Landcover Data Limits Our Understanding of Earth System Processes

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 functioning of the Earth System[1] is fundamentally connected to the characteristics of landcover (1), the physical constituents of the terrestrial surface. Human endeavors are both strongly governed by and shape landcover, whether it be felling ancient forests for timber production or burning savannas to stimulate a green flush for livestock, while landcover is a primary driver of climate and biogeochemical processes (2). The vastness of our 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 therefore essential to know the nature and distribution of landcover.

This importance is understood by a growing number of social, economic, and physical scientists, who increasingly use landcover data to advance understanding in Earth System research areas ranging from food security (3–5) to carbon cycling (6, 7), biodiversity loss (8, 9), and demographic shifts (10).

The validity of the knowledge resulting from such studies depends on the veracity of the landcover data upon which they are based, much as a building requires a solid foundation if it is to remain standing. Unfortunately, the evidence so far suggests that Earth System science is being built on a shaky foundation. The reason for this is that landcover data can only practically be derived from satellite imagery, but in many regions the cover types of interest are smaller (e.g. smallholder's farms 11) than the sensor resolution, or spectrally indistinct from neighboring covers, which translates into substantial mapping errors (12–14). 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 (15–17). These errors mean that we are still unable to obtain the granular, mechanistic understanding of global change processes that we need.

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

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

Fortunately, the recent, explosive growth in public and private initiatives to develop new Earth observing capabilities, which range from small drones[2] to new high resolution satellite arrays[3] and better mapping methods (15, 20, 25), are finally providing the means to comprehensively interrogate the accuracy and biases in the landcover products that have become commonplace in global change research–and which are often used to make policy decisions (26).

Reserved for Publication Footnotes

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

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

Overview of study area and analyses

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

We used this dataset as a reference layer for evaluating four landcover products representative of the type commonly used in Earth Systems research. The first was South Africa's own 30 m resolution 2009 National Landcover map (SA-LC)(30), which is typical of the higher-resolution, Landsat-based maps that are typically available only for individual countries (e.g. 34). The second and third were respectively the 300 m GlobCover 2009 (33) and 500 m resolution 2011 MODIS Landcover products, which are widely used global-scale products (e.g. 35, 36). The fourth dataset was the new 1 km GeoWiki hybrid-fusion cropland map for Africa (18), which incorporates the first three datasets and represents the current state-of-the-art in landcover mapping. For comparison, we converted all five datasets into 1 km gridded maps of cropland density, expressed as the pixel-wise percent cover.

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

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

Map quality

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

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

Error and landcover density. Given that landcover characteristics can influence landcover map quality (12, 14, 25), we set out to elucidate the relationship between cropland configuration and map error. We first calculated the average 1 km reference cropland percentage within each of South Africa's 354 magisterial districts (the finest administrative unit, averaging 3,445 km^2; SI), which yielded a landscape-scaled metric of characteristic cropland density, and similarly calculated the district-wise MAE for each test map.

