Errors in Landcover Data Limit Our Understanding of Global Change

Lyndon Estes [*][†], Peng Chen [‡], Stephanie Debats [†] , Tom Evans [‡] , Stefanus Ferreira [§], Gabrielle Ragazzo [¶], Justin Sheffield [†] , Adam Wolf [¶] , Eric Wood [†] , and Kelly Caylor [†]

[*]Woodrow Wilson School, Princeton University, Princeton, NJ USA,[†]Civil and Environmental Engineering, Princeton University, Princeton, NJ USA,[‡]Indiana University, Bloomington, IN USA,[§]GeoTerraImage, Pretoria, RSA, and [¶]Ecology and Evolutionary Biology, Princeton University, Princeton, NJ USA

Research into global environmental change is often built upon landcover maps, which can have substantial errors. How these errors impact our understanding of global change processes and associated policies is thus of great interest, but is difficult to fully quantify because we lack spatially comprehensive reference data, particularly in the World's most rapidly developing regions. We used a high quality, national-scale cropland map to assess bias and accuracy in current generation landcover maps and four examples of "downstream" (landcover-dependent) analyses, two related to physical processes (estimates of vegetative carbon stocks and evapotranspiration) and two to socio-economic processes (gridded crop production estimates and agent-based simulation of household food security), at resolutions ranging from 1-100 km. We found that the landcover maps have substantial errors below 25 km resolution, with substantial biases and inaccuracies, particularly from maps derived from coarse resolution sensors. In some cases these error metrics remain high even when maps are aggregated to 100 km. These errors can be substantially amplified in downstream analyses (e.g. the carbon and crop production analyses), but in cases where the values of different cover types that underpin the analysis are similar (e.g. evapotranspiration estimates) the results can be relatively insensitive. To avoid these errors, and thus minimize the risk of misunderstanding global change processes or misinforming policy, substantial map aggregation of landcover maps from their base resolution is often needed. We provide several guidelines that can help landcover map users select appropriate datasets and aggregation scales under different use cases.

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

T he 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 imaging, which has several important constraints–related to both physics and analytics–that propagate mapping errors. First, in many regions the cover types of interest are smaller than the sensor resolution (e.g. smallholder's farms 11–13), or spectrally indistinct from neighboring covers (14), which are factors that increase mapping complexity (15). Second, the act of defining a cover class can cause error, in that selected classes may have highly diverse spectral properties (e.g. agriculture (12, 16)) and thus be difficult for the classifier to distinguish, while discretizing a continuous cover type (e.g. dividing a forest into different canopy cover classes) can also promote classification error, particularly near class boundaries (17), as well as confusion about the actual extent of the cover type (18). Furthermore, class definitions often vary between maps, complicating inter-comparison (19). Third, landcover map series are often used to detect changes (e.g. 20), or may use within-class seasonality as a means for discriminating between classes (e.g. phenology of cropping versus surrounding vegetation (14)), but seasonal variability and landcover changes can be easily confused. Given these multiple sources of error, landcover maps are often inaccurate at finer scales and disagree widely between products, particularly in the world's most rapidly developing regions (21–23). These errors limit our ability to obtain granular, mechanistic understanding of global change processes.

These problems with landcover products is known (22–26), and there are a variety of map improvement efforts underway (e.g. 16, 27). 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. Fully answering this question depends on having comprehensive, spatially representative ground truth data, which are typically unavailable for Africa and other rapidly changing regions (24), which have also been fairly sparsely mapped (15). Our current under-

Reserved for Publication Footnotes

standing of map accuracy over such areas is therefore based on a small number of maps (15), which were assessed using either bottom-up tests made with a relatively small number of ground truth points, or from top-down "sanity checks" made in comparison to aggregated survey data. In terms of the former, there are well-developed, best-practice guidelines for selecting and using ground-truth samples to robustly quantify map error, but these are not commonly applied by the broader community of landcover map users (often due to lack of data), which makes it challenging to compare how error varies within and between maps (17, 28–30). The latter form of assessment are made at coarser administrative levels relative to reported statistics (e.g. 22), which can also be suspect (31, 32).

