A Full Accounting of Landcover Map Error and Bias and Their Impacts on Assessments 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

Blah blah.

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

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

The nature and distribution of landcover provides significant insight into econonomic processes (1) because human endeavors are so closely tied to how we transform land, whether it be the felling of ancient forests for farmland or converting that farmland into office parks. The vastness of our alteration of Earths landscapes suggests that landcover is a prime mediator of many environmental and social processes that drive or are affected by global change (1), such as agricultural production and food security (2–4), carbon cycling (5, 6), biodiversity loss (7, 8), and changes in human demography (9). Like any view into nature, resolution and fidelity at fine scales are the keys to unlocking more granular and mechanistic insights into these processes (10). It is therefore unsurprising to see the explosive growth in private sector initiatives to develop new Earth observing capabilities, which range from small hobbyist drones[1] to satellite arrays[2], in order to add value to industries such as agriculture, mining, and construction. This rapid growth in fine-scale landcover mapping capability is creating new opportunities to develop actionable information for traditionally public-sector concerns, such as agricultural development[3], drought and flood adaptation[4], and carbon cycle management[5]. But while the demand for more nuanced, landcover-based insights is growing, there is only now the opportunity to use finer-scaled imagery to comprehensively interrogate the accuracy and biases in the landcover products that have become ubiquitous in global change research.

Global landcover data can only practically be derived from satellite imaging, but in many regions the cover types of interest are smaller, on average, than the sensor resolution, or spectrally indistinct from other neighboring covers, and these factors propagate classification errors (10–12). The result is that landcover maps are generally inaccurate at finer scales and disagree substantially with one another, particularly in those parts of the world undergoing the most rapid land use changes (13–15).

Errors in landcover products are widely-acknowledged (10, 14–17), and there are a variety of efforts underway to improve landcover maps, particularly for agriculture (12, 18). What is less known is the degree to which these errors bias analyses derived from the distributional and areal information in landcover. Errors are hard to quantify because spatially extensive reference data are not available for most regions of the world–particularly over Africa and other developing regions. Error assessments therefore typically rely on a small number of ground truth points for a bottom-up assessment or aggregated survey data for a top-down sanity check. For this reason, we have a better understanding of discrepancies between landcover datasets in relation to country-level statistics (14, 15, 19), which offers little direction for how to arrive at a true number.

Being unable to fully quantify the errors in landcover maps of course makes it difficult, if not impossible, to quantify their impact on downstream analyses. There has been some work examining how such error influences climate simulations (20), agricultural land use patterns (21), and carbon flux (22) and human population estimates (9), but these either use simulated landcover errors (20) or compare relevant differences in estimates between different satellite-derived landcover maps (9, 22). One exception is a Belgian study (21) that used ground-collected farm parcel data to assess how landcover map errors bias measurements of agricultural land use patterns, but the study extent was fairly small and the validation data were discontiguous.

Just as a building needs a solid foundation, global change science needs to be based on sound landcover data. There is thus an urgent need to more precisely quantify landcover map errors and how these vary over large regions, particularly for the regions where landcover is changing most rapidly yet is most poorly known. We address this need here, using a unique, high accuracy agricultural landcover map for the entire country of South Africa to conduct a spatially comprehensive, bottom-up quantification of error in several latest generation landcover maps that are widely used in global change studies. We use these errors to assess the biases and accuracies of i) these landcover data, ii) how landcover properties influence map errors, iii) how bias and accuracy change with aggregation scale, with the specific goal of determining “safe” scales for drawing area-based inferences, and iv) how landcover error propagates through “downstream” analyses that represent several major global change research areas, including biogeochemical and land use change studies, food security

Reserved for Publication Footnotes

[1]e.g. 3DRobotics, DJIA
[2]Planet Labs, Skybox
[3]USAID’s Feed the Future
[4]Global Index Insurance Facility, www.indexinsuranceforum.org
[5]United Nations REDD+, www.un-redd.org/aboutredd

assessments, land surface hydrology and climatology, and human geography.

Study area and landcover data

South Africa comprises nearly 6% of sub-Saharan Africa's (SSA) landmass, and has a large, diverse agricultural sector, ranging from large commercial operations to smallholder farms (23, 24). This diversity suggests that the country's agricultural landcover spans the range of types that are found throughout the rest of SSA.

The South African government commissioned a whole-country cropland boundary map to enhance its annual collection of agricultural statistics (25). The map was made by trained workers who visually interpreted high resolution satellite imagery (<5 m SPOT imagery) and manually digitized field boundaries following a standardized mapping protocol. The resulting vectorized field maps, provide a unique, high quality reference dataset describing South African crop field distributions and size classes for the period 2009-2011, and are 97% accurate in distinguishing cropland from non-cropland at 200 m resolution. We intersected the field vectors with a 1 km grid, and calculated the percent of each cell occupied by fields to create a gridded cropland reference map.