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

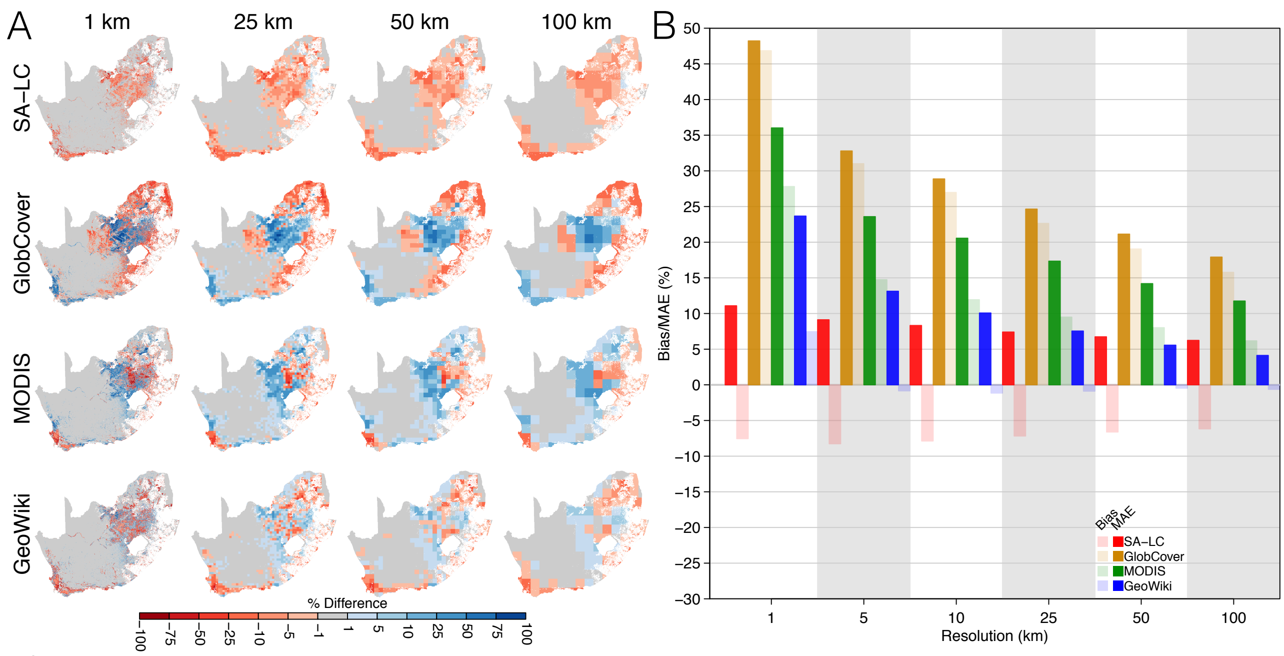

Fig. 2. (A) Errors in the percent cropland estimates resulting from each of the four test maps relative to the reference map at different scale of pixel aggregation. Rows indicate the test map being assessed (by subtraction from the reference map), while columns refer to resolution of aggregation. White 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.

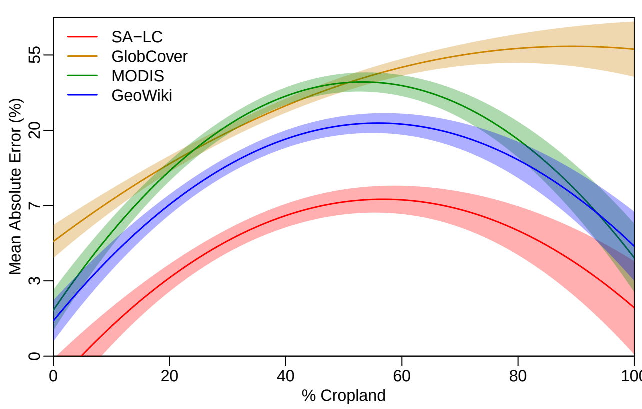

Fig. 1. The relationship between map accuracy (the mean absolute error) in test maps and the actual cropland cover within agricultural landscapes (reference map pixels having >0.5% cropland), here defined by the boundaries of magisterial districts (n = 345), as fit with a generalized additive model. Prediction curves are color-coded to the different test maps, with the solid line indicating predicted absolute bias, and the lighter shading the standard error of the coefficients.

The impact of map error on physical analyses

Estimating vegetative carbon stocks. To understand the carbon cycle and climate forcing due to land use change, it is important to have accurate, high resolution maps of vegetative carbon stocks (26). One widely used vegetative carbon dataset is that of (37), 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 (37).

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 (37)) to the non-cropland fraction of our maps. These hypothetical maps represented the range in potential carbon densities, and allowed us to investigate how carbon estimates can vary as a function of i) test map errors and ii) the properties of neighboring cover types. To assess carbon estimation error, we subtracted test map-derived carbon maps from those based on the reference map, and calculated bias and accuracy scores.

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

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

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

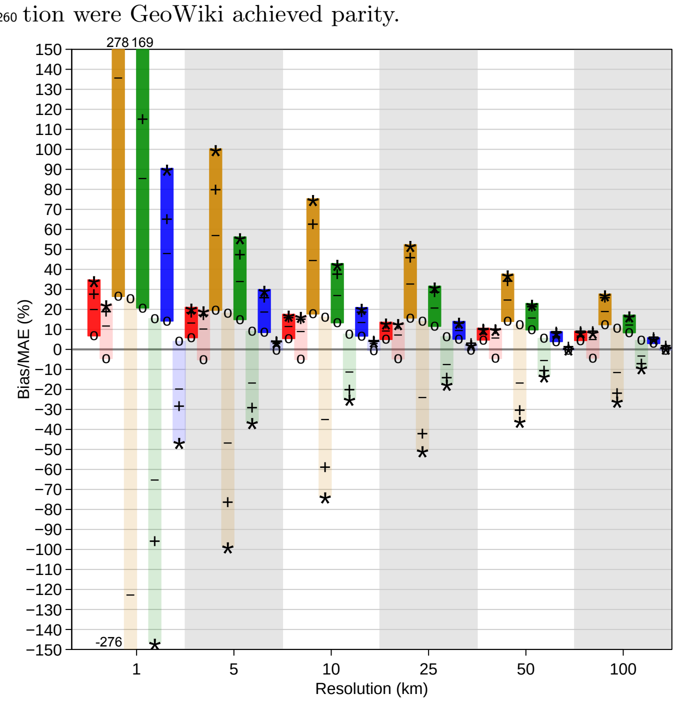

Fig. 3. Biases and accuracies (mean absolute errors) of carbon densities derived from cropland maps, calculated as percents relative to the reference map. Bias estimates (represented by symbols) fall within the semi-transparent floating bars, while accuracies are contained in the solid bars. Bar colors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