Since it is difficult to fully quantify landcover map errors, it is even more challenging to gauge their impact on downstream analyses, in which there is substantial risk of error amplification(19). There has been some work examining how map error influences climate simulations (33), agricultural land use patterns (34), carbon flux measurements (35), human population estimates (10), and species distribution models (36), but these either use simulated landcover errors (33) or compare relevant differences in estimates between different satellite-derived landcover maps (10, 35, 36). One exception is a Belgian study (34) 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 (12, 21, 27, 36), are finally providing the means to more 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 (37).

In this study, we take advantage of this recent growth in high resolution data to address the call to more thoroughly assess errors and in landcover maps (19, 30, 38), and further examine how these errors 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 both the physical and socioeconomic components of global change. Our objective is to provide scientists and policy-makers who use landcover maps an up-to-date assessment 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 (39). The resulting vectorized field boundaries provide highly accurate data on field sizes and distribution for the period 2009-2011. Measured in terms of its ability to distinguish crop fields from other cover types, this dataset is 97% percent accurate at 1 km resolution, with user's accuracies of 94% and 98% for the cropland and non-cropland classes, and producer's accuracies of 84 and 99% (16, SI).

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 (37), yet is notably lacking trustworthy maps of existing agriculture land (22). Moreover, South Africa's large agricultural sector represents the diversity of farming systems found throughout SSA, ranging from large commercial operations to smallholder farms (40, 41).

We used this dataset as a reference layer for evaluating four landcover products representative of the type commonly used in global change studies. The first was South Africa's own 30 m resolution 2009 National Landcover map (SA-LC)(42), which is typical of the higher-resolution, Landsat-based maps that are typically available only for individual countries (e.g. 43). The second and third were respectively the 300 m GlobCover 2009 (44) and 500 m resolution 2011 MODIS Landcover products, which are widely used global-scale products (e.g. 20, 45). The fourth dataset was the new 1 km GeoWiki hybrid-fusion cropland map for Africa (25), which incorporates the first three datasets and represents the current state-of-the-art in landcover mapping. Since GeoWiki provides a continuous bounded value of landcover (percent cropland cover), which cannot be feasibly converted to a categorical value, we converted all other datasets, including the reference vectors, into comparable 1 km gridded percent cropland estimates.

We first evaluated the quality of the landcover product-derived percent cropland maps (hereafter referred to as the "test maps"). Instead of the standard confusion matrix-based accuracy assessment metrics (29, 30), which apply to categorical landcover maps, we assessed the bias and accuracy of the test maps based on the gridded residuals that resulted when each test map was subtracted from the reference map. Here bias is the mean residual value, weighted by the density of reference cropland, and accuracy is the mean of the absolute values of residuals (also weighted by cropland density), thus accuracy is inversely related to this value. We calculated these metrics 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 bias and accuracy 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 global change research. The first two related to physical processes, namely the calculation of vegetative carbon stocks and the simulation of evapotranspiration. The second two were socio-economic in nature: the estimation of agriculture yield and production, and an agent-based model simulating household dynamics in dryland agroecosystems. 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 nearly 10% of the total study area in the 2009-2011 time period. 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