We compared the reference map with similar maps derived from four existing landcover datasets. We obtained South Africa's 30 m resolution National Landcover map (SA-LC) for 2009 (26), the 500 m resolution MODIS Landcover for 2011 (27, 28), the 300 m resolution GlobCover 2009 (29), and the new 1 km Geo-wiki hybrid-fusion cropland map for Africa (16). We chose these particular datasets because they are nearly contemporaneous with our reference data, and represent the major types of landcover products used by researchers: SA-LC typifies the higher resolution, Landsat-derived maps that are developed individually for many countries (e.g. 30), MODIS and GlobCover are widely used global-scale products (31, 32), while Geo-Wiki incorporates the first three datasets and represents the current state-of-the-art in landcover mapping. We extracted the cropland classes from the first three datasets and converted these to 1 km resolution percent cropland estimates, resulting in 4 maps (hereafter simply the "cropland maps") to compare to our reference map.

Quantifying Error and Bias

We first calculated errors in pixel-wise estimated cropland percentages by differencing the reference map with each cropland map at each of the five aggregation scales (1-100 km). We used these errors to calculate map bias (the mean pixel-wise error) and accuracy (the mean absolute error, and how these vary with scale. Next, we assessed the degree to which the average cropland cover in agricultural landscapes, a descriptor of landcover pattern, impacts map accuracy.

To investigate how landcover map error can impact downstream global change research, we quantified the biases and accuracies in five landcover-based analyses built upon our reference and cropland maps. The first three represented "first-order" analyses, in which a variable of interest is mapped onto a landcover type(s) using a simple empirical relationship. The first of these was the widely used International Panel on Climate Change's Tier-1 approach for mapping vegetative carbon stocks, as developed by (33). The second was maize yield maps derived by disaggregating district-scale agricultural census data for both maize yield and harvested area (following 34, 35), from which we calculated the third map, gridded maize production estimates. Maps based on these analyses underpin many assessments of crop productivity and production (e.g. 4, 36).

The other two analyses can be considered second-order, wherein a process model draws on the cover types' values to calculate an output value. For the first of these, we used the Variable Infiltration Capacity (37) land surface hydrology model to calculate monthly evapotranspiration, using the reference and cropland maps to adjust landcover-specific leaf area index (LAI) values that VIC uses to partition water vapor fluxes into their evaporative and transpirative components. In the second example, we examined how these map errors impact the land allocation process of an agent-based food security model (38).

Landcover maps

Bias and accuracy. We created the 1 km reference map after removing all field types classified as communal/smallholder agriculture (individual fields in this category were not delineated, thus they were removed to prevent potential cropland area overestimates, SI) or permanent tree-crops (SI), and calculated the total cropland extent in the remaining area (1,081,000 km^2, or 90% of South Africa). The 2011 reference map showed a cropland area of 104,304 km^2, which the SA-LC and GeoWiki maps overestimated by 31 and 10%, respectively, and GlobCover and MODIS underestimated by 18 and 23%.

We then aggregated the reference and each cropland map to 5, 10, 25, 50, and 100 km resolutions, and subtracted the four landcover maps from each reference (2007 and 2011) map at each scale of aggregation to assess error patterns (Fig. 1A). Negative pixels here represent overestimation error, while positive values indicate underestimates.