Evapotranspiration estimates. Accurate estimation of hydrological fluxes is critical to understanding how land-atmosphere interactions impact the climate system and runoff (41). Land surface hydrological models, such as the Variable Infiltration Capacity (41), 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 the landcover-specific leaf area index (LAI) values 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 (27), 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.

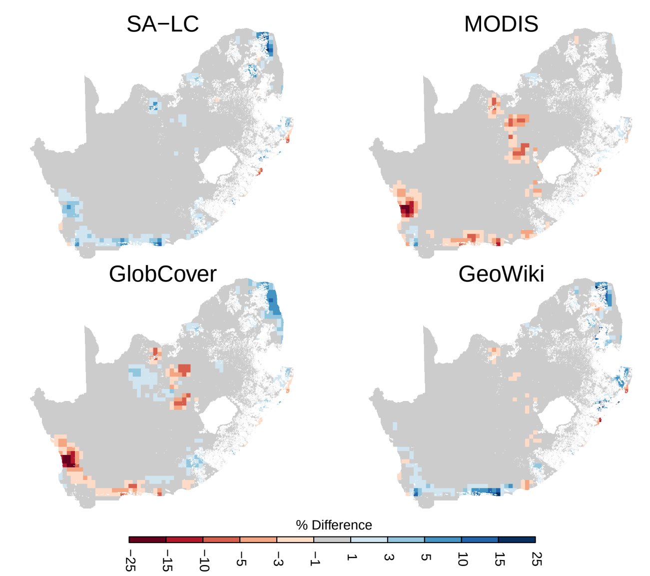

Fig. 4. Differences in annual mean evapotranspiration estimates from 29-year runs of the VIC land surface hydrology model when initialized with LAI response curves derived from the reference map, versus those from the four test maps.

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, 38). 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 (38, 39). These disaggregated datasets, which are constrained to match the agricultural statistics within the boundaries of the areas for which they are reported, have seen increasing use because they are considered to be more accurate than single source methods, and provide consistent data on which to base studies of global change (12, 39)

We used these same methods (38, 39) to disaggregate maize harvested area (South Africa's largest crop (15)) on top of our reference and cropland maps, followed by yields, which were assigned to cells have harvested areas greater than zero. We then used these two layers to calculate maize production, and further aggregated the yield and production grids to 5, 10, 25, 50, and 100 km resolutions before quantifying the bias and accuracy of each test map's yield and production values. In this case, we could not convert cell-wise errors into percentages of the reference map values (because many cells had zero values for one map but not the other), so calculated bias and accuracy from the map residuals and then normalized their values to the reference map means.

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

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

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

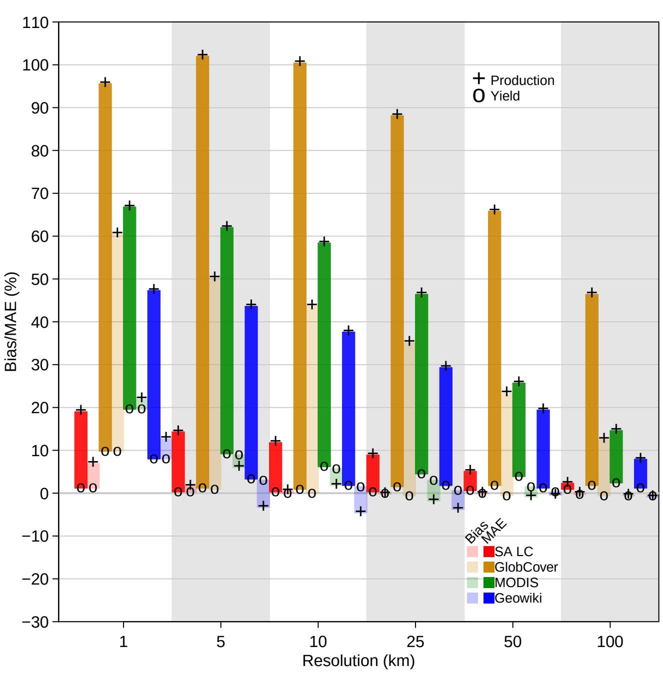

Fig. 5. Bias (mean error) and accuracy (mean absolute error [MAE]) in disaggregated maize yield and production estimates. Bias estimates (represented by symbols) fall within the semi-transparent bars, mean absolute errors in the solid bars, with bar colors coded to specific cropland maps. Symbols code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