---

[1] e.g. 3DRobotics, DJIA
[2] Planet Labs, Skybox

eastern and northern margins. These patterns translated into substantial map bias (Fig. 1B), with GlobCover and MODIS mean error 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 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 high heterogeneous error pattern, 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, 16, 24), we set 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.
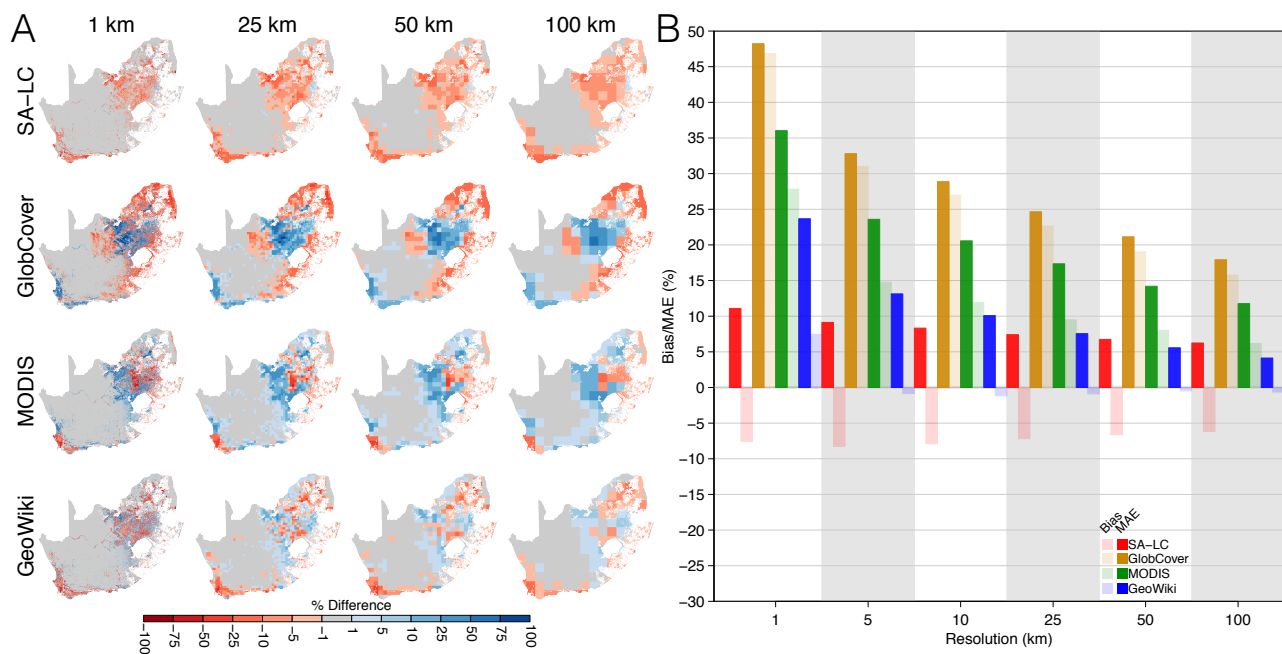

Fig. 1. (A) Errors in the percent cropland estimates resulting from each of the four test maps relative to the reference map at different scale of pixel aggregation. Rows indicate the test map being assessed (by subtraction from the reference map), while columns refer to resolution of aggregation. White 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.

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 vegeta-

tion, thus the maximum percentage was constrained by this mixture range.

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 (37). One widely used vegetative carbon dataset is that of (46), 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 (46).

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 (46)) to the non-
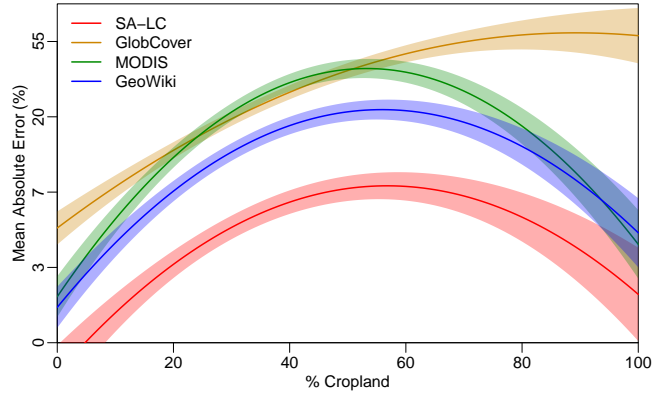

Fig. 3. The relationship between map accuracy (the mean absolute error) in test maps and the actual cropland cover within agricultural landscapes (reference map pixels having >0.5% cropland), here defined by the boundaries of magisterial districts (n = 345), as fit with a generalized additive model. 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.
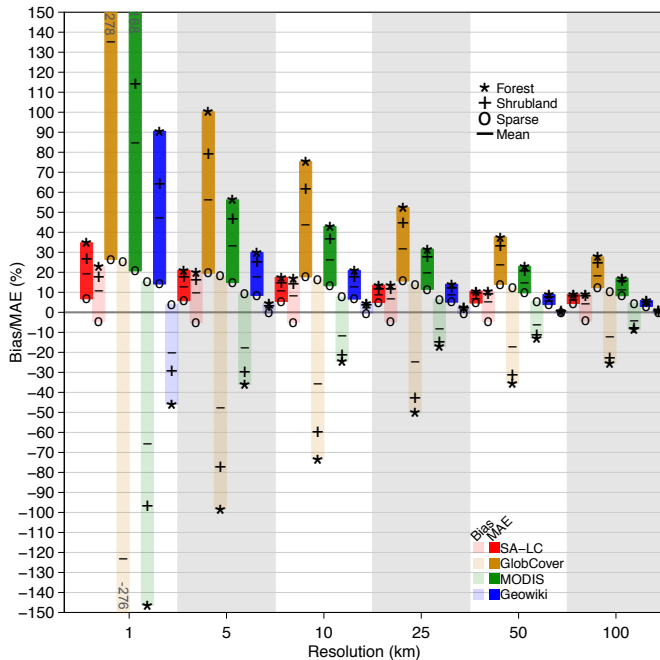