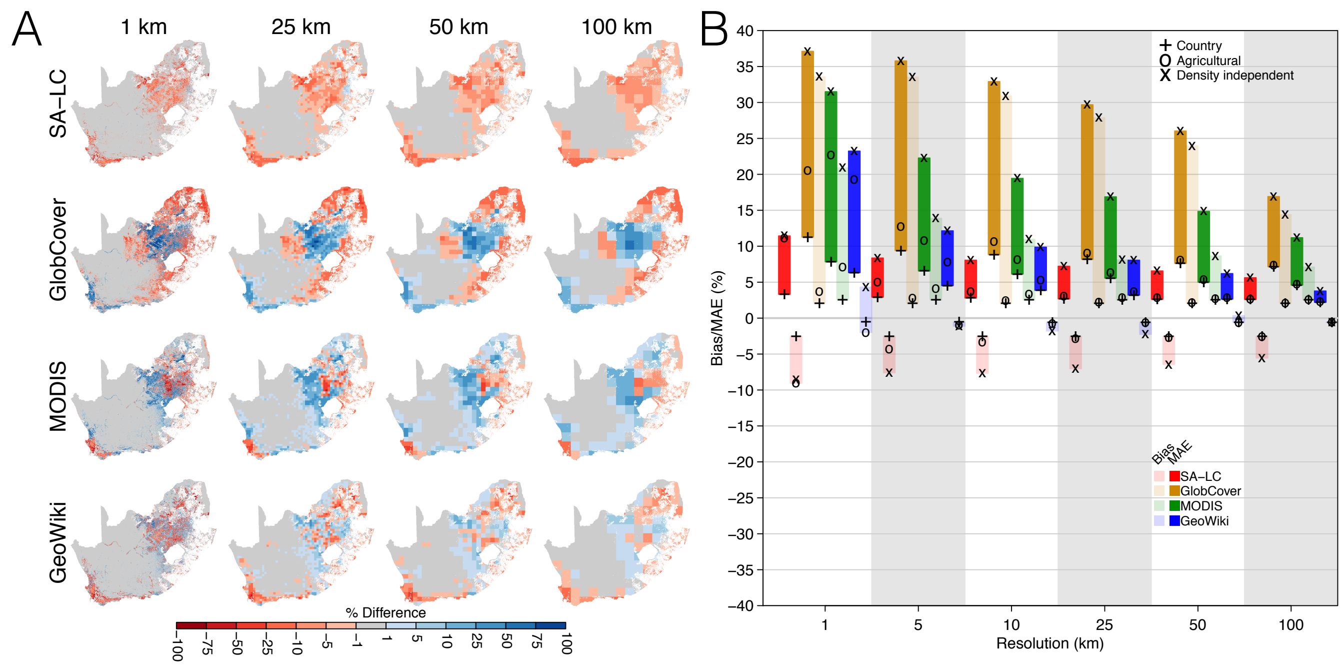

Fig. 1. (A) Errors in the percent cropland estimates resulting from each of the four cropland maps relative to the reference map at different scale of pixel aggegation. Rows indicate the landcover map being assessed (by subtraction from the reference map), while columns refer to resolution of aggregation. White indicates areas with no data where communal farmlands or plantation forests were removed from analysis. (B) The bias (mean error) and accuracy (mean absolute error [MAE]) of each cropland map at each scale of aggregation. 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. The symbols indicate the method of calculating each metric, either by averaging across the entire country, within the union of agricultural areas (cropland >0.05%) for each reference-cropland map combination, or as the average of the averages calculated per 5% bin of cropland cover, a measure of error that is independent of cropland density.

The most pronounced errors were in the MODIS and GlobCover maps, which both underestimated cropland extent by 10-75% in the center of the country (blue areas in Fig. 1A, and the dominant production region), and overestimated along the eastern to northern margins (red areas in Fig. 1A). All maps have less than +/-5% bias (crosses in semi-transparent bars, Fig. 1B), at all aggregation scales, when errors are averaged across the entire country, due to the extensive non-agricultural areas in the country that were effectively discerned by all maps. The whole country mean absolute error (the measure of map accuracy) were a few percent higher values in most cases, but still <10%, except for GlobCover (∼11%) at 1 km resolution. Biases increase only slightly when errors are averaged within the combined agricultural regions (areas with greater than 0.05% cropland, indicated by circles in transparent bars in Fig. 1B) of the reference and each cropland map, but accuracy decreases substantially, with mean absolute errors increasing to 10% for SA-LC and around 20% for MODIS, GlobCover, and GeoWiki at 1 km resolution (circles in solid bars in Fig. 1B), and remain between 5-10% for all three datasets even at 25 km of aggregation.

As a third measure, we calculated bias and mean absolute error by averaging their mean per-ventile values (i.e. the average bias or accuracy value for areas have 0-5% cropland, 5-10% cropland, etc.). These provide density independent measures of map performance because they remove the dominance any particular level of cover has on map accuracy, which in this case was the 0-5% cover class where map performance was best (SI). Calculated this way, bias and accuracy were substantially worse. The average bias jumped to 21% for MODIS and 34% for GlobCover at 1 km resolution (Fig. 1B), meaning that each map had a strong tendency to underestimate cropland. This decreased with pixel aggregation, falling to 8% at 50 km for MODIS, but remaining as high as 14% at 100 km for GlobCover. SA-LC bias was smaller in magnitude, but tended towards cropland overestimates of 5-10% at all aggregation scales. GeoWiki alone was relatively unbiased, but showed substantial inaccuracy with a density-independent mean absolute error of 23%, which dropped below 10% only above 25 km of aggregation. GlobCover was by far the most inaccurate, having >35% mean absolute error at 1 km and 17% at 100 km, following by MODIS (over 30% at 1 km and 11% at 100 km). SA-LC was the most accurate of all four maps with the density-independent measure, ranging from just above 10% mean absolute error at 1 km to 5% at 100 km.

Landscape characteristics and error. The density-independent measures indicate that error and bias increase with cropland density, which suggests that map error varies as a function of the typical abundance of cover in the landscape being imaged. To investigate the relationship between cover density and map accuracy, we calculated the mean density for agricultural pixels (>0.05% cropland) in the 1 km reference map within the boundaries of 354 magisterial districts (South Africa's finest administrative unit, averaging 3,445 km^2; SI Figure S3), providing a landscape-scale measure representing typical cropland density. We then extracted the cropland map errors for these same pixels, and calculated their district-wise mean absolute errors.

A generalized additive model fit to district-level mean absolute error (log-transformed) shows that error peaks at 50-60% cropland cover for all but the GlobCover map (which continued to increase with cropland cover), and is lowest when the landscape is dominated either by cropland or other cover types (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 cover class contributing to the percentage cropland estimate was a mixed class defined as 50-70% crops mingled with other vegetation, thus the maximum percentage was constrained by this mixture range.

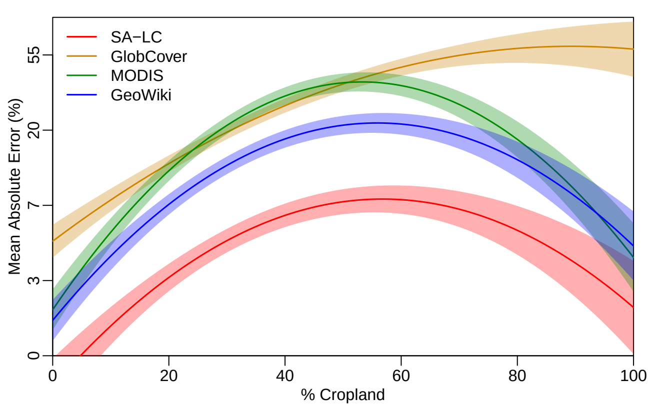

Fig. 2. The relationship between map accuracy (the mean absolute error) in cropland maps and the actual cropland cover within agricultural landscapes (reference map pixels having >0.5% cropland), here defined by the boundaries of magisterial districts (n = 345), as fit with a generalized additive model.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