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 (45). 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 (42). 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

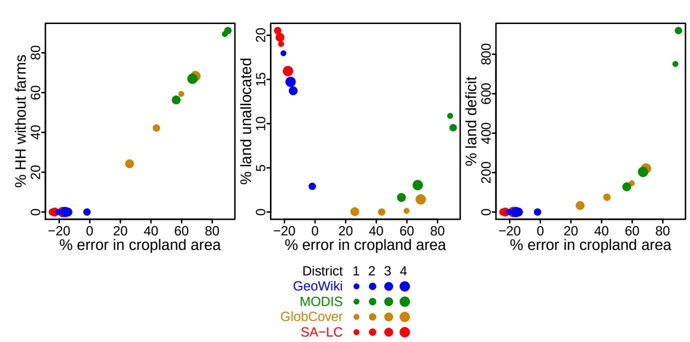

Fig. 6. Biases in agent-based model initialization relative to the district-wise errors (as a percent) in total cropland area, measured in terms of the percent of households having no cropland allocated (left), and the percent of cropland left unallocated (right). Dot sizes correspond to district numbers, colors represent the landcover map.

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 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 (16).

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 [4] 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 (17, 18). This method mirrors the ensemble averaging used by other fields (e.g. crop (46), climate (47), and ecological modeling (48)) to increase prediction confidence.

GeoWiki's statistical constraint procedure is similar to the one we used (following 39), 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 (49, 50). 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)

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

- 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 (14)
- 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 (8, 26), 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 (51, 52).
  - 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/non-cropland 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 (42). 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.

The South African government commissioned a whole-country cropland boundary map to enhance its annual collection of agricultural statistics (29). 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 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.

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 (37). The second was maize yield maps derived by disaggregating district-scale agricultural census data for both maize yield and harvested area (following 38, 39), 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. 5, 40).

For the first of these, we used the Variable Infiltration Capacity (41) 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 (42).

The reference dataset covered all of South Africa's field crop estate, from which we filtered out areas of communally owned farmlands and permanent tree crops to have a common basis of comparison across all landcover products, leaving us with map, after conversion to cropland percentage, that covered or 90% of South Africa (1,081,000 km^2), of which 104,304 km^2 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%.

from agricultural pixels ($>$0.05% cropland)

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), as the large uncropped area of South Africa (Fig. 1) dilutes the errors within the

The disaggregated yield and harvested area maps of (38) 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 39). 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 (39, 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 (43). To create these maps, we followed (38) by disaggregating district-level (n = 354, mean area = 3,445 km^2) agricultural census data (44) for maize (South Africa's largest crop by area, (43)) yield and harvested area, aggregated each set of maps, and multiplied the two to calculate production at each scale.