Fig. 4. Biases and accuracies (mean absolute errors) of carbon densities derived from cropland maps, calculated as percents relative to the reference map. Bias estimates (represented by symbols) fall within the semi-transparent floating bars, while accuracies are contained in the solid bars. Bar colors 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.

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% mean error 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 <10% at ≥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 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 were 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 (47). Land surface hydrological models, such as the Variable Infiltration Capacity (47), 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 (40), 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 country were either negligible to absent. All but MODIS underestimated ET by 5-15% in the northern tip of the country.
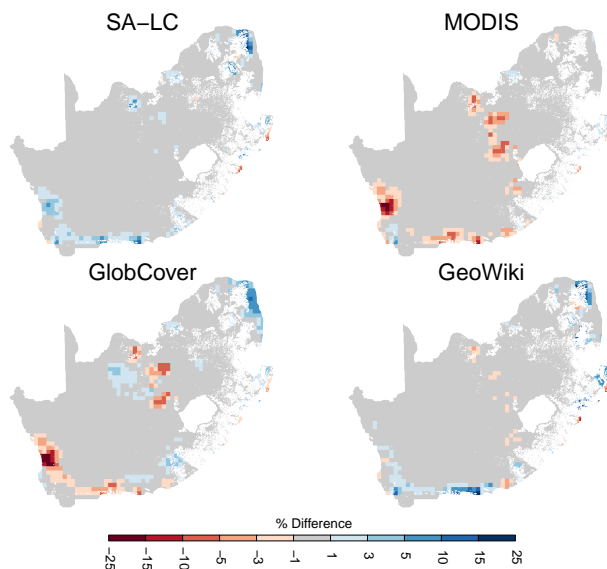

Fig. 2. Differences in annual mean evapotranspiration estimates from 29-year runs of the VIC land surface hydrology model when initialized with LAI response curves derived from the reference map, versus those from the four test maps.

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, 48). 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 land-cover data (48, 49). 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 (24, 49)

We used these same methods (48, 49) to disaggregate maize harvested area (South Africa's largest crop (21)) 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 we 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 <10% with aggregation.

Production biases were generally slightly higher, but still low, for most datasets, with the exception of GlobCover, which had a large 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.

In contrast, the accuracy of production estimates was poor. Here all datasets but SA-LC had MAE values of ≥30% below 25 km of aggregation (Fig. 5), reaching as high as 100% for GlobCover at 1 km, followed by 65% for MODIS and 45% for GeoWiki. 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 totals match reported statistics does relatively little to correct the map's underlying commission or omission errors, and this constraint in fact appears to shorten the distance between negative and positive residuals (SI), thereby increasing absolute errors.

