croplands around the world? *Global Ecology and Biogeography* 19(6):769–782.
6. Asner GP et al. (2010) High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences* 107(38):16738–16742.
7. Gaveau DLA et al. (2014) Major atmospheric emissions from peat fires in Southeast Asia during non-drought years: evidence from the 2013 Sumatran fires. *Scientific Reports* 4.
8. Newbold T et al. (2015) Global effects of land use on local terrestrial biodiversity. *Nature* 520(7545):45–50.
9. Luoto M, Virkkala R, Heikkinen RK, Rainio K (2004) Predicting bird species richness using remote sensing in boreal agricultural-forest mosaics. *Ecological Applications* 14(6):1946–1962.
10. Linard C, Gilbert M, Tatem AJ (2010) Assessing the use of global land cover data for guiding large area population distribution modelling. *GeoJournal* 76(5):525–538.
11. Jain M, Mondal P, DeFries RS, Small C, Galford GL (2013) Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. *Remote Sensing of Environment* 134:210–223.
12. See L et al. (2015) Improved global cropland data as an essential ingredient for food security. *Global Food Security* 4:37–45.
13. Lobell DB (2013) The use of satellite data for crop yield gap analysis. *Field Crops Research* 143:56–64.
14. Estes L et al. (2015) DIYlandcover: Crowdsourcing the creation of systematic, accurate landcover maps. *PeerJ PrePrints* 3:e1266.
15. Estes LD et al. (2013) Projected climate impacts to South African maize and wheat production in 2055: a comparison of empirical and mechanistic modeling approaches. *Global Change Biology* 19(12):3762–3774.
16. Fritz S, See L, Rembold F (2010) Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *International Journal of Remote Sensing* 31(9):2237–2256.
17. Fritz S et al. (2011) Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. *Geophysical Research Letters* 38:L04404.
18. Fritz S et al. (2015) Mapping global cropland and field size. *Global Change Biology* 21(5):1980–1992.
19. Verburg PH, Neumann K, Nol L (2011) Challenges in using land use and land cover data for global change studies. *Global Change Biology* 17(2):974–989.
20. Fritz S et al. (2012) Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software* 31:110–123.
21. Kaptu Tchuent AT, Roujean JL, De Jong SM (2011) Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. *International Journal of Applied Earth Observation and Geoinformation* 13(2):207–219.
22. Ge J et al. (2007) Impacts of land use/cover classification accuracy on regional climate simulations. *Journal of Geophysical Research: Atmospheres* 112(D5):D05107.
23. Schmit C, Rounsevell MDA, La Jeunesse I (2006) The limitations of spatial land use data in environmental analysis. *Environmental Science & Policy* 9(2):174–188.
24. Quaife T et al. (2008) Impact of land cover uncertainties on estimates of biospheric carbon fluxes. *Global Biogeochemical Cycles* 22(4):GB4016.
25. Debats S, Luo D, Estes L, Fuchs T, Caylor K (year?) A generalized computer vision approach to mapping agricultural fields in Sub-Saharan Africa. *Remote Sensing of Environment.*
26. Searchinger TD et al. (2015) High carbon and biodiversity costs from converting Africa's wet savannahs to cropland. *Nature Climate Change* 5(5):481–486.
27. Fourie A (2009) Better Crop Estimates in South Africa. *ArcUser Online* (1).
28. Hardy M, Dziba L, Kilian W, Tolmay J (2011) Rainfed Farming Systems in South Africa in *Rainfed Farming Systems*, eds. Tow P, Cooper I, Partridge I, Birch C. (Springer Netherlands), pp. 395–432.
29. Estes LD et al. (2014) Using changes in agricultural utility to quantify future climate-induced risk to conservation. *Conservation Biology* 28(2):427–437.
30. SANBI (2009) National Landcover 2009, (South African National Biodiversity Institute; National Department of Environmental Affairs and Tourism, Pretoria, South Africa), Technical report.
31. Fry J, Coan M, Homer C, Meyer D, Wickham J (2009) Completion of the National Land Cover Database (NLCD) 1992-2001 Land Cover Change Retrofit Product, (U.S. Geological Survey), USGS Numbered Series 2008-1379.
32. Arino O et al. (2012) *Global land cover map for 2009 (GlobCover 2009).* (European Space Agency & Universit Catholique de Louvain).
33. Gross D et al. (2013) Monitoring land cover changes in African protected areas in the 21st century. *Ecological Informatics* 14:31–37.
34. Shackelford GE, Steward PR, German RN, Sait SM, Benton TG (2015) Conservation planning in agricultural landscapes: hotspots of conflict between agriculture and nature. *Diversity and Distributions* 21(3):357–367.
35. Ruesch A, Gibbs HK (2008) New IPCC Tier-1 global biomass carbon map for the year 2000. *Carbon Dioxide Information Analysis Center (CDIAC), Oak Ridge National Laboratory, Oak Ridge, Tennessee. Available online at: http://cdiac. ornl. gov/epubs/ndp/global_carbon/carbon_documentation. html.*
36. Liang X, Lettenmaier DP, Wood EF, Burges SJ (1994) A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research* 99(D7):14415.
37. Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles* 22:GB1022.
38. Ramankutty N, Evan AT, Monfreda C, Foley JA (2008) Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochemical Cycles* 22:19 PP.
39. Berger T, Schreinemachers P (2006) Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society* 11(2):19.
40. Chen P, Plale B, Evans T (2013) *Dependency Provenance in Agent Based Modeling.* pp. 180–187.
41. Lambin E et al. (year?) Estimating the world's potentially available cropland using a bottom-up approach. *Global Environmental Change.*
42. Chen J et al. (2015) Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing* 103:7–27.
43. Carletto C, Jolliffe D, Banerjee R (2013) The emperor has no data! Agricultural statistics in Sub-Saharan Africa. *World Bank.*
44. FAO, Bank W (2013) *Action Plan of the Global Strategy to Improve Agricultural and Rural Statistics: For Food Security, Sustainable Agriculture and Rural Development.* (FAO, Rome).
45. Hansen MC et al. (2013) High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342(6160):850–853.
46. Cattaneo A et al. (2010) On international equity in reducing emissions from deforestation. *Environmental Science & Policy* 13(8):742–753.
47. Strassburg BB et al. (2010) Global congruence of carbon storage and biodiversity in terrestrial ecosystems. *Conservation Letters* 3(2):98–105.
48. Foley JA et al. (2011) Solutions for a cultivated planet. *Nature* 478(7369):337–342.
49. West PC et al. (2010) Trading carbon for food: Global comparison of carbon stocks vs. crop yields on agricultural land. *Proceedings of the National Academy of Sciences.*
50. Sacks WJ, Cook BI, Buenning N, Levis S, Helkowski JH (2008) Effects of global irrigation on the near-surface climate. *Climate Dynamics* 33(2-3):159–175.
51. Janssen MA, Ostrom E (2006) Empirically based, agent-based models. *Ecology and Society* 11(2):37.
52. Rulli MC, Saviori A, D'Odorico P (2013) Global land and water grabbing. *Proceedings of the National Academy of Sciences* 110(3):892–897.
53. Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43(6):1223–1232.
54. Sheffield J et al. (2013) A Drought Monitoring and Forecasting System for Sub-Sahara African Water Resources and Food Security. *Bulletin of the American Meteorological Society* 95(6):861–882.
55. Estes LD et al. (2013) Comparing mechanistic and empirical model projections of crop suitability and productivity: implications for ecological forecasting. *Global Ecology and Biogeography* 22(8):1007–1018.
56. Africa SS (2007) Commercial Census of Agriculture, South Africa.