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. Lambin EF, Geist HJ, Lepers E (2003) Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources* 28(1):205–241.
3. Lark TJ, Salmon JM, Gibbs HK (2015) Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environmental Research Letters* 10(4):044003.
4. Wright CK, Wimberly MC (2013) Recent land use change in the Western Corn Belt threatens grasslands and wetlands. *Proceedings of the National Academy of Sciences* 110(10):4134–4139.
5. Licker R et al. (2010) Mind the gap: how do climate and agricultural management explain the yield gap of croplands around the world? *Global Ecology and Biogeography* 19(6):769–782.
6. Asner GP et al. (2010) High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences* 107(38):16738–16742.
7. Gaveau DLA et al. (2014) Major atmospheric emissions from peat fires in Southeast Asia during non-drought years: evidence from the 2013 Sumatran fires. *Scientific Reports* 4.
8. Newbold T et al. (2015) Global effects of land use on local terrestrial biodiversity. *Nature* 520(7545):45–50.
9. Luoto M, Virkkala R, Heikkinen RK, Rainio K (2004) Predicting bird species richness using remote sensing in boreal agricultural-forest mosaics. *Ecological Applications* 14(6):1946–1962.
10. Linard C, Gilbert M, Tatem AJ (2010) Assessing the use of global land cover data for guiding large area population distribution modelling. *GeoJournal* 76(5):525–538.
11. Jain M, Mondal P, DeFries RS, Small C, Galford GL (2013) Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. *Remote Sensing of Environment* 134:210–223.
12. See L et al. (2015) Improved global cropland data as an essential ingredient for food security. *Global Food Security* 4:37–45.
13. Lobell DB (2013) The use of satellite data for crop yield gap analysis. *Field Crops Research* 143:56–64.
14. Estes L et al. (2015) DIYlandcover: Crowdsourcing the creation of systematic, accurate landcover maps. *PeerJ PrePrints* 3:e1266.
15. Estes LD et al. (2013) Projected climate impacts to South African maize and wheat production in 2055: a comparison of empirical and mechanistic modeling approaches. *Global Change Biology* 19(12):3762–3774.
16. Fritz S, See L, Rembold F (2010) Comparison of global and regional land cover maps with statistical information for the agricultural domain in Africa. *International Journal of Remote Sensing* 31(9):2237–2256.
17. Fritz S et al. (2011) Cropland for sub-Saharan Africa: A synergistic approach using five land cover data sets. *Geophysical Research Letters* 38:L04404.
18. Fritz S et al. (2015) Mapping global cropland and field size. *Global Change Biology* 21(5):1980–1992.
19. Verburg PH, Neumann K, Nol L (2011) Challenges in using land use and land cover data for global change studies. *Global Change Biology* 17(2):974–989.
20. Fritz S et al. (2012) Geo-Wiki: An online platform for improving global land cover. *Environmental Modelling & Software* 31:110–123.
21. Kaptu Tchuent AT, Roujean JL, De Jong SM (2011) Comparison and relative quality assessment of the GLC2000, GLOBCOVER, MODIS and ECOCLIMAP land cover data sets at the African continental scale. *International Journal of Applied Earth Observation and Geoinformation* 13(2):207–219.
22. Ge J et al. (2007) Impacts of land use/cover classification accuracy on regional climate simulations. *Journal of Geophysical Research: Atmospheres* 112(D5):D05107.
23. Schmit C, Rounsevell MDA, La Jeunesse I (2006) The limitations of spatial land use data in environmental analysis. *Environmental Science & Policy* 9(2):174–188.
24. Quaife T et al. (2008) Impact of land cover uncertainties on estimates of biospheric carbon fluxes. *Global Biogeochemical Cycles* 22(4):GB4016.
25. Debats S, Luo D, Estes L, Fuchs T, Caylor K (year?) A generalized computer vision approach to mapping agricultural fields in Sub-Saharan Africa. *Remote Sensing of Environment*.
26. Searchinger TD et al. (2015) High carbon and biodiversity costs from converting Africa's wet savannahs to cropland. *Nature Climate Change* 5(5):481–486.
27. 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.
28. Estes LD et al. (2014) Using changes in agricultural utility to quantify future climate-induced risk to conservation. *Conservation Biology* 28(2):427–437.
29. Fourie A (2009) Better Crop Estimates in South Africa. *ArcUser Online* (1).
30. SANBI (2009) National Landcover 2009, (South African National Biodiversity Institute; National Department of Environmental Affairs and Tourism, Pretoria, South Africa), Technical report.
31. 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.
32. 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.
33. Arino O et al. (2012) *Global land cover map for 2009 (GlobCover 2009).* (European Space Agency & Universit Catholique de Louvain).
34. 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.
35. Gross D et al. (2013) Monitoring land cover changes in African protected areas in the 21st century. *Ecological Informatics* 14:31–37.
36. 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.
37. 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.*
38. 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.
39. 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.
40. Foley JA et al. (2011) Solutions for a cultivated planet. *Nature* 478(7369):337–342.
41. 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.
42. Chen P, Plale B, Evans T (2013) *Dependency Provenance in Agent Based Modeling.* pp. 180–187.
43. 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.
44. Africa SS (2007) Commercial Census of Agriculture, South Africa.
45. Berger T, Schreinemachers P (2006) Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society* 11(2):19.
46. Asseng S et al. (2013) Uncertainty in simulating wheat yields under climate change. *Nature Climate Change* advance online publication.
47. 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.
48. Arajo M, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution* 22(1):42–47.
49. Carletto C, Jolliffe D, Banerjee R (2013) The emperor has no data! Agricultural statistics in Sub-Saharan Africa. *World Bank*.
50. 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).
51. Estes LD et al. (2014) Changing water availability during the African maize-growing season, 19792010. *Environmental Research Letters* 9(7):075005.
52. 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.