Agent-based simulation of households in dryland agroecosystems. 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 (50). 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 (50). In our example, we used an ABM of household food security that simulates 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 (51). The model is initialized such that each household is allocated its designated share of cropland based on aggregated farm-level statistics, and then calculates household-level agricultural production based on maize yields simulated by the DSSAT cropping system model (52). Yield estimates vary as a function of soil properties, management inputs, and weather. The former two factors differ between households, as do the total areas of cropland, thus total production can vary substantially between households. 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 farm-
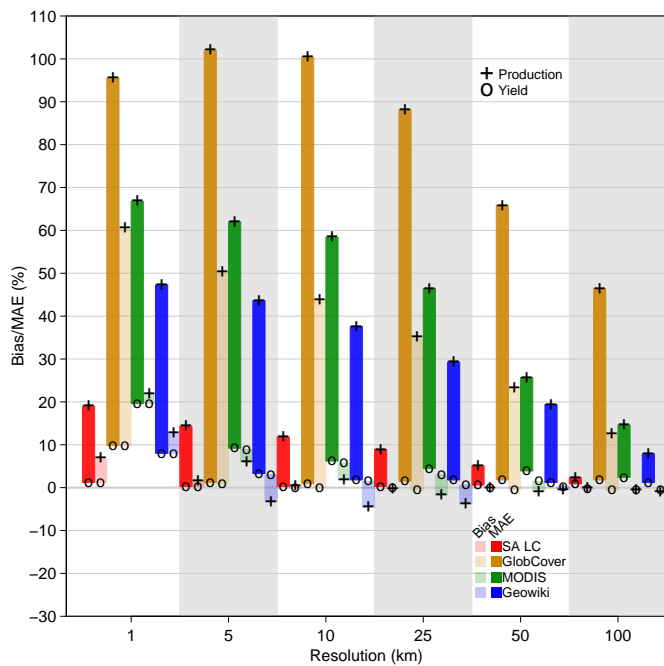
Fig. 5. Bias (mean error) and accuracy (mean absolute error [MAE]) in disaggregated maize yield and production estimates. Bias estimates (represented by symbols) fall within the semi-transparent bars, mean absolute errors in the solid bars, with bar colors coded to specific cropland maps. Symbols code the different variables (production and yield), normalized to their respective means.

ers. To match these computational characteristics, we selected four contiguous magisterial districts (ranging from 1,040-1,343 km^2, SI) 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 be well-calibrated when all households were allocated their appropriate area of cropland, and all cropland was allocated to households.
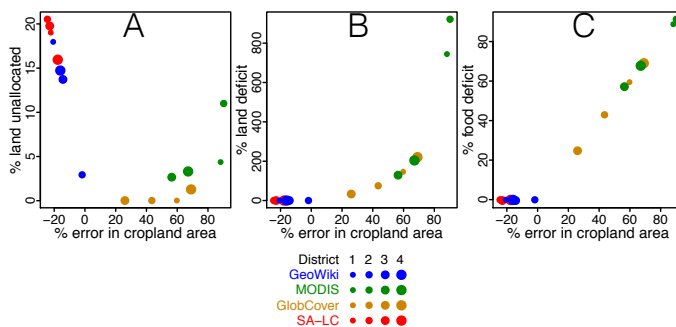

Fig. 6. Biases in agent-based model results relative to the district-wise errors (as a percent) in total cropland area, in terms of A) the percent of cropland in each district that was not allocated to any household, B) the land deficit, or the total area of cropland that should have been allocated to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

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 what share of the district's total cropland area was not assigned to any household, and is a measure of 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 non-linear 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, is calculated from the total area of cropland that should have been allocated to households in each district, but wasn't, due to the mismatch between the cropland map and the statistical estimates of what total cropland holdings should have been (given the number of households and the size class distributions of cropland in the district). Land deficit, expressed as a percentage of the total area of cropland in the district (thus values exceeded 100% when total deficit exceeded cropland area), 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 percentage shortfall in the average amount of food production 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 landcover map assessment provides new insight into the extent and causes of map error, and its consequences for understanding global change. The analysis was made possible by a spatially comprehensive, high accuracy dataset that likely provides the truest measure of cropland area and distribution currently available for this region. Although this reference map is not perfect, being affected by the map-makers' occasional interpretation errors (mostly of omission, SI), while temporal mismatches between the reference and test maps may account for some of the error we identified, our assessment (SI) suggests that such discrepancies do not appreciably impact our findings. Our results are also supported by previous, continental-scale assessments that revealed substantial errors and inconsistencies between landcover maps (20, 22). These similarities, combined with the larger extents of these studies, also suggest that our findings are applicable beyond South Africa's borders.

General guidelines for selecting and using landcover data. Our findings regarding bias and accuracy suggest several guidelines for selecting and using landcover data in global change research (Table 1), and suggested aggregation scales. These guidelines vary according to the intended application, what

error property the users wants to minimize or maximize, and what tradeoff the user is willing to make between error and resolution. For the purpose of clarity in these guidelines, we assume that a maximum error magnitude of 10% is acceptable, but the results of our analyses (Fig. 1, 3, 5) allow different thresholds to be evaluated.