First-order analyses

Tier-1 carbon estimates. Using the methods provided by (33), we calculated average carbon densities for African forests, secondary forests, shrublands, croplands, grasslands, and sparse habitats (semi-arid grasslands and low shrublands), and assigned cropland carbon values to map cells in proportion to their cropland cover. For the non-cropland proportions, we assigned the carbon value from each of the other types, creating five different carbon maps for each landcover map at each aggregation scale (Fig. S4), which allowed us to test how carbon estimates vary as a function of i) cropland map bias and ii) the characteristics of adjacent cover types.

The difference between total carbon stocks for the country made using any of the cropland maps were within +/-3% of those based on the reference map, regardless of which cover type was adjacent to cropland (Table S1), because the large of non-cropland in the country (~50-70%, Fig. 1) dilutes any map errors. Comparing total stocks between maps for just the agricultural area (30-50% of the country) reveals much greater differences (Table S1). SA-LC overestimated carbon stocks by just 2% when the adjacent cover type was forest, and up to 15% when it was sparse cover. MODIS ranged from negligible differences in denser carbon classes (forest, secondary forest, and shrublands) to 8-13% underestimates for the grassland and sparse classes. GeoWiki underestimated for all types, from <1% for sparse cover to 8% for forest. GlobCover grossly overestimated total carbon stocks for agricultural areas, varying from 64% for sparse lands to 162% for forest. The magnitude of this bias was due to false positives–GlobCover identified cropland in nearly 50% of pixels, compared to 30% for the other three cropland maps.

The spatial patterns of errors in carbon estimates (Fig. S4) reflect those of the cropland maps (Fig. 1). Where cropland was underestimated and the surrounding cover type was of higher carbon density than cropland, carbon density was overestimated. For lower density cover (grassland and sparse vegetation), carbon stocks were underestimated, but by small magnitudes. These tendencies were reflected in each map's biases, as calculated over the cropped areas of the country as jointly defined by the reference and each cropland map (Fig. 3). For example, MODIS and GlobCover bias was ~50% (overestimation) at 1 km resolution when forest was the cropland-adjacent cover (stars in semi-transparent green and gold bars, Fig. 3; Table S2). For sparse vegetation (open circles in Fig. 3), MODIS bias was 3% at all scales, whereas maps that overestimated cropland (e.g. SA-LC, semi-transparent red) overestimated carbon density for this cover type, because cropland has a higher carbon density (33). Overall, GeoWiki had the lowest bias, for all cover types and all resolutions. Its worst bias was a tendency to overestimate by 12% at 1 km when forest was adjacent, but at coarser scales this bias reduced to just a few percent (Fig. 3, Table S2). All maps' biases are within +/-10% bias after aggregation to 25 km.

The mean absolute error in carbon maps (solid colored bars in Fig. 3) generally followed the same patterns, but with higher magnitudes and a few important differences. The most notable is that GeoWiki, despite relatively low bias, was highly inaccurate at 1 km, with 27% mean absolute error across across cover types (line in solid blue bar, Fig. 3, Table S2), which is close to the 36-37% for GlobCover and MODIS. SA-LC had the lowest mean absolute error across scales, averaging (across cover types) 14% at 1 km to 3% at 100 km. (Fig. 3, Table S2). The decrease in GeoWiki's accuracy relative to SA-LC's can be attributed to the highly heterogeneous nature of its cropland errors (Fig. 1), which alternated between high magnitude positive and negative errors over relative short distances.

Gridded yield and production estimates. The disaggregated yield and harvested area maps of (34) are built upon cropland fraction maps where the total area is adjusted to match survey-derived cropland area statistics reported for administrative districts (provinces, in South Africa's case 35). To be consistent with this methodology, we first adjusted our cropland maps according to this procedure, using the reference map to calculate total cropland area for each of South Africa's nine provinces, then updating the pixel-wise cropland percentages in the four cropland maps so that the province-wise sums matched the reference areas (35, and see SI). Despite this statistical constraint, the updated cropland maps still had substantial errors that were similar in pattern (Fig. S5) to those in the unadjusted maps (Fig. 1), and we evaluated how these residuals affected gridded estimates of the yield and production of maize, South Africa's largest crop (39). To create these maps, we followed (34) by disaggregating district-level (n =

354, mean area = 3,445 km^2) agricultural census data (40) for maize (South Africa's largest crop by area, (39)) yield and harvested area, aggregated each set of maps, and multiplied the two to calculate production at each scale.

The yields disaggregated onto the cropland maps were markedly different to those on the reference map, particularly in the lower density cropland areas in the center of the country, where GlobCover overestimated yields and MODIS and GeoWiki (to a lesser extent) underestimated them at 1 km resolution (Fig. S6). However, only GlobCover showed a notable bias in yields at this resolution, which was equivalent to nearly 60% of the mean reference yield of 3.4 tons ha^{-1}. All other maps had biases of just +/-5% at 1 km (Fig. 5). Interestingly, GeoWiki and MODIS biases increased with aggregation, peaking at 10 km where both had underestimation biases of 30%, thereafter declining to 10% at 100 km. In contrast, GlobCover's yield bias declined linearly with aggregation (Fig. 5).