The first general guideline is that landcover products derived from coarse resolution sensors, such as MODIS and Glob-Cover, 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. If the objective is to minimize systematic errors in order to calculate an unbiased mean (e.g. average carbon density), then at least 25-50 km of aggregation is needed to reduce map bias to within +/-10% (Table 1, Fig. 1B & 3). If the user wants to maximize accuracy, which 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 (20). Outside of South Africa, where smallholder-driven agriculture predominates (53), 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% (54)). Our results also suggest that, despite their higher accuracy, they may 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 lowest bias across most scales, which is the largely the result of calibrating the cropland percentages to match reported agricultural statistics (23, 25). This procedure is similar to the one we used in disaggregating maize harvested areas (48, 49), 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 (31, 32), which is a vulnerability of the fusion approach (24). 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. To minimize these types of errors, different techniques are needed, such as creating a single consensus map from multiple landcover products (e.g. 22, 36). This second kind of data fusion was used in the first stage of creating the GeoWiki map (before statistical adjustment), and it accounts for GeoWiki's higher accuracy scores relative to MODIS and GlobCover (Fig. 1B, 3, & 5). It may be particularly effective in minimizing underestimation errors that result when a mixed class type becomes dominant (as with GlobCover), or when dealing with substantial sub-pixel heterogeneity (25, 36)

These last few points suggest two final guidelines. For high accuracy, users should select well-validated datasets from higher resolution sensors–this option should be increasingly

---

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

Table 1. Guidelines for selecting and using different classes of landcover product (based on the characteristics of the satellite sensor(s) and map-making methods), according to the general purpose of the application, the error minimization objective, and the range of aggregation needed to achieve the error objective (here we use a maximum acceptable error magnitude of 10%). The preferred landcover product class for each application is indicated in boldface.

Application	Objective	Source of Landcover Map	Minimum aggregation*
High frequency change detection	Maximize accuracy	Coarse resolution sensor (e.g. MODIS)	>100 km
		High resolution sensor (e.g. Landsat)	1-25 km
Mean values (e.g. carbon density)	Minimize bias	Coarse resolution sensor	25-50 km
		if mixed class dominates	>100 km
		Fusion product: statistical constraint	1-5 km
		High resolution sensor	1-10 km
Spatial comparisons (e.g. crop production)	Maximize accuracy	Coarse resolution sensor	25-100 km
		if mixed class dominates	>100 km
		Fusion product: map consensus**	10-50 km
		High resolution sensor	1-25 km

*Aggregation ranges account for variation in the results from the different map analyses.
**This assumes that consensus-based fusion products (e.g. 23, 25, 36) incorporate maps from coarse resolution sensors, making them less accurate at finer scales than products created using high resolution sensors.

available as computational power grows and high resolution image archives expand (54, 55); 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–1-25 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 products often represent just a single time period (e.g. 25, 36), thus applications requiring higher frequency (e.g. annual), large area landcover maps will depend on MODIS or similar sensors, 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 potential to skew our understanding of global change processes. For example, the global carbon carbon density map of (46), 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 ∼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 (56), which our results suggest could be substantially underestimated (Fig. 3). In contrast to these, an assessment of the congruence between carbon and biodiversity (57) aggregated the carbon densities to ∼110 km, where accuracy is much higher.

A related example can be seen in high resolution maps (10 km) showing the potential food production benefit of closing yield gaps (e.g. Figure 3 in 58). Such maps may 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 and maps the per hectare values, and the map-makers stress that their products should not be used at the grid cell scale[4], it is not inconceivable that users might make decisions based on their visual interpretations of maps. For this reason, it is advisable to aggregate presented maps until their accuracy reaches an acceptable standard (Table 1).

What then, of the impacts of landcover map error 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 (33). 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 and can substantially modify regional climate(59, 60). 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. (33).), 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 (61). 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. One of these lies in assessments of land availability for agricultural expansion, which could be particularly erroneous if they use a "residual approach" (53), 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 (62). 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 (37). This sort of analysis entails finding "win-win" land use solutions, which are essentially areas that deviate from an expected development cost-benefit relationship; for example, regions of unusually low carbon density but high crop fertility (37). The validity of such analysis are entirely dependent 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 (37). The validity of such analyses, which must be undertaken at the fine-scales where land use decisions are made (hectares rather than square kilometers), depends on having high quality map data, in order to

---

[4] www.earthstat.orgwp-contentuploadsMETADATA_HarvestedAreaYield175Crops.pdf

have confidence that the identified tradeoffs are based on real biophysical properties, not simply map error.

The way forward. Our analysis demonstrates that the land-cover 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 (Table 1) for selecting and using land-cover 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 (48, 49) using the GeoWiki map could substantially improve our knowledge about current global agricultural distribution and productivity. Similarly, reconstructing the IPCC's Tier 1 vegetative carbon map (46) using a combination of the GeoWiki and new global landcover consensus maps (e.g. GLC-Share or 36) the GLC-Share datasets would improve our confidence in regional to global assessments of vegetative carbon stocks. Alternatively, the new high-resolution, Landsat-based GLC30 (54), aggregated and converted to fractional cover classes at 1 km–and fused with high quality national landcover maps, where available (similar to 23, 25)–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 new high resolution global forest change maps (18, 55) are prominent examples of these newer methods. Another approach is that of the Mapping Africa project[5], which, inspired by quality of the reference map used in this study, enlists internet-based workers to manually digitize cropland boundaries (16). Merging this crowdsourced approach for creating vector maps of discrete cover types with newer computer vision-based mapping algorithms (12) 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 (16)) 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 create the next generation of landcover maps, so that we can gain a clearer, unbiased understanding of global change.

Materials and Methods

Methods. To develop the cropland reference map (see SI for accuracy assessment details), we intersected the original field boundary vector maps 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 (39), which posed a risk of overestimating cropland extent. We also excluded the permanent tree crop class, to avoid confusion with landcover maps that exclude 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, following the approach used in creating the GeoWiki maps (25) 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; 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 the coordinates of each district's centroid, and weighted the model by the number of agricultural pixels per district (63).

To develop the carbon maps, we followed the methods used by Ruesch and Gibbs (46) 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 (SI).

We used the Variable Infiltration Capacity (47) land surface hydrology model run with the Africa Flood and Drought Monitor's meteorological data (64) to produce monthly gridded evapotranspiration estimates for South Africa for the years 1979-2010. We adjusted VIC's native landcover scheme (derived from AVHRR) 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 Ramankutty et al. (49) 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 (48) for disaggregating maize (South Africa's largest crop 65) planted area and yields onto the reference and adjusted test maps. The necessary statistics were obtained from magisterial district-level agricultural censuses conducted for the year 2007 (66). 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) (51). 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 each district, 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 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, as simulated by the DSSAT model. We ran each ABM simulation 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.