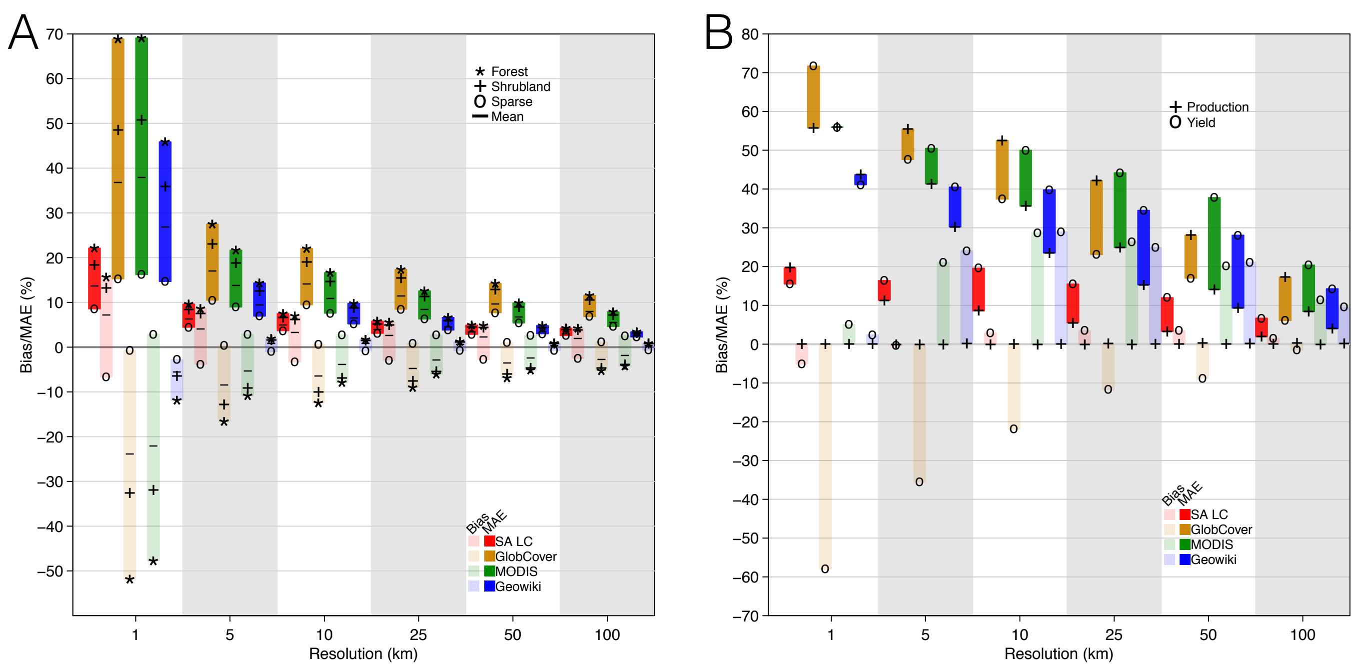

Fig. 3. Bias (mean error) and accuracy (mean absolute error [MAE]) in (A) carbon densities derived from cropland maps and (B) in disaggregated maize yield and production estimates. Bias estimates (represented by symbols) fall within the semi-transparent bars, mean absolute errors in the solid bars, with bar colors coded to specific cropland maps. For the carbon maps (A), symbols indicate the cover type used to calculate cropland-adjacent carbon density: the line represents the mean bias/MAE across all 5 cover types; shrubland and grassland values (near zero) and secondary forest (close to forest values) are not shown for clarity (see SI). For (B) symbols code the different variables (production and yield), normalized to their respective means. Values in A and B were calculated within the union of agricultural areas (cropland >0.05%) for each reference-cropland map combination.

Production errors were completely unbiased (Fig. 5). The statistical constraints on harvested and cropland areas resulted in the canceling out of spatial errors in production estimates, which is evident in the checkerboard-like pattern in maps of production biases (Figure S7). However, this reduction in bias comes at the cost of higher error magnitude, as the mean absolute error in production estimates were large, between 40 to 55% for GeoWiki, MODIS, and GlobCover at 1 km, and remained generally high (10-28%) even up to 50 km of aggregation (Fig. 5). SA-LC production biases were lowest across all spatial scales (20% at 1 km, dropping linearly to 2% by 100 km).

Absolute mean errors in yield were also substantial, and generally 10-15% larger than production biases across all aggregation scales, except for GlobCover where absolute production biases exceeded yield bias at 5-100 km of aggregation.

Second-order analyses

Evapotranspiration estimates. Compared to carbon and crop related examples, bias and accuracy in evapotranspiration (ET) calculated using the VIC model were small and averaged to less than +/-1%. However, there were several error hotspots in the resulting ET difference maps (Fig. 6). The most pronounced of these are the 5-15% overestimates in the center resulting from VIC when 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 (23), where growing season rainfall is approximately 200 mm.

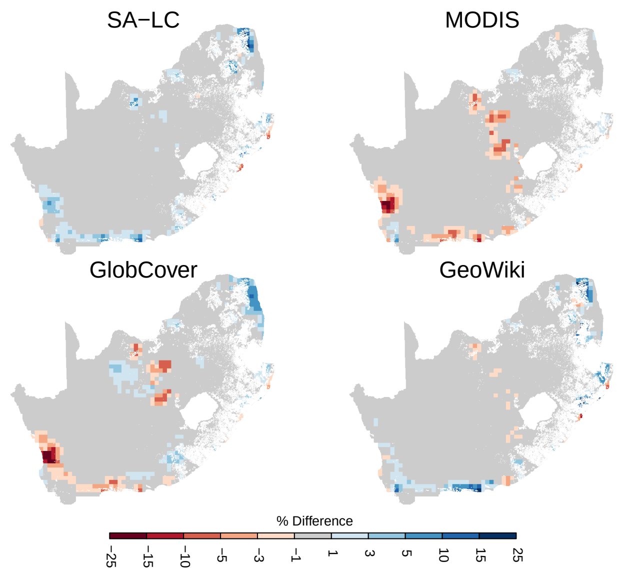

Fig. 4. Differences in annual mean evapotranspiration estimates from 29-year runs of the VIC land surface hydrology model when initialized with LAI response curves derived from the reference map, versus those from the four cropland maps.

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

Initializing an agent-based model. Spatially-explicit agent-based model (ABMs) are frequently employed to understand land use decision-making, often to facilitate improved policy, particularly in the arena of human development (41). A common feature of such ABMs is that they need to be calibrated against data describing the characteristics of land users, including an initialization step to assign land resources to "agents" representing the land users, wherein the simulated landscape pattern and distribution of agent resources matches those in the real world. In our example, we used an ABM of household food security that simulate the interactions between many individual farming households (the agents) and their environment over multiple seasons (38). We used cropland maps to provide the model the location and abundance of cropland, which is used to allocate an initial share of cropland to each simulated household. 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 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 is derived from survey data of communities where all cropland is owned. The model is thus considered adequately initialized when all households are allocated their appropriate area of cropland, and all cropland is occupied.

We used the reference map and each cropland map to separately initialize the model, and compared the agent allocation results to assess how cropland map errors impacted the initialization process. We examined three variable, the first being the number of agents that were not assigned fields, the second the amount of cropland left unallocated, and the third the area of land deficit, or the amount of land that should have been assigned to households but wasn't. For the first variable, there was a one-to-one relationship between the percentages by which cropland was underestimated and households that could not be assigned fields (Fig. 6, left panel). The most extreme examples occurred when MODIS cropland initialized the ABM in districts 1 and 2, where ∼85% of agents did not receive cropland. All households were assigned fields when total cropland area was overestimated (GeoWiki, SA-LC), but in these cases the area of cropland allocated to no one (the second variable) was proportional to the size of the overestimate (e.g. ∼20% for SA-LC, Fig. 6 right panel). Interestingly, the overall relationship between the percent of cropland allocated and percent cropland error was U-shaped, as the model also failed to give land to households when cropland was underestimated by more than 50% (Fig. 6, right panel). MODIS again provided the most pronounced results in districts 1 and 2, where 7-12% of cropland was left unallocated despite the fact that 85% of agents had no land. This curious relationship occurred because cropland tends to cluster, and when it is underestimated, the size of these clusters is small, resulting in islands of cropland that fall outside of the search radius (which is constrained by an absolute distance and the proximity of other agents) within which cropland is sought when agents are seeded onto the landscape.

The last measure, land deficit, increased exponentially in relation to cropland underestimation–reaching around 800% for MODIS in districts 1 and 2–and would become infinite in the case of a 100% underestimate.

Discussion

This spatially comprehensive, bottom-up assessment of landcover map bias and inaccuracy and provides unique insight into their extent, causes, and consequences for understanding global change processes, made possible by a unique, high accuracy dataset that likely provides the truest measure of total cropland area and distribution that is currently available for this region. This dataset is of course not perfect, being affected by the map-makers' occasional interpretation errors (mostly

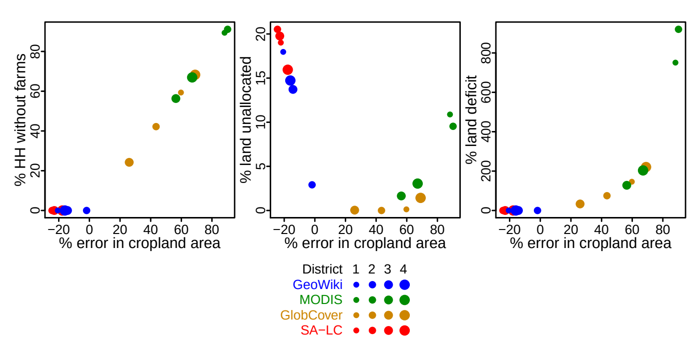

Fig. 5. Biases in agent-based model initialization relative to the district-wise errors (as a percent) in total cropland area, measured in terms of the percent of households having no cropland allocated (left), and the percent of cropland left unallocated (right). Dot sizes correspond to district numbers, colors represent the landcover map.

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

of omission), while some of the cropland map error we found may have been caused by the slight temporal mismatches between the reference data and the original landcover datasets we used. 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 large scale of disagreements between landcover map-based cropland area estimates and national inventory data (14).