---

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

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. Debats S, Luo D, Estes L, Fuchs T, Caylor K (2015) A generalized computer vision approach to mapping agricultural fields in Sub-Saharan Africa. *PeerJ PrePrints* 3:e1688.
13. Ozdogan M, Woodcock CE (2006) Resolution dependent errors in remote sensing of cultivated areas. *Remote Sensing of Environment* 103(2):203–217.
14. Sweeney S, Ruseva T, Estes L, Evans T (2015) Mapping Cropland in Smallholder-Dominated Savannas: Integrating Remote Sensing Techniques and Probabilistic Modeling. *Remote Sensing* 7(11):15295–15317.
15. Yu L et al. (2014) Meta-discoveries from a synthesis of satellite-based land-cover mapping research. *International Journal of Remote Sensing* 35(13):4573–4588.
16. Estes L et al. (2015) A platform for crowdsourcing the creation of representative accurate landcover maps. *PeerJ PrePrints* 3:e1844.
17. Foody GM (2002) Status of land cover classification accuracy assessment. *Remote sensing of environment* 80(1):185–201.
18. Sexton JO et al. (2015) Conservation policy and the measurement of forests. *Nature Climate Change* advance online publication.
19. Kuemmerle T et al. (2013) Challenges and opportunities in mapping land use intensity globally. *Current Opinion in Environmental Sustainability* 5(5):484–493.
20. Gross D et al. (2013) Monitoring land cover changes in African protected areas in the 21st century. *Ecological Informatics* 14:31–37.
21. 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.
22. 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.
23. Fritz S et al. (2011) Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. *Geophysical Research Letters* 38:L04404.
24. See L et al. (2015) Improved global cropland data as an essential ingredient for food security. *Global Food Security* 4:37–45.
25. Fritz S et al. (2015) Mapping global cropland and field size. *Global Change Biology* 21(5):1980–1992.
26. 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.
27. Fritz S et al. (2012) Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software* 31:110–123.
28. Stehman SV, Olofsson P, Woodcock CE, Herold M, Friedl MA (2012) A global land-cover validation data set, II: augmenting a stratified sampling design to estimate accuracy by region and land-cover class. *International Journal of Remote Sensing* 33(22):6975–6993.
29. Olofsson P, Foody GM, Stehman SV, Woodcock CE (2013) Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment* 129:122–131.
30. Olofsson P et al. (2014) Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment* 148:42–57.
31. Carletto C, Jolliffe D, Banerjee R (2013) The emperor has no data! Agricultural statistics in Sub-Saharan Africa. *World Bank*.
32. 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).
33. Ge J et al. (2007) Impacts of land use/cover classification accuracy on regional climate simulations. *Journal of Geophysical Research: Atmospheres* 112(D5):D05107.
34. 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.
35. Quaife T et al. (2008) Impact of land cover uncertainties on estimates of biospheric carbon fluxes. *Global Biogeochemical Cycles* 22(4):GB4016.
36. Tuanmu MN, Jetz W (2014) A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. *Global Ecology and Biogeography* 23(9):1031–1045.
37. 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.
38. Olofsson P et al. (2012) A global land-cover validation data set, part I: fundamental design principles. *International Journal of Remote Sensing* 33(18):5768–5788.
39. Fourie A (2009) Better Crop Estimates in South Africa. *ArcUser Online* (1).
40. 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.
41. Estes LD et al. (2014) Using changes in agricultural utility to quantify future climate-induced risk to conservation. *Conservation Biology* 28(2):427–437.
42. SANBI (2009) National Landcover 2009, (South African National Biodiversity Institute; National Department of Environmental Affairs and Tourism, Pretoria, South Africa), Technical report.
43. 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.
44. Arino O et al. (2012) *Global land cover map for 2009 (GlobCover 2009).* (European Space Agency & Universit Catholique de Louvain).
45. 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.
46. 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.*
47. 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.
48. 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.
49. 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.
50. Berger T, Schreinemachers P (2006) Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society* 11(2):19.
51. Chen P, Plale B, Evans T (2013) *Dependency Provenance in Agent Based Modeling.* pp. 180–187.
52. Jones J et al. (2003) Modelling Cropping Systems: Science, Software and Applications. *European Journal of Agronomy* 18:235–265. 3-4.
53. Lambin E et al. (year?) Estimating the world's potentially available cropland using a bottom-up approach. *Global Environmental Change.*
54. 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.
55. Hansen MC et al. (2013) High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* 342(6160):850–853.
56. Cattaneo A et al. (2010) On international equity in reducing emissions from deforestation. *Environmental Science & Policy* 13(8):742–753.
57. Strassburg BB et al. (2010) Global congruence of carbon storage and biodiversity in terrestrial ecosystems. *Conservation Letters* 3(2):98–105.
58. Foley JA et al. (2011) Solutions for a cultivated planet. *Nature* 478(7369):337–342.
59. 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.
60. Mueller ND et al. (2015) Cooling of US Midwest summer temperature extremes from cropland intensification. *Nature Climate Change.*
61. Janssen MA, Ostrom E (2006) Empirically based, agent-based models. *Ecology and Society* 11(2):37.
62. Rulli MC, Saviori A, D'Odorico P (2013) Global land and water grabbing. *Proceedings of the National Academy of Sciences* 110(3):892–897.
63. Wood SN (2001) mgcv: GAMs and Generalized Ridge Regression for R. *R News* 1(2):20–25.
64. 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.
65. 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.
66. Africa SS (2007) Commercial Census of Agriculture, South Africa.