Our results suggest several guidelines for using landcover data in global change research, and contain some important implications for how understanding of global change processes based on the data, and associated policy decisions, may be affected. In terms of developing a base landcover map, the first rule of thumb is that standard landcover products derived from coarse resolution sensors, such as MODIS and GlobCover, appear to be too biased to be useful without substantial aggregation. If we use the standard that bias within +/-10% is acceptable, then at least 25-100 km of aggregation is needed to sufficiently cancel out the errors in the base landcover data and subsequent first order estimates built on them (Fig. 3 & 4). The upper range of aggregation scale is necessary if a mixed pixel class becomes dominant, as in the case with GlobCover, because these lead to underestimation bias that will persist until the pixel size becomes substantially greater than the average area of landscapes that are dominated by the cover type of interest, which can be >1000 km^2 in some of South Africa's farming regions.

Maps derived from higher resolution sensors, such as the SA-LC dataset, if carefully done, do not have this mixed class problem, and are sufficiently unbiased for most applications with just 1-5 km of aggregation. However, such datasets are typically developed for specific countries, using varying methods, and can be hard to obtain. For broader scale analyses, the best option is to use newer generation maps such as GeoWiki (and the GLC-Share [6] datasets for other cover types) which is relatively unbiased at 1 km resolution. GeoWiki's lower bias comes from its process of evaluating consensus between several landcover datasets (including the other three in this study), resulting in cropland probabilities that are converted to percentages by calibrating to statistical data (15, 16). This method mirrors the ensemble averaging used by other fields (e.g. crop (42), climate (43), and ecological modeling (44)) to increase prediction confidence.

GeoWiki's statistical constraint procedure is similar to the one we used (following 35), which produced unbiased maize production estimates (Fig. 3) by eliminating bias in the adjusted cropland and harvested area maps that they were built upon. This result, together with GeoWiki's low bias, indicate the value of fusing inventory data with remote sensing. However, this method depends on the quality of inventory data, which are often suspect, particularly in Africa (45, 46). The statistical constraint also does not greatly improve map *accuracy*, as evidenced by GeoWiki's 23% mean absolute error in 1 km cropland percentage estimates (Fig. S1), which is only slightly more accurate than MODIS (31%) but worse than SA-LC (11%). GeoWiki is definitely most accurate among the large scale landcover products, but its improvement is related to the map consensus methods, which can correct for omission or commission errors made by the classifier. Statistical constraints only adjust map values at locations where cropland is identified, so their use it .

Map accuracy is perhaps more important than bias for landcover maps.

Broader regional implications - error higher elsewhere

Main points:

What we found, significance of study

- First large area quantification of spatial biases
- How large those biases are, for one of the most widely spread (spreading landcovers)
- Insight into causes of bias, and thus some understanding of where biases are likely to be greater or smaller
- How much progress made in reducing it
- Class type and bias

- Bias decreases as function of scale
- General bias patterns, appropriate use of landcover products, which landcover products
- Appropriate scales of inference, by type of product -
- Aggregation improves results for landcover, generally,
- Sensor resolution, statistical resolution, and merging products have high value
  - But don't remove spatial bias - absolute bias matters. Statistical constraints seems to just compress spatial biases to higher rates of turnover. Geo-wiki
  - But of course these types of data are then dependent on how accurate the statistical data are defining the constraint (cite emperor has no data)
- Mixed landscapes increase the chances of omission and commission errors by increasing the number of cover classes, or because such landscapes are less spectrally distinct (12)
- caveats: Only single country in South Africa. More commercial farming than many other countries, but results are still instructive. Analysis of error as function of landscape type suggests that areas where cropland is more mixed with natural vegetation have higher errors. These sorts of landscapes quite common in smallholder-dominated systems, thus suggests that biases may be even higher elsewhere on the continent.

Implications for understanding global change and policy:

Increasing awareness of need to have spatial assessments in global change analyses. Do things such as identify areas where yield gaps are high, or how much carbon or biodiversity will be lost to changes in land use, in order to try prioritize development (7, 47), or to understand coupled human-social dynamics, etc.

Our finding suggest:

- Area-based estimates only safe at coarser scales of aggregation for most types of global change analyses, and primarily with constrained products.
- 50-100 km scale of aggregation reduces bias sufficiently.
- Not so with unconstrained products
- Assessments of spatial variability unsafe, for all products, bar one - finer country-scale product. Here you look at absolute bias. This is high in many products even at higher scales of aggregation.
- This suggests that disaggregation approaches or paint by numbers approaches are nice maps, but can't give clear guidance about differences between grid cells, even when highly aggregated. [work on this]
- can lead to misinformed policies
  - E.g. Efforts to identify area where yield gaps are most pronounced and/or concentrated are likely to be highly misleading, leading to ineffective targeting of resources. Most informative simply to look at these areas at the political boundary resolution
  - Comparing carbon stocks against potential yield for tradeoff analysis, which may be done with conservation

planning to find areas with high benefit/low-cost. Also misleading.
  - Looking at land availability for cropland or biofuels expansion (look at biofuel paper for example)–land might not be as available as people think. Can lead to formulation of bad policy
- Analyses of higher order interactions, biogeochemistry, human decision-making, also misleading (maybe pair this with yield example).
  - Our example here, ET estimates not heavily biased, but in marginal areas of low rainfall some pronounced differences. These are areas where irrigation is more common, but VIC doesn't simulate this, so absolute bias in those zones likely to be underestimated, and such regions can have substantial impacts on altering climate (48, 49).
  - Can skew understanding of more advanced attempts to understand the human factors that go into driving agricultural productivity. Examples here

Way forward

- For now, use latest generation products fusion products or more detailed country-level products
- Avoid change detection based on landcover products, e.g. MODIS.
- But moving forward key will be developing new approaches to map landcover with much greater fidelity, e.g. scaling out approach that led to this dataset, combining with latest computer vision algorithms, etc.

Materials and Methods

Methods. Perhaps it is right **SI Materials and Methods**.

Describe weighted mean bias reasons.

We disaggregated the cropland percentages in all maps to binary cropland/noncropland cover types with 100 m resolution, which matches the typical field size (1 ha) for smallholder farmers in household survey data (collected in Zambia) used in developing the agent-based model (38). The surveys provided the mean cropland area per household (2.2 ha) and frequencies distribution of cropland area holdings across households (e.g. how many households have 1 ha, 2 ha, etc.). We used these statistics to calculate the "true" number of households per district by dividing reference cropland areas by the mean cropland area, and preserved the cropland area distributions by multiplying the total number of households by the frequencies. We then initialized the model, which takes a weighted (by cropland area frequency) random draw of 100 households and places these within the 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.

Digital RCD Analysis.

Appendix: App 1

Appendix

This is an example of an appendix without a title.

ACKNOWLEDGMENTS. I thank everyone tearfully.

References

1. Lambin EF (1997) Modelling and monitoring land-cover change processes in tropical regions. *Progress in Physical Geography* 21(3):375–393.
2. Lark TJ, Salmon JM, Gibbs HK (2015) Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environmental Research Letters* 10(4):044003.
3. 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.
4. 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.
5. 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.
6. 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.
7. Newbold T et al. (2015) Global effects of land use on local terrestrial biodiversity. *Nature* 520(7545):45–50.
8. 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.
9. 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.
10. See L et al. (2015) Improved global cropland data as an essential ingredient for food security. *Global Food Security* 4:37–45.
11. Lobell DB (2013) The use of satellite data for crop yield gap analysis. *Field Crops Research* 143:56–64.
12. Estes L et al. (2015) DIYlandcover: Crowdsourcing the creation of systematic, accurate landcover maps. *PeerJ PrePrints* 3:e1266.
13. 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.
14. 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.
15. Fritz S et al. (2011) Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. *Geophysical Research Letters* 38:L04404.
16. Fritz S et al. (2015) Mapping global cropland and field size. *Global Change Biology* 21(5):1980–1992.
17. 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.
18. Fritz S et al. (2012) Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software* 31:110–123.
19. 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.
20. Ge J et al. (2007) Impacts of land use/cover classification accuracy on regional climate simulations. *Journal of Geophysical Research: Atmospheres* 112(D5):D05107.
21. 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.
22. Quaife T et al. (2008) Impact of land cover uncertainties on estimates of biospheric carbon fluxes. *Global Biogeochemical Cycles* 22(4):GB4016.
23. 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.
24. Estes LD et al. (2014) Using changes in agricultural utility to quantify future climate-induced risk to conservation. *Conservation Biology* 28(2):427–437.
25. Fourie A (2009) Better Crop Estimates in South Africa. *ArcUser Online* (1).
26. SANBI (2009) National Landcover 2009, (South African National Biodiversity Institute; National Department of Environmental Affairs and Tourism, Pretoria, South Africa), Technical report.
27. DAAC) LPDAACL (2011) MODIS MCD12q1 Land Cover Type Yearly L3 Global 500 m SIN Grid. Version 5.01, (NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota), Technical report.
28. Friedl MA et al. (2010) MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment* 114(1):168–182.
29. Arino O et al. (2012) *Global land cover map for 2009 (GlobCover 2009)*. (European Space Agency & Universit Catholique de Louvain).
30. 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.
31. Gross D et al. (2013) Monitoring land cover changes in African protected areas in the 21st century. *Ecological Informatics* 14:31–37.
32. 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.
33. 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*.
34. 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.
35. 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.
36. Foley JA et al. (2011) Solutions for a cultivated planet. *Nature* 478(7369):337–342.

37. 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.
38. Chen P, Plale B, Evans T (2013) *Dependency Provenance in Agent Based Modeling*. pp. 180–187.
39. 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.
40. Africa SS (2007) Commercial Census of Agriculture, South Africa.
41. Berger T, Schreinemachers P (2006) Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society* 11(2):19.
42. Asseng S et al. (2013) Uncertainty in simulating wheat yields under climate change. *Nature Climate Change* advance online publication.
43. Giorgi F, Mearns LO (2002) Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the Reliability Ensemble Averaging (REA) method. *Journal of Climate* 15(10):1141–1158.
44. Arajo M, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution* 22(1):42–47.
45. Carletto C, Jolliffe D, Banerjee R (2013) The emperor has no data! Agricultural statistics in Sub-Saharan Africa. *World Bank*.
46. 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).
47. 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.
48. Estes LD et al. (2014) Changing water availability during the African maize-growing season, 19792010. *Environmental Research Letters* 9(7):075005.
49. 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.